UNIVERSITY OF CALIFORNIA,
IRVINE

The natural history of an abundant soil bacterium: Niche differentiation and adaptation of *Curtobacterium*

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Alberto Barron Sandoval

Dissertation Committee:
Professor Jennifer B. H. Martiny, Chair
Professor Adam C. Martiny
Professor Kailen A. Mooney

2026

DEDICATION

To

my parents, family and friends

in recognition of your unconditional support and love

"I do not study to write, much less to teach (which in me would be excessive pride), but only to see if by studying I might be less ignorant. This is my answer and this is how I feel."

Sor Juana Inés de la Cruz

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGMENTS

Thank you to my advisor, Jennifer Martiny, for your mentorship, support, patience and care. You have changed the way I look to the world, and I can say that it has been for the better. You are an inspiring person, and I will always be grateful for the opportunity and experience of working with you.

Thank you also to my committee members, Dr. Adam Martiny and Dr. Kailen Mooney for always showing enthusiasm about my work. Thank you for the critical feedback and guidance, your contributions made me a better ecologist and improved this dissertation.

To the members of the J. Martiny lab, thanks for being amazing labmates, always supportive and with accurate feedback that improved my work. Also, thanks to members of the Allison, Treseder, A. Martiny, Rodriguez-Verdugo, Rebolleda-Gomes and Hammer labs that contributed to this work with feedback, ideas and help with lab and field work. It was so fun sharing these years with you all.

Financial support was provided by the National Science Foundation (DEB-2113004) awarded to JBHM. A Mildred E. Mathias Graduate Student Research Grant from the UC Natural Reserve System awarded to ABS as well as the University of California, Irvine Graduate division for the Miguel Velez scholarship awarded to ABS

VITA

Alberto Barrón Sandoval

EDUCATION

2026 Ph.D., in Biological Sciences, University of California, Irvine (UCI)
Advisor: Jennifer Martiny

2023 M.S., Biological Sciences, University of California, Irvine (UCI)

2020 M.S., Biological Sciences, Universidad Nacional Autónoma de México (UNAM)
Advisor: Ana E. Escalante

2015 B.S., Biology, Universidad Nacional Autónoma de México (UNAM)
Advisor: Ana E. Escalante

FIELD OF STUDY

Ecology and Evolutionary Biology

PUBLICATIONS

Abbruzzini, T.F.; Avitia, M.; Carrasco-Espinosa, K.; Peña, V.; Barrón-Sandoval, A.; Salazar Cabrera, U.I.; Cruz-Ortega, R.; Benítez, M.; Escalante, A.E.; Rosell, J.A.; Wegier, A.; Campo, J. 2021. Precipitation Controls on Soil Biogeochemical and Microbial Community Composition in Rainfed Agricultural Systems in Tropical Drylands. Sustainability 13, 11848.

Avitia, M., Barrón-Sandoval, A., Hernández-Terán, A., Benítez, M., Barron-Gafford, G.A., Dontsova, K., et al. & Escalante, A.E. 2021. Soil microbial composition and carbon mineralization are associated with vegetation type and temperature regime in mesocosms of a semiarid ecosystem. FEMS Microbiology Letters 368(4), fnab012.

Barbour, K.M., Barrón-Sandoval, A., Walters, K.E. & Martiny, J.B.H. 2022. Towards quantifying microbial dispersal in the environment. Environmental Microbiology. doi:10.1111/1462-2920.16270

Barrón-Sandoval, A., Martiny, J.B.H., Pérez-Carbajal, T., Bullock, S.H., Leija, A., Hernández, G., Escalante A.E. 2023. Functional significance of microbial diversity in arid soils: biological soil crusts and nitrogen fixation as a model system. FEMS Microbiology Ecology 99(2), fiad009.

Carrasco-Espinosa, K.; Avitia, M.; Barrón-Sandoval, A.; Abbruzzini, T.F.; Salazar Cabrera, U.I.; Arroyo-Lambaer, D.; Uscanga, A.; Campo, J.; Benítez, M.; Wegier, A.; et al. 2022. Land-Use Change and Management Intensification Is Associated with Shifts in Composition of Soil Microbial Communities and Their Functional Diversity in Coffee Agroecosystems. Microorganisms 10, 1763.

Hacopian, M.T., Barrón-Sandoval, A., Romero-Olivares, A.L., Berlemont, R., and Treseder, K.K., 2025. Warming Is Associated With More Encoded Antimicrobial Resistance Genes and Transcriptions Within Five Drug Classes in Soil Bacteria: A Case Study and Synthesis. Environmental Microbiology 27(4): e70097.

López-Lozano, N.E., Escalante, A.E., Barrón-Sandoval, A., Perez-Carbajal, T. 2018. Terrestrial N Cycling in an Endangered Oasis. In: García-Oliva F., Elser J., Souza V. (eds) Ecosystem Ecology and Geochemistry of Cuatro Ciénegas. Cuatro Ciénegas Basin: An Endangered Hyperdiverse Oasis. Springer, Cham.

Fellowships

2025 Miguel Velez Scholarship — UC Irvine Graduate Division, $11,000
2021 Mildred E. Mathias Graduate Student Research Grant —UC Natural Reserves System, $3,500
2018 – 2020 CONACYT Graduate Fellowship for Graduate Students — 2 years full stipend, ~$12,000

ABSTRACT OF THE DISSERTATION

The Natural History of an Abundant Soil Bacterium: Niche Partitioning and Local Adaptation of *Curtobacterium*

by

Alberto Barrón Sandoval

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2026

Professor Jennifer B.H. Martiny, Chair

Microbial communities play an essential role in sustaining ecosystem processes and life on Earth. As such, the field of microbial ecology has gradually shifted from establishing taxonomic patterns of distribution to identifying the traits that dictate where different microorganisms live to isolating the eco-evolutionary processes that create and maintain these patterns. Yet our understanding of these matters still largely reflects broad taxonomic levels. For soil bacteria in particular, most studies rely on highly conserved markers (e.g., the 16S rRNA gene). However, genomic variation at these scales overlooks substantial trait divergence that is potentially ecologically relevant. As a result, our understanding of how closely related lineages, akin to bacterial species, partition the soil environment and assemble into communities remains limited. In this dissertation I aimed to address this gap by using the *Curtobacterium* genus, an abundant bacterial taxon in soil leaf litter, as a model organism and ask: 1) How does the distribution of different lineages of *Curtobacterium* change along environmental gradients? 2) Does the phenotypic characterization of representative isolates of these lineages align with the

biogeographic patterns observed? and 3) Are the biogeographic and phenotypic patterns observed a result of local adaptation to combinations of environmental conditions?

In the first section of this dissertation, I addressed the first two questions by conducting a large-scale survey on 24 locations across California to capture a wide environmental gradient in terms climate, ecosystems and vegetation types. I collected both grass and dominant-vegetation litter at each site to decouple substrate effects from climate, and isolated *Curtobacterium* strains from these samples for phenotypic characterization. Using metagenomic sequencing I characterized microbial communities, and applied random forest modeling, distance-based ordination, and lab-based thermal and pH performance curves to link lineages distributions to environmental drivers. Overall *Curtobacterium* abundance was primarily predicted by both litter chemistry (cellulose content) and climate. The four most abundant ecotypes (lineages) exhibited distinct biogeographic patterns: ecotype IIIA was more abundant in cooler, wetter sites with lignin-rich litter, while ecotypes IC and IVB were associated with warmer, drier conditions, and ecotype IIG showed broader thermal tolerance. These field patterns were corroborated by phenotypic data, with ecotype IC showing significantly higher temperature and pH optima than IIIA. These results demonstrate that Curtobacterium ecotypes undergo niche partitioning shaped by both climate and litter chemistry and suggest that accounting for heterogeneity in the soil substrate can reveal underlying climate signal.

To test whether these biogeographic and phenotypic patterns reflect adaptation, in the second section of this dissertation I addressed the third question by conducting a replicated reciprocal transplant anchored at a focal site in southern California and replicated across a subset of eight sites from the initial survey. I manipulated site (as a proxy of climate), leaf litter substrate and inoculum community origin, to tease apart the effect of these three factors on

compositional responses and asses the adaptive response of individual ecotypes to the experimental manipulations. Transplanted communities shifted in composition toward that of the native away community, with both climate and litter substrate chemistry independently driving convergence. These patterns were consistent across two phylogenetic scales: the whole bacterial community and the genus *Curtobacterium*, for which independent biogeographic and phenotypic evidence supports ecotypic differentiation along the same environmental axes. Specifically, the four most abundant *Curtobacterium* ecotypes shifted in relative abundance across transplant paths in directions predicted by their thermal preferences and biogeographic distributions, suggesting that community-level convergence reflects the sorting of evolutionarily differentiated lineages rather than stochastic assembly. Together, these results demonstrate that leaf litter bacterial communities are locally adapted to both climate and litter substrate chemistry and establish community-level convergence as a metric for detecting local adaptation in microbial systems where individual-level approaches remain infeasible. As microbial communities underpin critical ecosystem processes including decomposition and nutrient cycling, understanding the eco-evolutionary mechanisms that structure their diversity across environmental gradients has important implications for predicting how these processes will respond to ongoing environmental change.

INTRODUCTION

One of the main goals of ecology is to describe and understand how species are distributed in space and time, and the processes that drive these distributions. Historically, this endeavor was called the study of the natural history of organisms, and naturalists would observe and describe where a species lives and how it interacts with its surrounding environment. In 1807, Alexander von Humboldt published a cross-section drawing of a volcano. The *Naturgemälde* depicted the plant species along the Chimborazo elevational gradient, arranged by the elevation at which they grew, annotated with the temperature, humidity, and atmospheric pressure of each zone. Its relevance was not the catalogue of species but the relationship it asserted, that the distribution of life is structured by the physical environment. The questions that follow from that assertion: 1) where an organism lives, 2) what allows it to persist there, and 3) how that correspondence arose, have organized ecological and evolutionary research ever since.

For much of the following century, these questions were addressed through observation and comparison. It was not until 1947 that the foundational work by Clausen, Keck and Heisey changed the observational approach to an experimental one (Clausen et al., 1947). By working with North American plants, they moved individuals among sites differing in elevation and climate to compare their performance at home and away sites. This experiment demonstrated that populations of the same species differ heritably in ways corresponding to the climatic conditions of their site of origin. This pattern is what is known as local adaptation and is caused by differential selection regimes that favor different genotypes and phenotypes, such that some organisms outperform others under contrasting conditions. Where selection is sufficiently divergent and gene flow sufficiently limited, populations diverge, and the resulting differentiation is a fundamental process in the generation and maintenance of biodiversity

(Gavrilets, 2003). Reciprocal transplants remain the standard for testing local adaptation in the field (Johnson et al., 2022) and allow researchers to distinguish adaptation from other ecological processes such as dispersal limitation or drift.

Microorganisms represent an important gap in the study of natural history. They constitute the second largest reservoir of biomass on Earth after terrestrial plants (Bar-On et al., 2018), and their metabolism is responsible for the biogeochemical cycles that sustain all other life forms (Falkowski et al., 2008). Yet their natural history remain largely unresolved in comparison with their macrobial counterparts, even though this information is essential for understanding how they adapt in field settings (Kraemer & Boynton, 2017).

There are historical and technical reasons that explain this knowledge gap. Historically, the formulation that 'everything is everywhere, but the environment selects' assumes that microorganisms are not dispersal limited. Thus, all observed patterns of distribution are a result of contemporary deterministic environmental filtering. Nonetheless, there is substantial evidence inconsistent with those assumptions, showing that microorganisms (like macroorganisms) show biogeographic patterns (Martiny et al., 2006; Hanson et al., 2012). Yet demonstrating that distributions are structured is not the same as explaining why, and the second and third natural history questions (what traits underlie those distributions, and whether the correspondence reflects adaptation) have received far less attention. Technically, it is not possible to track individual microbes with current technologies in the field, and most of what is known about microbial distributions derives from highly conserved marker genes such as the 16S rRNA gene. These resolve broad taxonomic groupings but obscure genomic and phenotypic variation among closely related lineages, even though it is often at those finer scales that ecologically relevant differentiation frequently resides (Martiny et al., 2023). Surveying microbial diversity at

conventional resolution amounts to describing the natural history of birds using only the presence of feathers; real distinctions are recovered, but not the ones that explain why a given organism occurs where it does.

Addressing this gap requires a unit of analysis with a phylogenetic resolution fine enough to be ecologically relevant and coherent. The ecotype concept provides one. Introduced by Turesson (1922) for plant populations differentiated by habitat, and later reformulated for bacteria as a lineage within which periodic selection maintains cohesion and between which ecological differentiation persists (Cohan, 2002), an ecotype is a phylogenetic group whose members share physiological characteristics conferring predictable advantages under particular conditions. This concept can generate testable expectations; if lineages within a genus function as ecotypes, strains within a lineage should respond consistently to environmental variation, be distributed non-randomly along environmental gradients, and differ measurably in traits relevant to those gradients. This has been well documented for *Prochlorococcus* ecotypes partitioning the light and nutrient gradients of the ocean (Johnson et al., 2006).

In this dissertation I aim to write the natural history of an abundant soil bacterium. For this purpose, I build on the findings of a well-characterized bacterial genus that has been shown to be abundant in the leaf litter layer of soils, the *Curtobacterium* genus. Previous work determined that this taxon encompasses a wide genetic variation that cannot be resolved by the 16S rRNA gene, with well-defined but closely related lineages that show differential distributions along an elevational gradient and in their phenotypic traits (Chase et al., 2017, 2018).

In the first chapter of this dissertation, I address the first two questions that natural history poses: what drives the distribution of the *Curtobacterium* lineages along environmental axes and what

phenotypic traits that underlie such distribution. In the second chapter I address whether the correspondence of distribution and traits is due to adaptation with a reciprocal transplant experiment.

CHAPTER 1: Niche partitioning of soil bacterial ecotypes by climate and plant litter chemistry

ABSTRACT

Over the last two decades, microbial biogeography has well characterized distribution patterns at broad taxonomic levels in soil environments. However, how fine-scale genetic clades partition their environment in soil is far less understood. Here, we investigated the biogeography of the Curtobacterium genus, a model soil bacterium with defined ecotypes, to test whether niche partitioning is driven by climate, local litter chemistry, or their interaction. We sampled leaf litter at 24 sites spanning a wide climate gradient across California, collecting both grass and dominant-vegetation litter at each site to decouple substrate effects from climate. Using metagenomic sequencing we classified reads into 25 subclades (ecotypes) and applied random forest modeling, distance-based ordination, and lab-based thermal and pH performance curves to link ecotype distributions to environmental drivers. Overall Curtobacterium abundance was primarily predicted by both litter chemistry (cellulose content) and climate. In contrast, alpha and beta diversity were most strongly structured by the temperature–precipitation gradient. The four most abundant ecotypes exhibited distinct biogeographic patterns: ecotype IIIA was more abundant in cooler, wetter sites with lignin-rich litter, while ecotypes IC and IVB were associated with warmer, drier conditions, and ecotype IIG showed broader thermal tolerance. These field patterns were corroborated by phenotypic data, with ecotype IC showing significantly higher temperature and pH optima than IIIA. These results demonstrate that Curtobacterium ecotypes undergo niche partitioning shaped by both climate and litter chemistry and suggest that accounting for heterogeneity in the soil substrate can reveal underlying climate signal.

INTRODUCTION

Over the past twenty years, the field of microbial biogeography has gradually shifted from establishing taxonomic patterns (Fierer & Jackson, 2006; Martiny et al., 2006; Thompson et al., 2017) to identifying the traits responsible (Green et al., 2008; Martiny et al., 2015; Piton et al., 2023) to isolating the eco-evolutionary processes behind the patterns (Hanson et al., 2012; Nemergut et al., 2013; Barbour & Martiny, 2024). Yet our understanding of these matters still largely reflects broad taxonomic levels. For soil bacteria in particular, most studies rely on highly conserved markers (e.g., the 16S rRNA gene) or reference databases that are vastly incomplete (to classify metagenomic sequences) (Nayfach et al., 2021). However, genomic variation at these scales overlooks substantial trait divergence that is potentially ecologically relevant (Martiny et al., 2023). As a result, our understanding of how closely-related ecotypes (Cohan, 2001; Moore et al., 2002), akin to bacterial species, partition the soil environment remains limited.

In contrast to soil, bacterial biogeography at fine genetic scales has been well characterized in the marine environment, most notably for the cyanobacterium *Prochlorococcus.* The genus can be partitioned into ecotypes, defined as genetic clades with distinct trait differences (Z. I. Johnson et al., 2006; Moore et al., 1998). In the ocean, these ecotypes exhibit stark biogeographic patterns that correlate with light, temperature, and nutrient gradients, that correspond to the phenotypic differences observed in the lab (Z. I. Johnson et al., 2006; Larkin et al., 2023; Malmstrom et al., 2012). Together, these results demonstrate clear niche partitioning among *Prochloroccocus* ecotypes, and similar patterns were later found in other abundant groups of marine bacteria including *Synechococcus* and SAR11 (Carlson et al., 2009; Kent et al., 2019).

Like their marine counterparts, soil bacteria also exhibit fine-scale genetic diversity, or microdiversity, that seems to correspond to environmental preferences in either the lab and/or

field (Fuhrman et al., 1998; Larkin & Martiny, 2017; Schloter et al., 2000; Zimmerman et al., 2013). For instance, *Bacillus* ecotypes show preferences for specific soil texture and temperature regimes (Connor et al., 2010), non-symbiotic *Bradyrhizobium* ecotypes are structured by the organic and mineral soil layers (VanInsberghe et al., 2015), and *Streptomyces* lineages show differences in their thermal preferences that are related to their latitudinal distribution (Choudoir & Buckley, 2018). Nonetheless, we still lack a thorough account of the biogeography of soil bacterial ecotypes of any genus as we do in the ocean. Ideally, such an example would include: (i) a large-scale, systematic spatial sampling of genetic diversity within the genus, (ii) identification of the main drivers underlying differential ecotype distributions, and (iii) confirmation of ecotype differentiation corresponding to these drivers by physiological measurements of representative isolates.

To conduct such a study in soil, we investigated how diversity within the genus *Curtobacterium* (family Microbacteriaceae) partitions the soil environment. *Curtobacterium* is associated with live plants and also abundant in the leaf litter layer of surface soil (Chase et al., 2016). Isolates from the genus have been classified into five phylogenetic clades (I-V) that can be further subdivided into many more subclades with distinct phenotypic traits (Chase et al., 2017). The subclades (AAI >90%) have therefore been proposed to be ecotypes.

One challenge in understanding the biogeographic patterns of soil bacterial diversity is the inherent heterogeneity of the soil environment. For instance, while climate varies predictability across terrestrial ecosystems, temperature and moisture are much more heterogeneous at small (micrometer to meter) scales than in seawater, potentially obscuring biogeographic patterns. Similarly, while distinct soil categories can be mapped over large geographic scales, soil properties also vary greatly at small scales depending on the local

vegetation and disturbance regimes. Moreover, soil properties and climate are often decoupled whereas in the ocean, temperature and light strongly co-vary with nutrient availability and water column stratification (Cullen et al., 1982) despite microscale heterogeneity (Stocker, 2012). As a result, characterizing the niche distributions of soil bacterial ecotypes requires sampling strategies that span multiple spatial scales, integrating both the broader climatic context and microhabitat conditions. Indeed, in an initial survey of five sites along a regional climate gradient in southern California, *Curtobacterium* subclades varied in abundance, but their geographic distribution was not clearly related to environmental variables (Chase et al., 2018). However, a subsequent transplant experiment that held the litter substrate constant revealed that the subclades may be adapted to different climates (Chase et al., 2021). We therefore hypothesize that *Curtobacterium* ecotypes partition their niche along both climate and soil chemistry axes. Specifically, we expect that litter chemistry explains a large part of the variation in ecotype distribution that might otherwise mask the influence of climate on the biogeographic patterns.

To address our hypothesis and the challenge of soil heterogeneity, we sampled *Curtobacterium* using a nested sampling design. Within 24 locations across California, we collected samples from two distinct types of plant litter: (1) litter of the dominant vegetation at the site and (2) litter of grasses, some species of which were found at every site. This design allowed us to compare bacteria associated with two distinct litter chemistries experiencing the same climate conditions and to compare bacteria associated with similar litter chemistry (grasses) across a wide range of climate conditions. In this way, we aimed to tease apart these two niche axes that often covary in soil biogeography studies. Because the 16S rDNA region cannot resolve *Curtobacterium* ecotypes, we assessed *Curtobacterium* composition using metagenomic sequencing. We also isolated *Curtobacterium* from each site to build a large

genomic database to better classify the metagenomic sequences within the genus. We then used random forest modelling to identify the relative influence of climate and litter chemistry variables on ecotype distribution, accounting for potential non-linear relationships. Finally, we assayed temperature and pH preferences of *Curtobacterium* isolates in the lab to test the biogeographic predictions of an ecotype's niche preferences. By integrating biogeographic, phylogenetic, and physiological information, we were able to identify patterns of niche differentiation among bacterial ecotypes in this soil genus.

MATERIALS AND METHODS

Field sites and litter collection

We sampled leaf litter at 24 sites across California between November 2021 to July 2022. All sites were located within reserves managed by the University of California Natural Reserve System (UCNRS) (Supplementary table 1 – Sample metadata and coordinates). The 24 sites comprise a wide range of climatic variation and vegetation types. At each site we established two 10m × 10m plots, one primarily covered by grass(es) and the other by the dominant vegetation at the site and recorded the plant species observed within the plots. Within each plot, we collected ~5 g of surface leaf litter at the corners of the plots for a total of 4 samples for each of the two litter types (grass and dominant vegetation) at each site, for a total of 188 samples. For some of the downstream analyses, we pooled the data from samples of the same reserve and litter type to match the number of samples that were analyzed for litter chemistry (n= 94, see below).

Climate and litter chemistry variables

To characterize climate variables at each site, we retrieved the monthly averages of the maximum and minimum temperature and total precipitation for 1991-2020 for the coordinates

for each plot with a resolution of 800m^2 from the PRISM climate group data set (PRISM Group and Oregon State University, https://prism.oregonstate.edu, accessed 26 Feb 2024). With these data, we calculated the 19 bioclimatic variables as defined by WorldClim using the 'dismo' package (Hijmans et al., 2010) in R.

To characterize the biotic variables at each site, we measured pH, and carbohydrate composition of the plant litter samples. All samples were finely ground with coffee grinders (KitchenAid coffee grinder, give specs) before further analysis. pH measurements were performed as described in (Cornelissen et al., 2006) with slight modifications. Briefly, 1g of ground litter was added to 32ml of deionized water (1:32 volume ratio) in a 50ml falcon tube and shaken to 250 rpm for 60 min, then centrifuged at 4500 rpm for 5 min. The supernatant was transferred to a clean urine cup and pH was measured with a pH meter.

Oven-dried collected litter was sent to Cumberland Valley Analytical Services (Waynesboro, PA, USA) for near-IR spectroscopy to quantify non-ash organic compounds including cellulose (acid detergent fiber − lignin), hemicellulose (neutral detergent fiber − acid detergent fiber), lignin, crude protein, and structural carbohydrates (non-fiber carbohydrates − non-structural carbohydrates). These values are reported as percentage of the dry litter weight. We pooled pairs of samples of the same site and litter type for a final sample size of 94 samples. To assess differences in litter chemistry across ecosystems and litter types, we performed a PERMANOVA test with type III sum of squares under a reduced model with a centered and scaled Euclidean distance matrix, implemented in the PRIMER+ v6 software (Anderson, Gorley, and Clarke 2008).

Isolation and identification of Curtobacterium strains

To isolate *Curtobacterium* strains, we vortexed 0.2 g of ground litter in 5 ml of 0.9% saline (NaCl) solution for 5 min. Samples were serially diluted and plated on grassland leaf litter leachate media from two different sources, Loma Ridge grass litter and the grass litter collected at each site (Chase et al., 2016). Colonies were visually screened for phenotypic characteristics ascribed to *Curtobacterium* (Evtushenko & Takeuchi, 2006), streaked on Luria Broth (LB) media agar plates, transferred three times, and stored in glycerol solution at −80 C.

We confirmed *Curtobacterium* strains by whole genome sequencing. A single colony was transferred into 3 ml of liquid LB media, grown for 48-72 hours, and centrifuged to form a pellet. DNA extraction was performed with the PureLinkTM Pro 96 Genomic DNA purification kit (Invitrogen). Extracted DNA was quality assessed on a Nanodrop (Thermo Fisher; Waltham, MA) and diluted to 0.5 ng/µl for library preparation. Sequencing libraries were prepared using the Nextera XT Library Preparation kit (Illumina Inc., San Diego, CA). Samples were pooled in equimolar portions and assessed using the High Sensitivity Bioanalyzer. The pooled library was sequenced using an Illumina HiSeq4000 instrument (Illumina Inc., San Diego, CA) with 150 bp paired-end reads.

Demultiplexed sequence data were assembled using the SPAdes genome assembler (Bankevich et al., 2012) with a 'careful' iterative k-step ranging from k = 31 to 111. We assessed the quality of the assemblies by creating taxon annotated-GC-coverage (TAGC) plots using 'blobtools' (Challis et al., 2020). Specifically, we calculated coverage for each contig by mapping back the raw sequence data to assembled contigs using Bowtie2 (Langmead & Salzberg, 2012) and taxonomic assignments were done using MegaBLAST against the NCBI nucleotide database (Morgulis et al., 2008) with an e value of 1×10^{-5}. Based on these results we discarded contigs with coverage < 30, length < 500 bp and GC% < 55%. Additionally, genomes

with high number of contigs assigned to different genera were removed from further analysis. Next, we calculated the BUSCO score (Simão et al., 2015) for genome completeness and assessed contamination with CheckM (Parks et al., 2015). Genomes with <90% completeness and >5% contamination were also removed from further analysis. In total, we identified 259 high-quality *Curtobacterium* genomes that were further used to build a phylogenetic tree and the reference data base.

We created a *Curtobacterium* phylogeny using a multilocus sequence alignment. We included the newly isolated *Curtobacterium* genomes (n = 246), previously isolated and identified genomes from our lab (n = 70), and genomes downloaded from NCBI (n = 72) as well as a *Frigoribacterium* genome (RefSeq accession number: GCF_001424645.1) to serve as an outgroup. Each genome was translated using Prodigal (Hyatt et al., 2010) and screened for the presence of 21 marker genes (Chase et al., 2018) using HMMER with an E value of 1×10^{-10} (Finn et al., 2011). Each marker gene was aligned independently using ClustalO v.1.2 (Sievers et al., 2011) to create a concatenated protein alignment of 3971 aminoacids for phylogenetic analysis using RAxML v.9 (Stamatakis, 2014) under the PROTGAMMAWAG model for 100 replicates. We designated the five major clades within the genus as previous observed (Chase et al., 2018). Next, we identified finer taxonomic groupings by calculating pairwise average amino acid identity (AAI) across all genomes using 'fastAAI' (Konstantinidis et al., 2022). Genomes that clustered into subclades ≥ 90% AAI at the whole genome level were designated ecotypes (Chase et al., 2018). All subclade designations were supported by the phylogeny (Fig. S1.5).

Metagenomic sequencing and data processing

To characterize *Curtobacterium* community composition, we sequenced metagenomes from each litter sample. We extracted DNA from 0.05 g of ground litter using the ZymoBiomics

DNA kit (Zymo Research, Irvine, CA). Extracted DNA was diluted to 0.5 ng/µl and 1ng of DNA was used as input for the Nextera XT library Prep kit for sequencing on the Illumina HiSeq 4000 instrument with 150 bp paired-end reads at the Genomics & Research Technology Hub at the University of California, Irvine. Raw reads were quality filtered and adapters were removed with bbduk from the BBmap suite (Bushnell, 2016). We remove eukaryotic DNA by mapping quality filtered reads to reference genomes of an abundant grass (*Lolium perenne*; Accession: MEHO01000000) and fungus (*Pyrenophora teres*; Accession: NZ_AEEY00000000) using BWA (Li & Durbin, 2009). The remaining reads were then merged using BBmap to form paired end reads. If a read could not be merged merge with its counterpart, then only the forward read was included for further analysis. Paired-end shotgun metagenomic reads were deposited under NCBI BioProject PRJNA1499846.

Merged metagenomic reads were searched against a custom reference database of conserved single-copy marker genes from 388 *Curtobacterium* genomes plus 81 genomes of closely-related genera abundant in leaf litter (*Agreia*, *Bacillus*, *Clavibacter*, *Frigoribacterium*, *Frondihabitans*, *Herbiconiux*, *Leifsonia*, *Leucobacter*, *Microbacterium*, *Mycobacterium*, *Pseudoclavibacter*, *Pseudomonas*, *Rathayibacter* and *Sphingomonas*) that served as outgroups. To create the database, we screened the genomes for the orthologous protein groups identified previously by Chase et al. (2018). Briefly, assembled genomes were translated to protein sequences using Prodigal (Hyatt et al., 2010) and searched for the presences of the single-copy marker genes with reference HMMER profiles using HMMER with an *E* value of 1×10^{-20} (Finn et al., 2011). Next, we dereplicated gene sequences with CD-HIT (Fu et al., 2012) to avoid ambiguous read classifications, resulting in a total of 373,619 sequences representing 1075 single

copy marker genes. A blast-format data base was then constructed using the NCBI blast toolkit (Camacho et al., 2009).

To search against the custom database, we used HS-BLAST (Chen et al., 2015) using an *E* value of 1×10^{-50}, keeping only the best hit for each query sequence. The BLAST output files were parsed with in-house scripts to obtain total read counts for each genome in the database for each sample. To account for differences in the number and length of marker genes in the genomes, we normalized the total read counts for each genome in the database by the total base pairs (sum of the length of all genes) represented in the database. We report these values as read count per million bases mapped.

To calculate the relative abundance of the *Curtobacterium* genus within the whole bacterial community, we used Kraken2 (Wood et al., 2019) to classify metagenomic reads against the GTDB database (version R10-RS226, Parks et al., 2026) and absolute read count per taxa and sample were calculated with Bracken (Lu et al., 2017). Clade and subclade abundances in each sample were calculated by summing the read counts assigned to genomes within each clade and subclade.

Analysis of Curtobacterium community diversity and composition

To account for differences in sequencing depth, metagenomic read counts were rarefied to 1,900 reads per sample with 100 iterations. Alpha diversity metrics (Shannon index) were calculated on each rarefied replicate using the vegan R package (Oksanen et al., 2018) and reported as the mean across iterations. For beta diversity, a Bray–Curtis dissimilarity matrix was computed for each rarefied replicate and averaged across 100 iterations to produce a Bray–Curtis dissimilarity matrix for downstream analyses.

We tested for the effects of site and litter type on community composition using PERMANOVA (type III sums of squares; 9,999 permutations) with site included as a fixed effect and litter type treated as a random effect. To assess the influence of environmental predictors, we performed a distance-based redundancy analysis (dbRDA) on the averaged Bray–Curtis dissimilarity matrix. Explanatory variables included two climatic axes (PC1 and PC2 derived from 19 bioclimatic variables; variables were scaled and centered prior to PCA) and five leaf-litter chemistry predictors (pH, cellulose, hemicellulose, lignin, crude protein). We quantified the contribution of each predictor to the constrained and total variance using ANOVA-like permutation tests (9,999 permutations) implemented in vegan. Species (subclade) scores were added to the dbRDA plot as the abundance-weighted mean of the sample (site) scores, with weights given by the Hellinger-transformed relative abundance of that subclade across samples (i.e. the square root of within-sample proportional abundance), added post-hoc with vegan::sppscores()), which marks the weighted centroid of the samples in which the subclade is most abundant. PERMANOVA was performed with PRIMER+ software (version 6 (Clarke, 1993). All other analyses were performed in R (version 4.5.0; packages: vegan 2.7.1 (Oksanen et al. 2018)).

We used regression trees along with partial dependance plots (PDPs) to assess the relationship between climate PC axes and litter-chemistry measurements and *Curtobacterium* abundance (relative to all bacteria), alpha diversity (Shannon index), and subclade-level relative abundances. We fit 5000-tree regression forests in R with the randomForest package (Liaw & Wiener, 2002). First, we assessed overall model accuracy via out-of-bag (OOB) RMSE and its derived pseudo-R^2. We tested the significance of the cross-validated R^2 by permuting the response 1000 times under a 10-fold scheme using the A3 package (Fortmann-Roe, 2015),

yielding both an observed CV R^2 and a permutation *p*-value. Predictor significance was evaluated two ways: 1) multivariate OOB importances via rfPermute (Archer, 2011), which permutes each variable 500 times in turn and measures the resulting increase in OOB RMSE. The p-value for each variable was calculated as the proportion of permuted OOB RMSE meeting or exceeding the observed value. And 2) unit-scaled (root mean squared error) RMSE importances via the 'iml' package (Molnar et al., 2018), which repeated permutation of each predictor 100 times to compute mean ΔRMSE ± 95 % CI. Finally, we generated bootstrapped partial-dependence plots (100 bootstrap refits on a fixed grid) to visualize each significant predictor's marginal effect using the 'iml' package. To complement the Random Forest analysis and assess whether monotonical predictor-response relationships are consistent under a parametric framework, we fitted Gaussian generalised linear models (GLMs) for each clade and subclade independently, using the same response variables and predictors. Model performance was evaluated using 10-fold cross-validation R^2 and an overall F-test for model significance. Variable significance was determined from coefficient-level t-tests ($\alpha = 0.05$).

Thermal and pH performance curves of Curtobacterium isolates

Isolates were streaked onto LB agar plates from glycerol stocks and incubated at 28°C until colonies were visible. A single colony was picked and transferred into liquid LB and incubated for 48 hours at 28°C on a shaker at 200 rpm. Once liquid culture was dense enough it was transferred to M63 minimal media pH 6.8 on a ratio of 1:1000 and grown at 28°C for 48 hours. For both assays, culture density was standardized to a starting optical density of 0.1 in 100 ul.

To assess thermal performance, each isolate was inoculated in triplicate into 384 well plates with M63 minimal media (pH 6.8) (see supplementary methods). 10 replicate plates were

incubated at 10 temperatures (4, 9, 16, 22, 28, 30, 34, 40, 43 and 46°C) for 48 hours. Optical density (OD) was read every 2-3 hours on a plate reader (Biotek).

To assess pH performance, 24 medias were created by adjusting pH from 3.2 to 10.1 in 0.3 increments (see supplementary materials). The medias were loaded onto each 384 well plate such that each isolate was grown on all 24 media types on a single plate. The plates were incubated at 28ºC in the plate reader for 48 hours, and OD was automatically recorded every 15 minutes. Three replicates of each isolate were distributed across different plates.

We analyzed the growth of each isolate across the temperature and pH gradients by calculating the maximum growth rate (h^{-1}) and yield (maximum OD) with the package 'growthcurveR' (Sprouffske & Wagner, 2016). The OD of the blanks (media only) was first subtracted from the OD at each time point. Growth rates and yield was set to zero (0) if the yield did not exceed that of the blanks. Average growth rate and yield was calculated for the three technical replicates. Strains were excluded from the analysis if over half of the replicates were assigned a growth rate of zero.

To fit thermal performance curves, we used the rTPC package in R (Padfield et al., 2021) with growth rate as the metric of performance. We assessed the performance of six different models using AICc scores. We selected the Sharp-Schoolfield model (Schoolfield et al., 1981) as it was the best fitting model for most isolates. For each fitted curve, we extracted four parameters using the rTPC package: optimal temperature, lower and upper critical temperatures and breadth (defined as the range over which growth rate is >50% of the maximum growth rate).

To fit pH performance curves, we used yield as the metric of performance. At each pH level, we averaged the yield of the three independent replicates and fitted a loess-smoothed line for plotting. Optimal pH was defined as the pH at which isolates showed the largest yield,

maximum and minimum pH was defined as the pH level at which growth could still be detected and breadth is the range between the maximum and minimum pH. To test for differences in thermal and pH performance traits among subclades, we used Kruskal-Wallis tests. Pairwise comparisons were assessed with Dunn tests with the Benjamini–Hochberg correction.

RESULTS

Spatial variation in climate and litter chemistry

We designed this study's sampling to capture a wide degree of variation in climate and litter chemistry. Across the 24 sites, mean annual temperature (MAT) ranged from 5 to 22 ºC and mean annual precipitation (MAP) from 521 to 1187 mm (Table S1.1). As expected, climate differed significantly by ecosystem type (as defined by the UCNRS classification) (PERMANOVA: $F_{df:6} = 8.8$, $p < 0.001$) (Fig. 1.1A).

To reduce their dimensionality for subsequent analyses, we conducted an ordination of 19 climate variables. The first two PCA components explained most of the variation across the sites (Figure 1.1A). PC1 (45%) captured climatic variability. Sites with higher PC1 values experience relatively high variation in temperature, low variation in precipitation, and low isothermality (ratio of day-to-night versus summer-to-winter oscillations). In contrast, PC2 captures mean differences in temperature and precipitation (hereafter, termed the temperature/precipitation PC). Higher PC2 values experience warmer and dryer conditions and lower values, colder and wetter conditions. We use these two PCs to test the influence of climate on *Curtobacterium* abundance and composition below.

Litter chemistry also varied considerably among the samples (Fig. 1.1B), which included 41 different plant species. Litter chemistry varied significantly by litter type (PERMANOVA ~

Litter type: F = 11.8, p = 0.001). The grass litter contained higher amounts of cellulose and hemicellulose, whereas the non-grass litter samples were higher in lignin (Fig. S1.2B). In contrast, crude protein content and pH discriminated between litter samples on an orthogonal axis of variation. pH ranged greatly among the litter samples (3.75 to 7.32). Grass litter also varied significantly by ecosystem (F = 3.8, p = 0.001). Nonetheless, although pH and lignin varied orthogonally in the PCA ordination we observe a significant positive correlation (r = 0.36, Fig S1.2)

Curtobacterium abundance and alpha-diversity

Curtobacterium was the fifth most abundant bacterial genera in the leaf litter samples, averaging 2.0% of all bacteria across the samples and ranging from 0.4 - 6.7 % across the sites. The four more abundant genera were: *Sphingomonas* (5.8%), *Streptomyces* (4.8%), *Nocardioides* (2.3%) and *Hymenobacter* (2.0%). Together, the top five genera made up 17.0% of all bacteria.

Curtobacterium abundance (relative to all bacteria) was influenced both by litter chemistry and climate (Fig. 1.2), as indicated by the significant variables after the permutation test and importance values (RMSE) of the random forest model (Fig. 1.2B). Abundance increased in litter with higher cellulose content and less thermally variable sites (climate PC1) (Fig. 1.2A). This result is consistent with the pattern that *Curtobacterium* abundance tends to be higher in grass litter (Fig. S1.3A), which contained a higher concentration of cellulose than the other litter types (Fig. S1.3B).

The patterns of *Curtobacterium* alpha diversity (Shannon index) generally followed those of relative abundance, but in this case, both climate axes were stronger predictors than litter variables. Alpha diversity increased significantly with warmer and dryer climates (climate PC2) and decreased with measures of climatic variability (climate PC1) (Fig. 1.2C). Among the

significant litter chemistry variables, pH and hemicellulose emerged as the strongest predictors (Fig. 1.2D), with lower diversity observed in samples of intermediate pH and high hemicellulose content.

Biogeography of Curtobacterium ecotypes

We isolated 260 *Curtobacterium* strains across the 24 sites (Fig. S1.4) to create a marker gene database from a total of 388 *Curtobacterium* genomes (Fig. S1.5). Using this database, we classified the metagenomic reads into 25 subclades (potential ecotypes) within five larger clades (I-V) (Fig. S1.5). All subclades were detected in each sample, varying in their relative abundance across sites and between litter types (Fig. S1.6). Each of the four most abundant clades was dominated by a single subclade (Fig. S1.6). Based on previous work, we use the terms subclade and ecotypes interchangeably for these dominant subclades.

The composition of all *Curtobacterium* subclades was significantly influenced by the categorical variables reflected in the sampling design: site, litter type (grass versus dominant vegetation), and their interaction (PERMANOVA: $F = 1.83$, $p = 0.001$; Table S1.2). To quantify the variance in community composition explained by the continuous environmental variables (climate and litter chemistry variables), we conducted a distance-based redundancy analysis (dbRDA) (Fig. 1.3). Climate and litter chemistry explained 27.41% of the total variation in subclade composition, with climate explaining 10.35% versus litter variables, 17.07% (Fig. 1.3B). Climate PC2 and hemicellulose content emerged as the most influential variables, while climate PC1, lignin, and pH were not significant contributors (Fig. 1.3B).

The dbRDA also revealed that subclades IIIA and IIA generally showed opposing correlations with the explanatory variables than IC and IVB (Figure 1.3A). Thus, we aimed to disentangle the specific preferences of these four abundant ecotypes. The random forest models

also showed distinct distributions of the ecotypes, particularly contrasting IIIA against IC and IVB (Fig.1. 4). Ecotype IIIA was most abundant in wetter and colder locations (low climate PC2) and on litter with high lignin and intermediate pH, whereas ecotypes IC, IVB, and IIG were more abundant in drier and warmer locations (high climate PC2). Among litter chemistry variables, lignin had the strongest influence on ecotypes IC and IIIA, while cellulose and hemicellulose were the dominant predictors for ecotype IVB, which was most abundant on litter rich in these compounds (Fig. 1.4B). However, the cross-validated R^2 ranged from 0.08 (IIG) to 0.32 (IVB) (Table S1.3), indicating that a substantial portion of the variation in distribution patterns remained unexplained. In general, similar trends were observed using linear models rather than random forests except for the non-linear relationship between IIIA abundance and pH (Fig. S1.8).

The remaining, less abundant subclades also showed significant relationships with climate and litter chemistry variables, as evidenced by high CV-R2 values (Table S1.3). Even though the strength of these patterns is low (they do not explain a high fraction of community variation, given their low abundance), it is interesting to note that the direction of the patterns sometimes differed within the same clade (Fig. S1.7A). For example, subclade IIIB increased in warmer and drier sites (climate PC2) in contrast to IIIA. Similarly, subclades IVC and IVE increased their abundances at low litter pH, whereas IVB increased its abundance at high pH.

Phenotypic characterization of Curtobacterium ecotypes

We next characterized the performance of isolates of the four most abundant ecotypes across a range of temperatures and pH in the lab. *Curtobacterium* isolates grew at temperatures ranging from an average minimum temperature of 1.5ºC to an average maximum temperature of 41.6ºC and at pH values ranging from an average minimum pH of 3.9 to an average maximum

pH of 8.7. Generally, IC isolates performed best in the laboratory conditions, displaying the highest growth rates and yields, while IVB isolates performed the worst (Fig. 1.5A and 5B).

The abundant ecotypes displayed significant differences in their thermal and pH performance curves (Fig. 1.5, S1.9). Ecotype IC had the highest temperature optimum (39.1ºC), higher than that of ecotypes IIIA and IVB by 12.5ºC and 6.7ºC, respectively (Fig. 1.5A; pairwise tests: $p < 0.05$). In contrast, ecotype IIG showed a significantly wider temperature breadth than ecotypes IC and IIIA ($p < 0.05$). Similarly, the pH optimum of ecotype IC was 6.4, 0.76 units lower and 0.66 pH units higher than ecotypes IIG and IIIA, respectively (Fig. 1.5B, S1.9; pairwise tests: $p < 0.05$). Notably, the maximum pH of ecotype IC was 9.2, much higher than those of ecotypes IIIA and IVB (7.9 and 7.7, respectively; $p < 0.05$).

Ordinations of the performance curve parameters confirm the overall differences between ecotypes (PERMANOVA: temperature: $p < 0.001$, pH: $p < 0.01$), while highlighting within-ecotype variability (Fig. 1.5C and D). Overall, the patterns indicate that ecotype IC prefers warmer temperatures compared to ecotypes IIIA and IVB, whereas IIG is a thermal generalist. Similarly, ecotype IC is a pH generalist whereas IIG and IIIA have a narrower breadth and more acidic pH preferences.

DISCUSSION

By focusing on a widespread and abundant soil genus, we found support for the hypothesis that soil bacterial ecotypes partition niche space along both climate and litter chemistry axes. The biogeographic patterns of the dominant *Curtobacterium* ecotypes indicate that they differentiate between cool-wet versus warm-dry conditions and lignin- versus cellulose-rich substrate. Specifically, ecotype IIIA is most abundant in colder, wetter sites on lignin-rich

litter with an intermediate pH (~5-6). In contrast, ecotype IC and IVB are more abundant in warmer, drier sites, with IC preferring low-lignin litter and IVB, high-cellulose litter. However, rather than dramatic changes in the rank order of the ecotypes, relative ecotype abundance shifted gradually over space. These patterns are in stark contrast to the distribution of *Prochlorococcu*s ecotypes that change in dominance across environmental gradients in the ocean (Z. I. Johnson et al., 2006). We suspect that these more gradual biogeographic patterns are due to the high levels of soil environmental heterogeneity at small spatial scales. Such microhabitats may allow many ecotypes to coexist even as average conditions vary (Vos et al., 2013; Curd et al., 2018).

In the lab, *Curtobacterium* ecotypes differed in their thermal preferences, and these results generally supported the biogeographic patterns of the ecotypes. Isolates of ecotype IC had a dramatically higher thermal optimum than those of ecotype IIIA, mirroring the higher abundances of IC in warmer, drier sites and IIIA, in cooler, wetter ones. An informative exception is ecotype IVB. Like IC, this ecotype reached its highest abundance in warmer sites, but it had the lowest lab optimum (26.5 °C). However, the culture medium may be biasing these results. IVB grew relatively slow in the lab compared to the other ecotypes, making it difficult to characterize the ecotype's performance (Figure 5). This was further compounded by the small number of isolates available (n = 9), which itself was a result of low isolation success and poor growth in culture.

Curtobacterium ecotypes also differed in their pH preferences, suggesting niche partitioning at this phylogenetic resolution and consistent with the idea that soil bacteria specialize on a narrow pH range (Rousk et al., 2010). Indeed, the optimal pH of ecotype IIIA (~5.6) corresponded with its greater field abundance at intermediate pH. However, while ecotype

IC exhibited this highest pH optimum, its distribution was not significantly correlated with field pH. Similarly, the relative abundance of ecotypes IVB and IIG were also not significantly correlated with field pH. We can think of at least two reasons for this disconnect. First, despite distinct preferences in the lab, other factors besides pH may be much more limiting to growth in the field. Second, our field measurements may not reflect the pH experienced by *Curtobacterium*, perhaps because of small scale heterogeneity.

In addition to ecotype distribution, our data also provide insights into the overall preferences of the *Curtobacterium* genus. The strongest predictors of *Curtobacterium* abundance (relative to all other bacteria) were cellulose content and climatic variability (PC1), where abundance was highest in cellulose-rich litter such as grass and more stable climates, as observed in previous surveys (Chase et al., 2016; Glassman et al., 2018). In contrast, *Curtobacterium* alpha diversity was primarily driven by climate with litter pH and hemicellulose concentration playing a secondary role. In particular, pH explained less than a third as much variation in alpha diversity as climate (PC1 and PC2 combined). This result contrasts with the alpha diversity of broader bacteria taxa (e.g., 16S diversity), where pH is the dominant driver in soil (Fierer & Jackson, 2006; Lauber et al., 2009; Zhou et al., 2020).

On a more technical note, the study demonstrates the value of combining metagenomic surveys with targeted cultivation efforts to resolve fine-scale microbial diversity in complex soil systems. Indeed, the accuracy of taxonomic assignments of metagenomic reads (or assemblies) remains limited by reference databases, especially for soil bacteria, which are highly underrepresented (Nayfach et al., 2021). Our genomic database created from newly isolated *Curtobacterium* strains enabled ecotype-level classification of metagenomic reads (Nayfach et al., 2016). Still, the dbRDA model explained only 27.41% of the variation in ecotype

composition and the random forest models, only 30.4%. This large unexplained variation indicates that important ecological drivers of ecotype distribution remain unmeasured and/or reflect temporal differences across the sampling window (winter through summer) (Langenheder et al., 2012; Lauber et al., 2013).

No single biotic or abiotic factor explains biogeographic variation of soil microbiomes, but generally soil chemistry (including pH and carbon quality and quantity) is thought to be more influential than climatic factors such as moisture and temperature (Fierer, 2017). While this pattern held true here as well, climate variables also contributed to a substation degree of variation explained (10%) in ecotype composition as well as overall *Curtobacterium* abundance and alpha-diversity. In general, we lack information about how soil bacterial traits related to climate preference vary at such a fine genetic resolution and their contribution to the maintenance of bacterial diversity in terrestrial ecosystems. By integrating phylogenetic, ecological, and functional perspectives, this approach provides a more complete understanding of how microbial diversity is maintained in soil.

ACKNOWLEDGMENTS

We thank everyone in the Martiny, Allison and Treseder labs who assisted with field work and sample processing, and Alex Chase for his guidance with the bioinformatic analyses. We also thank Adam Martiny and Kailen Mooney for their valuable feedback on earlier versions of this manuscript. This work was performed (in part) at the following reserves of the University of California Natural Reserve System: Blue Oak Ranch Reserve (doi:10.21973/N3ZT08), Hastings Natural History Reservation (doi:10.21973/N33Q0G), Bodega Marine Reserve (doi:10.21973/N32Q05), Jepson Prairie Reserve (doi:10.21973/N3D082), McLaughlin Natural

Reserve (doi:10.21973/N3W08D), Quail Ridge Reserve (doi:10.21973/N30T0K), Steele/Burnand Anza-Borrego Desert Research Center (doi:10.21973/N3Q94F), White Mountain Research Center (doi:10.21973/N3KM2J), Merced Vernal Pools and Grassland Reserve (doi:10.21973/N3FT0M), Sierra Nevada Research Station – Yosemite Field Station (doi:10.21973/N3V36C), Emerson Oaks Reserve (doi:10.21973/N34H2G), James San Jacinto Mountains Reserve (doi:10.21973/N3KQ0T), Motte Rimrock Reserve (doi:10.21973/N31T0W), Sweeney Granite Mountains Desert Research Center (doi:10.21973/N3S942), Elliott Chaparral Reserve (doi:10.21973/N36H23), Scripps Coastal Reserve (doi:10.21973/N3QH2R), Coal Oil Point Natural Reserve (doi:10.21973/N3Z07N), Kenneth S. Norris Rancho Marino Reserve (doi:10.21973/N37943), Sedgwick Reserve (doi:10.21973/N3C08R), Valentine Eastern Sierra Reserve – Sierra Nevada Aquatic Research Laboratory (doi:10.21973/N3966F), Valentine Eastern Sierra Reserve – Valentine Camp (doi:10.21973/N3JQ0H), Fort Ord Natural Reserve (doi:10.21973/N3GT0X), Landels-Hill Big Creek Reserve (doi:10.21973/N3NH24), and Younger Lagoon Reserve (doi:10.21973/N3894D). This work was supported by the National Science Foundation (DEB-2113004 to JBHM) and a Mildred E. Mathias Graduate Student Research Grant from the UC Natural Reserve System to ABS.

FIGURES AND TABLES

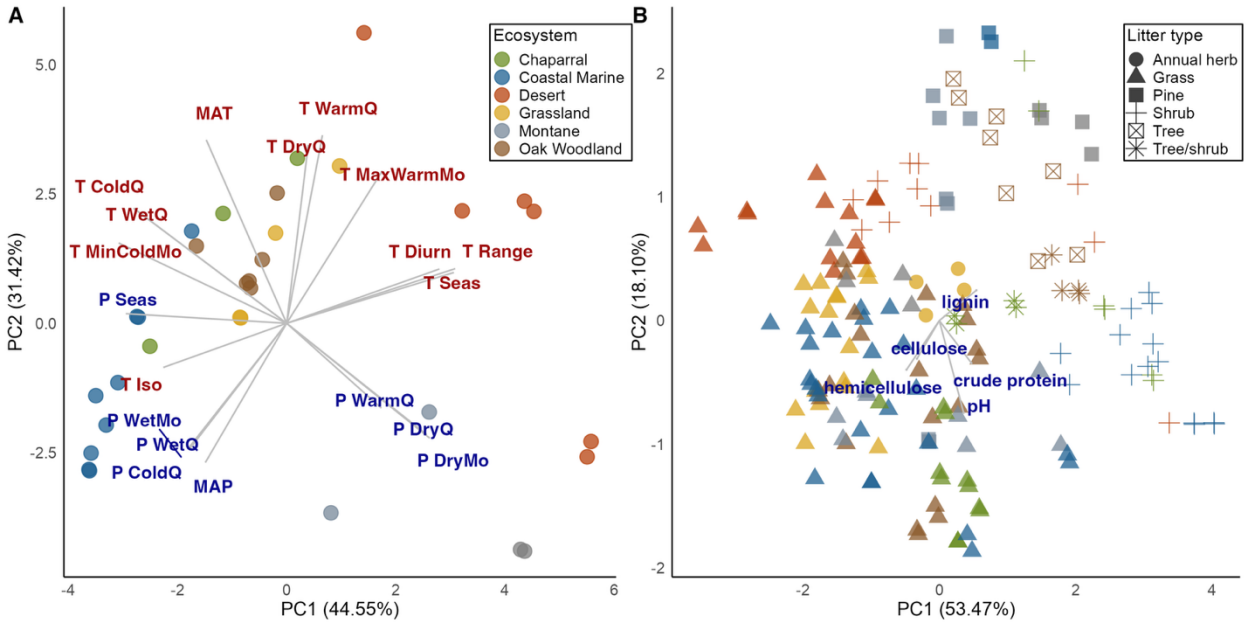

Figure 1.1. Ordinations of (A) the 24 sites based on 19 climatic variables and (B) the litter samples based on five chemistry variables. Temperature (T) variables are indicated by red labels, and precipitation (P) variables, by blue labels. See Table S1.1 for full names of the variables and mean values and range. In (A), there are six sites that include two points as the distance between the grassland and dominant vegetation plots are represented by different climatic data.

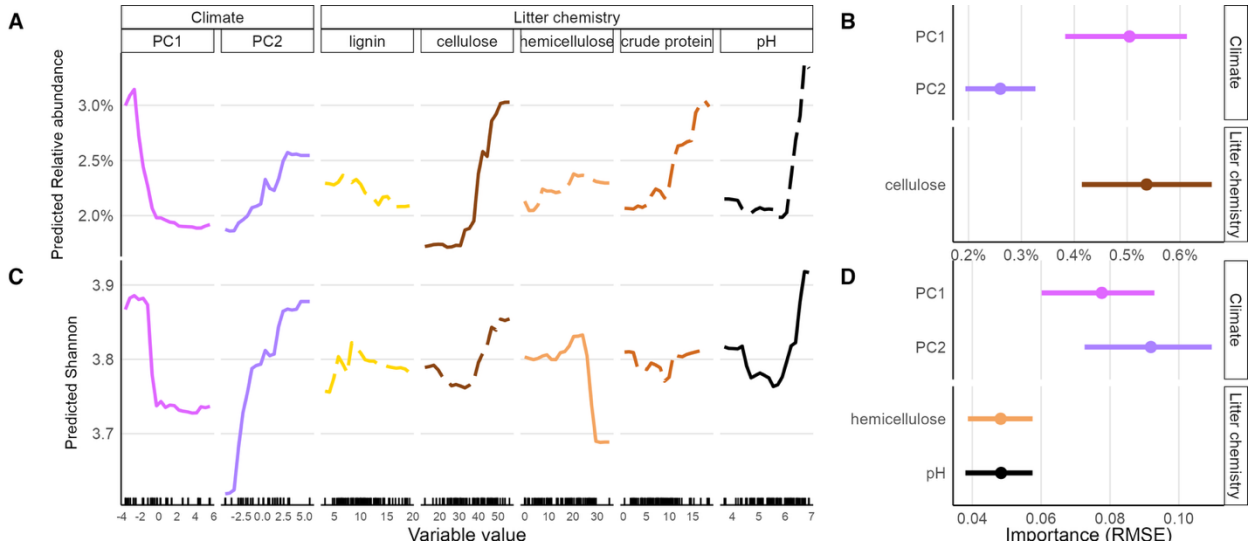

Figure 1.2. Partial dependence plots (PDP) of the predicted effects of climatic and litter chemistry variables on (A) the relative abundance of *Curtobacterium* (the percentage out of total prokaryotic reads) and (C) *Curtobacterium* alpha diversity (Shannon diversity index). Partial dependence plots show bootstrapped mean predicted values (n = 100 resamples), with tick marks on the x-axis representing the observed distribution of each variable. Solid lines indicate predictors of statistically significant variable importance ($p < 0.05$), and dashed lines indicate non-significant predictors ($p \geq 0.05$). Importance of each variable in the random forest model as determined by RMSE for (B) relative abundance and (D) alpha diversity. The climate variables correspond to the PCA axes in Figure 1A. Litter variable values (except pH) are percentage of dry weight mass.

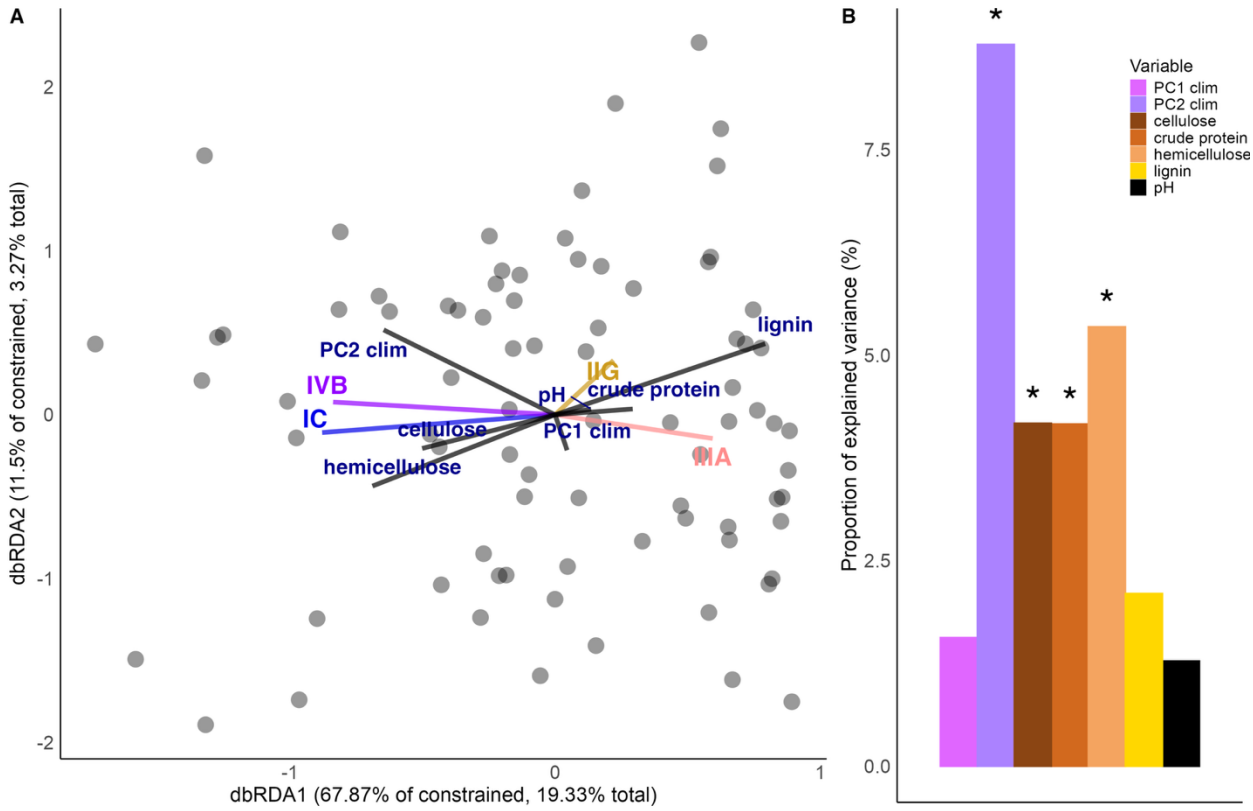

Figure 1.3 Environmental drivers of *Curtobacterium* ecotype composition. (A) dbRDA ordination of Bray–Curtis dissimilarities constrained by the two climate PCs (PC1, PC2) and five litter chemistry predictors. Points indicate ecotype composition of the litter samples (n = 94), and the colored vectors indicate correlations of individual ecotypes with the constrained axes. The predictor vectors indicate direction and strength of correlation with the constrained axes. (B) Proportion of constrained variance explained by each predictor (ANOVA-like permutation test, 9,999 permutations), * indicates variables that are significant ($p < 0.05$) in the permutation text.

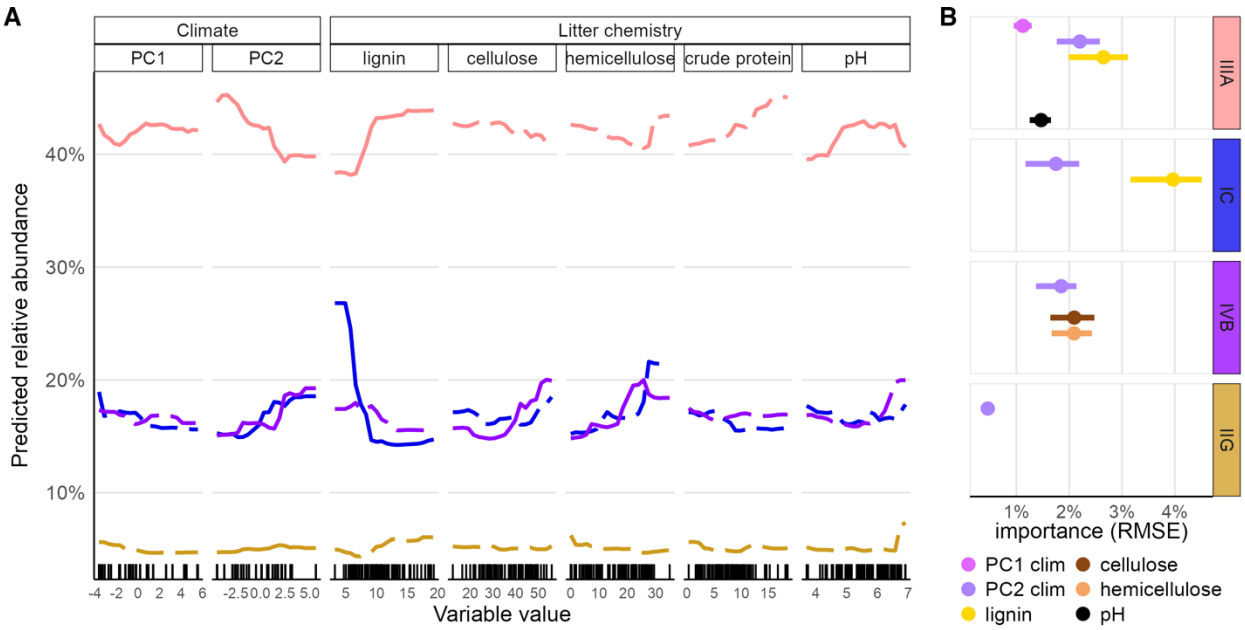

Figure 1.4. A) Partial dependence plots (PDP) for the four most abundant *Curtobacterium* subclades. Solid lines indicate predictors with statistically significant variable importance ($p < 0.05$), and dashed lines indicate non-significant predictors. Partial dependence plots show bootstrapped mean predicted values (n = 100 resamples), with tick marks on the x-axis representing the observed distribution of each variable. B) Importance of variables for each subclade.

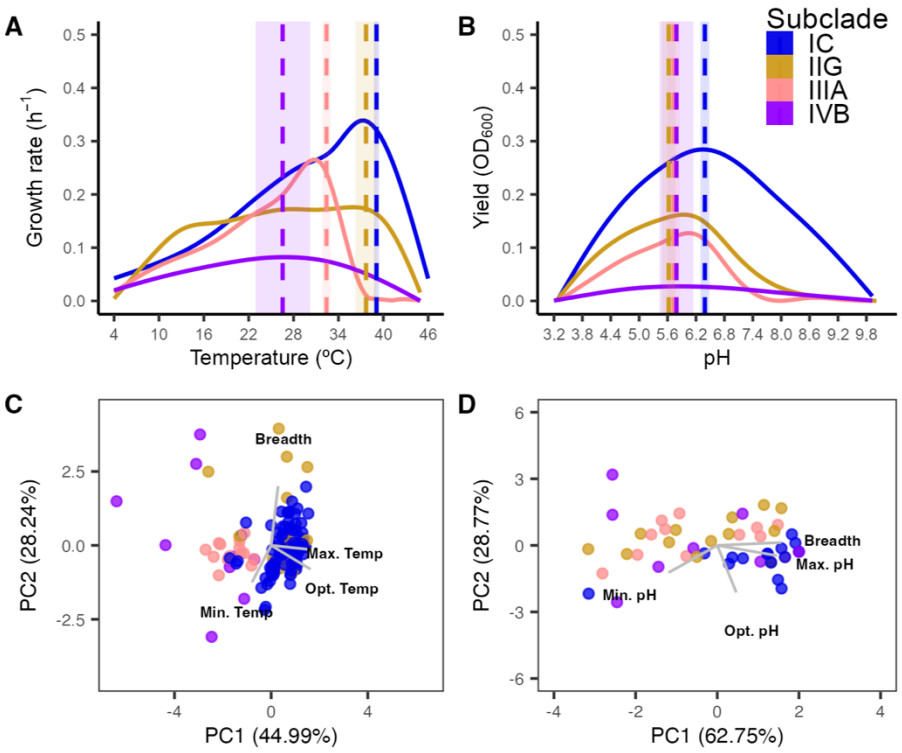

Figure 1.5. A) Temperature and B) pH performance curves of the top four most abundant ecotypes. Dashed lines indicate mean optima and shaded areas around the lines, the standard error. For visualization, we generated smooth predictions on a fine temperature grid (50 temperatures over the thermal gradient) and averaged the individual predicted curves at each temperature over the grid and displayed a loess-smoothed line. PCA ordinations of C) temperature and D) pH performance curve parameters for individual isolates colored by subclade. Vectors indicate loading of parameters with direction indicating increasing values.

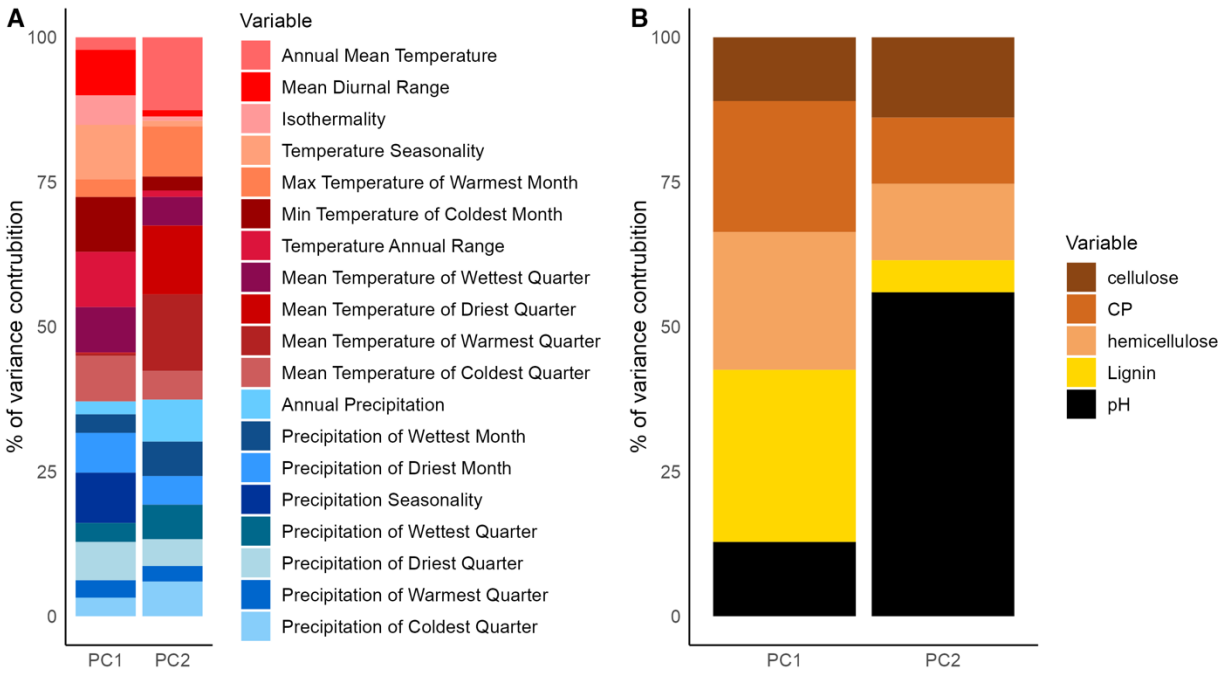

Figure S1.1. Variance contribution of each variable on the first two principal components in the PCA ordination for A) climatic variables and B) litter chemistry variables

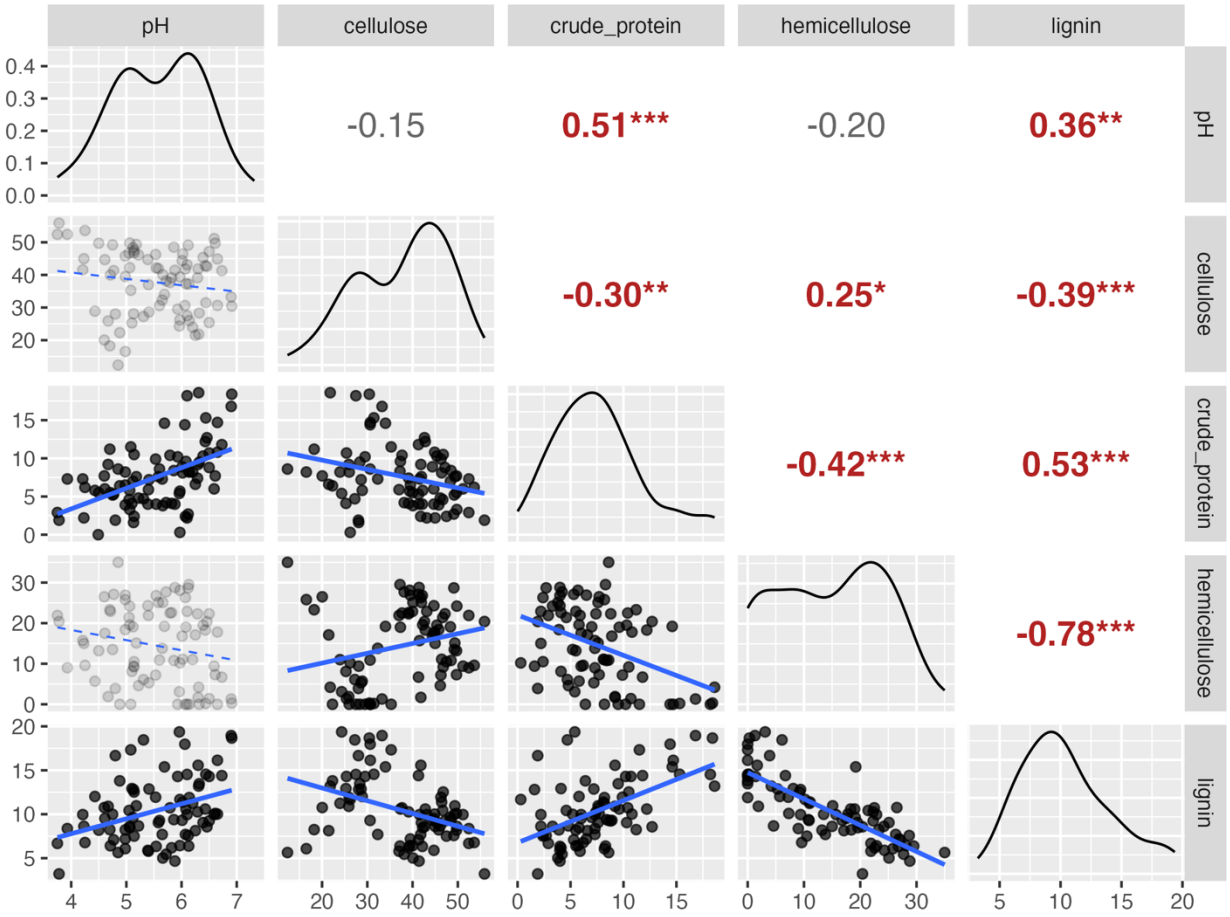

Figure S1.2. Spearman correlations among all litter chemistry variables. Upper triangle show correlation coefficients, lower triangle shows data points between pairs of variables. The diagonal shows the distribution of data for each variable. Significance is shown with * for $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

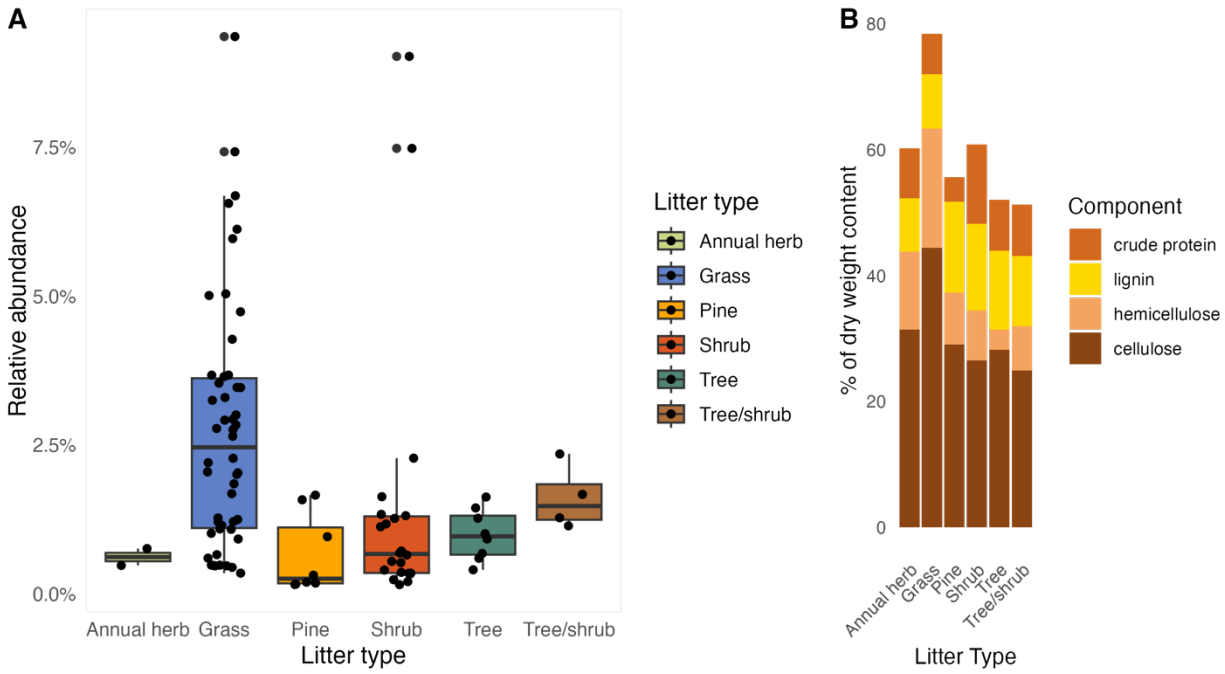

Figure S1.3. A) Relative abundance of *Curtobacterium* metagenomic reads in leaf litter types and B) chemical composition of litter types. ANOVA: Relative Abundance ~ Litter type: $F(5) = 3.34$, $p < 0.01$. No significant differences after post-hoc HSD Tukey test.

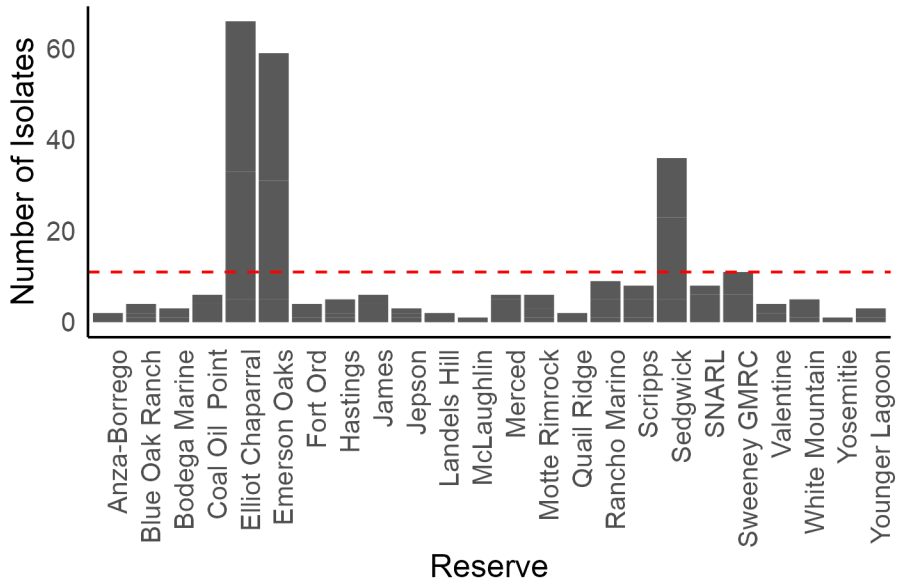

Figure S1.4. Number of *Curtobacterium* isolates by the sample UC Reserves (site). The red dashed line indicates the average number of isolates per reserve.

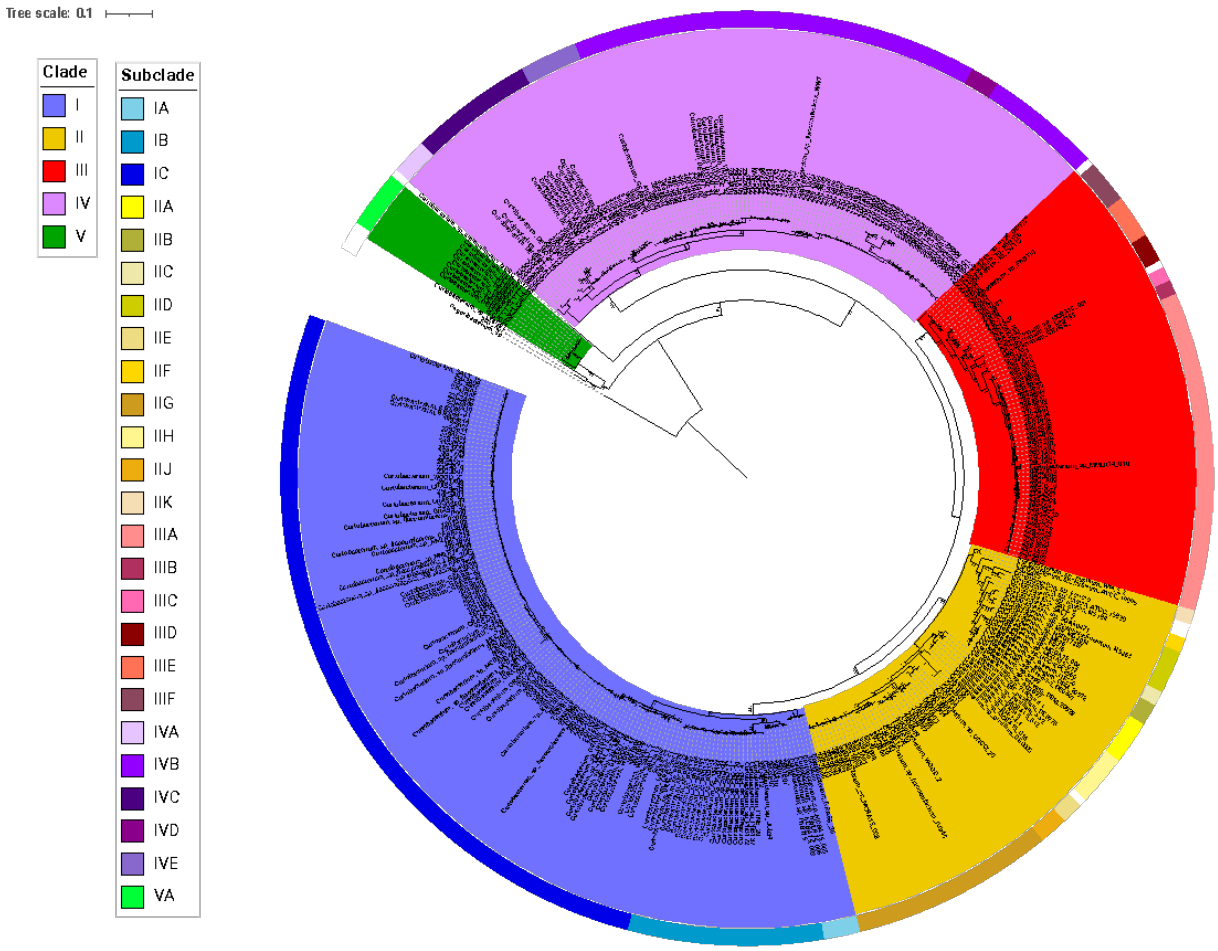

Figure S1.5. Phylogenetic tree of 389 *Curtobacterium* genomes. Colored areas represent clade and subclade (outer ring) classifications.

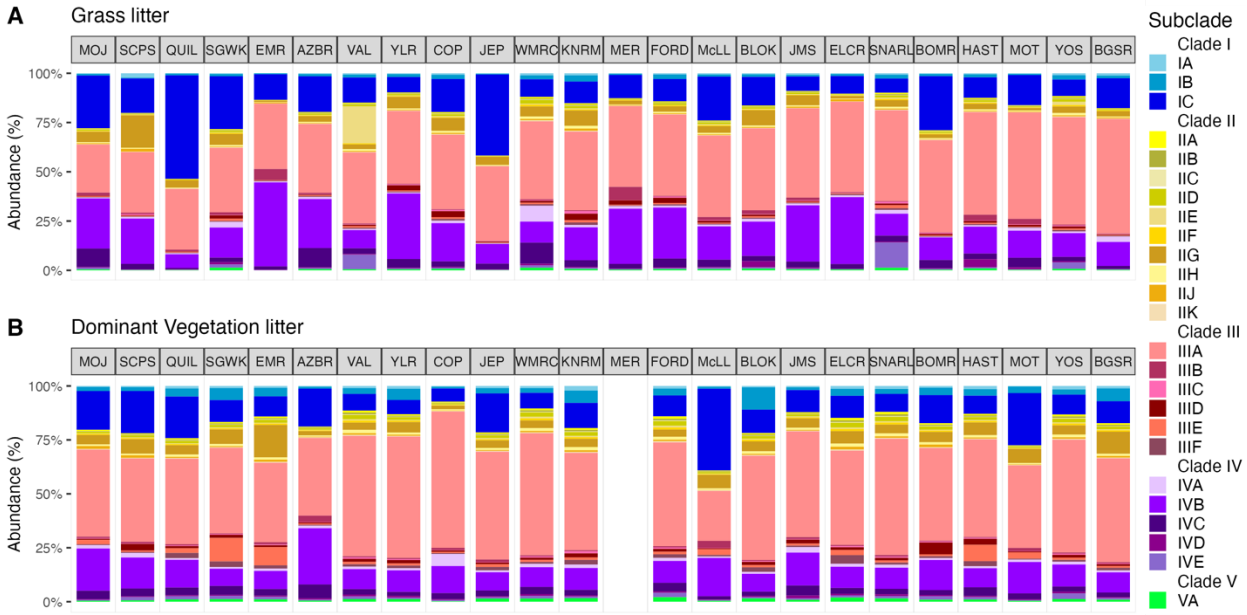

Figure S1.6. Relative abundance of the *Curtobacterium* subclades in A) grass litter and B) dominant vegetation samples across the 24 sites. We were unable to collect other type of vegetation in the MER site (UC Merced Natural Reserve).

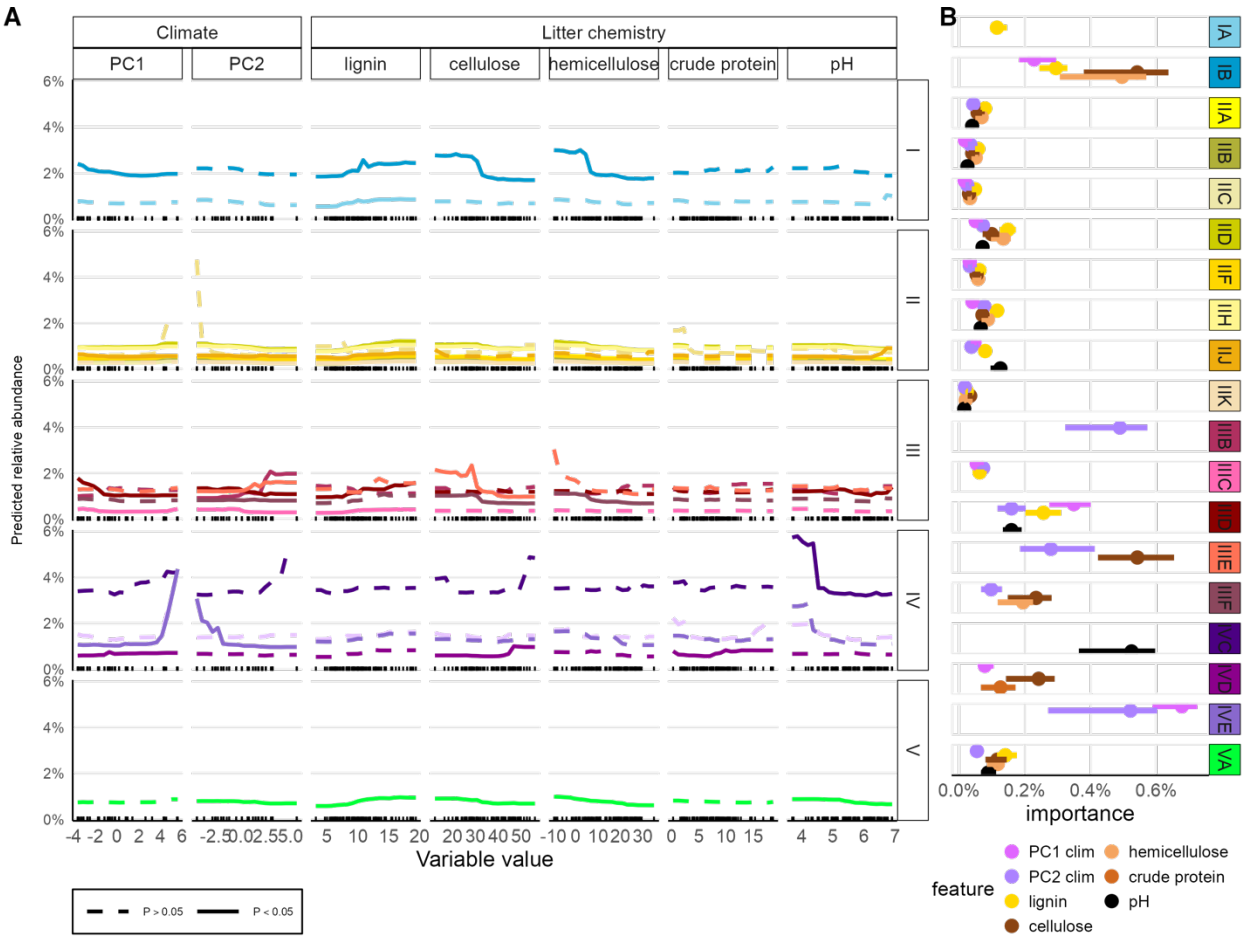

Figure S1.7. A) Partial dependance plots (PDP) for each of the low abundance subclades. B) Importance of variables for each subclade.

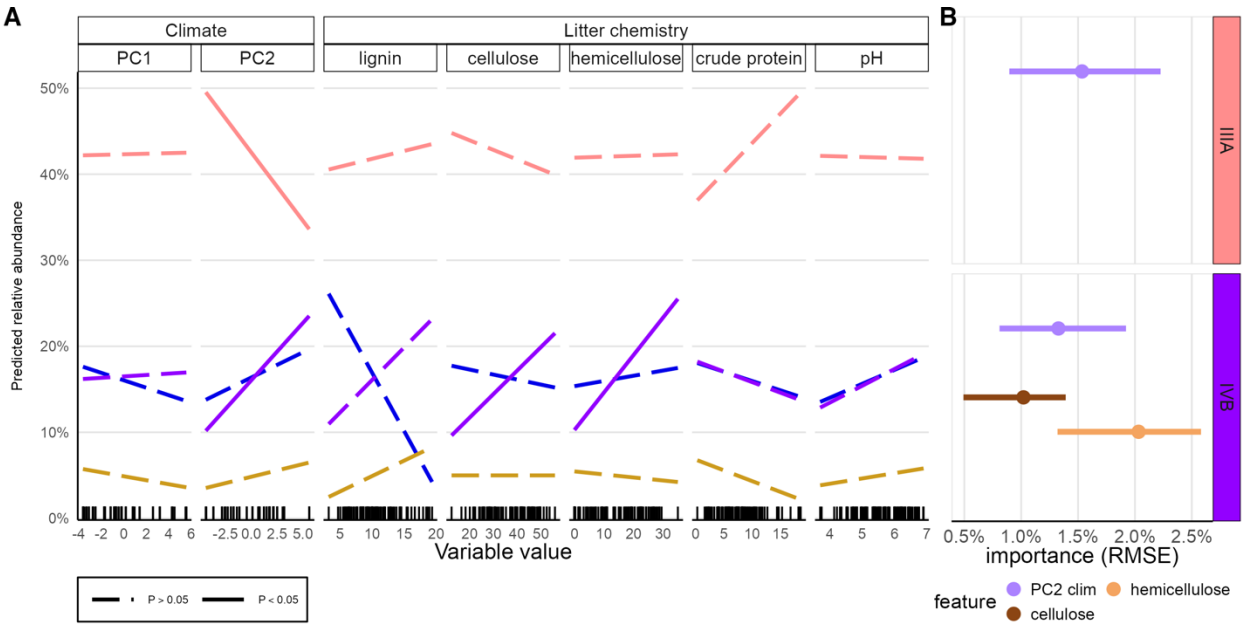

Figure S1.8. To complement the Random Forest analysis and assess whether monotonical predictor-response relationships are consistent under a parametric framework, we fitted Gaussian generalized linear models (GLMs) for each clade and subclade independently. A) Partial dependence plots (PDP) for the 4 most abundant Curtobacterium subclades. B) Importance of variables for each subclade. Solid lines indicate predictors with statistically significant coefficients ($p < 0.05$), as determined by coefficient-level t-tests from a Gaussian generalized linear model (GLM) fitted independently for each subclade. Variable importance was quantified as the permutation-based increase in root mean squared error (RMSE) upon randomization of each predictor (500 repetitions), with error bars representing the 5th–95th percentile range across permutations. Dashed lines indicate non-significant predictors ($p \geq 0.05$). Partial dependence plots show bootstrapped mean predicted values (n = 100 resamples), with tick marks on the x-axis representing the observed distribution of each variable.

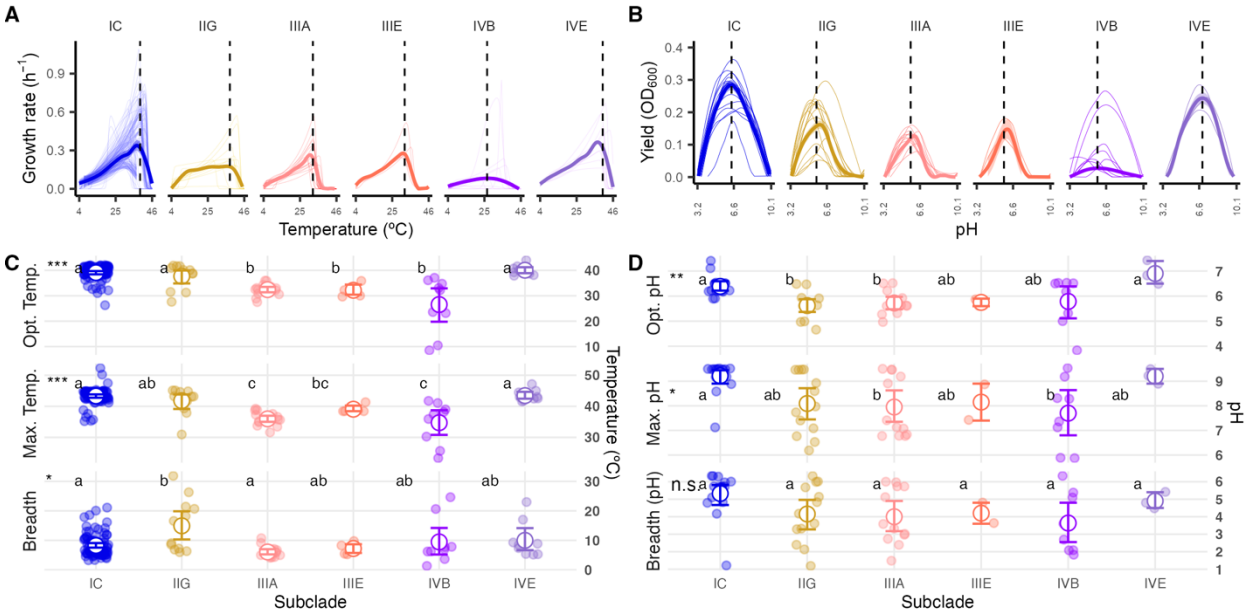

Figure S1.9. Average (bold lines) and individual strain curves (dimmed lines) for A) thermal and B) pH performance by subclade/ecotype. Dashed lines indicate the average optima for each subclade/ecotype. Average and 95% Cis for each curve parameter and subclade/ecotype for C) thermal and D) pH performance curves. Kruskall-Wallis test significance is indicated with * for $p < 0.05$, ** <0.01 and *** < 0.001. Letters indicate significant results of Dunn test pairwise comparisons at a p adjusted value < 0.05.

Table S1.1. Mean, maximum and minimum values of the environmental variables analyzed.

Variable	Mean	Max.	Min.
Annual Precipitation - MAP (mm)	521.78	1187.12	113.20
Mean Annual Temperature - MAT (ºC)	14.36	22.41	5.24
Precipitation of Coldest Quarter - P ColdQ (mm)	302.85	628.91	58.72
Precipitation of Driest Month - P DryMo (mm)	1.48	7.17	0.00
Precipitation of Driest Quarter - P DryQ (mm)	8.60	28.41	1.91
Precipitation Seasonality - P Seas	91.74	105.85	59.87
Precipitation of Warmest Quarter -P WarmQ (mm)	12.42	41.29	1.91
Precipitation of Wettest Month - P WetMo (mm)	110.60	237.43	24.00
Precipitation of Wettest QuarterP - WetQ (mm)	304.68	628.91	58.72
Mean Temperature of Coldest Quarter -T ColdQ (ºC)	8.96	13.87	-2.37
Mean Diurnal Range -T Diurn (ºC)	12.68	19.38	6.83
Mean Temperature of Driest Quarter - T DryQ (ºC)	19.70	24.75	13.73
Temperature Isothermality - T Iso (%)	51.52	63.66	35.62
Max. Temp. of Warmest Month -T MaxWarmMo (ºC)	28.40	41.60	17.20
Min. Temp. of Coldest Month -T MinColdMo (ºC)	2.99	8.80	-10.10
Temperature Annual Range -T Range (ºC)	25.41	40.50	10.80
Temperature Seasonality -T Seas (MAT SD * 100)	474.76	779.99	146.35
Mean Temperature of Warmest Quarter - T WarmQ (ºC)	20.49	31.58	13.75
Mean Temperature of Wettest Quarter - T WetQ (ºC)	9.06	13.87	-2.37
Cellulose (%)	37.71	55.89	12.36
Crude protein (%)	7.52	18.60	0.00
Hemicellulose (%)	13.94	35.00	0.00
Lignin (%)	10.42	19.37	3.21
pH	5.57	7.32	3.75

Table S1.2. PERMANOVA results testing the effect of categorical environmental variables on community composition of *Curtobacterium* ecotypes (subclades).

Source	df	SS	MS	Pseudo-F	*p*	% Variance Explained
Site	23	1.57	0.07	1.78	0.009	21.87%
Plant type	5	0.57	0.11	3.08	0.002	22.33%
Site x Plant type	17	0.63	0.04	1.83	0.001	21.86%

Table S1.3. Performance metrics for Random Forest models predicting bacterial diversity and abundance, by taxonomic level

Response	Taxonomic level	OOB MSE	OOB pseudo-R^2	10-fold CV R^2	Permuted CV p-value
Relative Abundance	Genus	0.000109	0.302	0.337	<0.001
Shannon	Genus	0.066608	0.337	0.349	<0.001
I relative abundance	Clade	0.008874	0.233	0.213	<0.001
II relative abundance	Clade	0.002814	0.152	0.19	<0.001
III relative abundance	Clade	0.009614	0.112	0.152	<0.001
IV relative abundance	Clade	0.006529	0.293	0.287	<0.001
V relative abundance	Clade	0.000612	0.021	0.049	0.05
IA relative abundance	Subclade	0.000038	-0.012	-0.016	0.15
IB relative abundance	Subclade	0.000198	0.356	0.333	<0.001
IC relative abundance	Subclade	0.009468	0.293	0.251	<0.001
IIA relative abundance	Subclade	0.000004	0.456	0.472	<0.001
IIB relative abundance	Subclade	0.000002	0.489	0.507	<0.001
IIC relative abundance	Subclade	0.000002	0.375	0.406	<0.001
IID relative abundance	Subclade	0.000016	0.436	0.466	<0.001
IIE relative abundance	Subclade	0.00156	-0.076	-0.037	0.2
IIF relative abundance	Subclade	0.000004	0.444	0.472	<0.001
IIG relative abundance	Subclade	0.001356	0.051	0.082	0.05
IIJ relative abundance	Subclade	0.000012	0.312	0.317	<0.001
IIK relative abundance	Subclade	0.000001	0.367	0.362	<0.001
IIH relative abundance	Subclade	0.00001	0.447	0.452	<0.001
IIIA relative abundance	Subclade	0.009765	0.145	0.177	<0.001
IIIB relative abundance	Subclade	0.000232	0.138	0.17	<0.001
IIIC relative abundance	Subclade	0.000005	0.305	0.333	<0.001
IIID relative abundance	Subclade	0.000089	0.223	0.252	<0.001
IIIE relative abundance	Subclade	0.000287	0.214	0.233	<0.001
IIIF relative abundance	Subclade	0.000047	0.275	0.297	<0.001
IVA relative abundance	Subclade	0.000321	-0.09	-0.03	0.15
IVB relative abundance	Subclade	0.005674	0.302	0.323	<0.001
IVC relative abundance	Subclade	0.000709	0.001	0.06	<0.001
IVD relative abundance	Subclade	0.000054	0.192	0.253	<0.001
IVE relative abundance	Subclade	0.000483	0.228	0.238	<0.001
VA relative abundance	Subclade	0.000019	0.331	0.334	<0.001

Response: the target variable modeled (relative abundance or Shannon index); **Taxonomic level**: genus or clade-level; **OOB MSE**: out-of-bag mean squared error; **OOB pseudo-R^2:** variance explained by OOB

predictions; **10-fold CV R^2**: cross-validated R^2; **Permuted CV p-value**: permutation test p-value for CV R^2 (A3 package).

CHAPTER 2: Testing adaptation to the local environment in leaf litter bacterial communities

ABSTRACT

Adaptation is a fundamental process structuring biodiversity, but testing adaptation of microorganisms to their environment in natural settings is difficult. Isolating individual strains, reintroducing them at ecologically realistic densities, and tracking their fitness in situ remains largely unfeasible in complex soil communities. It is also not clear what constitutes the best experimental unit when the candidates range from single isolates to whole taxa within an assemblage. Here we propose and demonstrate a community-level modification of the reciprocal transplant framework, using a metric of compositional convergence toward the native community as a measure of an adaptive response. We established eight replicated reciprocal transplants of soil microbial communities in the surface litter anchored at a focal grassland site in southern California. We manipulated the microbial inoculum, leaf litter substrate, and transplant location to partition the effects of climate, resource chemistry, and community origin on compositional outcomes. Transplanted communities converged toward the composition of native away communities, with both climate and litter chemistry contributing to convergence. These patterns were consistent at two phylogenetic scales: the whole bacterial community and the genus *Curtobacterium*, for which independent biogeographic and phenotypic evidence supports ecotypic differentiation along the same environmental axes. The four most abundant *Curtobacterium* ecotypes shifted in relative abundance in directions predicted by their thermal and chemical preferences, indicating that community-level convergence reflects the sorting of evolutionarily differentiated lineages rather than stochastic assembly. Convergence provides a

tractable test of microbial adaptation to local conditions in systems where individual-level approaches remain infeasible.

INTRODUCTION

Spatial variation in biotic and abiotic conditions imposes differential selection regimes that favor different genotypes and phenotypes. Consequently, some organisms may outperform others under a set of local conditions, a pattern often referred as local adaptation (Wadgymar et al., 2022) and a fundamental process in the generation and maintenance of biodiversity (Gavrilets, 2003). While local adaptation is typically studied at the population or species level, its consequences extend beyond individual lineages to shape community assembly and structure, influencing species' coexistence, abundances and functional roles (Leibold et al., 2022; Vellend, 2010). Studies of microbial adaptation have typically approached this question using laboratory experiments designed to emulate natural conditions (Belotte et al., 2003; Kraemer & Kassen, 2015), and/or biogeographic studies that infer adaptation from strong correlations of abundance of genotypes with environmental gradients (Johnson et al., 2006). Both approaches have important limitations. Laboratory testing cannot recreate the full ecological complexity of real communities and environmental variability. Biogeographic studies, in turn, cannot exclude confounding effects of dispersal limitation and ecological drift, as well as covarying environmental variables. Thus, an ideal approach would separate the effects of genotype and environment while preserving ecological realism.

For macroorganisms, reciprocal transplant experiments have become the gold standard for testing adaptation in nature. In these experiments, individuals of the same species are moved between at least two contrasting environments, and local performance is compared against that of

foreign genotypes in a fully crossed design (Clausen et al., 1940; L. C. Johnson et al., 2022). The sympatric (home) vs allopatric (away) comparison directly tests for the genotype by environment interactions while controlling environmental variation, such that when local genotypes consistently outperform foreign ones across sites, this provides strong evidence of local adaptation. However, applying this approach to microorganisms presents both a technical and conceptual challenge. Technically, isolating individual strains from natural communities, reintroducing them at ecologically realistic densities, and tracking their performance in situ remains largely infeasible with current technologies, particularly in complex soil communities where thousands of taxa co-occur. Conceptually, there is not a straightforward delimitation of what constitutes an "individual" or "experimental unit" in microbial systems in the context of reciprocal transplants. The term could be reasonably applied to single isolates, strains, populations, ecotypes or the community as a whole. This ambiguity is not merely semantic; the discrepancy between the scale at which we can experimentally manipulate and assess processes, and the relevant scale for the unit of study risks obscuring meaningful biological signal. This motivates the search for alternative approaches to assess adaptation in microbial communities.

Together, these challenges mean that classical reciprocal transplant designs developed for plants and macroorganisms cannot be directly applied to investigate microbial adaptation in natural environments. Typically, the use of reciprocal transplants of microbial communities has been applied to investigate the influence of composition on ecosystem functioning, showing that different communities may not perform equally under similar environments (Reed & Martiny, 2007; Allison et al., 2013; Hawkes et al., 2017; Waring & Hawkes, 2018). These experiments isolate microbial communities from the surrounding environment using "microbial cages", typically made of membrane barriers that prevent the dispersal of microorganisms but allow the

flow of nutrients in and out of the cage. Here, we propose and demonstrate the use of reciprocal transplants to investigate the degree of adaptation by soil bacteria within a community. If taxa are, on average, adapted to their local environment, then communities transplanted from another environment will shift in composition toward that of the local community over time (i.e., they will converge). This compositional convergence, measured as changes in the relative abundance of taxa following transplantation, provides a tractable test of adaptation that does not require tracking the fitness of individual strains. Analogous to classical reciprocal transplants, just as local genotypes are expected to outperform foreign genotypes in their home environment, taxa within a community that are better adapted to local conditions are expected to increase in abundance when placed in a more preferable environment while non-adapted taxa will decline.

To detect adaptation through community-level convergence, we use leaf litter bacterial communities as our study system. Leaf litter – the surface layer of soil – is characterized by high heterogeneity in abiotic conditions, such as temperature and moisture, and resource chemistry, largely determined by the identity of the plant species producing the litter. Leaf litter bacterial communities have been shown to respond to both axes of environmental variation, making them a promising system in which to detect signatures of adaptation (Glassman et al., 2018; Chase et al., 2021; Finks et al., 2021; Barbour et al., 2022). To confirm and refine the community-level interpretation, we additionally focus on the genus *Curtobacterium*, for which we have independent evidence of ecotypic differentiation and adaptation (Chase et al., 2017, 2018; Scales et al., 2022). The focus on the whole community and *Curtobacterium* allows us to test whether adaptive patterns are consistent across phylogenetic scales and consistent with prior biogeographic and phenotypic characterization. Specifically, we hypothesize that bacterial taxa are generally adapted to different substrate chemistry and/or the abiotic environment

(temperature and moisture). For *Curtobacterium*, we further expect that the experimental results will align with results based on biogeographic patterns and phenotypic assays (Chapter 1). To address this hypothesis, we established a reciprocal transplant experiment anchored at a focal grassland site in southern California. We manipulated the microbial inoculum, litter substrate, and environment (transplant location) across eight sites along a climatic gradient in California. This design allowed us to partition the effects of climate, resource, and the initial community on compositional outcomes, and to assess whether transplanted communities converge toward local community composition.

MATERIALS AND METHODS

Study sites

The focal site of the experiment is located at the Loma Ridge Global Change Experiment, a southern California grassland ecosystem 5 km north of Irvine, California, USA (338440 N, 1178420 W, 365 m elevation), dominated by exotic annual grasses and forbs (Potts et al. 2012; M. L. Goulden, G. C. Winston, S. Parker, K. Suding, and D. Potts, unpublished manuscript). The grassland is dominated by non-native annual grasses (*Bromus diandrus*, *Avena fatua*) and the native forb *Deindra fasciculata*. The climate is mediterranean with a mean annual temperature of 17.8 C, mean annual precipitation of 325 mm, and little rainfall between April and October. The other 8 sites against which this focal site is being compared are located within reserves managed by the University of California Natural Reserve System (UCNRS) (Supplementary table S2.1). These sites comprise a wide range of climatic variation and vegetation types.

Experimental setup

We selected a focal site (hereafter, referred to as Loma) and established reciprocal transplants between Loma and each of 8 sites in California (hereafter, Away sites, Figure 2.1A & B). Leaf litter for the experiments was collected in October of 2023. At each site we collected ~200 g of leaf litter of the most dominant plant or vegetation type (Supplementary table 1 – Sample metadata and coordinates). We used this litter to create the inoculum and as a sterile substrate for the transplants. To prepare the substrate, leaf litter was clipped to a size of one inch and placed into nylon membrane bags with 0.22-μm pores (cat. no. SPEC17970; Tisch Scientific). This pore size was selected to allow for the movement of water and nutrients but prevent the dispersal of exogenous microbial cells into the bags (Allison et al., 2013). The litterbags were then gamma irradiated with at least 22 kGy (Sterigenics). To manipulate the microbial community, homogenized leaf litter from each site was ground in coffee grinders (KitchenAid), and 50 mg of inoculum was added to the irradiated litter bags.

For each reciprocal transplant, we manipulated site, litter substrate, and inoculum with 4 replicates for each treatment combination (2 sites × 2 inocula × 2 litter substrates × 4 replicates = 32 litter bags). Because our purpose was to test a well characterized community where the *Curtobacterium* genus is abundant (citations about LOMA), we excluded the treatment where away communities were placed on the Loma site and litter (Fig. 2.1A and B). Additionally, we included a set of litter bags with sterile LOMA leaf litter but no inoculum as a control to assess the effect of the irradiation alone on the emerging communities. Finally, we included two sets of bags at all nine locations containing clipped Loma or away leaf litter made of 2-mm mesh window screening (open bags) to assess changes in community composition after dispersal from the surrounding communities into the irradiated litter. In total, we deployed a total of 28 bags

(closed + open) at each of the away sites (224 total), and 108 bags in Loma (332 total). Litter bags were deployed in November 2023 and retrieved after 6 months (May 2024) except for the litter bags deployed in the SNARL site, which were collected in June 2024 after the snow pack cleared.

Characterization of microbial communities

To characterize bacterial community composition, we created and sequenced metagenomic libraries from each litter bag. We extracted DNA from 0.05 g of ground litter using the ZymoBiomics DNA kit (Zymo Research, Irvine, CA). Extracted DNA was diluted to 0.5 ng/µl and 1ng of DNA was used as input for the Nextera XT library Prep kit for sequencing on the Illumina HiSeq 4000 instrument with 150 bp paired-end reads at the University of California Irvine Genomics & Research Technology Hub (GRT Hub). Illumina raw reads were quality filtered, and adapters removed with bbduk from the BBmap suite (Bushnell, 2016), and filtered to remove eukaryotic DNA by mapping quality filtered reads to reference genomes of an abundant grass (*Lolium perenne*; Accession: MEHO01000000) and fungus (*Pyrenophora teres*; Accession: NZ_AEEY00000000) using BWA (Li & Durbin, 2009). Filtered reads were then merged using BBmap to form paired end reads. If a read could not be merged merge with its counterpart, then only the forward read was included for further analysis. Next, merged reads were searched against the *Curtobacterium* custom database (Barron-Sandoval et al. in progress) using HS-BLAST (Chen et al., 2015) using an *E* value of 1×10^{-50}, keeping only the best hit for each query sequence. Next, the BLAST output files were parsed with in-house scripts to obtain total read counts for each genome in the data base for each sample. Next, we adjusted read count by dividing the total read count for each genome in each sample by the total base pairs (sum of the length of all genes) represented in the data base for each genome and multiplied this to a

million to get a count reads per million bases mapped. We classified metagenomic reads against the against the GTDB database (version R10-RS226, Parks et al., 2026) at the genus level with Kraken2 (Wood et al., 2019). Absolute read count per genera and sample were calculated with Bracken (Lu et al., 2017). For the *Curtobacterium* community, clade and subclade abundances were calculated by summing sample-level read counts assigned to genomes within each clade/subclade in the custom reference database (see above).

Community analysis

To account for differences in sequencing depth, metagenomic read counts were rarefied to 435 reads for the *Curtobacterium* community and 20,833 reads per sample for the whole bacterial community and repeated 100 independent iterations to reduce stochastic effects. Alpha diversity metrics (observed richness and Shannon index) were calculated on each rarefied replicate using the vegan R package (Oksanen et al., 2018) and final values are reported as the mean across iterations. For beta diversity, a Bray–Curtis dissimilarity matrix was computed for each rarefied replicate and the resulting matrices were averaged across the 100 iterations to produce a single, average Bray–Curtis dissimilarity matrix for downstream analyses. We tested for effects of the experimental manipulations (site × inoculum × litter substrate) on community composition using PERMANOVA (type III sums of squares; 9,999 permutations), with all factors as fixed effects using the Bray-Curtis dissimilarity matrix using the PRIMER+ software (version 6.0; Anderson et al. 2006).

To calculate the convergence metric (see below) we used an averaged Euclidean matrix across 100 iterations. We used Euclidean distance on the top 400 most abundant genera and the top four most abundant subclades in the *Curtobacterium* community for the convergence metric. While Bray-Curtis dissimilarity is often preferred for snapshot ordination of community

composition (owing to its handling of joint absences) it is not ideal as a convergence metric for three reasons: (1) it is not a true metric, violating the triangle inequality and undermining geometric reasoning about directional change; (2) its bounded, non-linear nature compresses differences near convergence and conflates metric behavior with biological change. Euclidean distance, being linear and a proper metric, ensures that decreasing distances correspond proportionally to compositional convergence. The joint-absence concern was mitigated by restricting analyses to the most abundant taxa, which represent the dominant compositional signal.

Using Euclidean distances, we compared treatment groups within each reciprocal-transplant pair (Loma-Away pair, where Away is one of the 8 other sites), We defined groups according to the experimental treatment (H_0, I or A_0; Fig. 2.1C), under consideration and calculated the mean pairwise Euclidean distances between groups. For any two groups, A and B, this quantity was computed as the arithmetic mean of all pairwise Euclidean distances between samples in A and samples in B:

$$\bar{d}(\mathrm{A},\mathrm{B}) = \frac{1}{n_\mathrm{A} n_\mathrm{B}} \sum_{i \in A}^{n} \sum_{j \in B}^{n} d_{\mathrm{ij}}$$

here d_{ij} is the Euclidean distance between samples i and j. Within-group distance was calculated as the mean of all unique pairwise distances within a group. To evaluate convergence towards the native away-community state (A_0), we defined, for each reciprocal-transplant pair, a starting community state at Loma home (H_0), intermediate states (I), and the native away reference (A_0) (Fig. 1C). Likewise, we quantified the convergence of the away communities (A_0) toward the Loma native reference (H_0) in a similar manner where possible. We quantified the convergence (S) of an intermediate state (I) toward the native away reference state as:

$$S = 1 - \frac{d(\mathrm{I}, A_0)}{d(H_0, A_0)}$$

This measures how far a community has travelled towards the Away reference community. Under this formulation, positive values indicate convergence toward the native away state, values near zero indicate no change in compositional distance to the away community, and negative values indicate divergence from the away community. This normalized metric allowed direct comparison among reciprocal-transplant pairs that differed in their initial compositional separation.

To assess whether microbial communities converged toward the reference community (A_0) along each experimental path, we tested whether the mean convergence metric (S) differed significantly from zero using a bootstrap hypothesis test. For each path × community level combination (five paths × two community levels = ten tests), we computed the observed mean S across the eight replicated reserve pairs. To construct a null distribution that does not assume symmetry around zero (an assumption we considered inappropriate given the directional nature of the metric), we first centered the eight values at zero by subtracting their mean, then resampled with replacement from the centered values 9,999 times, computing the mean of each resample. The p-value was defined as the proportion of bootstrap means whose absolute value equaled or exceeded the absolute value of the observed mean. P-values were corrected for multiple comparisons within each community level using the Benjamini–Hochberg false discovery rate (FDR) procedure.

To test whether convergence differed among paths (omnibus test), we used a within-pair permutation F-test. Because the same eight reserve pairs contributed one S value per path, we used a two-way ANOVA framework (S ~ path + pair) that partitions out pair-level variation before testing the path effect. To generate the null distribution, path labels were permuted

independently within each pair across 9,999 iterations, preserving the repeated-measures structure while breaking any systematic path signal. The p-value was the proportion of permuted F-statistics equaling or exceeding the observed F-statistic. All analyses were conducted in R v4.5.0.

RESULTS

To evaluate how bacterial communities responded to the transplants, we examined compositional changes after six months in the field. We first considered all the replicates inoculated with the Loma community and asked how site and litter substrate altered community composition when holding the inoculum constant. As expected, both site and litter substrate influenced the composition of the whole bacterial community (Figure 2.2A) and *Curtobacterium* ecotypes (Fig. 2.2B). The amount of composition variation explained was slightly higher for the bacterial community than that of *Curtobacterium* ecotypes. The total compositional variance explained by site, litter and their interaction was 70.5% for the whole bacterial community and 63.8% for the *Curtobacterium* community (Table S2.2). This difference was due to litter substrate explaining slightly more variation in the whole community.

We next analyzed all samples from the 8 transplant experiments, including the effect of the inoculum. Overall, the experimental factors (site, litter, inoculum and their interactions) explained on average 77.4% of the compositional variation for the whole bacterial community after all 8 experiments (Fig. 2.3A). These results indicate that both axes of environmental variation manipulated in the experiment (climate, as proxied by site, and resource chemistry, as proxied by litter substrate) structure community composition, providing the necessary foundation for the convergence analysis that follows. In addition, the inoculum explained a smaller but

significant proportion of compositional variation (mean: 9.25%) (Fig. 2.3A and Table S2.3), signifying that a legacy of the initial inoculum remains at the end of the experiment.

For the *Curtobacterium* community, the experimental factors explained on average 59.4% of the compositional variation (Fig. 2.3B, Table S2.3). In this case, the inoculum origin alone explained a smaller proportion of the compositional variation (mean: 1.35%) than for the whole community. In fact, this main effect of inoculum was only significant in one of the eight reciprocal transplant experiments, suggesting reduced legacy effects of the initial inoculum on *Curtobacterium* composition.

Next, we quantified whether the transplanted communities shifted in composition towards that of the native communities over the course of the experiment. Considering each transplant pair separately in ordination space (Figure 2.4), the intermediate steps (that is, when communities were moved to either a different site, litter or both) generally fall within the ordination space between the Loma home state (Loma communities in the Loma site inoculated onto Loma litter; shown in the ordination with dark-green solid circles), and the Away native state (away communities in the away site inoculated onto away litter; shown in the ordination with purple solid circles). For instance, this pattern is evident in figure 2.4A panel MOJ and 4C panel BOMR. Quantification with the convergence metric, confirms this pattern. With the except of one treatment in one experiment, all transplants showed positive convergence towards the away native community, both for all bacteria and for *Curtobacterium* community (Figure 2.4B and D). This pattern held true for all three intermediate treatments. When the Loma microbial community was moved to a new climate (site) but kept on Loma litter, the composition of the whole bacterial community and *Curtobacterium* ecotypes converged towards the Away native reference community (dark green bars in figure 2.4B and D). Similarly, when the Loma

community was shifted onto Away litter but kept in the Loma climate, it shifted towards the Away native reference community (intermediate green bars). Finally, when both climate and litter were changed simultaneously, the Loma community converged more towards the Away community (lightest green bars).

The effect of the experimental manipulations tended to be larger for the whole bacterial community than the *Curtobacterium* community, with mean convergence values for the climate + litter path of 0.381 and 0.306 for the bacterial and *Curtobacterium* community, respectively (Fig. 2.4B and D, Table S2.4). Permutation tests confirm that the observed mean values are significantly greater than zero (Table S2.4). Convergence was not even across the eight replicated experiments and community level of analysis. For the bacterial community, samples transplanted to the SNARL site showed the largest average convergence across paths (mean S = 0.386), while for the *Curtobacterium* community the samples transplanted to the BOMR site showed the largest average convergence across paths (mean S = 0.34, Fig. S2.1, Table S2.5).

Convergence was not even across paths, nonetheless, this was only observed for the bacterial community level (Observed F = 2.84, p = 0.047). Finally, transplants of the Away communities into Loma (purple bars in Figure S2.2A and B) were similarly influenced by both climate and litter, although the degree of variability in this direction was wider and overall lower convergence than the Loma communities (Table S2.4).

To better understand these patterns of convergence, we plotted changes in the relative abundances of the four most abundant *Curtobacterium* ecotypes across the three transplant paths (Fig. 2.1C) from Loma to the Away sites (Figure 2.5). Consistent with the positive convergence metrics observed, ecotype composition shifted systematically following transplantation, with the abundance of each ecotype typically increasing or decreasing monotonically towards the Away

native reference community (Figure 2.5). Only in a handful of cases was this not the case (for instance, subclades IC and IVB at the intermediate step in the climate path when moving the Loma community to SNARL; Fig. 2.5 LOMA vs SNARL by Climate panel).

Finally, we assessed whether the ecotype shifts observed in the transplant experiment were consistent with the biogeographic distribution and phenotypic traits of the *Curtobacterium* ecotypes. Examining the samples inoculated with the Loma community on Loma litter (to isolate the effect of site; climate path in figure 2.1C), we found that ecotype relative abundances varied predictably with mean annual temperature across sites (Fig. 2.6). Ecotype IIIA declined in relative abundance with increasing MAT, while ecotypes IC and IIE increased. These patterns align closely with the biogeographic and phenotypic data obtained in parallel studies, where we observe the same trends across a larger geographic area and thermal preference traits (Chapter 1). Particularly, we have shown that subclade IC increases in relative abundance along the temperature/precipitation gradient while subclade IIIA decreases (Chapter 1, Fig. 1.4A). Consistent with this result, isolates from subclade IC shows a higher optimal temperature than those from subclade IIIA (Chapter 1, Fig. 1.5A). These results indicate that community-level convergence is consistent with the hypothesis that *Curtobacterium* ecotypes are adapted to different temperatures.

DISCUSSION

We found support for our hypothesis that bacterial taxa within litter communities are adapted to both climate and litter chemistry. Communities transplanted to a new litter substrate or a new climate converged to the native reference communities. Further, these shifts were detectable at two phylogenetic scales, suggesting that the compositional shifts are consistent and

directional. These trends were stronger for broader taxa within the whole bacterial community than for ecotypes within the *Curtobacterium* genus. Likewise, convergence was greater for the transplants that moved communities to both the away site and away litter, indicating that both axes exert ecological filtering simultaneously. Our interpretation was confirmed by analyzing changes in the relative abundance of the *Curtobacterium* ecotypes in transplants that held the litter substrate constant but were moved to another climate; these results yielded a pattern similar to that observed in a biogeographic survey. Ecotype IC increased in abundance as it was moved to sites with warmer temperatures whereas ecotype IIIA decreased in abundance. These results are consistent with the hypothesis that the *Curtobacterium* ecotypes are adapted to different combinations of climatic conditions and resource chemistry.

Unlike transplants of single individuals, moving whole communities means that the biotic interactions are incorporated into the analysis. While this adds ecological realism, interactions among the taxa may influence compositional outcomes as well as differential preferences of abiotic or resource conditions (Kraft et al., 2015; Scheuerl et al., 2020). Whether the inclusion of interactions is advantageous or problematic remains unclear. Classical transplant studies of plants, for instance, cannot separate fitness effects arising from environmental differences from those stemming from altered interspecific interactions. For instance, in a reciprocal transplant of culturable seaweed epibacteria between host-derived media, isolates showed higher fitness in their native host environment. However, the authors note that isolates were assayed independently of both their wider community and the living host, so interactions within the microbiome and with the host could not be captured (Corr et al., 2025).

The enclosure of communities should be considered while interpreting the convergence metric. Since new cells cannot enter the bags, any taxon favored and increasing in abundance must have been present in the inoculum. Convergence is thus restricted by any taxa that are abundant at an away site but absent from the Loma inoculum, such that the metric is a conservative estimate of the adaptive response. That said, the convergence metric was calculated using the four most abundant *Curtobacterium* ecotypes that are detected at all locations (Barron-Sandoval et al., Chapter 1). The whole-community analyses were similarly restricted to the top 400 most abundant taxa such that we do not expect the convergence metric to be restricted by missing taxa.

A community could therefore converge on the native composition because conditions at the away site happen to favor taxa that were already abundant in the transplanted community, without any of those taxa being locally adapted in the strict sense. Indeed, in a laboratory experiment, soil microcosms were inoculated microbial communities originating from either soil or a freshwater lake and both developed towards a composition resembling those of soil communities (Causevic et al., 2025). Because the lake communities have no plausible history of selection in soil, their convergence toward a soil-like composition does not reflect adaptation to those taxa to soil. Convergence alone is therefore not evidence of a particular taxon's adaptation to the specific habitat. However, our design differs from this type of experiment in two aspects. First, their inocula originated from unrelated habitat types and were tracked at coarse taxonomic resolution in closed microcosms, whereas ours are litter communities compared across a climate gradient within a single habitat, for which independent trait data are available for a specific taxon at finer phylogenetic resolution. Second, that finer resolution allows the two explanations to be distinguished. Filtering on physiological tolerance alone predicts that taxa better suited to the

new conditions increase in abundance, without specifying which. In contrast, sorting among differentially adapted lineages predicts that particular taxa increase in directions matching preferences characterized independently of the experiment. The *Curtobacterium* ecotype shifts that we observe follow the latter expectation.

What our metric captures, then, is aggregated sorting of individual taxa, measured at the community level. It is also worth distinguishing which taxa can establish at a site from the abundances they reach, convergence in composition reflects the latter, and requires that responding taxa not merely persist but change in relative abundance in a consistent direction. We interpret positive convergence as indicating that, on average, the taxa increasing in the transplanted community preferentially grow and survive under the new conditions relative to other members of that community. Whether those responses reflect adaptation rests on the *Curtobacterium* results, where the direction of change can be compared against independently characterized environmental preferences.

Finally, in a synthetic community experiment, when replicate communities assembled from the same set of species are propagated in parallel under identical conditions, they diverge into several alternative states rather than converging on a common one, indicating that community assembly trajectories can experience stochastic effects even in the absence of environmental variation (Celiker & Gore, 2014). This indicates that detecting directional convergence toward the native away community across eight independent transplants is therefore unlikely to arise from stochasticity alone. Aggregated species-level sorting is what our metric is designed to detect, and the *Curtobacterium* results indicate that at least part of that sorting reflects adaptive differentiation among closely related taxa.

Several aspects of the experiment should be considered while interpreting our results. First, litter bags do not fully recreate the natural conditions outside them. They prevent light penetration and may alter moisture and temperature relative to the surrounding litter layer. Because bag construction was identical for all treatments, any effects of the bag are shared across treatments rather than assuming them as negligible. Furthermore, the effects of the experimental manipulations are being compared against the home and away reference communities that underwent the same manipulations. Therefore, the responses we observe are unlikely to arise from differential effects of the bags across treatments. Second, our characterization of litter substrate captured cellulose, hemicellulose and lignin content but did not resolve finer aspects of litter quality, C:N ratio, or secondary compounds released during decomposition that are known to influence microbial metabolism and could contribute to the substrate effects we attribute to the measured variables (Rousk & Bååth, 2007; Chomel et al., 2016). Our conclusions therefore identify litter chemistry as a driver of convergence without isolating which of its components is responsible. Third, the design was asymmetric: we transplanted Loma communities to away sites but did not establish the reciprocal treatment of transplanting away communities to Loma site and litter. This was a deliberate choice, since Loma is a well-characterized site where *Curtobacterium* abundance and ecotypic structure have been documented in detail (Chase et al., 2017; Scales et al., 2022), but it means we cannot test whether show the reciprocal pattern at the Loma site. Finally, litterbags containing sterile litter and no inoculum allowed us to assess background colonization and sterilization efficacy; these controls remained compositionally distinct from inoculated bags throughout the experiment, confirming that the communities we characterized generally reflected the experimental inoculum, rather than recolonization from the surrounding environment (Figs. S2.3 and S2.4).

CONCLUSION

Reciprocal transplants have long served as the standard approach for testing adaptation in plants and other macroorganisms, but their application to microorganisms has been limited by the difficulty of manipulating individual strains in field conditions. By shifting the response variable from individual performance to compositional convergence, we aimed to quantify the degree to which bacterial lineages are adapted to local conditions. This convergence measure is flexible in regard to the focal unit, for instance, here we focused on both 16S-defined taxa and pre-defined ecotypes within a genus. It also allows adaptation to be assessed simultaneously at different phylogenetic scales. The same sequencing effort that characterizes whole communities can resolve responses of individual lineages within it. Furthermore, ecologically meaningful variation in soil bacteria is often partitioned below the level distinguished by 16S rRNA gene OTUs. Our results confirm that these lineages are also adapted to both litter substrate and climate. Understanding how microbial communities respond to global change, and implications for ecosystem processes, may then be improved by detecting the adaptive differentiation below the conventional marker gene resolution.

ACKNOWLEDGMENTS

We thank everyone in the Martiny, Allison, Treseder and Rebolleda-Gomez labs who assisted with field work and sample processing, especially Nickole Villabona and Jose Murua-Royo for their invaluable help on the field work for this project. We also thank Adam Martiny and Kailen Mooney for their valuable feedback on earlier versions of this manuscript. This work was performed (in part) at the following reserves of the University of California Natural Reserve System: Blue Oak Ranch Reserve (doi:10.21973/N3ZT08), Bodega Marine Reserve

(doi:10.21973/N32Q05), McLaughlin Natural Reserve (doi:10.21973/N3W08D), Emerson Oaks Reserve (doi:10.21973/N34H2G), Sweeney Granite Mountains Desert Research Center (doi:10.21973/N3S942), Coal Oil Point Natural Reserve (doi:10.21973/N3Z07N), Sedgwick Reserve (doi:10.21973/N3C08R), Valentine Eastern Sierra Reserve – Sierra Nevada Aquatic Research Laboratory (doi:10.21973/N3966F), This work was supported by the National Science Foundation (DEB-2113004 to JBHM) and a Mildred E. Mathias Graduate Student Research Grant from the UC Natural Reserve System to ABS. We also thank the Irvine Ranch Conservancy and UCI Environmental Collaboratory for facilitating research at Loma Ridge. The LRGCE is located on the unceded land of the Tongva/Kizh and Acjachemen Indigenous peoples.

FIGURES AND TABLES

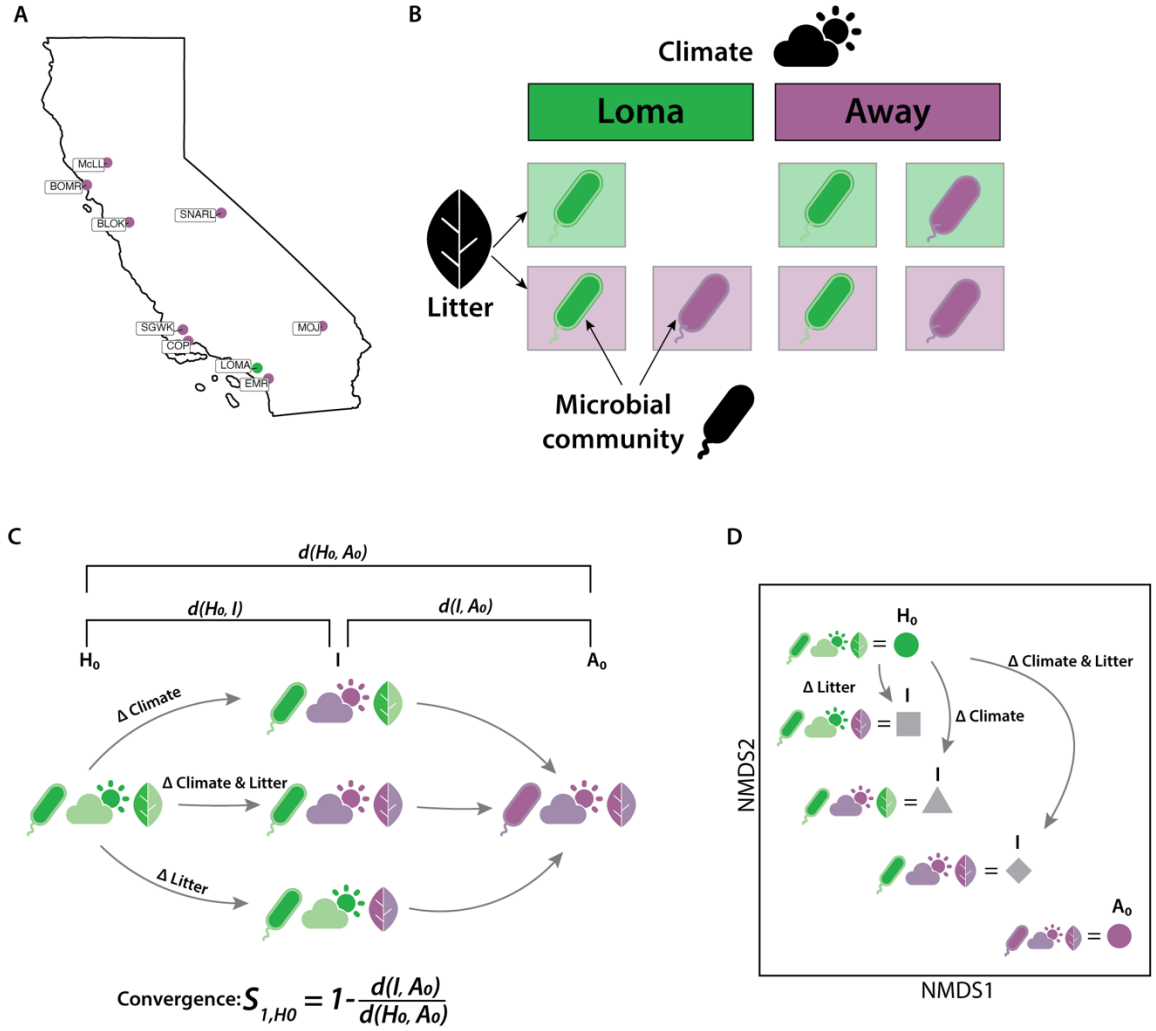

Figure 2.1. A) Distribution of sites across California and B) Experimental design of the replicated reciprocal transplant experiment across the 8 Away sites. C) The convergence metric and convergence paths analyzed. H_0: Loma community on its home environment; I: Intermediate state after experimental manipulation: A_0: Away community on tis home environment. Brackets show the calculated distance between experimental treatments. Different color shapes represent 'Loma' with green and 'away' with purple. D) Hypothetical ordination results depicting how the experimental treatments in panel (C) alter community composition. The symbols correspond to those used in figure 4, where the green circle represents the Home (Loma) community, the purple symbol, the Away reference community, and the three gray symbols, the intermediate pathways.

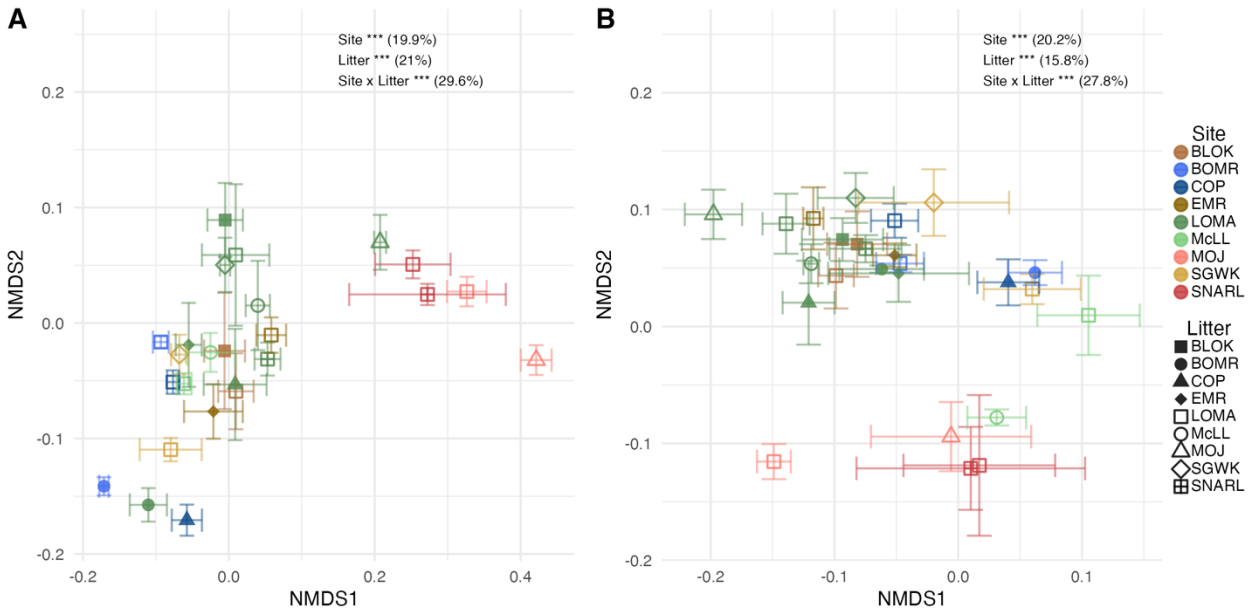

Figure 2.2. Ordination of samples inoculated with the Loma community for (A) the bacterial community at the genus level and (B) the *Curtobacterium* community calculated based on subclade abundances. Points show the centroids for each combination of site and litter. Error bars are 1 ± SE on both axes. Text on the plot shows significance from the PERMANOVA tests and proportion of variance explained by each term.

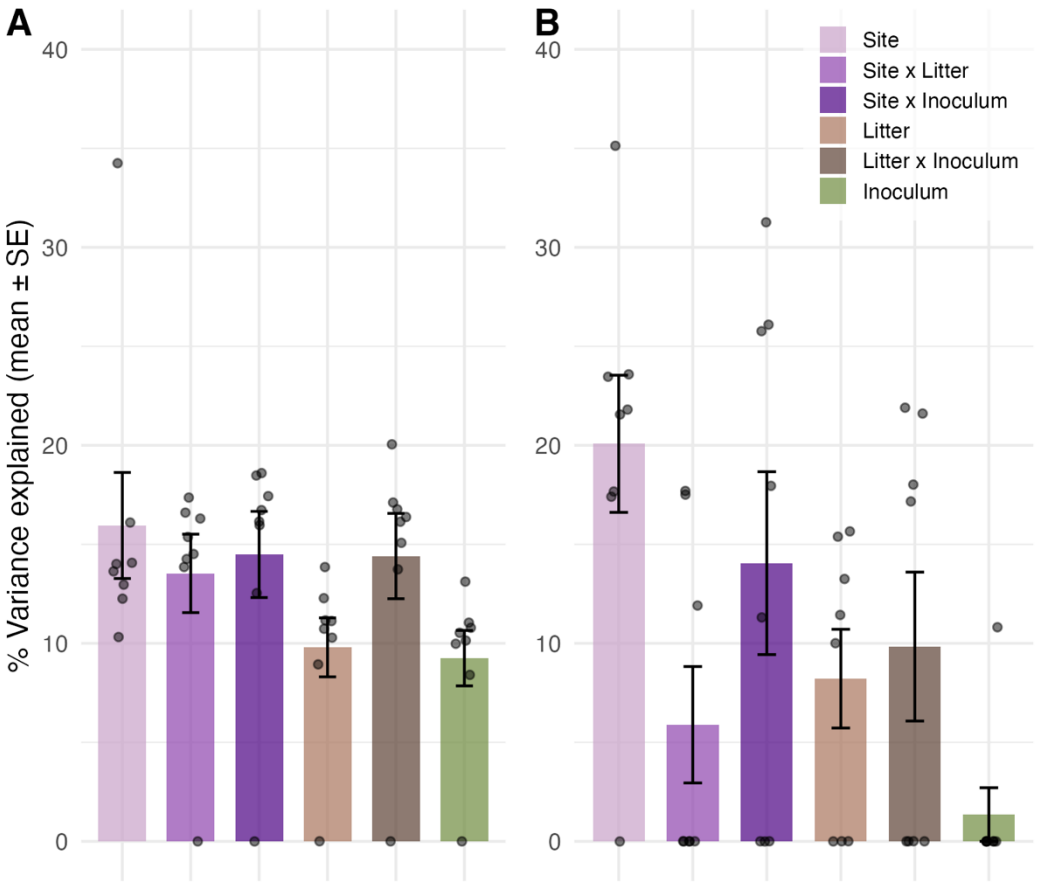

Figure 2.3. Variance explained by each term and their interactions from the PERMANOVA tests for (A) the whole bacterial community and (B) the *Curtobacterium* community (Table S2.3). Points show the observed values for each source of variation. Values observed at 0 represent cases where the PERMANOVA test was not significant (p >0.05) for that term. Means and standard errors are calculated including those cases where variance explained is 0.

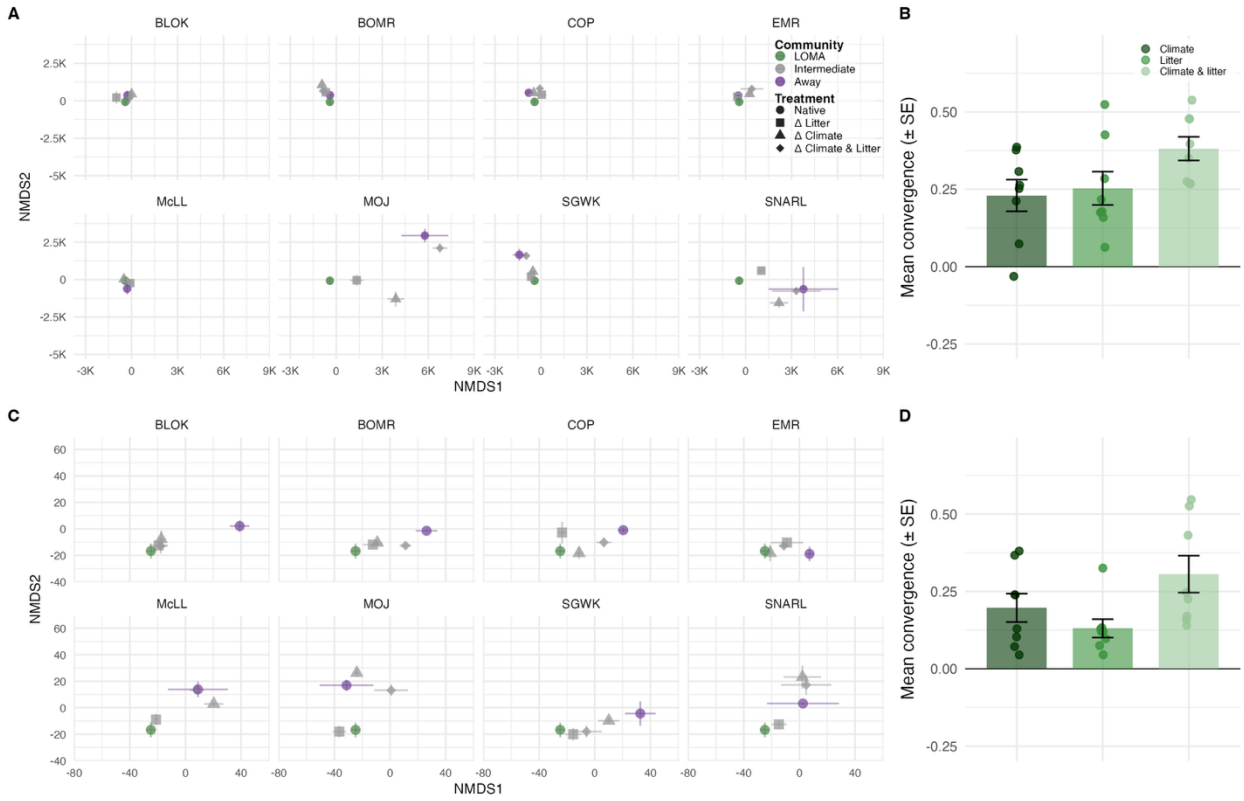

Figure 2.4. Ordination of the eight replicated transplant experiments of samples with the Loma inoculum for (A) the whole bacterial community analyzed at the genus level and (C) the *Curtobacterium* community analyzed at the subclade level. Points show the centroids for each treatment; error bars show the SE on both axes. Community convergence by each path on the experimental design (Fig. 2.1C) for (B) all bacteria and (D) the *Curtobacterium* community.

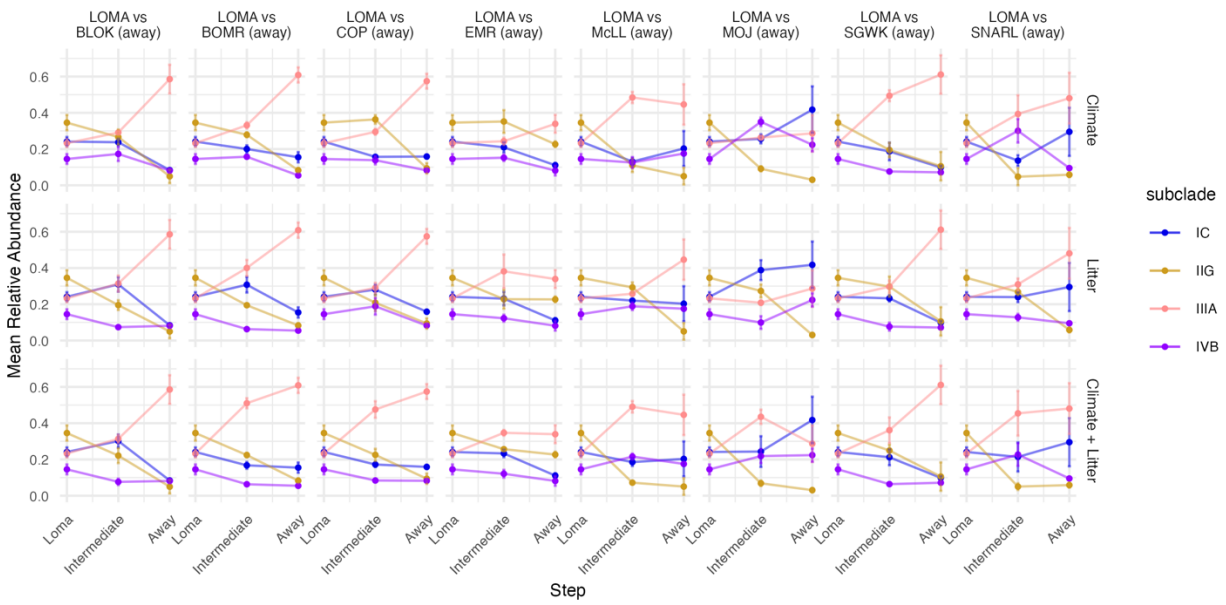

Figure 2.5. Mean relative abundances for each of the top four most abundant subclades/ecotypes of the Loma *Curtobacterium* community for each step in the paths defined by the experimental set up. Climate path shows changes when communities are moved to the away climate (site) but remain in the same site litter. Litter path show changes when communities are moved to the away litter but remain in the same climate (site). Climate + litter path shows changes when communities are moved to another climate (site) and litter. Error bars show SE around the mean. Notice that in all cases the Loma community has the same subclade relative abundances. Likewise, for all paths across each of the pairs, the away communities have the same subclade relative abundances.

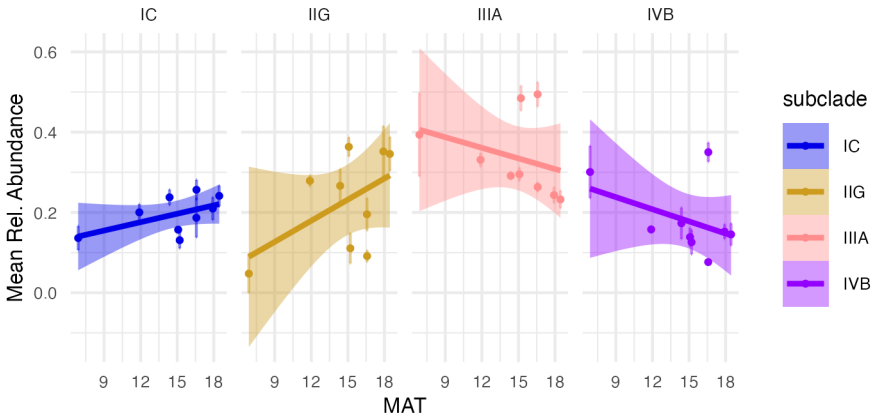

Figure 2.6. Mean relative abundance of the top four most abundant subclades along the MAT (mean annual temperature) across the 9 different sites. Only samples with Loma inoculum and Loma leaf litter are shown to maintain consistency across the comparison.

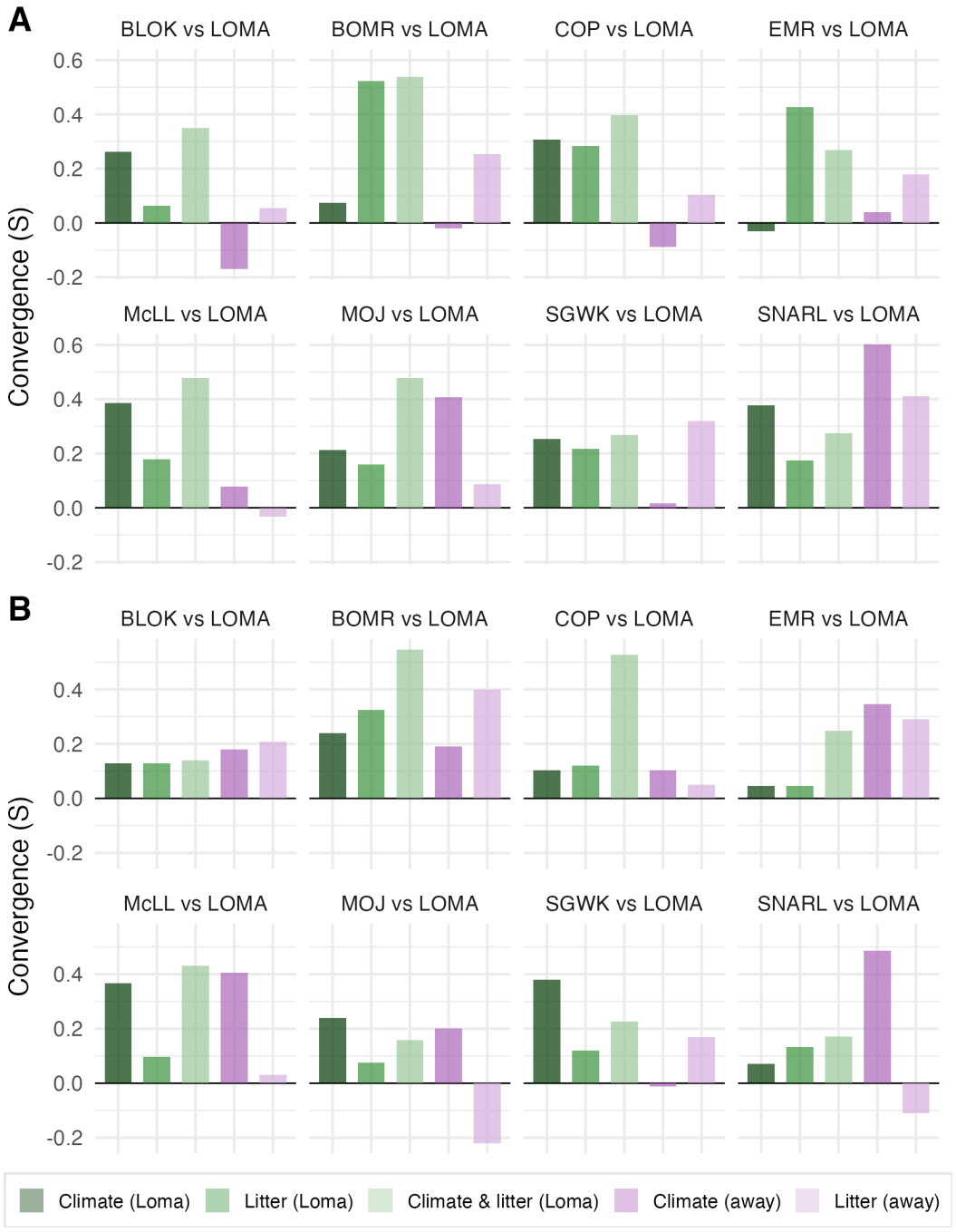

Figure S2.1. Convergence values by pair and path for A) the bacterial community and B) *Curtobacterium* community.

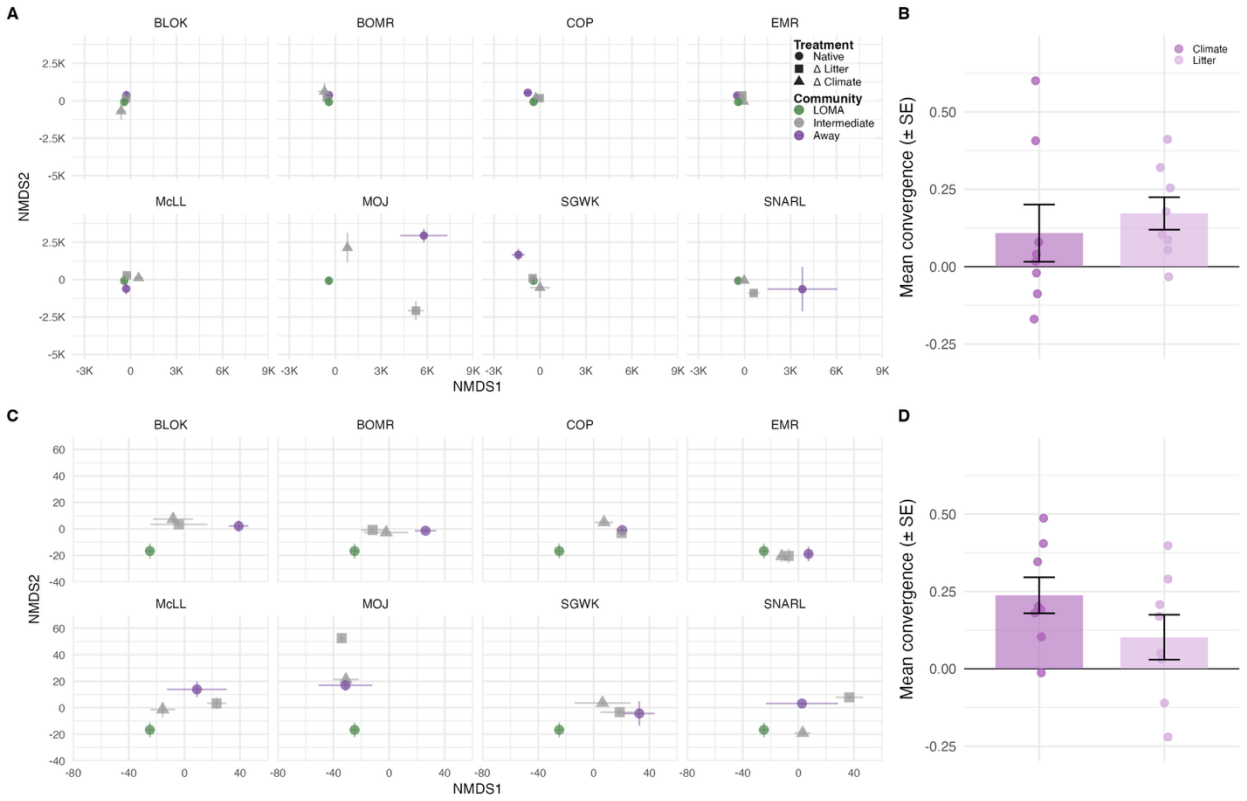

Figure S2.2. Ordination of the eight replicated transplant experiments of samples with the Away inoculum for (A) the whole bacterial community analyzed at the genus level and (C) the *Curtobacterium* community analyzed at the subclade level. Points show the centroids for each treatment, error bars show the SE on both axes. Community convergence by each path on the experimental design (Fig. 1C) for (B) all bacteria and (D) the *Curtobacterium* community.

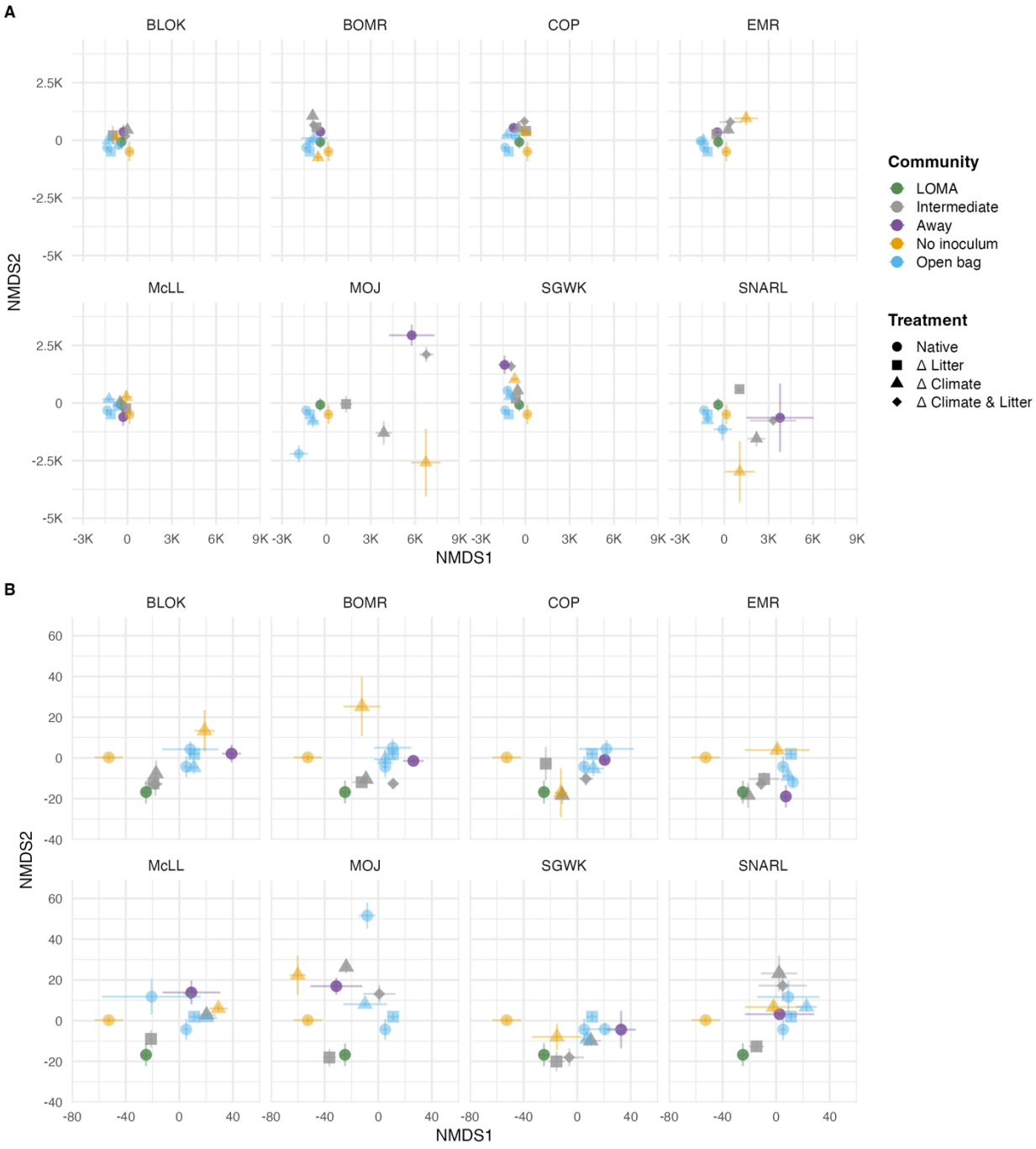

Figure S2.3. Ordination of the eight replicated transplant experiments of samples with the Loma inoculum for (A) the whole bacterial community analyzed at the genus level and (C) the *Curtobacterium* community analyzed at the subclade level. Points show the centroids for each treatment, error bars show the SE on both axes. Experimental controls are shown with blue and yellow symbols. No inoculum represents samples with sterile litter but no inoculum in bags that prevent dispersal. Open bag represents samples for which dispersal was allowed.

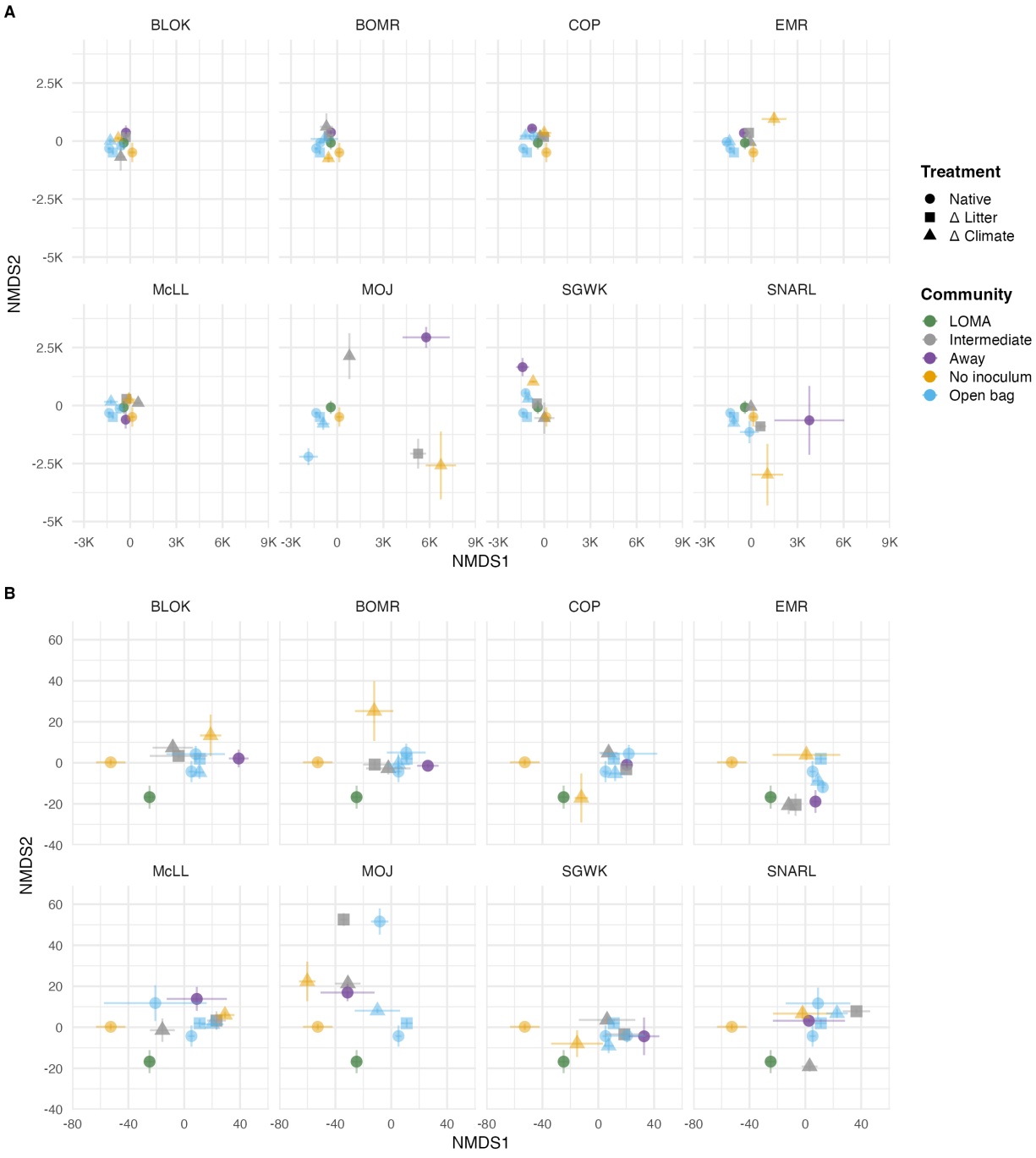

Figure S2.4. Ordination of the eight replicated transplant experiments of samples with the Away inoculum for (A) the whole bacterial community analyzed at the genus level and (C) the *Curtobacterium* community analyzed at the subclade level. Points show the centroids for each treatment, error bars show the SE on both axes. Experimental controls are shown with blue and yellow symbols. No inoculum represents samples with sterile litter but no inoculum in bags that prevent dispersal. Open bag represents samples for which dispersal was allowed.

Table S2.1. Metadata for the 9 sites in the experiment.

Site code	Reserve name	Ecosystem	Plant leaf litter species	Latitude	Longitude	Annual Mean Temperature (°C)	Annual Precipitation (mm)
BLOK	Blue Oak Ranch Reserve	Oak Woodland	*Quercus douglasii*	37.3833	-121.7371	14.37	581.24
BOMR	Bodega Marine Reserve	Coastal Marine	*Artemisa tridentata*	38.3174	-123.0625	11.9	823.14
COP	Coal Oil Point Natural Reserve	Coastal Marine	*Larrea tridentata*	34.4168	-119.88	15.05	497.34
EMR	Emerson Oaks Reserve	Oak Woodland	*Quercus berberidifolia*	33.4689	-117.043	17.92	408.71
LOMA	Loma Ridge Global Change Experiment	Shrubland/Grassland	*Bromus diandrus, Avena fatua*	33.7418	-117.7038	18.45	370.02
MOJ	Sweeney Granite Mountains Desert Research Center	Desert	*Larrea tridentata*	34.7828	-115.6516	16.58	205.26
McLL	McLaughlin Natural Reserve	Grassland	*Avena barbata, Bromus diandrus*	38.8686	-122.4142	15.24	944.66
SGWK	Sedgwick Reserve	Oak Woodland	*Quercus douglasii*	34.693	-120.0479	16.57	543.69
SNARL	Sierra Nevada Aquatic Research Laboratory	Desert	*Artemisa tridentata*	37.6131	-118.831	6.95	333.49

Table S2.2. PERMANOVA results for the effect of site and litter on samples inoculated with the Loma community.

Source	df	SS	MS	Pseudo-F	P(perm)	% variance explained
PERMANOVA: Bacterial community						
Site	8	0.69422	8.68E-02	3.4033	**0.001**	20.16%
Litter	8	0.50394	6.30E-02	2.4705	**0.001**	15.82%
Site x Litter	8	0.66685	8.34E-02	3.2691	**0.001**	27.78%
Res	72	1.8359	2.55E-02			
PERMANOVA:*Curtobacterium* community						
Site	8	2.2175	0.27718	4.5653	**0.001**	19.93%
Litter	8	2.4033	0.30041	4.9478	**0.001**	21.04%
Site x Litter	8	2.3862	0.29828	4.9127	**0.001**	29.61%
Res	72	4.3715	6.07E-02			

Table S2.3. Mean variance explained from the PERMANOVA results by the each experimental treatment and its interactions for both community levels.

Source	Mean % variance explained	Community level
Climate	20.08	*Curtobacterium*
Climate x Litter	5.89	*Curtobacterium*
Climate x Inoculum	14.05	*Curtobacterium*
Litter	8.22	*Curtobacterium*
Litter x Inoculum	9.84	*Curtobacterium*
Inoculum	1.35	*Curtobacterium*
Climate	15.95	Bacterial community
Climate x Litter	13.53	Bacterial community
Climate x Inoculum	14.49	Bacterial community
Litter	9.80	Bacterial community
Litter x Inoculum	14.41	Bacterial community
Inoculum	9.25	Bacterial community

Table S2.4. Mean convergence by path and adjusted p values of bootstraping test testing weather the observed mean value is significantly different from 0.

Community	Path	Mean convergence (S)	Adjuste *p* (FDR)
Bacterial community	Climate (LOMA)	0.2302	0.0002
	Climate + litter (LOMA)	0.3815	0.0002
	Litter (LOMA)	0.2534	0.0002
	Litter (away)	0.1721	0.0006
	Climate (away)	0.1083	0.2138
Curtobacterium	Climate (LOMA)	0.1969	0.0001
	Climate (away)	0.2375	0.0001
	Climate + litter (LOMA)	0.3059	0.0001
	Litter (LOMA)	0.1308	0.0001
	Litter (away)	0.102	0.1366

Table S2.5. Mean convergence and standard error (SE) by pairs and community levels.

Pair	Bacterial community	*Curtobacterium* community	Bacterial community SE	*Curtobacterium* community SE
BLOK vs LOMA	0.112	0.157	0.091	0.016
BOMR vs LOMA	0.274	0.34	0.114	0.063
COP vs LOMA	0.201	0.181	0.086	0.087
EMR vs LOMA	0.176	0.195	0.081	0.063
MOJ vs LOMA	0.269	0.091	0.075	0.082
McLL vs LOMA	0.218	0.266	0.095	0.084
SGWK vs LOMA	0.215	0.177	0.052	0.064
SNARL vs LOMA	0.368	0.15	0.071	0.097

REFERENCES

Allison, S. D., Lu, Y., Weihe, C., Goulden, M. L., Martiny, A. C., Treseder, K. K., & Martiny, J. B. H. H. (2013). Microbial abundance and composition influence litter decomposition response to environmental change. *Ecology*, *94*(3), 714–725. https://doi.org/10.1890/12-1243.1

Archer, E. (2011, August 6). rfPermute: Estimate Permutation p-Values for Random Forest Importance Metrics. In *CRAN: Contributed Packages*. Comprehensive R Archive Network. https://doi.org/10.32614/CRAN.package.rfPermute

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Barbour, K. M., & Martiny, J. B. H. (2024). Investigating eco-evolutionary processes of microbial community assembly in the wild using a model leaf litter system. *The ISME Journal*, *18*(1), wrae043. https://doi.org/10.1093/ismejo/wrae043

Barbour, K. M., Weihe, C., Allison, S. D., & Martiny, J. B. H. (2022). Bacterial community response to environmental change varies with depth in the surface soil. *Soil Biology and Biochemistry*, *172*, 108761. https://doi.org/10.1016/j.soilbio.2022.108761

Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences*, *115*(25), 6506–6511. https://doi.org/10.1073/pnas.1711842115

Belotte, D., Curien, J. B., Maclean, R. C., & Bell, G. (2003). AN EXPERIMENTAL TEST OF LOCAL ADAPTATION IN SOIL BACTERIA. *Evolution*, *57*(1), 27–36. https://doi.org/10.1111/j.0014-3820.2003.tb00213.x

Bushnell, B. (2016). *BBMap Short Read Aligner* [Computer software].

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 1–9. https://doi.org/10.1186/1471-2105-10-421/FIGURES/4

Carlson, C. A., Morris, R., Parsons, R., Treusch, A. H., Giovannoni, S. J., & Vergin, K. (2009). Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *The ISME Journal*, *3*(3), 283–295. https://doi.org/10.1038/ISMEJ.2008.117

Causevic, S., Tackmann, J., Sentchilo, V., Malfertheiner, L., von Mering, C., & van der Meer, J. R. (2025). Habitat filtering more than microbiota origin controls microbiome transplant outcomes in soil. *The ISME Journal*, *19*(1), wraf162. https://doi.org/10.1093/ismejo/wraf162

Celiker, H., & Gore, J. (2014). Clustering in community structure across replicate ecosystems following a long-term bacterial evolution experiment. *Nature Communications*, *5*(1), 4643. https://doi.org/10.1038/ncomms5643

Challis, R., Richards, E., Rajan, J., Cochrane, G., & Blaxter, M. (2020). BlobToolKit – Interactive Quality Assessment of Genome Assemblies. *G3 Genes|Genomes|Genetics*, *10*(4), 1361–1374. https://doi.org/10.1534/G3.119.400908

Chase, A. B., Arevalo, P., Polz, M. F., Berlemont, R., & Martiny, J. B. H. (2016). Evidence for ecological flexibility in the cosmopolitan genus Curtobacterium. *Frontiers in Microbiology*, *7*(NOV), 1–11. https://doi.org/10.3389/fmicb.2016.01874

Chase, A. B., Gomez-Lunar, Z., Lopez, A. E., Li, J., Allison, S. D., Martiny, A. C., & Martiny, J. B. H. (2018). Emergence of soil bacterial ecotypes along a climate gradient. *Environmental Microbiology*, *20*(11), 4112–4126. https://doi.org/10.1111/1462-2920.14405

Chase, A. B., Karaoz, U., Brodie, E. L., Gomez-Lunar, Z., Martiny, A. C., & Martiny, J. B. H. (2017). Microdiversity of an abundant terrestrial bacterium encompasses extensive variation in ecologically relevant traits. *mBio*, *8*(6), 1–11. https://doi.org/10.1128/mBio.01809-17

Chase, A. B., Weihe, C., & Martiny, J. B. H. (2021). Adaptive differentiation and rapid evolution of a soil bacterium along a climate gradient. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(18). https://doi.org/10.1073/pnas.2101254118

Chen, Y., Ye, W., Zhang, Y., & Xu, Y. (2015). High speed BLASTN: an accelerated MegaBLAST search tool. *Nucleic Acids Research*, *43*(16), 7762–7768. https://doi.org/10.1093/NAR/GKV784

Chomel, M., Guittonny-Larchevêque, M., Fernandez, C., Gallet, C., DesRochers, A., Paré, D., Jackson, B. G., & Baldy, V. (2016). Plant secondary metabolites: A key driver of litter decomposition and soil nutrient cycling. *Journal of Ecology*, *104*(6), 1527–1541. https://doi.org/10.1111/1365-2745.12644

Choudoir, M. J., & Buckley, D. H. (2018). Phylogenetic conservatism of thermal traits explains dispersal limitation and genomic differentiation of Streptomyces sister-taxa. *ISME Journal*, *12*(9), 2176–2186. https://doi.org/10.1038/s41396-018-0180-3

Clausen, J., Keck, D. D., & Hiesey, W. M. (1940). Experimental studies on the nature of species. I. Effect of varied environments on western North American plants. *Experimental Studies on the Nature of Species. I. Effect of Varied Environments on Western North American Plants.*, (520).

Clausen, J., Keck, D. D., & Hiesey, W. M. (1947). Heredity of Geographically and Ecologically Isolated Races. *The American Naturalist*, *81*(797), 114–133.

Cohan, F. M. (2001). Bacterial Species and Speciation. *Systematic Biology*, *50*(4), 513–524. https://doi.org/10.1080/10635150118398

Cohan, F. M. (2002). What are Bacterial Species? *Annual Review of Microbiology*, *56*(1), 457–487. https://doi.org/10.1146/annurev.micro.56.012302.160634

Connor, N., Sikorski, J., Rooney, A. P., Kopac, S., Koeppel, A. F., Burger, A., Cole, S. G., Perry, E. B., Krizanc, D., Field, N. C., Slatont, M., & Cohan, F. M. (2010). Ecology of speciation in the genus Bacillus. *Applied and Environmental Microbiology*, *76*(5), 1349–1358. https://doi.org/10.1128/AEM.01988-09

Cornelissen, J. H. C., Quested, H. M., Van Logtestijn, R. S. P., Pérez-Harguindeguy, N., Gwynn-Jones, D., Díaz, S., Callaghan, T. V., Press, M. C., & Aerts, R. (2006). Foliar pH as a new plant trait: Can it explain variation in foliar chemistry and carbon cycling processes among subarctic plant species and types? *Oecologia*, *147*(2), 315–326. https://doi.org/10.1007/S00442-005-0269-Z/TABLES/4

Corr, S., Lowe, C., & Vos, M. (2025). Reciprocal transplantation experiments reveal local adaptation of seaweed-associated bacteria. *ISME Communications*, *5*(1), ycaf205. https://doi.org/10.1093/ismeco/ycaf205

Cullen, J. J., Reid, F. M. H., & Stewart, E. (1982). Phytoplankton in the surface and chlorophyll maximum off southern California in August, 1978. *Journal of Plankton Research*, *4*(3), 665–694. https://doi.org/10.1093/PLANKT/4.3.665

Curd, E. E., Martiny, J. B. H., Li, H., & Smith, T. B. (2018). Bacterial diversity is positively correlated with soil heterogeneity. *Ecosphere*, *9*(1), e02079. https://doi.org/10.1002/ecs2.2079

Evtushenko, L. I., & Takeuchi, M. (2006). The Family Microbacteriaceae. In *The Prokaryotes* (pp. 1020–1098). Springer, New York, NY. https://doi.org/10.1007/0-387-30743-5_43

Falkowski, P. G., Fenchel, T., & Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science (New York, N.Y.)*, *320*(5879), 1034–1039. https://doi.org/10.1126/science.1153213

Fierer, N. (2017). Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, *15*(10), 579–590. https://doi.org/10.1038/nrmicro.2017.87

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

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), 00124. https://doi.org/10.1525/elementa.2021.00124

Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research*, *39*(suppl_2), W29–W37. https://doi.org/10.1093/NAR/GKR367

Fortmann-Roe, S. (2015). Consistent and Clear Reporting of Results from Diverse Modeling Techniques: The A3 Method. *Journal of Statistical Software*, *66*(7), 1–23. https://doi.org/10.18637/JSS.V066.I07

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, *28*(23), 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Fuhrman, J. A., Campbell, U., & Campbell, L. (1998). Microbial microdiversity. *Nature*, *393*(6684), 410–411. https://doi.org/10.1038/30839

Gavrilets, S. (2003). Perspective: Models of Speciation: What Have We Learned in 40 Years? *Evolution*, *57*(10), 2197–2215. https://doi.org/10.1111/j.0014-3820.2003.tb00233.x

Glassman, S. I., Weihe, C., Li, J., Albright, M. B. N. 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 of the United States of America*, *115*(47), 11994–11999. https://doi.org/10.1073/pnas.1811269115

Green, J. L., Bohannan, B. J. M., & Whitaker, R. J. (2008). Microbial biogeography: From taxonomy to traits. *Science (New York, N.Y.)*, *320*(5879), 1039–1043. https://doi.org/10.1126/science.1153475

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), 497–506. https://doi.org/10.1038/nrmicro2795

Hawkes, C. V., Waring, B. G., Rocca, J. D., & Kivlin, S. N. (2017). Historical climate controls soil respiration responses to current soil moisture. *Proceedings of the National Academy of Sciences*, *114*(24), 6322–6327. https://doi.org/10.1073/pnas.1620811114

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2010, July). dismo: Species Distribution Modeling. In *CRAN: Contributed Packages*. https://doi.org/10.32614/CRAN.package.dismo

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 119. https://doi.org/10.1186/1471-2105-11-119

Johnson, L. C., Galliart, M. B., Alsdurf, J. D., Maricle, B. R., Baer, S. G., Bello, N. M., Gibson, D. J., & Smith, A. B. (2022). Reciprocal transplant gardens as gold standard to detect local adaptation in grassland species: New opportunities moving into the 21st century. *Journal of Ecology*, *110*(5), 1054–1071. https://doi.org/10.1111/1365-2745.13695

Johnson, Z. I., Zinser, E. R., Coe, A., McNulty, N. P., Woodward, E. M. S., & Chisholm, S. W. (2006). Niche partitioning among Prochlorococcus ecotypes along ocean-scale environmental gradients. *Science*, *311*(5768), 1737–1740. https://doi.org/https://doi.org/10.1126/science.1118052

Kent, A. G., Baer, S. E., Mouginot, C., Huang, J. S., Larkin, A. A., Lomas, M. W., & Martiny, A. C. (2019). Parallel phylogeography of Prochlorococcus and Synechococcus. *The ISME Journal*, *13*(2), 430–441. https://doi.org/10.1038/S41396-018-0287-6

Konstantinidis, K., Ruiz-perez, C., Gerhardt, K., Rodriguez-r, L., Tiedje, J., & Cole, J. (2022). FastAAI : E cient Estimation of Genome Average Amino Acid Identity and Phylum-level relationships using Tetramers of Universal Proteins. *Research Square*, 0–22. https://doi.org/10.21203/rs.3.rs-1459378/v1

Kraemer, S. A., & Boynton, P. J. (2017). Evidence for microbial local adaptation in nature. *Molecular Ecology*, *26*(7), 1860–1876. https://doi.org/10.1111/mec.13958

Kraemer, S. A., & Kassen, R. (2015). Patterns of Local Adaptation in Space and Time among Soil Bacteria. *The American Naturalist*, *185*(3), 317–331. https://doi.org/10.1086/679585

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, *29*(5), 592–599. https://doi.org/10.1111/1365-2435.12345

Langenheder, S., Berga, M., Östman, Ö., & Székely, A. J. (2012). Temporal variation of β-diversity and assembly mechanisms in a bacterial metacommunity. *The ISME Journal*, *6*(6), 1107–1114. https://doi.org/10.1038/ismej.2011.177

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Larkin, A. A., Hagstrom, G. I., Brock, M. L., Garcia, N. S., & Martiny, A. C. (2023). Basin-scale biogeography of Prochlorococcus and SAR11 ecotype replication. *ISME Journal*, *17*(2), 185–194. https://doi.org/10.1038/s41396-022-01332-6

Larkin, A. A., & Martiny, A. C. (2017). Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. *Environmental Microbiology Reports*, *9*(2), 55–70. https://doi.org/10.1111/1758-2229.12523

Lauber, C. L., Hamady, M., Knight, R., & Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and Environmental Microbiology*, *75*(15), 5111–5120. https://doi.org/10.1128/AEM.00335-09

Lauber, C. L., Ramirez, K. S., Aanderud, Z., Lennon, J., & Fierer, N. (2013). Temporal variability in soil microbial communities across land-use types. *The ISME Journal*, *7*(8), 1641–1650. https://doi.org/10.1038/ismej.2013.50

Leibold, M. A., Govaert, L., Loeuille, N., Meester, L. D., & Urban, M. C. (2022). Evolution and Community Assembly Across Spatial Scales. *Annual Review of Ecology, Evolution, and Systematics*, *53*(Volume 53, 2022), 299–326. https://doi.org/10.1146/annurev-ecolsys-102220-024934

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18–22.

Lu, J., Breitwieser, F. P., Thielen, P., & Salzberg, S. L. (2017). Bracken: Estimating species abundance in metagenomics data. *PeerJ Computer Science*, *3*, e104. https://doi.org/10.7717/peerj-cs.104

Malmstrom, R. R., Rodrigue, S., Huang, K. H., Kelly, L., Kern, S. E., Thompson, A., Roggensack, S., Berube, P. M., Henn, M. R., & Chisholm, S. W. (2012). Ecology of uncultured Prochlorococcus clades revealed through single-cell genomics and biogeographic analysis. *The ISME Journal*, 1–15. https://doi.org/10.1038/ismej.2012.89

Martiny, J. B. H., 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., Ovreås, 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, J. B. H., Jones, S. E., Lennon, J. T., & Martiny, A. C. (2015). Microbiomes in light of traits: A phylogenetic perspective. *Science*, *350*(6261), aac9323–aac9323. https://doi.org/10.1126/science.aac9323

Martiny, J. B. H., Martiny, A. C., Brodie, E., Chase, A. B., Rodríguez-Verdugo, A., Treseder, K. K., & Allison, S. D. (2023). Investigating the eco-evolutionary response of microbiomes to environmental change. *Ecology Letters*, *26*(November 2022), 81–90. https://doi.org/10.1111/ele.14209

Molnar, C., Casalicchio, G., & Bischl, B. (2018). iml: An R package for Interpretable Machine Learning. *Journal of Open Source Software*, *3*(26), 786. https://doi.org/10.21105/joss.00786

Moore, L. R., Post, A. F., Rocap, G., & Chisholm, S. W. (2002). Utilization of different nitrogen sources by the marine cyanobacteria Prochlorococcus and Synechococcus. *Limnology and Oceanography*, *47*(4), 989–996. https://doi.org/10.4319/LO.2002.47.4.0989

Moore, L. R., Rocap, G., & Chisholm, S. W. (1998). Physiology and molecular phylogeny of coexisting Prochlorococcus ecotypes. *Nature*, *393*(6684), 464–467. https://doi.org/10.1038/30965

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schäffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics*, *24*(16), 1757–1764. https://doi.org/10.1093/bioinformatics/btn322

Nayfach, S., Rodriguez-Mueller, B., Garud, N., & Pollard, K. S. (2016). An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Research*, *26*(11), 1612–1625. https://doi.org/10.1101/GR.201863.115/-/DC1

Nayfach, S., Roux, S., Seshadri, R., Udwary, D., Varghese, N., Schulz, F., Wu, D., Paez-Espino, D., Chen, I. M., Huntemann, M., Palaniappan, K., Ladau, J., Mukherjee, S., Reddy, T. B. K., Nielsen, T., Kirton, E., Faria, J. P., Edirisinghe, J. N., Henry, C. S., … Eloe-Fadrosh, E. A. (2021). A genomic catalog of Earth's microbiomes. *Nature Biotechnology*, *39*(4), 499–509. https://doi.org/10.1038/S41587-020-0718-6;SUBJMETA=114,326,631;KWRD=COMPUTATIONAL+BIOLOGY+AND+BIOINFORMATICS,MICROBIOLOGY

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

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., & Wagner, H. (2018). vegan: Community Ecology Package. In *R package version 2.4-6* [Computer software]. http://cran.r-project.org/package=vegan

Padfield, D., O'Sullivan, H., & Pawar, S. (2021). rTPC and nls.multstart: A new pipeline to fit thermal performance curves in r. *Methods in Ecology and Evolution*, *12*(6), 1138–1143. https://doi.org/10.1111/2041-210X.13585

Parks, D. H., Chaumeil, P.-A., Mussig, A. J., Rinke, C., Chuvochina, M., & Hugenholtz, P. (2026). GTDB release 10: A complete and systematic taxonomy for 715 230 bacterial and 17 245 archaeal genomes. *Nucleic Acids Research*, *54*(D1), D743–D754. https://doi.org/10.1093/nar/gkaf1040

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Piton, G., Allison, S. D., Bahram, M., Hildebrand, F., Martiny, J. B. H., Treseder, K. K., & Martiny, A. C. (2023). Life history strategies of soil bacterial communities across global terrestrial biomes. *Nature Microbiology*, *8*(11), 2093–2102. https://doi.org/10.1038/s41564-023-01465-0

Reed, H. E., & Martiny, J. B. H. (2007). Testing the functional significance of microbial composition in natural communities. *FEMS Microbiology Ecology*, *62*(2), 161–170. https://doi.org/10.1111/j.1574-6941.2007.00386.x

Rousk, J., & Bååth, E. (2007). Fungal and bacterial growth in soil with plant materials of different C/N ratios. *FEMS Microbiology Ecology*, *62*(3), 258–267. https://doi.org/10.1111/j.1574-6941.2007.00398.x

Scales, N. C., Chase, A. B., Finks, S. S., Malik, A. A., Weihe, C., Allison, S. D., Martiny, A. C., & Martiny, J. B. H. (2022). Differential Response of Bacterial Microdiversity to

Simulated Global Change. *Applied and Environmental Microbiology*, *88*(6). https://doi.org/10.1128/AEM.02429-21/SUPPL_FILE/AEM.02429-21-S0001.PDF
Scheuerl, T., Hopkins, M., Nowell, R. W., Rivett, D. W., Barraclough, T. G., & Bell, T. (2020). Bacterial adaptation is constrained in complex communities. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-14570-z
Schloter, M., Lebuhn, M., Heulin, T., & Hartmann, A. (2000). Ecology and evolution of bacterial microdiversity. *FEMS Microbiology Reviews*, *24*(5), 647–660. https://doi.org/10.1016/S0168-6445(00)00051-6
Schoolfield, R. M., Sharpe, P. J. H., & Magnuson, C. E. (1981). Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. *Journal of Theoretical Biology*, *88*(4), 719–731. https://doi.org/10.1016/0022-5193(81)90246-0
Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*, 539. https://doi.org/10.1038/MSB.2011.75/SUPPL_FILE/MSB201175.REVIEWER_COMMENTS.PDF
Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351
Sprouffske, K., & Wagner, A. (2016). Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*, *17*(1), 1–4. https://doi.org/10.1186/S12859-016-1016-7/FIGURES/1
Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/BIOINFORMATICS/BTU033
Stocker, R. (2012). Marine microbes see a sea of gradients. *Science*, *338*(6107), 628–633. https://doi.org/10.1126/SCIENCE.1208929/ASSET/FE73DB31-FDFF-49D5-89A6-5316AA35287D/ASSETS/GRAPHIC/338_628_F3.JPEG
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., Zech Xu, Z., Jiang, L., … Knight, R. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463. https://doi.org/10.1038/nature24621
Turesson, G. (1922). The Genotypical Response of the Plant Species to the Habitat. *Hereditas*, *3*(3), 211–350. https://doi.org/10.1111/j.1601-5223.1922.tb02734.x
VanInsberghe, D., Maas, K. R., Cardenas, E., Strachan, C. R., Hallam, S. J., & Mohn, W. W. (2015). Non-symbiotic Bradyrhizobium ecotypes dominate North American forest soils. *The ISME Journal*, *9*(11), 2435–2441. https://doi.org/10.1038/ISMEJ.2015.54
Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, *85*(2), 183–206. https://doi.org/10.1086/652373
Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. *FEMS Microbiology Reviews*, *37*(6), 936–954. https://doi.org/10.1111/1574-6976.12023

Wadgymar, S. M., DeMarche, M. L., Josephs, E. B., Sheth, S. N., & Anderson, J. T. (2022). Local Adaptation: Causal Agents of Selection and Adaptive Trait Divergence. *Annual Review of Ecology, Evolution, and Systematics*, *53*(Volume 53, 2022), 87–111. https://doi.org/10.1146/annurev-ecolsys-012722-035231

Waring, B., & Hawkes, C. V. (2018). Ecological mechanisms underlying soil bacterial responses to rainfall along a steep natural precipitation gradient. *FEMS Microbiology Ecology*, *94*(2), fiy001. https://doi.org/10.1093/femsec/fiy001

Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, *20*(1), 257. https://doi.org/10.1186/s13059-019-1891-0

Zhou, Z., Wang, C., & Luo, Y. (2020). Meta-analysis of the impacts of global change factors on soil microbial diversity and functionality. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-16881-7

Zimmerman, A. E., Martiny, A. C., & Allison, S. D. (2013). Microdiversity of extracellular enzyme genes among sequenced prokaryotic genomes. *ISME Journal*, *7*(6), 1187–1199. https://doi.org/10.1038/ismej.2012.176

APPENDIX 1: Supplementary Methods for Chapter 1

Culture media preparation and pH adjustment for thermal and pH performance assays of *Curtobacterium* isolates

Media preparation

Growth assays were performed in a modified M63 minimal medium ("synthetic grass medium"; adapted from A. Rodríguez-Verdugo). All stock and base solutions were sterilized by filtration through a 0.22 μm PES membrane and light-sensitive components (iron, glucose, micronutrients) were stored at 4 °C protected from light. The complete 1× medium consisted of 1× M63 base, 0.02% (w/v) peptone, 4 mM glucose as the sole carbon source, and 1 mM $MgSO_4$. The 10× M63 base medium contained, per liter, 107.2 g $K_2HPO_4 \cdot 3H_2O$, 52.4 g KH_2PO_4, and 20 g $(NH_4)_2SO_4$, supplemented with 5 mL of an iron solution (1 mg mL^{-1} $FeSO_4 \cdot 7H_2O$ in 0.01 M HCl) and 50 mL of an SPV-4 micronutrient solution (per liter: 1 mg $CaCl_2 \cdot 2H_2O$, 5 mg $FeCl_3 \cdot 6H_2O$, 1 mg $MnSO_4 \cdot H_2O$, 2 mg $ZnCl_2$, 0.2 mg $CuSO_4 \cdot 5H_2O$, 0.4 mg $CoCl_2 \cdot 6H_2O$, and 0.2 mg $Na_2MoO_4 \cdot 2H_2O$). The base medium was filter-sterilized, stored at 4 °C in the dark, and used within two months.

pH adjustment

Thermal performance media

For thermal performance curves, the complete 1× medium was buffered with potassium phosphate (as formulated in the 10× M63 base) and adjusted to pH 6.8 with sterile NaOH prior to filter sterilization.

pH performance media

For pH performance curves, 24 media spanning pH 3.2 to 10.1 in 0.3-unit increments were prepared using a phosphate-free 10× M63 base, so that the added potassium phosphate buffer was the sole phosphate source and the only intrinsic determinant of medium pH. For each target pH, monobasic (KH_2PO_4) and dibasic (K_2HPO_4) phosphate solutions (each 0.2 M) were combined at the proportions tabulated in Ausubel et al. (1999; Appendix 1, Table A.1.3), scaled to a 50 mL final volume and a final phosphate concentration of 0.1 M.

Each 50 mL medium was assembled by adding the phosphate buffer at the pH-specific ratio, followed by phosphate-free 10× M63 base, glucose, and $MgSO_4$, and brought to a final volume of 50 mL with sterile DI water. Because the tabulated phosphate ratios buffer only across approximately pH 5.7–8.0, target pH values in the acidic (<5.7) and alkaline (>8.0) ranges were reached by titrating with KOH or HCl. The pH of each completed medium was measured at its final 50 mL volume to confirm that the target value was retained after all components were added, immediately before filter sterilization.

Reference

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K. (Eds.). (1999). *Short Protocols in Molecular Biology* (4th ed.). John Wiley & Sons.