DISSERTATION

UNRAVELING KEY DRIVERS OF MICROBIAL COMMUNITY ASSEMBLY AND IMPACTS ON MICROBIAL FUNCTION

Submitted by

Jennifer Doyle Rocca

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2015

Doctoral Committee:

Advisor: Matthew D. Wallenstein

M. Francesca Cotrufo
Alan K. Knapp
Melinda D. Smith

ABSTRACT

UNRAVELING KEY DRIVERS OF MICROBIAL COMMUNITY ASSEMBLY AND IMPACTS ON MICROBIAL FUNCTION

Microorganisms are the key drivers of biogeochemical cycling in soils, and consequentially their activity constrains both soil nutrient availability and ecosystem productivity. Despite their importance, we still lack basic understanding of the key factors that structure microbial communities and how this in turn impacts their function. The high taxonomic and functional diversity, microscopic size, rapid turnover, and elusive culturability of many microorganisms has hampered progress on identifying these influential factors. However, rapid declines in DNA sequencing costs have allowed for more thorough and widespread accounting of soil microbial communities. The primary goal of my research was to examine the influence of various aspects of scale on soil bacterial community structure and function. Specifically, I address the influence of taxonomic scale on indices of community structure, which are used as proxies for community assembly. Secondly, I assess the impact of the scale of aboveground plant community structure and its consequential impacts on soil function, and finally I assess the relationship strength of genetic functional capacity of microorganisms and whether it can scale up as a useful index of actual functional output.

First, I assessed the sensitivity of two indices of phylogenetic community dispersion to the taxonomic scale of estimation. Beyond just asking ‘who is where’ as a natural history-based tally of what microorganisms reside where, microbial ecologists are now identifying the specific mechanisms regulating the assembly of microbial communities. The main purpose of examining

community dispersion in phylogenetic tree space is to understand the relative influence of biotic and abiotic assembly processes on shaping community structure. Phylogenetic clustering of taxa is indicative of environmental filtering, such as salt tolerance or anaerobic conditions, while phylogenetic overdispersion suggests potentially strong competitive interactions and niche differentiation. The influence of taxonomic scale on dispersion indices was previously assessed in woody plant communities, in computer-generated communities, and in microbial communities over a small range of taxonomic scales or phylogenetic inclusiveness. To test the impact of taxonomic scale over a broader range of scales, I used soil bacterial communities from a field site exposed to seven years of elevated CO_2 and temperature. Overall there was an increase in phylogenetic clustering of communities with increased taxonomic scale. Overdispersion was rare in my dataset, and occurred mainly in fine taxonomic scales, like the bacterial family to class range. The implications for this study are that the typical domain-scale (i.e. Bacteria, Eukaryotic, or Archaebacteria) assessment of phylogenetic dispersion estimations of microbial communities is likely too broad for valid interpretation. Additionally, because microorganisms comprise vast functional diversity, it is conceivable that different microbial clades will differentially respond to habitat conditions. The assembly processes governing microbial community assembly are likely clade specific, and cannot simply be assessed with a single assessment at the domain taxonomic scale.

I used the information developed from the impact of taxonomic scale on phylogenetic dispersion to identify primary assembly processes regulating soil bacterial clades using the same PHACE soil bacterial communities employed in the taxonomic scale chapter. This study provided an excellent setup for examining the impacts of long-term elevated CO_2, warming, and interannual variability in soil moisture availability on bacterial community structure. I found

little influence of the experimentally implemented factors, CO_2 and warming, but collection date had a strong impact on community structure. Other studies have consistently demonstrated the importance of moisture as a key driver of soil microbial community structure, but this was not directly supported by my study. Soil moisture, averaged separately over week, month, and year prior to soil harvest had little to no effect on the structure of the bacterial communities. This indicates that some other aspect of seasonality, or indirectly by soil moisture availability, is driving these bacterial communities at PHACE. Since southern Wyoming is situated in a habitat exposed to strong seasons, the plant communities are likely to shift their functionality in response to the season, like: the timing of root exudation of photosynthates, plant uptake of soil nutrients, and plant senescence. In this way, the influence of forecasted elevated CO_2 and warming may not have direct impacts on soil microbial communities, but the climate-mediated effects on plant functionality could be substantial controllers of soil microorganisms.

Next, I examined both inter- and intraspecific variability in leachate chemistry from PHACE, and its consequential impacts on soil microbial community structure and function. Plants are well known drivers of soil microorganisms, through root exudation and direct plant tissue inputs, but less is known about the impact of the upwards of 16% of leaf tissue that permeates the soil profile as water-soluble leachate. Forecasted climate change may impose a two-fold shift in plant leachate chemistry that enters the soil via intraspecific shifts in chemistry and shifts in the overall plant community structure. In a series of laboratory incubations, I addressed the inter- and intraspecific differences in leachate from plants at the Prairie Heating and Carbon Dioxide Enrichment (PHACE) site (Cheyenne, WY), where northern mixed grassland was exposed to seven years of elevated CO_2 and warming. Overall, I found evidence that shifts in plant chemistry affected soil microbial function and bacterial community structure.

Specifically, I found that the invasive forb species, *Linaria dalmatica*, exhibited distinct leachate chemistry compared to the other plant species, differentially impacted soil function, and imposed stronger influences on the bacterial community structure relative to other species' leachates. Though plant litter leachate chemistry did not vary significantly, the separation in the chemistry was primarily driven by a handful of metabolites: porphine, fumaric acid, shikimic acid and galacturonic acid. The main driver of bacterial community structure of the soil incubation was laboratory incubation time, though bacterial communities treated with *L. dalmatica* leachate also separated from the other communities in ordination space. This consistent shift in bacterial community by incubation length is likely due to the drastic impact of altering soil physical structure and water availability imposed by the incubation setup itself. The distinct impact of the invasive *L. dalmatica* leachate on soil microbial structure and function may highlight the potential influence of plant invasions on microbial communities and how they can present substantial soil functional consequences under future climate. The interspecific differences in plant leachate chemistry and consequential soil function were not significant. Therefore, the PHACE treatments do not directly influence within plant species leachate chemistry such that soil functional impacts are detectable. However, because there are differences among species, and my collaborators have found significant impact of the PHACE treatments on plant community structure, the impacts of forecasted climate could impact soil function through plant community shifts. Broadly, these studies assess the importance of plant leachate in regulating the soil functional response to current and future climate.

I also tested the influence of late season senesced plant litter to see if interspecific differences continued after the initial leachate was lost from the plant biomass. Using a similar set of plant species from adjacent plots near the PHACE field site, I assessed interspecific

variability among leachate and potential soil functional impacts. There was a minimal correlation between leachate chemistry and plant relatedness, yet variance among the plant leaf tissue and among leachate never exceeded 1%. The effects of the senesced plant litter leachate on soil function was consequently not significantly different among plant species. Therefore, while I found significant differences among plant species' leachate derived from fresh leachate, this effect dissipates throughout senescence, and consequently the climate-induced shifts in plant community leachate and the soil functional impacts may only persist early in senescence and dissipate with time.

Finally, I conducted a meta-analysis to assess the scalability of microbial functional gene abundance as a valid proxy for corresponding function. For any enzyme-catalyzed reaction to occur, the matching gene and transcript are necessary prerequisites. It is therefore reasonable to assume a positive relationship between the abundance of protein-encoding genes and/or transcripts for enzymes catalyzing biogeochemical processes and the process itself. I tested this assumption by conducting a meta-analysis and found that of 415 studies that quantified protein-encoding genes or transcripts abundance, only 59 reported both gene or transcript abundance and the corresponding biogeochemical process. This result suggests that the relationship between molecular information and corresponding biogeochemical rate is commonly assumed yet rarely evaluated. Of the 224 individual relationships between gene or transcript abundance and process that were evaluated, from the 59 studies, I found a significant positive correlation, yet extensive variability amongst individual effect sizes were not explained by study site characteristics, suggesting that other ecological and methodological factors may affect this relationship, or that this type of molecular data does not directly reflect process flux and rate. These findings contradict how gene and transcript abundance information are currently being interpreted, and

highlight the need for fundamental research on the factors that control transcription, translation, and enzyme function in natural systems in order to better link genomic and transcriptomic data to ecosystem processes.

The main goal of my research was: to identify key factors that drive soil microbial community structure and function, assess the impact of taxonomic scale on phylogenetic dispersion, and examine the correlation between functional gene abundance and function. I found that the phylogenetic dispersion of soil bacterial communities is influenced by taxonomic scale, and seasonality is the primary driver of these bacterial communities at PHACE. Calculating phylogenetic dispersion patterns throughout tree space will be a useful exploratory tool for understanding what microbial clades are primarily structured by environmental filtering vs. competition under various habitat conditions. In assessing the influence of plant community on soil microbial community structure and function, I found interspecific differences in leachate on soil. In particular, one invasive plant exhibited distinct chemistry as well as strong impacts on soil microbial community structure and function. Scaling this effect up to the habitat level, there is a potential link between climate-mediated shifts in plant community and soil function. Therefore, future research could focus on unraveling the mechanisms behind why seasonality has a strong effect on grassland soil microorganisms and examining the effects of other invasive species, as there may be something unique about invasive plant litter chemistry in general. Finally, microbial functional gene abundance does not consistently scale up to the corresponding function, and consequently the relationship between these factors is not as straightforward as is commonly assumed.

ACKNOWLEDGEMENTS

I am grateful to many people, from various aspect of my life, who have helped tremendously during my doctoral research. My advisor, Matthew Wallenstein, was integral to all aspects of my degree. When I first arrived, he guided me through the initial determination of what I planned to do, and then prevented me from executing unattainably large laboratory soil incubations. I have a tendency to want to try out new methods of analyzing community data, but Matt was good to remind me to finish what I had already started. As a mentor, he also served as an excellent example of how to balance work and life, and succeed at both while retaining an easy-going attitude.

My other committee members, Francesca Cotrufro, Melinda Smith, and Alan Knapp were also instrumental in guiding my research and helping to synthesize the overall cohesion to the dissertation itself. Though not a formal member of my PhD committee, Ed Hall was essentially my fifth mentor here at CSU. He was helpful in throwing ideas around for my phylogenetic scaling chapter, and helped me immensely with the meta-analysis, as part of the USGS Powell Center working group 'The next generation of ecological indicators: defining which microbial properties matter most to ecosystem function and how to measure them'.

I am thankful for a being a part of a suite of helpful and extremely knowledgeable lab group members. Early on in my PhD, Sarah Evans, Jessica Ernakovich, and Barbara Fricks were excellent collaborators for exchanging our ideas at lab meetings, and also for providing much needed friendship at the beginning of my time here in Fort Collins. Newer lab members, Laurel Lynch, Peter Baas, Megan Machmuller, and Lauren Mancini have been great for discussing 'nodies', albeit probably too much at times, but they always helped me think though whatever

dataset was being discussed at our lab meetings. In addition to Wallenstein lab meetings, the weekly meetings of the Belowground Group were excellent for me to discuss my community-based analyses to a more biogeochemical-based lab group.

I would also like to thank Nishanth Tharayil, Francisco Calderon, and Guy Beresford for helping with various aspects of chemical and DNA sample preparations, and for helping me submit my samples to their respective analytical machines. My research was funded through the National Science Foundation, as a collaborative effort to analyze the impacts of the long-term PHACE treatments on soil function. Thank you to Matt for allowing me to use this funding source during my PhD program.

Finally, I am very thankful to my friends and family for emotional support and life perspective throughout graduate school. To Annie Hildebrand, Blythe Karow, Vicky Huang, Kirsten Holfelder, and Neely Clapp, you guys provided the emotional backbone and comic relief that made this all possible. To my brother and best bud, Daniel, you have been a good sounding board for thinking about the broader impacts of my research, and good friend/best brother throughout. Finally, to my Mom and Dad, you guys have been there for me for my whole life, providing unconditional love, support and feedback on life and science.

TABLE OF CONTENTS

Chapter 1: Introduction

Microorganisms are crucial drivers of soil biogeochemical cycling, and their activity regulates productivity up to the ecosystem scale. Despite their fundamental importance, microbial ecologists are still identifying key factors that structure microbial communities and understanding how these structural shifts in turn impact microbial function. The immense taxonomic and functional diversity, microscopic size, rapid turnover, and elusive culturability of many microorganisms has hampered progress on identifying these influential factors. However, the increased accessibility of DNA sequencing cost has allowed for more thorough and widespread accounting of soil microbial communities.

The primary goal of my research was to examine various aspects of scaling on the structuring of microbial communities. Understanding the influence of taxonomic scale, or phylogenetic breadth, on estimates of community dispersion is key in parsing out what factors primarily contribute to microbial community assembly and how these processes vary in their relative importance among microbial clades. With this assessment of taxonomic scale, I was able to extend this framework to determine the relative influence of abiotic factors on controlling soil microbial community structure and function at the habitat scale. Then, I assess a different aspect of scaling, by determining the influence of biotic factors, namely plant community structure, on soil bacterial community structure and function with a series of laboratory incubations. Finally, I assessed the scalability of the abundance of microbial functional genes for inferring corresponding functional capacity using a meta-analysis.

I assessed two indices of phylogenetic dispersion using a soil bacterial community dataset to examine the influence of taxonomic scale, or phylogenetic inclusiveness, on the dispersion

estimates. Microbial ecologists often use phylogenetic dispersion as a proxy for the relative influence of community assembly processes, including environmental filtering and competition. However, the default domain-level taxonomic scale may be too broad for valid interpretation. Competitive interactions likely operate at much finer taxonomic scales, while habitat filtering may operate at both fine and broad taxonomic scales. Using the information generated from the assessment of taxonomic scale, I applied this to a case study of soil bacterial communities from a field site exposed to forecasted climate change conditions, including long term elevated CO_2 and temperature. I was able to assess differences in the diversity of the bacterial communities, but also the prominent assembly driver of the various bacterial phylogenetic groups. This examination allows us to move beyond just what organisms are present, and how they vary by community, but what ecological processes are driving the structuring of these bacterial communities.

To assess the biotic factors, including the aboveground plant community, that may control soil microbial community structure and function; I conducted a series of laboratory soil incubations. Plant water-soluble leachate may be an important driver of soil microbial community structure and function. It is well known that plants influence their surrounding environment through competition for nutrients, but also through the contribution of plant compounds via root exudation of photosynthates or through litterfall and dead root inputs. However, less is known about the substantial portion of water-soluble plant metabolites from plant litter fall that enters the soil profile. It is well established that plant species often differ in their chemical signatures, and if this translates to differences in leachate chemistry, then plant community structure could strongly influence soil microbial structure and function through shifts in the chemistry of the metabolites entering the soil.

In addition to differences among plant species tissue chemistry, abiotic conditions could influence intraspecific variation in plant tissue chemistry. Forecasted increased atmospheric CO_2 will likely have direct impacts on plant WUE, with scaffolding effects on their tissue chemistry. Increased CO_2 results in dampened plant water stress and increased soil water availability, which could also influence plant chemistry. Increased temperature has an opposing impact on soil water availability. Water stressed plants are known to produce more reactive polyphenolic compounds in response to drought stress, which in turn could impact soil microbial function if these shifts change in the leachate chemistry as well.

Elevated CO_2 and temperature may also have direct impacts on soil microbial structure and function. Enzyme kinetics is constrained by temperature, and the CO_2-mediated increase in soil water availability could increase microbial community connectivity as well as microbial viability. Interannual variability in water availability has been shown to overshadow the soil impacts that induced elevated CO_2 and warming have on soil microbial community structure and function. Therefore, we expect patterns of water availability to be a prominent driver of soil microbial structure.

Finally, in addition to soil microbial structure, microbial ecologists are trying to link community structure with function. One idea that has been widely accepted, yet not rigorously tested, is the use of microbial protein-encoding gene abundance as a proxy for corresponding functional potential. I conducted a meta-analysis of studies that assessed both protein-encoding gene abundance and corresponding function to assess the validity of this correlation. This is an important assessment of the scaling of processes. It is widely accepted that there is some degree of linearity between DNA transcription and downstream enzymatic activity, but many factors may obscure the correlation.

Therefore, I addressed the following main questions with my research:

1. Does taxonomic scale influence estimations of phylogenetic dispersion in soil bacterial communities?
2. What are the impacts of long term elevated CO_2 and warming, and interannual variability in water availability on soil bacterial community structure?
3. Does plant chemistry vary interspecifically and does this impact soil function?
4. What are the impacts of altered abiotic conditions on plant leachate, and how do these changes translate to shifts in soil microbial community structure and function?
5. Is microbial functional gene abundance a valid proxy for corresponding function?

In the following chapters, I address these questions of scale by examination of field collections of intact soil microbial communities, through a series of two laboratory soil incubations, and a meta-analytic assessment of other studies in determining key drivers that structure microbial communities and the consequences for microbial function.

Chapter 2: Taxonomic breadth matters: assessment of bacterial phylogenetic community structure at high taxonomic inclusion obscures fine-scale patterns of dispersion

As microbial sequence data from a range of environments continues to amass, we have the opportunity to unravel the assembly processes shaping microbial community structure using tools developed to study the community ecology of macroorganisms. Phylogenetic community ecology largely centers on a theory that closely-related taxa possess a higher degree of niche overlap compared to more distantly related taxa. The incorporation of phylogeny into community analyses enables assessment of the relative influences of biotic and abiotic assembly processes on community structure. However, we hypothesize that the taxonomic scale, or phylogenetic breadth, typically used in microbial studies is too broad for valid interpretation when applied to bacterial sequence data. The primary objective of our study was to understand the influence of taxonomic scale on estimations of bacterial community structure. We used the phylogenetic community indices of dispersion, Net Relatedness Index (NRI) and Nearest Taxon Index (NTI), from Phylocom (V 4.2, Webb *et al* 2002) to assess patterns of phylogenetic dispersion throughout tree space. Overall, taxonomic scale strongly influenced dispersion estimates. We found an asymptotically positive relationship between taxonomic scale and NRI, and a strong positive correlation with NTI. At the bacterial domain and phylum scale, clustering prevailed, with 95% and 100% occurrence for NRI and NTI, respectively. In contrast, at finer scales, approximately corresponding to the bacterial family and genus levels, overdispersion was more prevalent. There was a significant positive correlation between NRI and log-transformed community size for nodes comprising less than 200 taxa ($r^2=0.26$, $p<0.0001$). Not surprisingly, if niche conservatism holds, competition should be most influential at finer taxonomic scales,

whereas clustering, associated with habitat filtering, might be expected when more taxa are included in the dispersion estimate. Finally, NTI, which only uses terminal branch lengths in the estimates of dispersion, produces more than two-fold overestimates of dispersion at larger taxonomic scales, such as a bacterial phylum or the domain. This overestimation is likely a mathematical artifact of not including deeper branch information in the calculation, but it does challenge the continued usage of this index at broad taxonomic scales to assess microbial community phylogenetic structure. Our results demonstrate that broad assessment of phylogenetic community structure can obscure fine-scale contrasting patterns. Importantly, our work highlights the importance of considering taxonomic scale when assessing microbial phylogenetic community structure.

Introduction

Soil microbial communities are among the most diverse communities on Earth. Both bacteria and fungi exhibit biogeographical patterns (Hughes-Martiny *et al* 2006, Fierer & Jackson 2006, Angel *et al* 2010), which suggest that the microorganisms present in any particular environment are a select subset of the total global biodiversity. Habitat filtering, dispersal limitations, and biotic interactions determine the specific subset of microorganisms in a local community relative to the global microbial pool (Hughes-Martiny *et al* 2006).

Abiotic factors are key drivers of microbial biogeography. pH is the primary regulator of microbial community structure in soils (Fierer & Jackson 2006, Lauber *et al* 2009). Water (Williams & Rice 2007, Collins *et al* 2008, Angel *et al* 2010, Bell *et al* 2014), temperature (DeAngelis *et al* 2015), and nutrient availability (Marschner *et al* 2003, Thebault *et al* 2014) are also key environmental filters to microbial community structure and function. However, at the

local scale, biotic interactions are likely more important determinants of microbial community structure. Mutualistic relationships exist between many plants and their root-associated microbiomes (Aleklett *et al* 2015, Bell *et al* 2015), and specific microbiome-nectar-pollinator relationships determine local microbial communities (Vannette *et al* 2012). Competitive interactions among microorganisms (Monod 1949, Fredrickson & Stephanopoulos 1981, Hibbing *et al* 2010, Koeppel & Wu 2014) and neighboring plants (Rousk *et al* 2007, Schimel *et al* 1989) can also influence local microbial community structure. However, given their immense functional and genetic diversity, microscopic size and rapid turnover, unraveling the relative roles of abiotic and biotic community assembly processes is challenging.

The exponential decrease in DNA and RNA sequencing costs has enabled deep assessment of soil microbial communities. To gain insight into the relative influence of specific drivers of microbial communities using nucleic acid datasets, microbial ecologists have adopted phylogenetic community ecological principles, rooted in plant community ecology, to infer mechanisms controlling microbial community assembly (Webb *et al* 2002, Cavender-Bares *et al* 2004, 2009). These approaches hinge on the concept of 'niche conservatism' (Losos *et al* 2008), which posits that closely related organisms have, on average, more overlap in traits of specific competitive ability compared to distantly related organisms (Mayfield & Levine 2010). These phylogenetically conserved traits result in more functional overlap, which in turn leads to stronger competitive interactions. Competitive exclusion of sister taxa, due to high overlap of niche space and competitive abilities, primarily results in overdispersion patterns (or evenness) of community structure (Weiher & Keddy 1995, Webb *et al* 2002).

In contrast to phylogenetic overdispersion, clustering, or a clumped distribution of taxa, is often attributed to environmental filtering or habitat tracking, where only certain organisms can

prevail under the habitat conditions (Webb *et al* 2002). For instance, woody plant communities of north central Florida have clustered phylogenetic community structure, likely attributed to a prevalence of trait conservatism at this broad taxonomic scale (Cavender-Bares *et al* 2006).

Recently, the opposing effects of competition and habitat filtering on phylogenetic structure have been amended to include the confounding effects of competition on phylogenetic clustering (Mayfield & Levine 2010, HilleRisLambers *et al*, 2012). Depending on the nature of the conserved trait(s), competitive interactions can lead to phylogenetic clustering or overdispersion (Mayfield & Levine 2010). For example, plants possessing phylogenetically conserved traits for nutrient availability will exhibit overdispersion as sister taxa are competitively excluded. In contrast, conserved traits that confer competitive ability, such as plant height, would result in phylogenetic clustering, where only tall plant clades prevail as they outcompete for sunlight (Mayfield & Levine 2010).

The primary goal of distributing species data of communities onto phylogenetic space is to understand the influence of biotic and abiotic assembly processes on shaping community structure. To detect the relative role of these processes, indices were developed to determine whether an intact community is non-randomly distributed in phylogenetic tree space, relative to a random distribution of taxa (Webb *et al* 2002; e.g., net relatedness index (NRI) and nearest taxon index (NTI), in the Phylocom software). Abundance-weighted NRI is calculated as 1 - 'mean pairwise distance' (MPD) index, which is the average phylogenetic distance between two individuals in a community compared to those from a random distribution null model. Abundance-weighted NTI is calculated as 1-'mean nearest taxon distance' (MNTD) distance, which is the average distance between each individual in the community and it's nearest relative, adjusted to an average nearest distance from a null model distribution (Webb *et al* 2002). As NTI

only accounts for average phylogenetic distance to the nearest individual or taxa, and does not account for ancestral tree space, it is suitable for narrow taxonomic scales (Swenson *et al* 2006, Kraft *et al* 2007, Swenson 2010). These phylogenetic dispersion indices, while primarily developed for application to plant communities, have more recently been utilized for assessing microbial community structure and assembly processes, where massive microbial biodiversity hampers examination of individual interspecific interactions.

While phylogenetic dispersion analyses have great potential to advance our understanding of microbial community assembly, we argue that the default taxonomic scale of many analyses to date are likely far too broad for meaningful interpretation. The majority of microbial community studies examine NRI and/or NTI at the bacterial domain level. Several studies have examined these indices at narrower taxonomic scales, like phylum and class level (Pontarp *et al* 2013, Evans & Wallenstein 2014, Brown *et al* 2015, Waring & Hawkes 2015). We propose that these levels may even still be too broad. At the domain level, a macroorganismal analogy is one where assembly processes are assessed among Eukaryotes, including a grass, a manatee, and a yeast, which collectively have minimal niche overlap. Thus, such level of taxonomic scale may not allow for distinguishing among mechanisms driving community assembly.

Our primary objective was to examine the influence of taxonomic scale on measures of community dispersion using a case study of soil bacteria (Figure 2.1). We used the NRI and NTI phylogenetic dispersion indices, included in the Phylocom 4.2 software package, to examine the potential influence of taxonomic scale on phylogenetic structuring in soil bacterial communities (Figure 2.1a). The issue of taxonomic scale and its possible influence has been extensively examined on theoretical datasets (Kraft *et al* 2007), Florida woody plant communities (Cavender-Bares *et al* 2006), and at broad taxonomic scales in bacterial communities (Horner-

Devine & Bohannan 2006). However, I expanded on this issue of taxonomic scale in my case study by fully examining phylogenetic dispersion values from the broadest, domain-level, scale, out to more distal nodes of the bacterial phylogeny. This is one of the first studies to fully examine dispersions throughout the *entire* bacterial domain tree space. Taxonomic scale was assessed in two different manners: one by regional pool size (i.e. the number of terminal nodes in each subtree) and by individual community abundance (i.e. the number of taxa present in each community within each subtree) (Kraft *et al* 2007). Based on previous examinations of taxonomic scales, we hypothesized that overdispersion should be more prevalent at narrower taxonomic scales, since the degree of functional overlap of taxa included in the calculation should increase with finer taxonomic scale. Coarse taxonomic scales capture large phylogenetic breadth of bacterial diversity. As such, the pattern at this scale should more indicative of environmental filtering. Therefore, we expect an overall positive relationship between phylogenetic community dispersion and taxonomic scale. Finally, we also examined the potential influence of subtree density, measured as community abundance/regional species pool, on phylogenetic dispersion.

Materials and Methods

Study site and sampling

The samples examined in this case study are soils from the Prairie Heating and CO_2 Enrichment (PHACE) experiment roughly 15 km west of Cheyenne, WY (41.198700 N, -104.887048 W) that was subjected to elevated CO_2 and warming for seven years (Dijkstra *et al* 2010). The PHACE site is characterized as a northern mixed grassland prairie, dominated by *Bouteloua gracilis* (Willd. ex Kunth), *Hesperostipa comata* (Trin. & Rupr.), and *Pascopyrum*

smithii (Rydb.), with mean annual precipitation of 384 mm, and the soil is characterized as Mollisols with an average pH of 7.0 (Morgan *et al* 2011).

Soils were collected during the spring and summer seasons, from 2007-2009, within native grassland soil on the PHACE plots. The soils were sieved to 2 mm, and immediately stored to (-)80 ^{o}C to preserve the intact microbial community at time of collection. Genomic DNA was extracted from soil samples with the MoBio Power Soil kit (MoBio, Carlsbad, CA). The bacterial V4 hypervariable region of the 16S rRNA gene (515F and 806R with 15-mer Illumina barcodes, Caporaso *et al* 2012) was amplified from genomic DNA using the Kapa2G Fast PCR kit (Kapa Biosystems, Wilmington, MA), and cleaned with the MoBio UltraClean PCR Cleanup kit to remove PCR residuals, including dNTPs and $MgCl_2$ (MoBio, Carlsbad, CA). The amplicons were quantified with Quanti-iT PicoGreen Assay kit (Life Technologies, Grand Island, NY) to combine equimolar quantities of DNA across soil samples into a single multiplex sample. The Research Technology Support Facility at Michigan State University sequenced the combined equimolar samples on a paired-end 2x250bp formatted Illumina run, using a 500 cycle Miseq V2 flow cell (Illumina, San Diego, CA) and an internal PhiX control standard. Base calling was done using Illumina Real Time Analysis (RTA) V1.18.54. RTA output was then demultiplexed and converted to FASTQ format using Illumina Bcl2fastq V1.8.4.

Preprocessing of DNA sequences

The demultiplexed FASTQ files were processed through an initial quality control using portions of the Qiime pipeline (V1.9.1, Caporaso *et al* 2010) by trimming Illumina-specific barcodes and removing low quality reads with the following specifications: fastqual < 25, homopolymer acceptance of < 6 bp, and zero tolerance for ambiguous base calls. High quality

sequences were clustering to the standard 0.97 sequence similarity using USEARCH (Edgar 2010), which creates 'seed' starting sequences, from which clusters are determined based on the percentage identity thresholds to the seed sequence. USEARCH was also used to identify and remove singleton and putative chimeric sequences, which were identified by comparing the data against the curated Silva111 16S rRNA Reference Database (Quast *et al* 2013). 4102 representative sequences were compiled using *pick_rep_set.*py, which selects the 'seed' of each OTU cluster. Abundance/absence tables were generated using the OTU cluster text file, and subsequently reduced to the sample with lowest sequencing depth, resulting in ten rarefied biom-formatted OTU tables, containing 1260 OTUs each (Caporaso *et al* 2010). Finally, the biom and the representative sequence files were used for all downstream analyses.

Phylogenetic analysis

As these community dispersion indices rely on efficacy of the phylogenetic tree, a series of alignments and phylogenies were simultaneously estimated on the 4102 representative sequences using the Practical Alignment using SATé and TrAnsitivity (PASTA, Mirarab *et al* 2014), which builds on the Simultaneous Alignment and Tree Estimation software (SATé, Liu *et al*, 2011) to provide accurate trees and alignments on large datasets of up to 200,000 sequences. The default settings in PASTA were implemented on each run (MAFFT aligner, MUSCLE merger, FastTree phylogeny estimation with GTR+G20 nucleotide substitution model) to generate 1000 iterations on each of three starting alignments to ensure reconstruction of the best tree estimation. To approximate the location of various standard taxonomic thresholds, the terminal nodes of the final phylogeny were paired with their taxonomic assignments, identified in Qiime against the curated Silva111 16S rRNA Reference Database (Quast *et al* 2013), and the 4101 internal nodes

were numerically labeled using a modification of the <*makenodelabels*> command in the Cran-R package, 'ape' (V3.2.0, Paradis *et al* 2004, 2015) and the phylogeny was converted to a newick-formatted output file.

To assess the influence of taxonomic scale, the trees and corresponding biom tables had to be subset at each internal phylogenetic node. Therefore, the 4101 internal nodes were identified using <*getNode*> from which each newick-formatted subtrees were extracted with <*subset*> from the regional bacterial domain tree using the Cran-R package, 'phylobase' (V3.2.0, Michonneau *et al* 2014). Correspondingly, the 10 rarefied biom tables were subset to reflect each subtree using *filter_otus_from_otu_table.py* within Qiime. The biom files were converted to text files and reformatted in Cran-R V3.2.0 for input to Phylocom. Each subtree and its corresponding set of rarefied OTU sub-tables were imported into the Cran-R V3.0.2 package, 'picante' (V1.6.2, Kembel *et al* 2014), the R-implementation of the Phylocom 4.2 software (Webb *et al* 2008), to estimate phylogenetic dispersion at every node. The dispersion indices were averaged across the 10 biomes at each node to increase confidence in the estimates. Phylogenetic distances within each subtree were calculated with the <*cophenetic*> function in Phylocom. The phylogenetic distances were then used to estimate abundance-weighted mean pairwise distance (MPD) and abundance-weighted mean nearest taxon distance (MNTD), which are equivalent to 1 - NRI and 1- NTI, respectively. In total, 86 soils were assessed on 2050 nodes (less those containing < 4 taxa) over the 10 rarefied biom tables, resulting in 3.53 million MPD and MNTD calculations, and corresponding p-values, which were then averaged over the rarefactions to 352,600 data points for downstream analyses. The default null model, 'taxa labels' in Picante (or 'phylogeny shuffle' in Phylocom) was implemented in all MPD and MNTD

calculations, which swaps taxa labels across the OTU table, while maintaining species richness and species turnover constant within a given community (Gotelli 2000).

Examination of general trends between taxonomic scale and NRI or NTI was visualized using the Cran-R V3.0.2 graphical package, 'hexbin' (V1.27.1, Carr *et al* 1991). The estimation of a linear model was fit to the NRI dataset for subtrees with <150 community membership, approximately corresponding to taxonomic classification below the bacterial phylum level, to assess the main objective of the potential influence of taxonomic scale on NRI.

To assess the potential influence of tree density, the ratio of individual community abundance to regional species pool was compared to the corresponding NRI and NTI value for each node, and a linear regression was fit to the dataset. A similar comparison was tested for the difference between NTI and NRI, to detect at what taxonomic scale the two indices differed.

Finally, two case study soil bacterial communities were identified for further graphical display. Soil A is from a control PHACE plot, collected in spring 2008. Soil B is from a heated PHACE plot, collected from summer 2008. Soil sample-specific Hillis plots were generated in the web-based iTOL software (V3.0 Letunic & Bork 2011) using predetermined branch label colors specific to each node.

Results

Taxonomic scale influenced the phylogenetic dispersion values. For both NRI and NTI, phylogenetic clustering increased with broader subtree (Figure 2.2, 2.3). Overall, the relationship between subtree size and NRI value was asymptotically positive, and fit a nonlinear model ($NRI \sim TreeSize + TreeSize^{0.3}$). There was a significantly positive linear relationship between NRI and community size less than 150 taxa (Figure 2.3)($r^2 = 0.27$, p-value < 0.0001). The

relationship between taxonomic scale and NTI fit a power function, increasing degrees of clustering with broader subtree size. Significant phylogenetic overdispersion (NRI or NTI <= 1.96) was primarily restricted to small subtrees (Figure 2.2, 2.3), and comprised 0.06% of NRI and 0.08% of NTI values. In contrast, significant clustering (NRI or NTI >= 1.96) made up 19.91% of NRI and 21.65% of NTI dispersion values. The relationship between phylogenetic dispersion and taxonomic scale was essentially the same with respect to regional pool size at each node (Figure 2.3).

There was no significant influence of tree density on NRI or NTI (Figure 2.3). The majority of subtrees had a tree density of 7-30% with similar average NRI and NTI range of -1 to 2 (Figure 2.4). The difference between NTI and NRI values consistently increases with local community abundance (Figure 2.4). At the bacterial domain scale, the average NTI values were 2.3X greater than the average estimated NRI values (NTI = 8.502 vs. NRI = 3.727). However, at fine-scale taxonomic scales, the average values among the dispersion indices did not differ significantly (NTI=0.429 vs. NRI=0.425) (Figure 2.5). Significant departure between calculated NRI and NTI values begins within the taxonomic scale of bacterial phylum, and continues to separate with increased subtree size (Figure 2.6).

The distribution of clustering and overdispersion signals varied throughout phylogenetic tree space and between the two example soil bacterial communities, denoted by red and blue branches, respectively (Figure 2.3). Soil A has a significant clustering signal at the bacterial domain taxonomic scale (NRI: 2.27), while Soil B had a nonsignificant dispersion signal at the domain level (NRI: 0.83) (Figure 2.3). While Soil A had a higher overall dispersion signal, and more clustering throughout tree space (21% clustered in Soil A vs. 8% in Soil B), it comprised more significant overdispersion throughout tree space (3% overdispersion in Soil A vs. 0.03% in

Soil B). Soil B only has one clade of overdispersion within the Bacteroidetes phylum on the left side of the Hillis plot. Soil A also has overdispersion within the Bacteroidetes, but also within other subclades of the Actinobacteria, Gemmatimonadetes, Planctomycetes, and Proteobacteria.

Discussion

The primary goal of this study was to examine the potential impact of taxonomic scale on phylogenetic dispersion in bacterial communities. We expected phylogenetic dispersion, assessed with NRI and NTI, to positively correlate with taxonomic scale. We hypothesized that overdispersion would prevail at narrow taxonomic scales and clustering to dominate the broader taxonomic scales. Our case study of soil bacterial communities supported these conjectures for both types of taxonomic scale: local community abundance and regional pool size. Overdispersion, while rare (0.3%), occurred more frequently at narrower taxonomic scales and clustering dominated the dispersion signals in larger trees, including the bacterial phyla and domain.

Our data support previous studies examining the influence of taxonomic scale on phylogenetic dispersion, including a theoretical dataset (Kraft *et al* 2007) and woody tree communities (Cavender-Bares *et al* 2006). Our analysis extended this over a broader range of taxonomic scales to include the default scales used by numerous microbial studies. The influence of taxonomic scale has also been examined at broad taxonomic scales in bacterial communities (Horner-Devine & Bohannan 2006), where they found minimal support for a positive influence of taxonomic scale on NRI and NTI. However, they note that the influence of taxonomic scale may operate at narrower scales than what they examined (Horner-Devine & Bohannan 2006). As the correlation flattens out around the bacterial phylum level in our dataset (Figure 2.2a, d), the

influence of scale may have been obscured in their study as well. Most studies have found that significant overdispersion is very rarely observed in microbial communities. The elusive overdispersive signal in microbial communities is likely due in part to the taxonomic scale of assessment. Within our case study, overdispersion was very rare, and primarily restricted to very narrow taxonomic scales. The standard OTU threshold for 16S rRNA is 97% may also contribute to the paucity of overdispersion. Microbial datasets delineated to 99% tend to reveal more overdispersion (Koeppel & Wu 2013, 2014), which is not surprising given that a 97% difference in conserved regions, like 16S, can span large distributions of phylogenetic diversification. Computational limitations, primarily at the construction of phylogenies, have hindered use of the 99% delineation, but competitive interactions are more likely to occur at this threshold. Lower NRI values with the small subtrees indicate that competition is an assembly process primarily operating at fine-scale taxonomic scales. Not surprisingly, clustering dominated at broad taxonomic scales, as only certain bacterial clades prevail at a given soil location. Dispersal limitations and environmental filtering may provide an explanation for the clustering at large taxonomic scale. In contrast, the significant overdispersion in the small trees supports the idea that competitive exclusion from niche overlap of sister taxa operates on fine taxonomic scales (Figure 2.3).

The departure in estimations of phylogenetic community dispersion between NTI and NRI is notable in that NTI grossly overestimates degree of clustering in large bacterial phylogenies, relative to NRI values at comparable tree size (Figure 2.4). Despite using a fully resolved phylogeny, the influence of only considering distal branch lengths in determining the abundance-weighted mean nearest individual distance results in overestimation of clustering as tree size increases. This mathematical artifact was previously examined in much smaller phylogenies (750

and 320 taxa) from a different domain (multicellular Eukaryotes; Swenson *et al* 2007), and confirmed with our larger examination of the bacterial phylogeny (4102 terminal nodes).

Finally, the assessment of taxonomic scale does not only influence NRI values generally, but specific patterns exist in phylogenetic tree space among bacterial communities (Figure 2.5). This is important because by simply assessing NRI at one scale, differences in clustering and overdispersion among different bacterial clades are obscured in the singular measurement. Instead, conducting a node-by-node assessment reveals patterns of overdispersion and clustering throughout tree space. Given the immense genetic and functional diversity of the bacterial domain of life, it is not surprising that for a given soil condition, certain bacterial clades might be more influenced by competitive interactions, relative to another soil (Figure 2.5). Specifically, our simple case study of Soil A and Soil B indicate that certain subclades of Actinobacteria and Planctomycetes are primarily assembled through competitive interactions under Soil A conditions. In contrast, more of the Proteobacteria in Soil A are overdispersed, relative to the same subclades of Soil B. The conditions present at Soil A result in a stronger environmental filter for Proteobacteria, compared to the conditions of Soil B. Extending this beyond just two case study soils, the node-by-node assessment of phylogenetic dispersion patterns becomes a clearly important link with understanding more about the niche conditions of bacterial taxa.

Our case study examination of soil bacterial communities highlights the importance of considering the ecological implications of assessing phylogenetic community dispersion at a given scale. Additionally, we highlight the mathematical artifact that vastly overestimates degree of clustering when estimated with NTI, relative to NRI. A priori knowledge of the threshold of where the NTI algorithm overestimates clustering is important to determine.

In moving forward, it is important to consider the scale of phylogenetic community structure. Even if studies find strong clustering at the bacterial domain, finer-scale patterns of overdispersion and/or clustering likely exist throughout tree space, and more importantly, among microbial communities. Defining the 'proper' taxonomic scale is not the goal of our examination. Instead, we propose a more holistic, node-by-node examination of microbial communities. This allows for specific clades to be classified as primarily driven by either environmental filtering or competition. The utility of phylogenetic community structure is to infer the influence of ecological processes on community assembly, but these processes may operate on different taxonomic scales (density-dependent interaction at finer scales, and environmental filtering at broader scales).

Figures

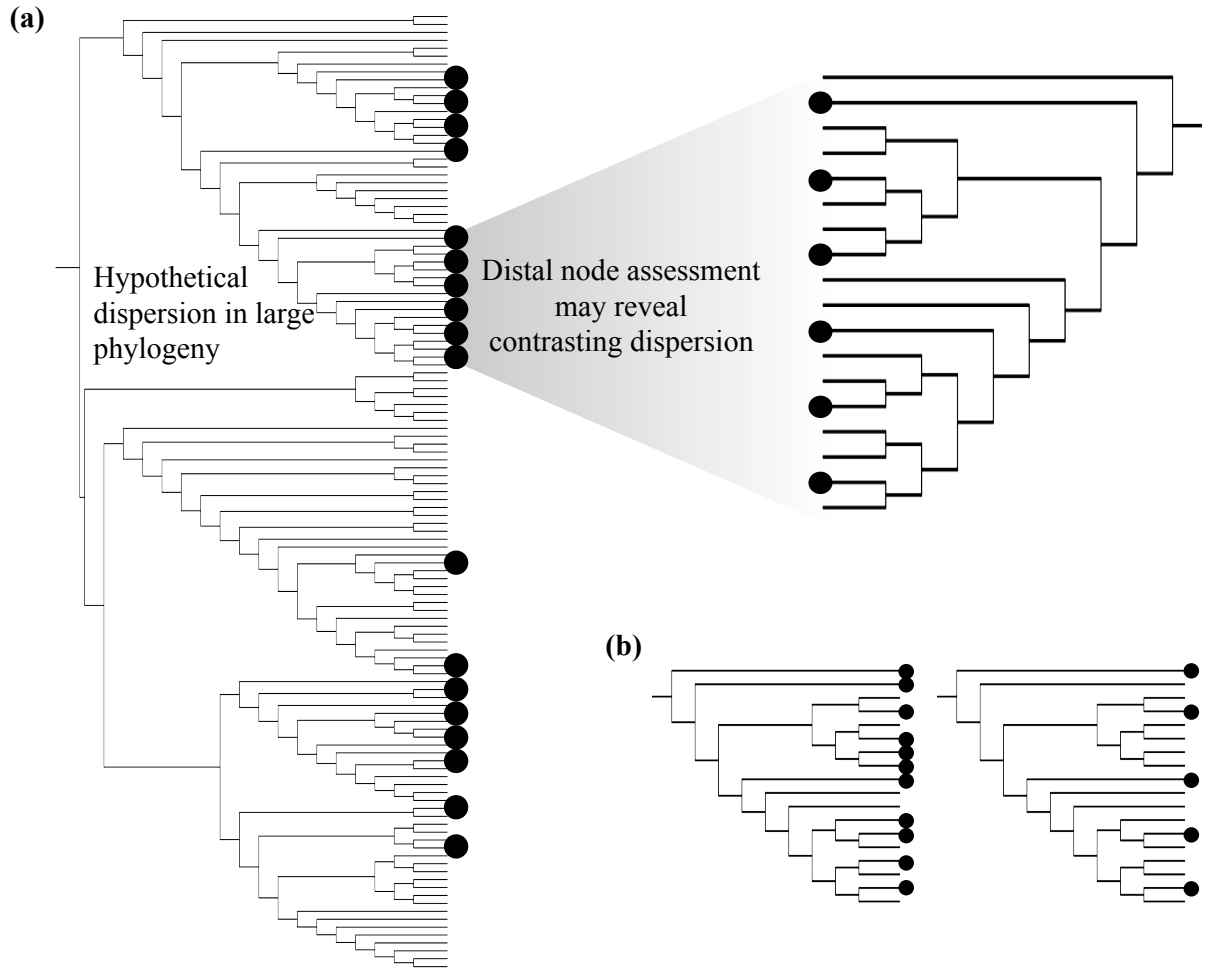


Figure 2.1. Does taxonomic scale matter? Are estimates of phylogenetic community dispersion influenced by the taxonomic scale of assessment? (a) Phylogenetic dispersion may be scale-dependent, where the distribution of clustering and overdispersion occur at various nodes in tree space. For example, clustering could be estimated from a given super tree scale, while distal subtrees may reveal different dispersion signals. (b) Density, here defined as the relative number of local community membership in the regional species pool at a given node, may also influence phylogenetic dispersion.

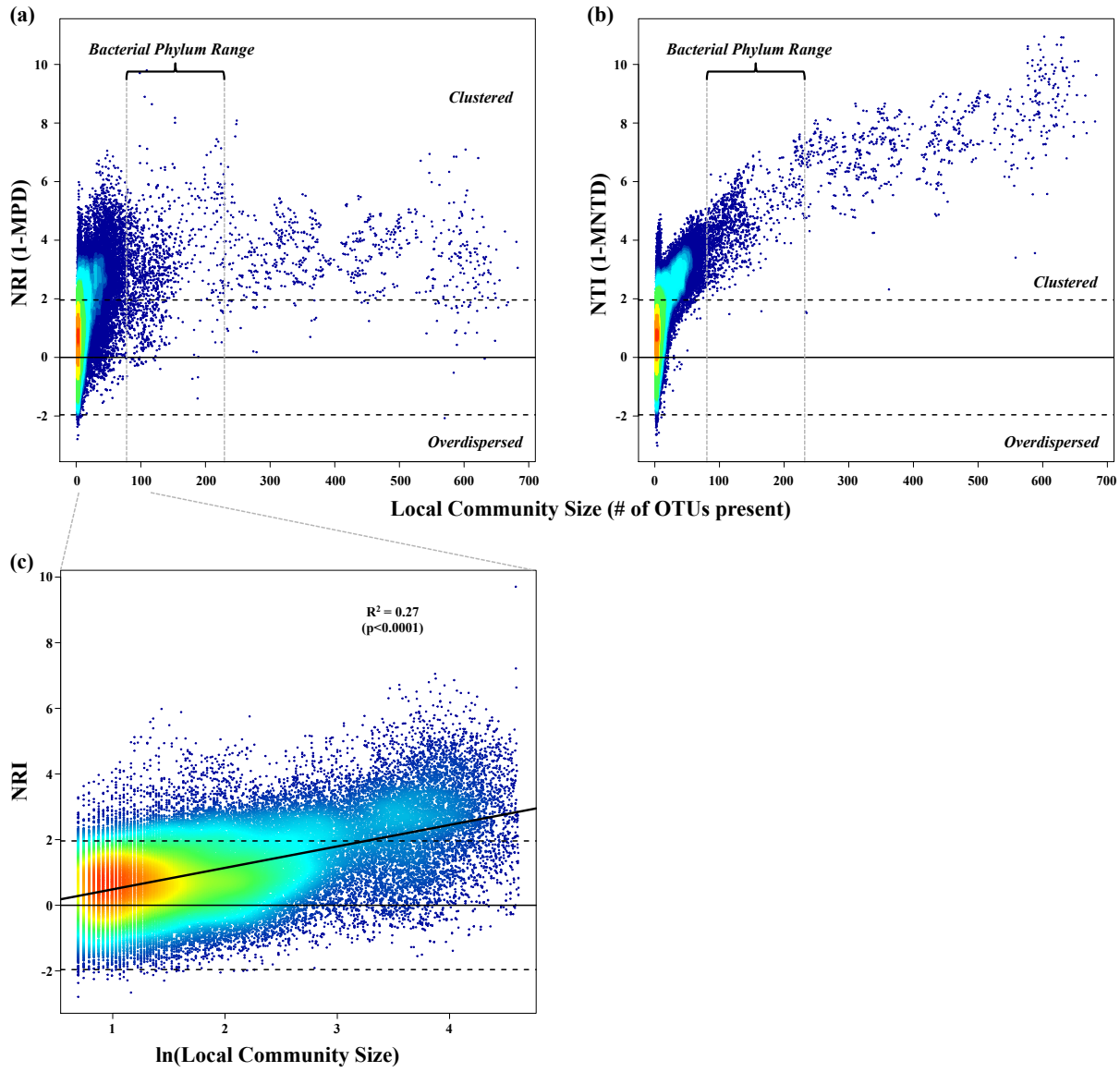


Figure 2.2. Relationship between phylogenetic dispersion and local community size: (a) Net Relatedness Index (NRI) and (b) Nearest Taxon Index (NTI) vs. local community size. Zero is marked with a horizontal solid line, and the horizontal dotted lines mark significance thresholds ($p<0.025$), where dispersion effects >1.96 correspond to significant phylogenetic clustering, and <1.96 as significant overdispersion. Vertical dotted lines mark the approximate range of bacterial phyla.

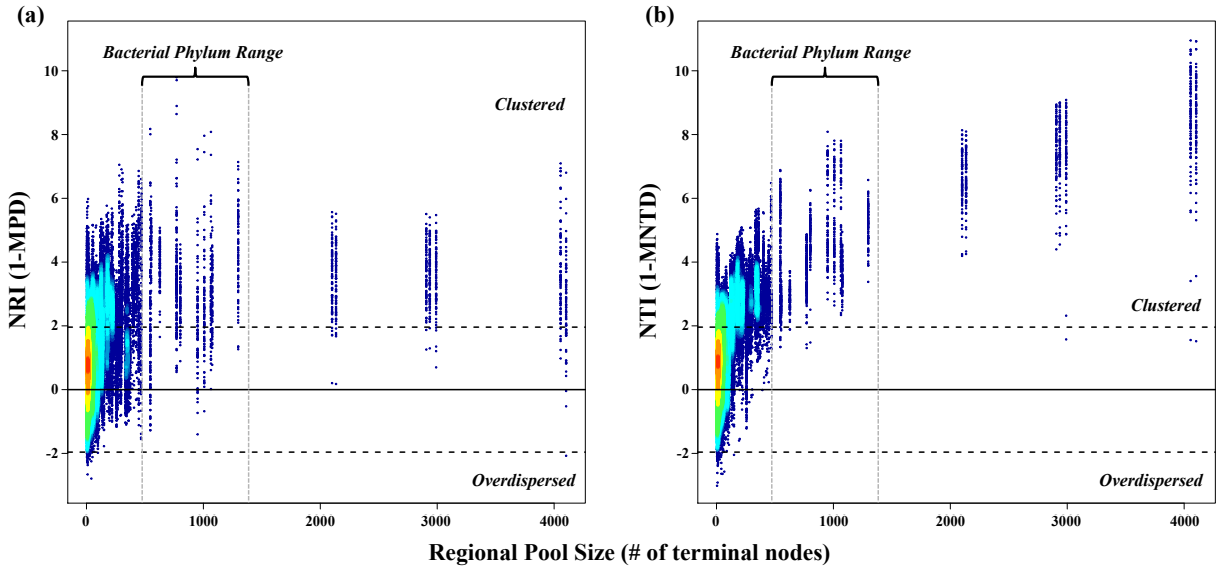


Figure 2.3. Relationship between phylogenetic dispersion and regional pool size: (a) Net Relatedness Index (NRI) and (b) Nearest Taxon Index (NTI) vs. regional pool tree. Zero is marked with a horizontal solid line, and the horizontal dotted lines mark significance thresholds ($p<0.025$), where dispersion effects >1.96 correspond to significant phylogenetic clustering, and <1.96 as significant overdispersion. Vertical dotted lines mark the approximate range of bacterial phyla.

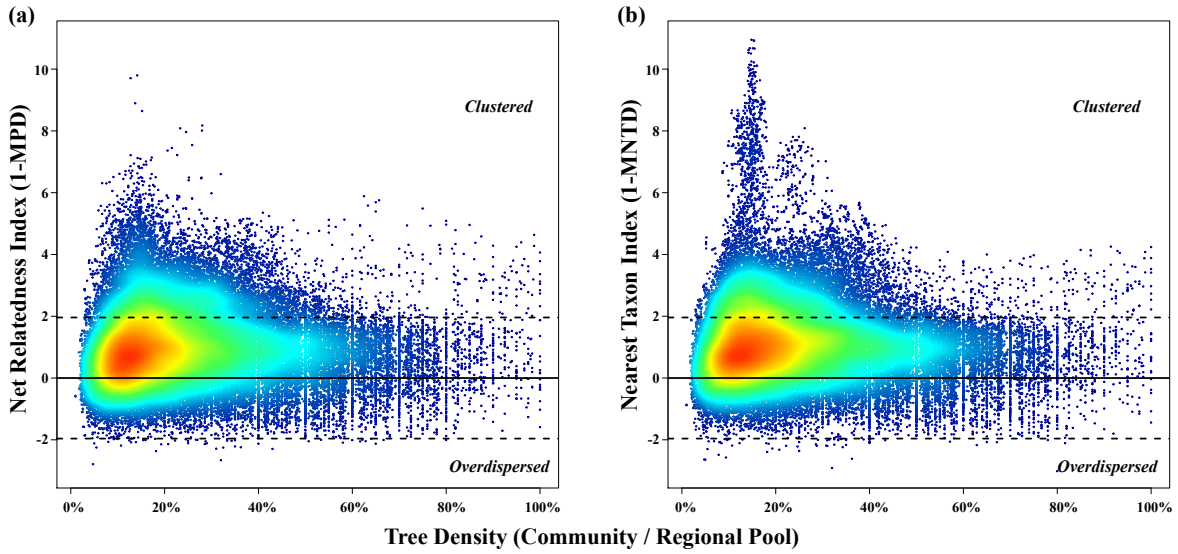


Figure 2.4. No influence of tree density on phylogenetic dispersion signals. Tree density, calculated as local community abundance relative to the total regional pool size for each node, had no significant influence on NRI and NTI. Zero is marked with a horizontal solid line, and the horizontal dotted lines mark significance thresholds ($p<0.025$), where dispersion effects >1.96 correspond to significant phylogenetic clustering, and <1.96 as significant overdispersion. Vertical dotted lines mark the approximate range of bacterial phyla.

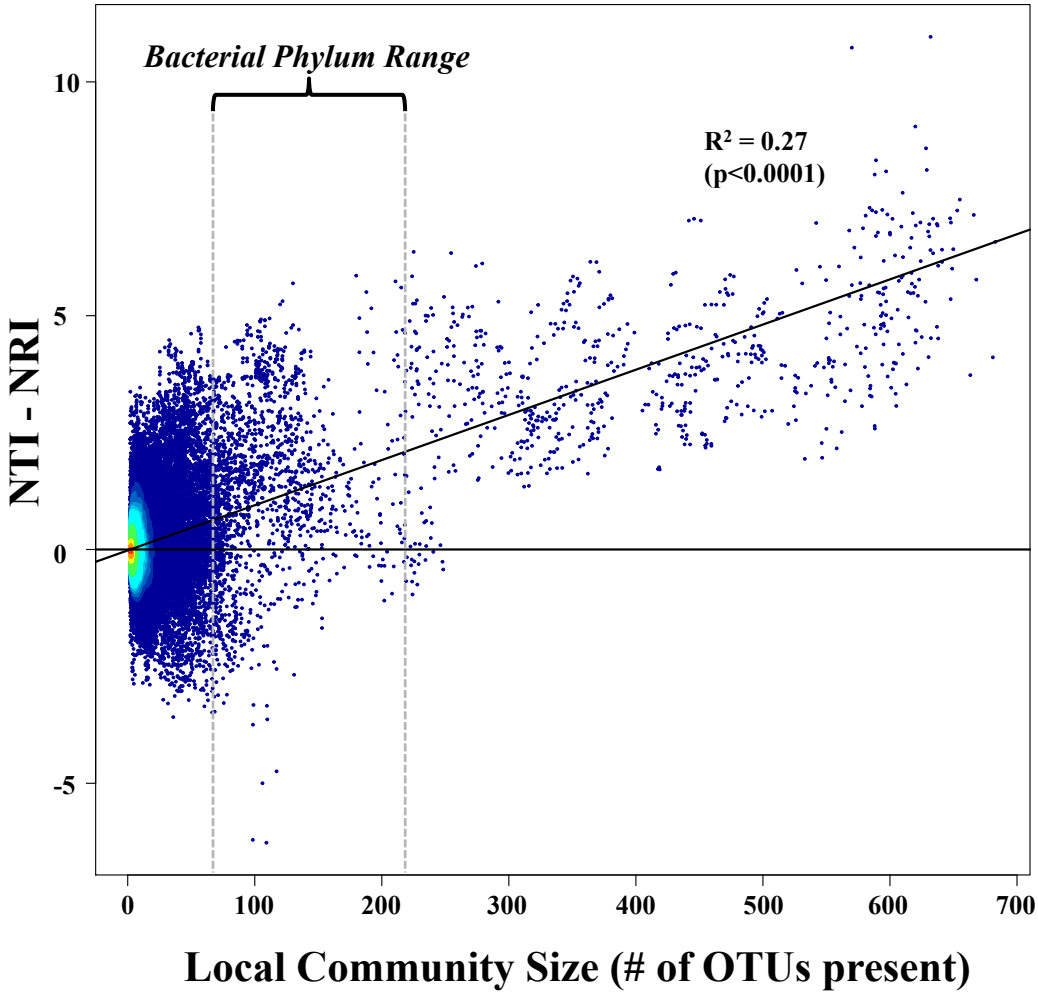
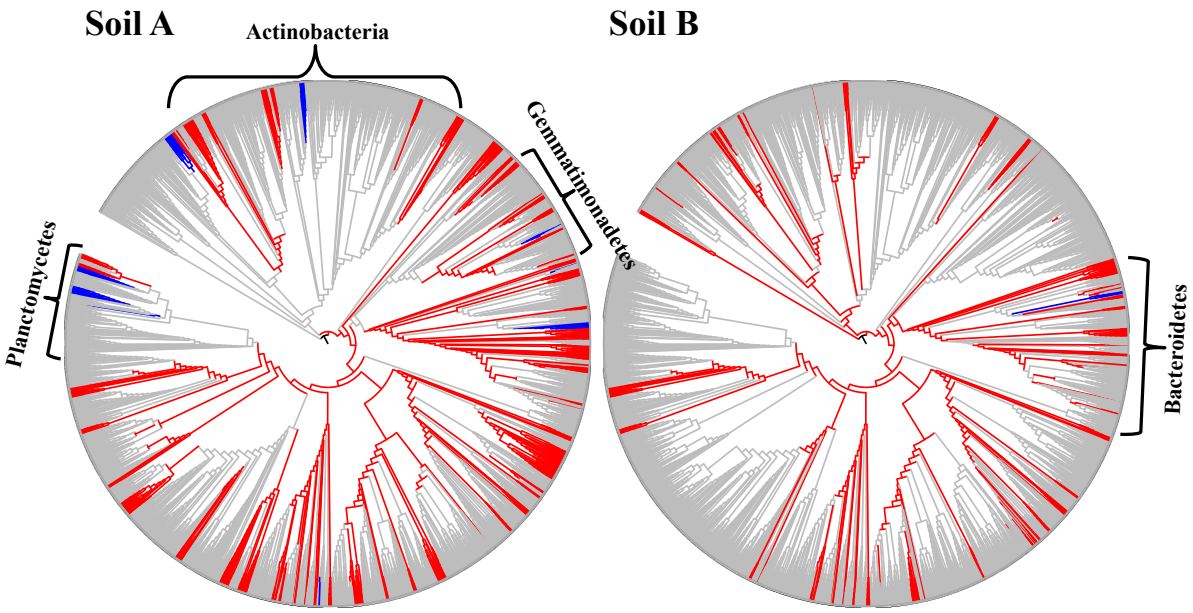


Figure 2.5. Difference between NTI and NRI values vs. taxonomic scale. NTI minus NRI values versus the local community size. Horizontal line marks no difference between NTI and NRI, and the positive line represents the linear regression fit for Difference~Local Community Size (R^2 = 0.27, p-value < 0.0001).



	Domain NRI	Overdispersed %	Clustered %
Soil A	2.27	3	21
Soil B	0.83	0.3	8

Figure 2.6. Phylogenetic dispersion varies throughout tree space and between two sample soil bacterial communities. A node-by-node assessment of NRI reveals that a broad-scale estimation of phylogenetic dispersion can obscure fine-scale patterns, as demonstrated here where Soil A has a significant clustering signal at the domain level, but has more overdispersion within the community compared to Soil B with lower NRI. Blue and red portions of the tree represent nodes of significant overdispersion and clustering of distal OTUs, respectively ($p<0.025$).

References

Aleklett, K., Leff, J. W., Fierer, N., and Hart, M. (2015). Wild plant species growing closely connected in a subalpine meadow host distinct root-associated bacterial communities. Peer J. 3:e804. doi: 10.7717/peerj.804

Angel R, Soares MI, Ungar ED, Gillor O. (2010) Biogeography of soil archaea and bacteria along a steep precipitation gradient. ISME J 4(4): 553-563.

Bell CW, Tissue DT, Loik ME, Wallenstein MD, Acosta-Martinez V, Erickson RA, Zak JA. (2014) Soil microbial and nutrient response to 7 years of seasonally altered precipitation in a Chihuahuan Desert grassland. GCB 20(5): 1657-1673.

Bell CW, Asao S, Calderon F, Wolk B, Wallenstein MD. (2015) Plant nitrogen uptake drives rhizosphere bacterial community assembly during plant growth. Soil Biol. Biochem. 85: 170-182.

Brown SP, Jumpponen A (2015) Phylogenetic diversity analyses reveal disparity between fungal and bacterial communities during microbial primary succession. Methods Ecol. Evol. 5: 1225-1235.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6: 1621-1624.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al*. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7(5): 335-336.

Carr DB. (1991) Looking at Large Data Sets Using Binned Data Plots, pp. 7–39 in Computing and Graphics in Statistics; Eds. A. Buja and P. Tukey, Springer-Verlag, New York.

Cavender-Bares J, Ackerly DD, Baum DA, Bazzaz FA.(2004)Phylogenetic Overdispersion in Floridian Oak Communities. Am. Nat.163(6): 823-843

Cavender-Bares, J., A. Keen, and B. Miles. 2006. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. Ecology 87:S109-S122.

Cavender-Bares, J., K. Kozak, P. Fine, and S. Kembel. (2009). The merging of community ecology and phylogenetic biology. (Review and Synthesis) Ecology Letters, 12:693-715.

Collins SL, Sinsabaugh RL, Crenshaw C, Green L, Porras-Alfaro A, Stursova M. (2008) Pulse dynamics and microbial processes in arid land ecosystems. J. Ecol. 96: 413-420.

DeAngelis KM, Pold G, Topcuoglu BD, van Diepen LTA, Varney RM, Blanchard JL, Melillo J, Frey SD. (2015) Long-term forest soil warming alters microbial communities in temperate forest soils. Front. Microbiol. 6: 104.

Edgar RC. (2010) Search and clustering order of magnitude faster than BLAST. Bioinformatics 26(19): 2460-61.

Evans SE, Wallenstein MD, Burke IG. (2014) Is bacterial moisture niche a good predictor of shifts in community composition under long-term drought? Ecology 95: 110-122

Fierer, N. and R.B. Jackson. (2006) The diversity and biogeography of soil bacterial communities. PNAS. 103(3): 626-631.

Fritschie KJ, Cardinale BJ, Alexandrou MA, Oakley TH. (2014) Evolutionary history and the strength of species interactions: testing the phylogenetic limiting similarity hypothesis. Ecology 95(5): 1407-1417.
Fukami T. (2015) Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. Annual. Rev. 46:
Gotelli NJ (2000) Null model analysis of species co-occurrence patterns. Ecology. 81(9): 2606-2621.
Hibbing ME, Fuqua C, Parsek MR, Peterson SB. (2010) Bacterial competition: surviving and thriving in the microbial jungle. Nat. Rev. Microbiol. 8: 15-25.
HilleRisLambers J, Adler PB, Harpole WS, Levine MS, Mayfield MM. (2012) Rethinking community assembly through the lens of coexistence theory. Annu Rev Ecol Evol 43: 227-248.
Horner-Devine MC, Bohannan BJ. (2006) Phylogenetic clustering and overdispersion in bacterial communities. Ecology 87(7):
Hughes Martiny JB, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine C, Kane M, Adams Krumins J, Kuske CR, Morin PJ, Naeem S, Ovreas L, Reysenbach A, Smith VH, Staley JT. (2006) Microbial biogeography: putting microorganisms on the map. Nat. Rev. Microbiol. 4: 102-112.
Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. (2010) Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26: 1463-64.
Koeppel A, Wu M (2013) Surprisingly extensive mixed phylogenetic and ecological signals among bacterial Operational Taxonomic Units. Nucleic Acid Research 41(10): 5175-5188.
Koeppel A, Perry EB, Sikorski J, Krizanc D, Warner A, Ward DM, *et al*. (2008) Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. Proc. Natl. Acad. Sci. U.S.A. 105: 2504-2509.
Koeppel A, Wu M (2014) Species matter: the role of competition in the assembly of congeneric bacteria. ISME J. 8: 531-540.
Kraft NJB, Cornwell WK, Webb CO, Ackerly DD. (2007) Trait evolution, community assembly, and the phylogenetic structure of ecological communities. Am. Nat. 170(2): 271-283.
Lauber CL, 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. Appl Environ. Microbio. 75(15): 5111-20.
Letunic I, Bork P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acid Res.
Liu K, Linder CR, Warnow T (2011) RAxML and FastTree: Comparing two methods for large-scale maximum likelihood phylogeny estimation. PLoS One. 6(11): 1-11.
Losos, J.B. 2008. Rejoinder to Wiens (2008): Phylogenetic niche conservatism, its occurrence and importance. Ecology Letters 11: 1005-1007
Marschner P, Kandeler E, Marchner B. (2003) Structure and function of the soil microbial community in a long-term fertilizer experiment. Soil Biol. Biochem. 35(3): 453-461.
Mayfield MM, Levine JM. (2010) Opposing effects of competitive exclusion on the phylogenetic structure of the communities. Ecol Lett 13: 1085-1093.
Mirarab S, Nguyen N, Warnow T. (2014) PASTA: ultra-large multiple sequence alignment. Proceedings of RECOMB.
Monad J (1949) The growth of bacterial cultures. Annu. Rev. Microbiol. 3: 371-394.

Morgan JA, LeCain DR, Pendall E, Blumenthal DM, Kimball BA, Carrillo Y, Williams DG, Heisler-White J, Dijkstra FA, West M. (2011) C4 grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland. Nature. 476: 202-205.

Naugton HR, Alexandrou MA, Oakley TH, Cardinale BJ. (2015) Phylogenetic distance does not predict competition in green algal communities. Ecosphere. 6(7).

Paradis E, Claude J, Strimmer K.(2004) APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20(2): 289-290.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. (2013) The SILVA ribosomal RNA database project: improved data processing and web-based tools. Nucl. Acids Res. 41: D590-D596.

Rousk J, Baath E, Goransson H, Fransson A. (2007) Assessing plant-microbial competition for 33P using uptake into phospholipids. Appl. Soil Ecol. 36(2): 233-237.

Schimel JP, Jackson LE, Firestone MK. (1989) Spatial and temporal effects on plant-microbial competition for inorganic nitrogen in a California annual grassland. Soil Biol. Biochem. 21(8): 1059-1066.

Swenson NG. (2009) Phylogenetic scale and quantifying the phylogenetic diversity and dispersion of communities. PLoS One. 4(2): 1-8.

Swenson NG. (2013) The assembly of tropical tree communities – the advances and shortcomings of phylogenetic and functional trait analyses. Ecography. 36(3): 264-276.

Thébault A, Clément J, Ibanez S, Roy J, Geremia RA, Pérez CA, Buttler A, Estienne Y, Lavorel S. (2014) Nitrogen limitation and microbial diversity at the treeline. Oikos 123(6): 729-740.

Violle C, Nemergut DR, Pu Z, Jiang L. (2011) Phylogenetic limiting similarity and competitive exclusion. Ecol. Lett. 14(8): 782-787.

Vannette RL, Gauthier ML,Fukami T. (2014) Nectar bacteria, but not yeast, weaken a plant–pollinator mutualism. Proc R Soc. B. 290:20122601

Waring BG & Hawkes CV. (2015) Short-Term Precipitation Exclusion Alters Microbial Responses to Soil Moisture in a Wet Tropical Forest. Micro. Ecol. 69(4):843-854

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. (2002) Phylogenies and community ecology. Annu. Rev. Ecol. Syst. 33: 475-505.

Webb CO, Ackerly DD, Kembel SW. (2008) Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics 18: 2098-2100.

Weiher E and Keddy PA. (1995) Assembly Rules, Null Models, and Trait Dispersion: New Questions from Old Patterns. Oikos 74(1): 159-164

Williams MA, Rice CW. (2007) Seven years of enhanced water availability influences the physiological, structural, and functional attributes of a soil microbial community. App. Soil Ecol. 35: 535-545.

Chapter 3: Collection date overshadows elevated CO_2 and warming as an important driver of soil bacterial community structure

Microorganisms are the main players of nutrient cycling in soils and contribute heavily to ecosystem productivity, yet we still do not know all the determining factors that structure microbial communities. It is important to establish which environmental factors control microbial community structure and how forecasted changes in climate may influence the functional capacity of these key players of ecosystem function. The goal of this study was to examine influence of elevated CO2, temperature, and soil moisture availability on structuring soil bacterial communities. The soils were collected from the Prairie Heating and CO2 Enrichment (PHACE) experimental site in southern Wyoming to address long-term impacts of these environmental factors on soil bacterial community membership and structure. There was no strong influence of either of the experimentally imposed treatment, CO2 or temperature, on the bacterial community. However, season and collection year were strong influences (p-val < 0.001) on both soil bacterial community membership and structure. Average soil moisture for the year prior to each harvest was also a significant contributor to both community membership (p-val: 0.024) and structure (p-val: 0.009). The strong influence of collection year and season on the soil microbial communities may be due to overall plant response to yearly climate. Finally, indicator species analysis by season showed that a substantial number of indicator taxa, and phylogenetic nodes, are from the Actinobacteria phylum in spring soils, and shifts to more ubiquitous distribution of indicator taxa for the summer soils. In the spring soil sampling, plants are likely performing different functions, like high photosynthetic rates and uptake of available nutrients, as compared to the summer, which may include increased root exudation of

photosynthates. The seasonality of plant functions and consequential impacts for soil nutrient availability may be contributing to the strong correlation between season and bacterial community structure. To conclude, the low impact of elevated CO2 and temperature on soil bacterial community structure does not necessarily indicate that these factors do not influence microbial communities, but that the seasonality of the system overshadows the influence of these factors.

Introduction

Soil microorganisms are crucial players in biogeochemical cycling of nutrients and in ecosystem functioning as a whole. Yet, because of their immense diversity, rapid turnover, and minute size, we are still learning about which factors drive microbial community structure and function. In the context of forecasted climate change, it is crucial to understand how microbial communities might respond directly to shifts in abiotic conditions, as well as climate-induced shifts in soil nutrient availability and changes in the plant community (Tharayil *et al* 2011, Suseela *et al* 2015), which interact closely with soil microbial communities (Hobbie 1992, Bais *et al* 2006, Freschet *et al* 2013).

At the continental scale, pH is the primary determinant of microbial community membership (Fierer & Jackson 2006, Hughes-Martiny *et al* 2006, Lauber *et al* 2009). Soil moisture availability (Williams & Rice 2007, Collins *et al* 2008, Angel *et al* 2010, Bell *et al* 2014), soil nutrient availability (Marschner *et al* 2003, Thebault *et al* 2014), temperature (Schindlbacher *et al* 2011, DeAngelis *et al* 2015), and plant community are also prominent drivers of microbial community structure in soils. However, less is known about the relative influences of these environmental factors on structuring soil microbial communities.

The primary objective of this study was to examine the impact environmental factors on soil bacterial community structure, using soils from the Prairie Heating and CO_2 Enrichment (PHACE) experimental study site in southern Wyoming that was subjected to seven years of elevated CO_2 and warming to reflect forecasted climate change for the area (Morgan *et al* 2011). Because of the long-term implementation of these field treatments, we expect the elevated CO_2 and temperature to have some effect on the structure of the soil microbial community structure, either directly or indirectly through changes in soil moisture. Elevated temperature can both hinder and help soil microbial function and in turn impact community structure. Warmer soil temperatures allow enzymatic activity to increase, which in turn increases nutrient cycling, and also likely favors a different suite of soil microorganisms. However, increased soil temperature also increases evaporation, decreasing the soil moisture availability, which in turn can negatively impact soil microbial function. In contrast, there are minimal direct impacts of increased CO_2 on soil microbial communities. Aside from autotrophic microorganisms residing on the soil surface, the majority of microorganisms are indirectly impacted by elevated CO_2 through the activity of the surrounding plant community. Elevated CO_2 allows plants to close their stomata earlier as their C demands for photosynthesis are met, which in turn decreases evapotranspiration loss of water. The plants become more water use efficient, which has scaffolding effects on soil moisture availability because the plant roots have a decreased demand for water from the soil. Therefore, elevated CO_2 often positively correlated with higher soil moisture availability, which positively impacts soil microbial function and differentially influences soil microbial groups. Together elevated CO_2 and temperature have somewhat counteractive effects, which likely impact the soil microbial community structure.

The importance of considering the assembly of soil microbial communities hinges on the early research of plant community ecologists. Organisms with different niche space will be differentially successful in a suite of patch conditions. If these organisms also fill distinct functional roles in the community, then it follows that understanding the community structure may inform us of the overall functionality of the community, as well as how the community will respond to environmental and biotic changes. As such, the soils from the PHACE site were assessed for their bacterial community structure using the 16S rRNA conserved region. Community indices, including: alpha- and beta-diversity, phylogenetic diversity, and phylogenetic dispersion indices where estimated from each soil bacterial community. These community indices highlight what soil samples are distinct in their overall community structure, as well as allow us to determine which bacterial groups are responding strongly to the imposed field conditions - elevated CO_2 and temperature.

Materials and Methods

Study site and sampling

The soil bacterial communities examined are from April and July harvest from 2007 to 2009 from the Prairie Heating and CO2 Enrichment (PHACE) field experimental site, which is located 15 km west of Cheyenne, WY (41.198700 N, -104.887048 W). The PHACE site and the surrounding habitat is characterized as a northern mixed grassland prairie, dominated by *Bouteloua gracilis* (Willd. ex Kunth), *Hesperostipa comata* (Trin. & Rupr.), and *Pascopyrum smithii* (Rydb.), with mean annual precipitation of 384 mm, and the soil is characterized as Mollisols with an average pH of 7.0 (Morgan *et al* 2011). The PHACE experimental consisted of a full two-factor field experiment, with imposed elevated CO_2 and warming for seven years

(Morgan *et al* 2011). Each circular PHACE plot consisted of a CO_2 dispersing ring near ground level and infrared heating lamps roughly 1 m off the ground, as well as non-functioning negative controls of each factor. Each plot was split in half by a native grassland side and a side subjected to invasion by plants, including *Linaria dalmatica* (L.) and *Centaurea diffusa* (Lam.).

Soils were collected, as a part of the 'belowground collection team', from the native aboveground portion of each PHACE plot for the following dates: July 2007, April 2008, July 2008, and April 2009. Five 1 cm cores were harvested from each plot, and sieved to 2 mm immediately after harvest at the PHACE site. The soil samples were immediately stored at 4 ^{o}C, and a 2 g subset was stored at (-)80 ^{o}C once back at the lab.

Field data

At each PHACE plot, continuous records of soil moisture were determined from in-plot soil moisture probes. To capture some temporal variability in soil moisture availability as it varied by PHACE plot, moisture data was extracted from the microbial activity database by averaging the soil moisture over a few time spans: average soil moisture for the week prior to the soil harvest from the PHACE plots, the month prior average, and the year prior average soil moisture.

Processing molecular data

Ninety-two soil samples were processed for molecular data, comprised of four field replicates of the four PHACE field treatments (ambient CO_2 & temp (ac), ambient CO_2 & elev. temp (aH), elev. CO_2 & temp (EH), and elevated CO_2 & ambient temp (Ec)) from the native grassland side of each PHACE plot, over four harvest dates spanning two season (July and

April). The soil samples were extracted for total DNA and eventual processing on the Illumina sequencing platform. Total microbial genomic DNA was extracted from the 2 g soil subsets, using the MoBio Power Soil kit and following standard protocol (MoBio, Carlsbad, CA). The bacterial-specific V4-hypervariable region of the conserved 16S rRNA gene was amplified from the total genomic DNA pool, with the standard 515F and 80R oligos with 15-mer sample-specific barcode and Ilumina tag (Caporaso *et al* 2012), using the Kapa2G Fast PCR kit (Kapa Biosystems, Wilmington, MA). V4 amplicons were purified of PCR reagents and residuals, including $MgCl_2$, extra dNTPs, and primer dimers with the MoBio UltraClean PCR Cleanup kit (MoBio, Carlsbad, CA). To combine each the PCR product of each soil sample equimolarly for even DNA sequencing effort, the V4 amplicons were quantified with the Quanti-iT PicoGreen Assay kit (Life Technologies, Grand Island, NY). Each sample was combined to bring the total to 1 µg DNA, and this sample was sent to the Research Technology Support Facility at Michigan State University for highthroughput sequencing on a paired-end 2x250bp formatted Illumina run, using a 500 cycle Miseq V2 flow cell (Illumina, San Diego, CA) and an internal PhiX control standard. The RTSF generated nucleic acid base calls with the Illumina Real Time Analysis (RTA) V1.18.54. The output was then demultiplexed by soil sample, using the 15-mer predefined barcodes, and converted to FASTQ format using Illumina Bcl2fastq V1.8.4.

Bacterial community analysis

Of the 92 original molecularly processed soil samples, 86 were analyzed. The decrease in analyzed samples was due to the fact that six community samples had to be discarded once the sequence database was returned as these were failed Illumina runs, all containing more than 200 sequences in total, which was far less the average per sample sequencing depth of 1500

sequences. Using the Quantitative Insights into Microbial Ecology (Qiime) sequence processing pipeline (V1.9.1, Caporaso *et al* 2010), low quality reads were discarded using the following criteria: fastqual < 25, homopolymer acceptance of < 6 bp, and zero tolerance for ambiguous base pair calls. The remaining high quality sequences were clustered based on 97% sequence similiarity using USEARCH (Edgar 2010), which generates starter 'seed' sequences, from which other sequences that match at 97% are aggregated into operational taxonomic units, or OTUs. The 97% sequence similarity threshold of the 16S rRNA roughly corresponds to the bacterial genus taxonomic level. After OTU clusters were determined, singleton sequences and artificially created sequence strands from cross over during PCR reactions, or putative chimera, were removed from the dataset using USEARCH against a curated Silva111 16S rRNA Reference Database (Quast *et al* 2013). In total, 4101 OTUs were identified across all soil samples, and representative sequences were extracted and compiled into a representative set of DNA sequences, using the *pick_rep_set.py* command in Qiime. An OTU table, which has the abundance patterns for each of the 4101 taxa among the 86 soil samples, was generated using the *make_otu_table.py* command in Qiime. The OTU table was then rarefied to soil sample with the lowest sequencing depth to produce ten biom-formatted tables of per sample sequencing depth of 1260 sequences (Caporaso *et al* 2010), even though several measures of community structure are robust to sequencing sampling depth (McCoy & Matsen 2013). The biom tables and representative sequence file were used for downstream analyses of diversity, correlation with environmental variables, and estimations of phylogenetic dispersion.

To assess phylogenetically-informed alpha and beta diversity, the representative sequences were aligned, using the *'Practical Alignment using SATé and TrAnsitivity'* (PASTA, Mirarab *et al* 2014), building on the *'Simultaneous Alignment and Tree Estimation'* algorithm (SATé, Liu *et*

al, 2011) to iteratively estimate accurate alignments and phylogenies on large datasets of up to 200,000 sequences. The default settings in PASTA were implemented on each run (MAFFT aligner, MUSCLE merger, FastTree phylogeny estimation with GTR+G20 nucleotide substitution model) to generate 1000 iterations on each of three starting alignments to ensure reconstruction of the best tree estimation.

Three indices of alpha diversity were calculated for each soil sample, across the ten rarefied biom tables using *alpha_diversity.py* with the methods: Chao1, Faith's Phylogenetic Distance, and Simpson index. The output diversity estimates were combined into a table to examine significant differences in alpha diversity among the soil bacterial community samples and their corresponding categorical and continuous treatments.

Beta-diversity was computed as the weighted (W) and unweighted (UW) UniFrac distances, which uses an input tree to estimate phylogenetic relatedness among communities (Lozupone & Knight 2005). Unweighted UniFrac estimates the difference in community membership, as it only accounts for presence-absence of OTUs, while weighted UniFrac distances are community structure-based differences among communities by accounting for OTU abundances per soil sample. Principle coordinates were determined from the output UniFrac distance matrices, using *principal_coordinates.py* in Qiime, which generates eigenvalues for each sample community. The eigenvalue files were then input to *make_2d_plots.py* command in Qiime to generate two-dimensional PCoA plots, which were separately identified by the environmental factors, including: PHACE treatment combination, CO_2 treatment, heating treatment, collection date, year, season; and the week, month and year average soil moistures.

The relative influence of each environmental factor was also assessed using the adonis statistical method (McArdle & Anderson 2001). Adonis is a nonparametric statistical method,

similar to permutational multivariate analysis of variance, except it can handle both categorical and continuous environmental variables. Adonis is also similar to distance-based redundancy analysis, except that adonis assesses variation in the distance matrices that can be non-dimensional (McArdle & Anderson 2001). The output adonis tests of each environmental variable, including the p-value and R^2 effect size estimation of variance explained were compiled into a table, and used for deciding which categories to use for identifying potential indicator OTU taxa. The Cran-R package, 'indicspecies' (V1.7.5, Caceres & Jansen 2015) was implemented for the identification of indicator OTUs for each of the statistically significant environmental factors. The *multipatt* function, which builds on the methodology of the *Indicator Value,* (Dufrene & Legendre 1997), reported a list of OTUs along with corresponding indicator value and p-value. Taxonomic identity of each indicator OTU was determined with *assign_taxonomy.py* command in Qiime against the Silva111 reference database. Additionally, the OTUs were mapped onto phylogenetic tree space using the Interactive Tree of Life (iTOL) online display tools (Letunic & Bork 2011).

Phylogenetic dispersion of each soil bacterial community was estimated on each of the 4102 internal tree nodes, including the standard bacterial domain assessment, using the study-wide phylogeny. All 4102 internal tree nodes were identified and corresponding subtrees were extracted, using *getNode* and *subset* from the Cran-R package, 'phylobase' (V3.2.0, Michonneau *et al* 2014), along with the ten rarefied biom tables. The biom tables were converted to text formatted OTU tables and input into the Cran-R V3.2.0 implementation of the Phylocom software (Webb *et al* 2008), 'picante' (V1.6.2, Kembel *et al* 2014). Phylogenetic distance matrices were estimated on the subtree of each node, using *cophenetic* function. The distance matrices were then used to compute phylogenetic dispersion on the OTU tables using two

phylogenetic indices: abundance-weighted mean pairwise distance (MPD) and abundance-weighted mean nearest taxon distance (MNTD), which are equivalent to 1 - NRI and 1- NTI, respectively. The null model, 'taxa labels' in Picante (or 'phylogeny shuffle' in Phylocom) was used to calculate the MPD and MNTD effect sizes. The 'taxa labels' null model swaps terminal phylogenetic labels across the OTU table, while maintaining alpha richness constant within a given community (Gotelli 2000). The 86 soil samples were assessed over 2050 internal nodes (minus those nodes containing less than for OTUs) averaged over the ten rarefied OTU tables, resulting in ~350k dispersion points for downstream examination. Examination of trends of phylogenetic dispersion among the PHACE soils was graphically displayed in a heat-map like figure, where nodes were sorted vertically by bacterial phylum, and the data was horizontally sorted by PHACE soil treatment and harvest date to identify overall patterns. Additionally, OTU-like tables were created from the dispersion dataset to identify potential indicator nodes of dispersion among the environmental variables. These tables were setup like the OTU tables, except that instead of per OTU abundance information, dispersion signals (separate tables for NRI and NTI) were reported for each phylogenetic node, relative to each soil sample. Finally, significance values for all statistical assessments were considered significant at $p<0.05$, alpha=0.01.

Results

The experimentally imposed PHACE treatments were marginally significant contributors to soil bacterial community structure (Table 3.1, Figure 3.1). Elevated CO_2 was not a significant contributor to differences in beta diversity among the soil bacterial community membership (p-val: 0.14) or community structure (p-val: 0.056). Warming did not significantly impact bacterial

community membership or structure (p-values: 0.286 and 0.354, respectively). The interactive effect of both elevated CO_2 and temperature were significant contributors to community structure (p-val: 0.041), but was not significant for bacterial community structure (p-val: 0.185). PCoA plots, colored by CO_2 and separately by temperature treatment reveal overlapping distributions of bacterial communities in ordination space (Figure 3.1).

All indices of collection time (year, month-year, and season of the soil harvest) showed significant (p-val < 0.001) impacts on beta-diversity among both soil bacterial community membership and structure (Table 3.1, Figure 3.2). In the PCoA plots colored by season and collection year (Figure 3.2), the patterns of separation in ordination space are very similar for both unweighted and weighted UniFrac beta-diversity distances. The PCoA colored by collection date shows clear separation of the final harvest (April 2009) from the other harvest dates along axis 1 (7.68% and 51.15% variation explained by unweighted and weighted UniFrac distances, respectively)(Figure 3.2b, d). The same pattern is evident in the PCoA colored by season, where summer samplings cluster, and the two spring collections cluster within harvest date but separate strongly in ordination space among the spring collection years (Figure 3.2a, c).

Ordinations of beta diversity as impacted by the three soil moisture temporal scales show no strong trends by week and month average soil moisture (Figure 3.3a, 3.3b, 3.3d, 3.3e), and this is statistically supported by marginal significance from the adonis tests (Figure 3.1). The week average soil moisture had nonsignificant impacts on community membership (p-val: 0.129) and marginal significance on the beta-diversity of community structure (p-val: 0.041). Month average soil moisture had no significant impacts on the bacterial communities (p-val: 0.29, 0.302, for unweighted and weighted, respectively), but year average did have significant impacts

on both community membership (p-val: 0.024) and structure (p-val: 0.009) (Table 3.1 and Figure 3.3).

The adonis tests revealed harvest date, and in particular, season, had strong influence on the soil bacterial communities. Therefore, subsets of taxa were identified as unique indicator OTUs to the spring versus summer harvests. These OTUs were identified with indicspecies Cran-R package (Table 3.2), and plotted onto the study-wide phylogeny in iTOL (Figure 3.4). The indicator species for the last two collection time points (Figure 3.4) shows a more or less exclusive distribution of indicator taxa of the spring collection versus summer collection for soil bacterial communities. Specifically, the indicator taxa for spring bacterial communities are primarily restricted to the Actinobacteria phylum, while the summer distribution is more ubiquitous throughout the bacterial phyla and high presence in the Bacteroidetes phylum (Figure 3.3). Table 3.1 reports the subset of the most significant (p-val < 0.05) indicator taxa separated by harvest date. Most of the indicator taxa in the April 2009 soils are in the Actinobacteria, while the other soils have more evenly distributed taxa among the bacterial phyla.

The assessment of bacterial community dispersion revealed similar trends to beta-diversity patterns. Large portions of the Actinobacteria and a portion of the Proteobacteria were significantly clustered (marked in red) with respect to null distributions of taxa that separate the April 2009 soils from the rest (Figure 3.5). Additionally, the other soil samples have significant clustering in the Acidobacteria, which is dampened in the April 2009 soils. Verrucomicrobia were primarily overdispersed (marked in blue) throughout all the samples, and most of the phylogenetic nodes. Using the indicspecies R package, significant 'indicator nodes' nodes were identified between spring and summer soil bacterial communities (Figure 3.6). Like the indicator taxa, the spring samples have most nodes of unique dispersion in the Actinobacteria.

Finally, alpha diversity only showed significant trends by two environmental factors. Faith's phylogenetic distance was marginally significantly different for the harvest date (p-val: 0.02) with higher alpha diversity in the July 2008 sampling, and by the year average soil moisture (p-val: 0.02, R^2=0.05). Additionally, the Chao1 index was significantly higher for the July 2008 collection time (p-value: 0.002).

Discussion

Despite seven years of substantial supplemental CO_2 and elevated temperature, any impacts of the experimentally imposed treatments at PHACE were obscured by the strong influence of seasonality (Table 3.1, Figure 3.1, 3.2). Previous studies have demonstrated that temporal heterogeneity patterns in soil moisture can be the key driver of soil microbial community structure and function (Fierer *et al* 2003, Hawkes *et al* 2011, Evans *et al* 2014). However, in this case, soil moisture was not the prominent driver of seasonal shifts in microbial community. Because the shift in microbial communities spans all PHACE treatment combinations, the prominent driver of this seasonal shift must be something that operates at the spatial level of the habitat. Possible drivers include drastic shifts in soil temperature, above the influence of the PHACE heating treatment. The seasonal driver may also be biotic. In southern Wyoming, the spring collection in April corresponds to leaf green up, where plants are beginning to photosynthesize and grow substantially after the winter. In contrast, the summer soil collections in July occur plants are at peak biomass, and/or just beginning to senesce. Plant root exudation, or other influence of the plants interactions with the soil may be more important than the direct impacts of elevated CO_2 and/or temperature. In fact, elevated CO_2 has been shown to increase

the exudation of plant-derived amino acids (Phillips *et al* 2006). Elevated CO_2 and temperature have been shown to impose substantial changes on plant community structure and function at the PHACE site (Morgan *et al* 2011, Suseela *et al* 2014), so the impacts of forecasted climate on soil microorganisms may just more strongly through plant-soil interactions. However, other studies have found no significant impacts of changes in aboveground community on soil microbial communities (Carey *et al* 2015). Either way, the PHACE treatments, and more broadly, forecasted climate change may have substantial impacts on soil microbial community structure and function through indirect impact on changing plant biomass structure and function.

The distinct, consistent shift in soil bacterial community structure with harvest season highlights some new hypotheses about why season matters for bacterial community structure in this study. Because the PHACE site is situated in a northern mixed grassland prairie, subjected to strong seasonality, it is not too surprising that season is a prominent driver of bacterial community structure.

The primary goal of this study was to understand the bacterial community structure of soils at the PHACE field site, and to examine relative influence of the environmental factors, both experimentally imposed and inter-annual variability in plot conditions. Seasonality obscured any influence of the PHACE treatments on structuring soil bacterial community structure and phylogenetic dispersion. Seven years of altered abiotic conditions is a long time from the standpoint of microbial turnover, which operates on much faster temporal scales. In the context of climate change, short-term stochasticity of microbial communities driven by inter-annual variability in habitat conditions may impose stronger influences on the long-term trajectory of the soil microbial community, its functional capacity and ecosystem as a whole.

Tables

Table 3.1. Influence of factors on bacterial community. Statistical assessment of the correlation of each factor with bacterial community structure. Significance was determined with Adonis tests of each factor, where the p-value is reported and the effect size, represented as R^2 as a measure of variation explained by each factor (McArdle & Anderson 2001).

	Unweighted UniFrac		Weighted UniFrac	
	p-value	*r2*	*p-value*	*r2*
PHACE Treatments				
2-Factor	0.041*	0.03966	0.185	0.04668
Temperature	0.286	0.01229	0.354	0.0115
CO2	0.014*	0.0155	0.056.	0.02625
Collection Time				
Year	0.001***	0.04912	0.001***	0.30971
Season	0.001***	0.03442	0.001***	0.14568
Month-Year	0.001***	0.11298	0.001***	0.51106
Soil moisture				
Week Average	0.129	0.01324	0.041*	0.03153
Month Average	0.29	0.01222	0.302	0.01273
Year Average	0.024*	0.01517	0.009**	0.04347

Table 3.2. Indicator OTUs for each collection date. Bacterial taxa identified as significant indicators (p-val: 0.005) of each collection date.

	OTU	*Phylum*	*Class*	*Order*	*Family*	*Genus*
July 2007	OTU_390	Actinobacteria	Thermoleophilia	Solirubrobacterales		
	OTU_455	Actinobacteria	Rubrobacteria	Rubrobacterales	Rubrobacteraceae	Rubrobacter
	OTU_558	Actinobacteria	Thermoleophilia	Solirubrobacterales		
	OTU_577	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	
	OTU_951	Actinobacteria	Rubrobacteria	Rubrobacterales	Rubrobacteraceae	Rubrobacter
	OTU_1130	Actinobacteria	Thermoleophilia	Gaiellales	Gaiellaceae	
	OTU_1207	Actinobacteria	Actinobacteria	Actinomycetales		
	OTU_226	Bacteroidetes	Saprospirae	Saprospirales	Chitinophagaceae	
	OTU_411	Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	
	OTU_2110	Chloroflexi	TK10	B07		
	OTU_1544	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales		
	OTU_256	Planctomycetes	Phycisphaerae	WD2101		
	OTU_3334	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	Gemmata
	OTU_532	Proteobacteria	Alphaproteobacteria			
	OTU_732	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium
April 2008	OTU_1191	Actinobacteria	Actinobacteria	Actinomycetales		
	OTU_3400	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	Pseudonocardia
	OTU_995	Gemmatimonadetes	Gemmatimonadetes			
	OTU_829	Proteobacteria	Betaproteobacteria			
July 2008	OTU_572	Acidobacteria	iii1-8	DS-18		
	OTU_2934	Acidobacteria	Acidobacteria-6	iii1-15		
	OTU_3727	Acidobacteria	Acidobacteria-6	iii1-15		
	OTU_1768	Actinobacteria	Thermoleophilia	Solirubrobacterales	Solirubrobacteraceae	
	OTU_826	Bacteroidetes				
	OTU_1340	Bacteroidetes	Saprospirae	Saprospirales	Chitinophagaceae	
	OTU_2937	Bacteroidetes	At12OctB3			
	OTU_1404	Chloroflexi	Ellin6529			
	OTU_839	Gemmatimonadetes	Gemmatimonadetes			
	OTU_2662	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales		
	OTU_2715	Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	Gemmata
	OTU_859	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	
	OTU_1029	Proteobacteria	Betaproteobacteria	SC-I-84		
	OTU_2554	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	
	OTU_3027	Proteobacteria				
April 2009	OTU_899	Acidobacteria	iii1-8	DS-18		
	OTU_926	Acidobacteria	Chloracidobacteria	RB41		
	OTU_2304	Acidobacteria	Chloracidobacteria	RB41		
	OTU_10	Actinobacteria	Actinobacteria	Actinomycetales		
	OTU_275	Actinobacteria	Thermoleophilia	Solirubrobacterales		
	OTU_630	Actinobacteria	Actinobacteria	Actinomycetales	Sporichthyaceae	
	OTU_863	Actinobacteria	Thermoleophilia	Solirubrobacterales	Solirubrobacteraceae	
	OTU_870	Actinobacteria	Rubrobacteria	Rubrobacterales	Rubrobacteraceae	Rubrobacter
	OTU_1161	Actinobacteria	Thermoleophilia	Solirubrobacterales		
	OTU_1432	Actinobacteria	Rubrobacteria	Rubrobacterales	Rubrobacteraceae	Rubrobacter
	OTU_2209	Actinobacteria	Actinobacteria	Actinomycetales	Mycobacteriaceae	Mycobacterium
	OTU_2909	Actinobacteria	Actinobacteria	Actinomycetales	Nocardiaceae	Rhodococcus
	OTU_3373	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Kribbella
	OTU_3713	Actinobacteria	Actinobacteria	Actinomycetales	Frankiaceae	
	OTU_432	Armatimonadetes	Chthonomonadetes	Chthonomonadales	Chthonomonadaceae	
	OTU_824	Armatimonadetes	0319-6E2			
	OTU_1748	Chloroflexi	Chloroflexi	AKIW781		
	OTU_538	Gemmatimonadetes	Gemmatimonadetes			
	OTU_53	Proteobacteria	Deltaproteobacteria	Myxococcales		
	OTU_1744	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	
	OTU_3876	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	
	OTU_397	Verrucomicrobia	Spartobacteria	Chthoniobacterales	Chthoniobacteraceae	DA101
	OTU_4054	Verrucomicrobia	Spartobacteria	Chthoniobacterales	Chthoniobacteraceae	Ellin506

Figures

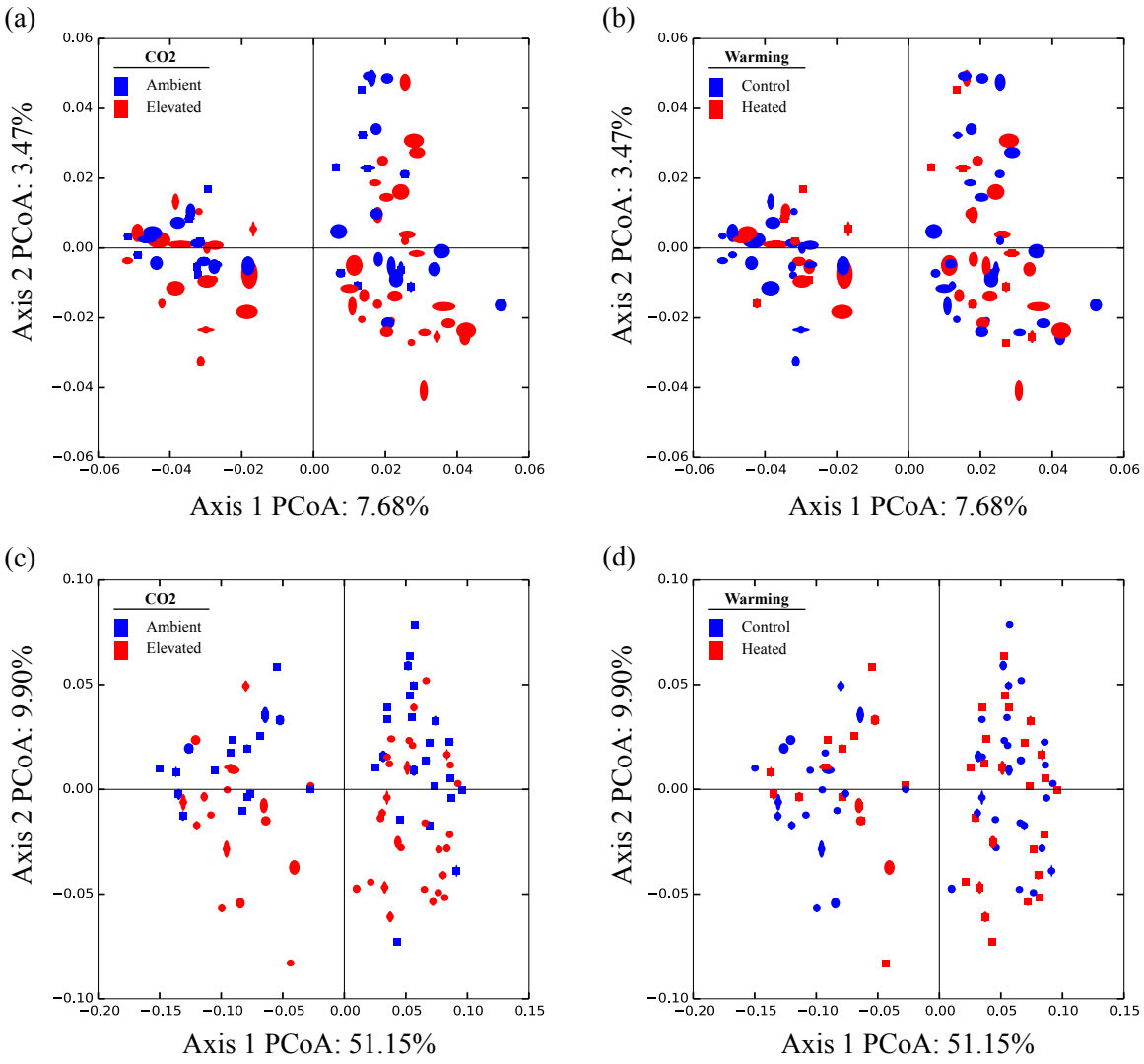


Figure 3.1. Ordination of UniFrac pairwise distances by each PHACE treatment PCoA plots, where each sphere represents a jackknifed (1000x) estimation of beta-diversity of bacterial community structure among the PHACE soils, using unweighted UniFrac (a, b), and abundance-weighted UniFrac (c,d). The representation of these bacterial communities are colored by PHACE CO_2 treatment (a,c) with ambient as blue and elevate CO_2 as red; and PHACE heating (b, d) with control as blue and heated as red.

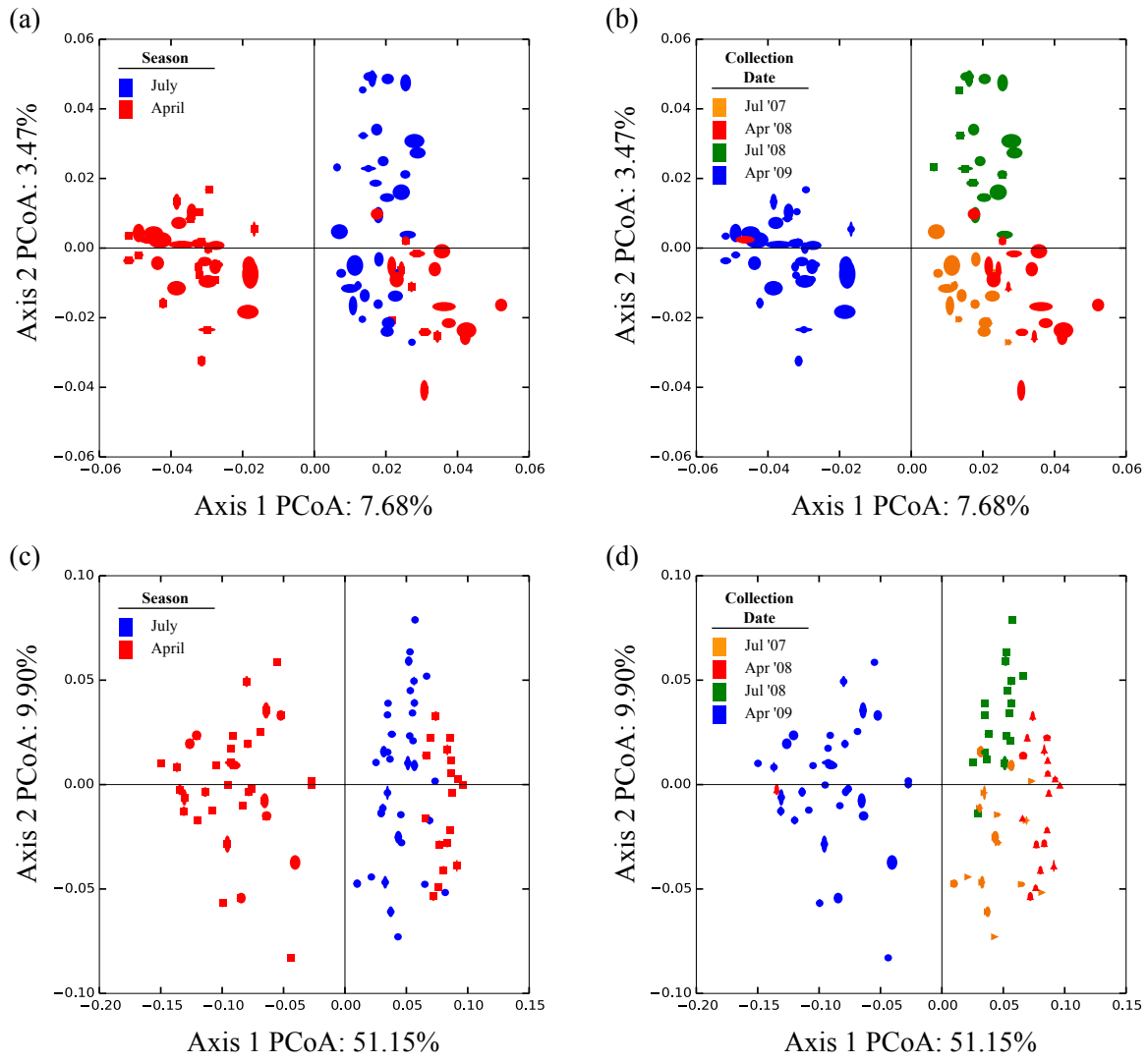


Figure 3.2. Ordination of UniFrac pairwise distances by season and collection date. PCoA plots, where each point represents a jackknifed estimation of beta-diversity of bacterial community structure among the PHACE soils, using unweighted UniFrac (a, b), and abundance-weighted UniFrac (c,d). The representation of these bacterial communities are colored by season (a,c) with July as blue and April as red; and collection date (b, d) colored by the four collection times.

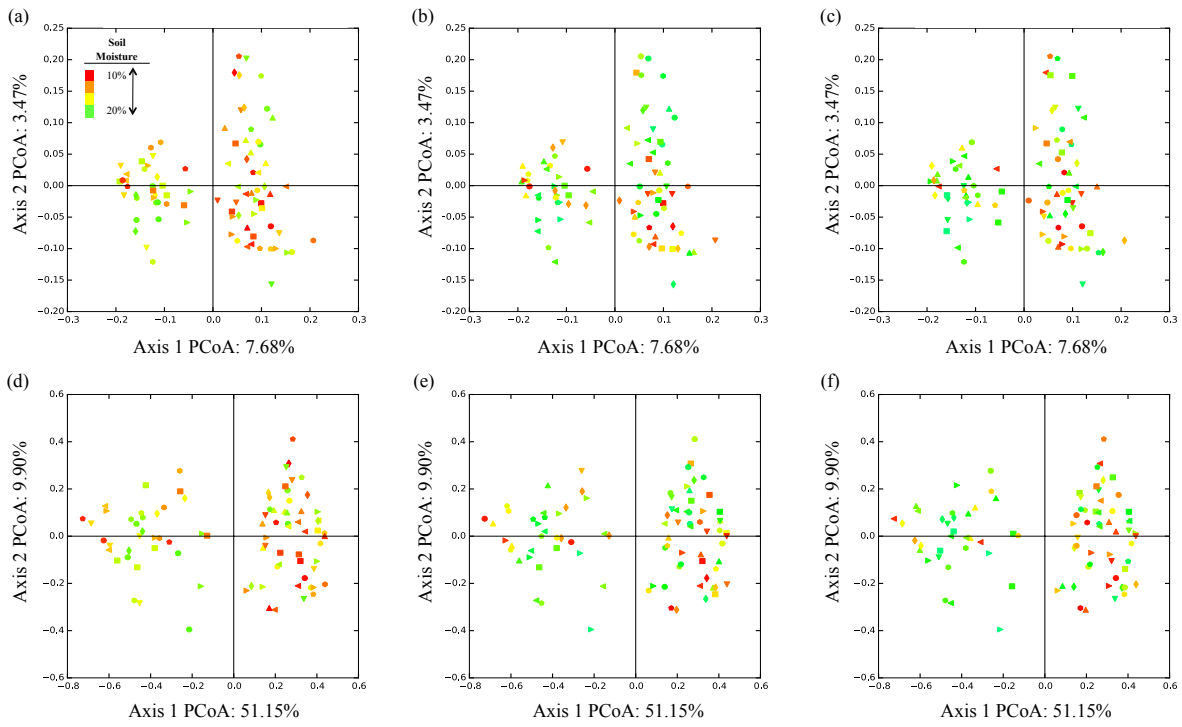


Figure 3.3. Ordination of UniFrac pairwise distances by season and collection date. PCoA plots, where each point represents an estimation of beta-diversity of bacterial community structure among the PHACE soils, using unweighted UniFrac (a, b, c), and abundance-weighted UniFrac (d, e, f). The representation of these bacterial communities are gradient colored by soil moisture %, with week prior average (a,d), month prior average (b, e), and year prior average (c, f).

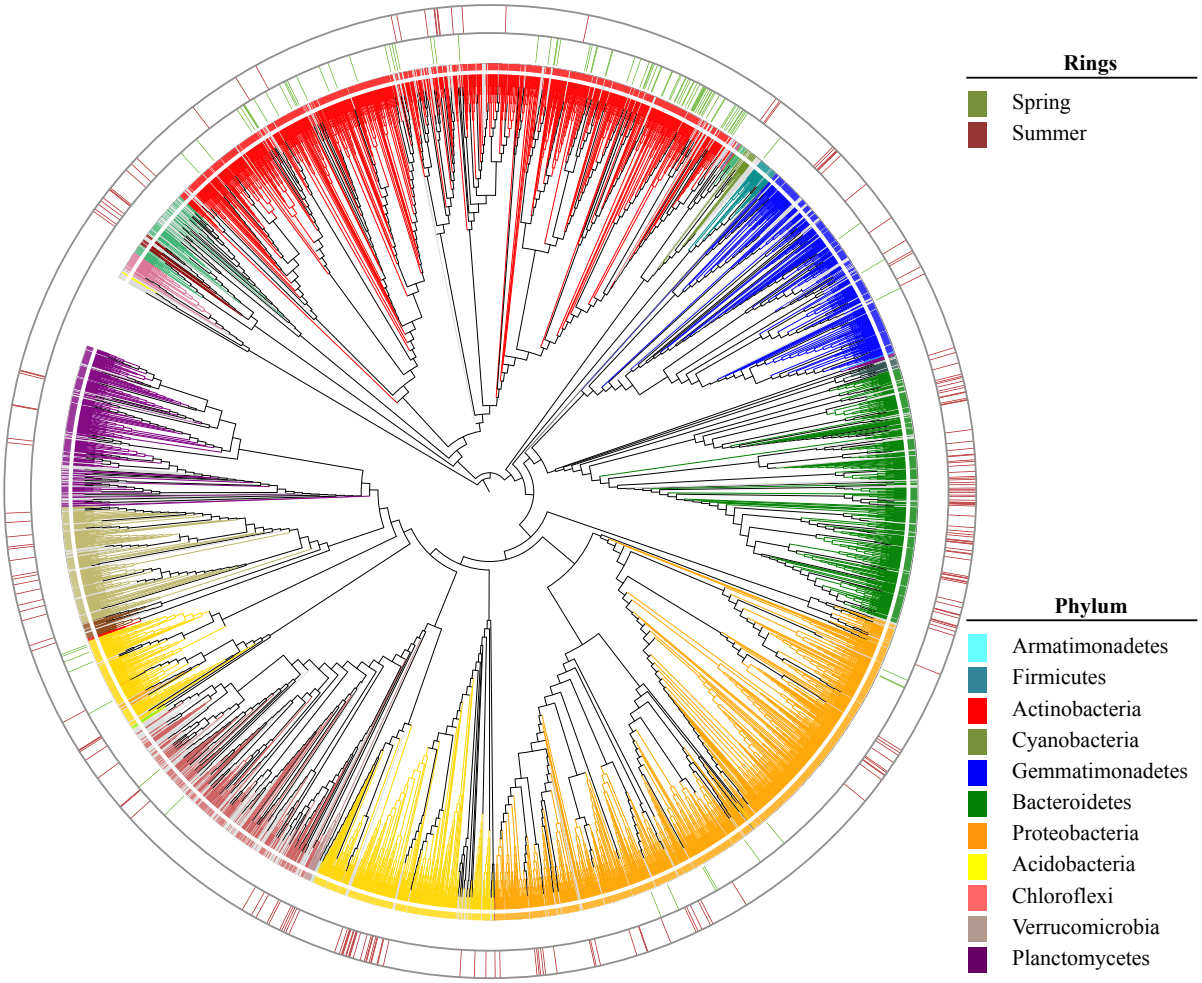


Figure 3.4. Indicator OTUs for each season mapped onto a study wide phylogeny. The outer rings (green: spring, red: summer) identify the distribution of bacterial OTUs, marked as putative indicator species, in phylogenetic space. The phylogeny is colored by bacterial phylum.

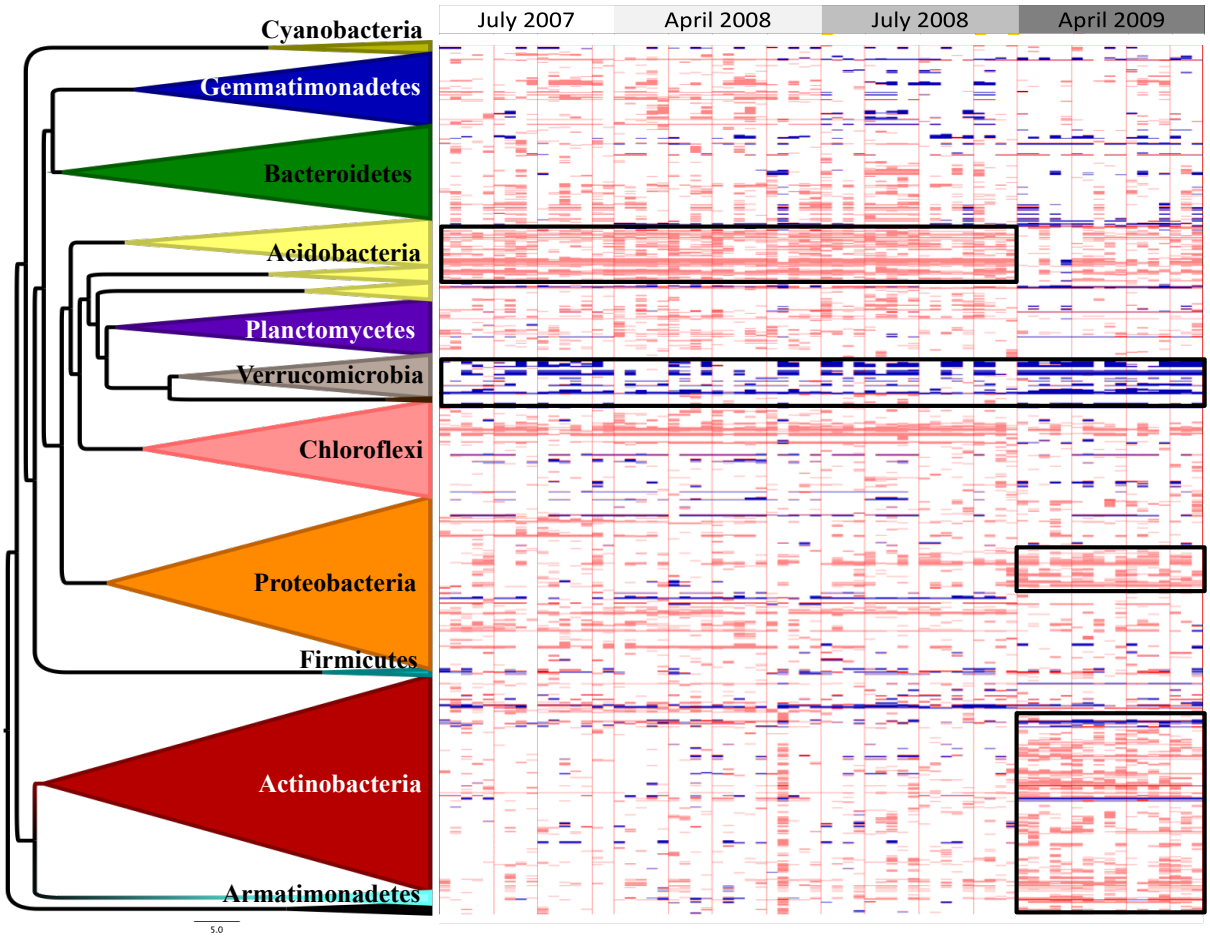
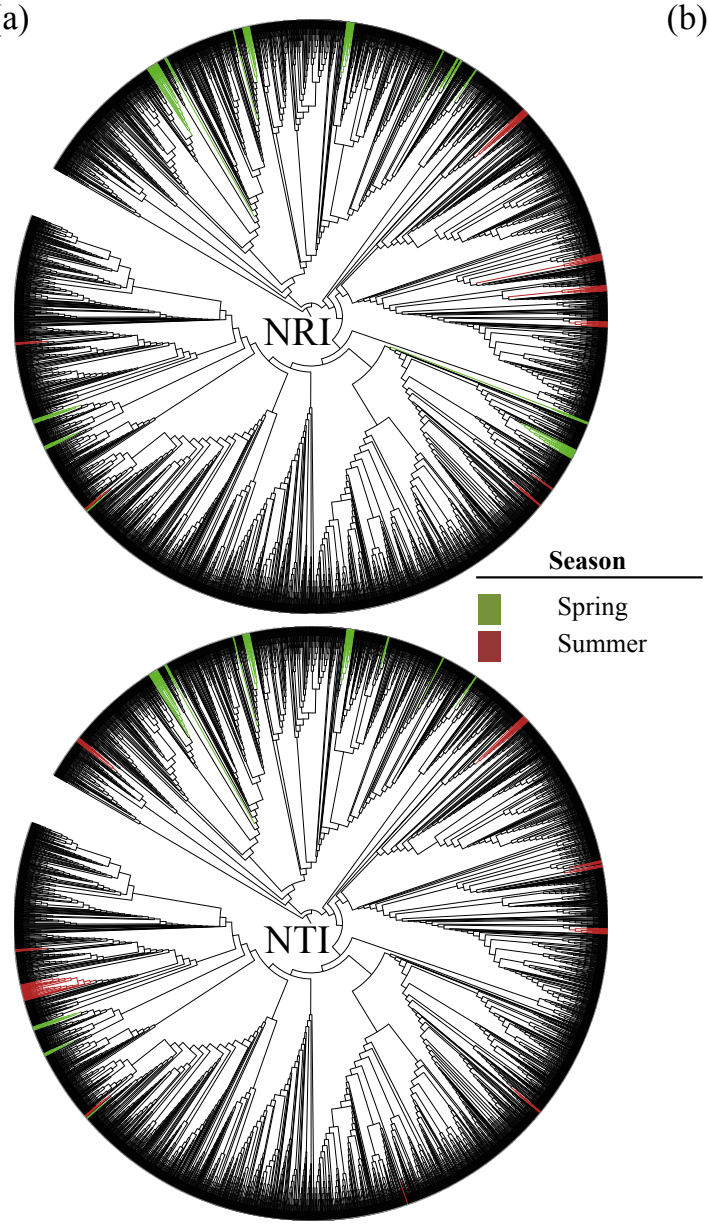


Figure 3.5. Map of phylogenetic dispersion, vertically ordered by collection date and horizontally by bacterial phylum. Each row of the map represents the dispersion patterns for one phylogenetic internal node. Each column represents the phylogenetic dispersion patterns for each soil sample. The red portions of the map represent phylogenetic clustering and blue by phylogenetic overdispersion, estimated from NRI. On the left the study-wide phylogeny, with nodes collapsed to bacterial phylum.

(a)



(b)

		Net Relatedness Index	Nearest Taxon Index
	node	*p-value*	*p-value*
Spring			
	n29	0.005**	0.005**
	n52	0.005**	
	n58	0.015*	
	n74	0.005**	0.005**
	n1479	0.03*	
	n1784	0.005**	
	n2541	0.015*	
	n2542	0.02*	
	n2544	0.005**	0.005**
	n2545	0.005**	0.005**
	n2590	0.01**	0.03*
	n2591	0.005**	
	n2592	0.005**	0.01**
	n2593	0.005**	
	n2850	0.015*	0.005**
	n3560	0.005**	0.005**
	n3570	0.01**	0.005**
	n3624	0.005**	0.005**
	n3675	0.005**	0.01**
	n3691	0.005**	0.005**
	n3692	0.005**	0.005**
	n3693	0.005**	0.005**
	n3695		0.005**
	n3698	0.005**	
	n3912		0.04*
	n4008	0.015*	
	n4009	0.005**	0.005**
Summer			
	n277	0.01**	0.005**
	n278	0.01**	
	n282	0.005**	0.005**
	n284	0.005**	
	n330	0.01**	0.01**
	n331	0.01**	0.005**
	n374	0.01**	
	n588	0.005**	0.005**
	n592		0.005**
	n597	0.005**	0.005**
	n1000	0.005**	0.01**
	n1112		0.02*
	n1516	0.01**	
	n1549	0.025*	0.045*
	n2383		0.005**
	n2446	0.015*	0.005**
	n2794	0.02*	0.04*
	n3224		0.005**

Figure 3.6. Indicator phylogenetic nodes of dispersion for each season. (a) Putative indicator nodes identified in the study-wide phylogeny for NRI (top) and NTI (bottom) and colored by season (spring: green, summer: red). (b) Table displaying the indicator nodes, and level of significance for NRI and/or NTI.

References

Angel R, Soares MI, Ungar ED, Gillor O. (2010) Biogeography of soil archaea and bacteria along a steep precipitation gradient. ISME J 4(4): 553-563.

Bais, H.P., T.L. Weir, L.G. Perry, S. Gilroy, J.M. Vivanco. (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. Annu Rev Plant Biol 57: 233-266.

Bell C, Carillo Y, Boot C, Rocca J, Pendall E, Wallenstein M (2014)Rhizosphere stoichiometry: are C : N : P ratios of plants, soils, and enzymes conserved at the plant species-level? 201(2):505-517.

Caceres M. and Jansen F. (2015) Associations between species and groups of sites: indices and statistical inference. Ecology 90(12): 3566-3574.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6: 1621-1624.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al*. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7(5): 335-336.

Carey CJ, Beman JM, Eviner VT, Malmstrom CM, Hart SC. (2015) Soil microbial community structure is unaltered by plant invasion, vegetation clipping, and nitrogen fertilization in experimental semi-arid grasslands. Front. Microbiol. 6: 466.

Collins SL, Sinsabaugh RL, Crenshaw C, Green L, Porras-Alfaro A, Stursova M. (2008) Pulse dynamics and microbial processes in arid land ecosystems. J. Ecol. 96: 413-420.

DeAngelis KM, Pold G, Topcuoglu BD, van Diepen LTA, Varney RM, Blanchard JL, Melillo J, Frey SD. (2015) Long-term forest soil warming alters microbial communities in temperate forest soils. Front. Microbiol. 6: 104.

Dufrene M. and Legendre P. (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol. Monog. 67(3): 345-366.

Edgar RC. (2010) Search and clustering order of magnitude faster than BLAST. Bioinformatics 26(19): 2460-61.

Evans SE, Wallenstein MD, Burke IG. (2014) Is bacterial moisture niche a good predictor of shifts in community composition under long-term drought? Ecology 95: 110-122.

Fierer N, Schimel JP, Holden PA. (2003) Influence of drying-rewetting frequency on soil bacterial community structure. Microbiol. Ecol. 45: 63-71.

Fierer, N. and R.B. Jackson. (2006) The diversity and biogeography of soil bacterial communities. PNAS. 103(3): 626-631.

Freschet, G.T., W.K. Cornwell, D.A. Wardle, T.G. Elumeeva, W. Liu, B.G. Jackson, V.G. Onipchenko, N.A. Soudzilovskaia, J. Tao, J.H.C. Cornelissen. (2013) Linking litter decomposition of above- and below-ground organs to plant-soil feedbacks worldwide. J Ecol 101(4): 943-952.

Gotelli NJ (2000) Null model analysis of species co-occurrence patterns. Ecology. 81(9): 2606-2621.

Hawkes CV, Kivlin SN, Rocca JD, Huguet V, Thomsen MA, Suttle KB. (2011) Fungal community responses to precipitation. Global Change Biol. 17(4): 1637-1645.
Hobbie S. (1992) Effects of plant species on nutrient cycling. Trends in Ecol & Evol. 7:336–339.
Hughes Martiny JB, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine C, Kane M, Adams Krumins J, Kuske CR, Morin PJ, Naeem S, Ovreas L, Reysenbach A, Smith VH, Staley JT. (2006) Microbial biogeography: putting microorganisms on the map. Nat. Rev. Microbiol. 4: 102-112.
Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. (2010) Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26: 1463-64.
Lauber CL, 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. Appl Environ. Microbio. 75(15): 5111-20.
Letunic I, Bork P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acid Res.
Liu K, Linder CR, Warnow T (2011) RAxML and FastTree: Comparing two methods for large-scale maximum likelihood phylogeny estimation. PLoS One. 6(11): 1-11.
Lozupone, C. and Knight R. (2005) UniFrac: A new phylogenetic method for comparing microbial communities. Appl. & Env. Microbiol. 71(12): 8228-8235.
Marschner P, Kandeler E, Marchner B. (2003) Structure and function of the soil microbial community in a long-term fertilizer experiment. Soil Biol. Biochem. 35(3): 453-461.
McArdle BH, Anderson MJ. (2001) Fitting multivariate models to community data: A comment on distance-based redundancy analysis. Ecology 82: 290-297.
McCoy CO, Matsen FA. (2013) Abundance-weighted phylogenetic diversity measures distinguish microbial community states and are robust to sampling depth. PeerJ Comp. Sci.
Michonneau *et al* (2015). Base package for phylogenetic structure and comparative data. Cran-R package.
Mirarab S, Nguyen N, Warnow T. (2014) PASTA: ultra-large multiple sequence alignment. Proceedings of RECOMB.
Morgan JA, LeCain DR, Pendall E, Blumenthal DM, Kimball BA, Carrillo Y, Williams DG, Heisler-White J, Dijkstra FA, West M. (2011) C4 grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland. Nature. 476: 202-205.
Phillips DA, Fox TC, Six J. (2006) Root exudation (net efflux of amino acids) may increase rhizodeposition under elevated CO_2. Global Change Biol. 12(3): 561-567.
Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. (2013) The SILVA ribosomal RNA database project: improved data processing and web-based tools. Nucl. Acids Res. 41: D590-D596.
Schindlbacher A, Rodler A, Kuffner M, Kitzler B, Sessitsch A, Zechmeister-Boltenstern S (2011) Experimental warming effects on the microbial community of a temperate mountain forest soil. Soil Biol. Biochem. 43(7): 1417-1425.
Suseela V., Triebwasser-Freese D., Linscheid N., Morgan J.A., Tharayil N. (2014) Litters of photosynthetically divergent grasses exhibit differential metabolic responses to warming and elevated CO2. Ecosphere. 5(9): 106.
Suseela, V., N. Tharayil, B. Xing, J.S. Dukes. (2015) Warming and drought differentially influence the resorption of elemental and metabolic nitrogen pools in Quercus rubra. Global Change Biology. (In Press)

Tharayil, N. V. Suseela, D.J. Triebwasser, C.M. Preston, P.D. Gerard, J.S. Dukes. (2011) Changes in the structural composition and reactivity of Acer rubrum leaf litter tannins exposed to warming and altered precipitation: climatic-stress-induced tannins are more reactive. New Phyt. 191: 132-145.

Thébault A, Clément J, Ibanez S, Roy J, Geremia RA, Pérez CA, Buttler A, Estienne Y, Lavorel S. (2014) Nitrogen limitation and microbial diversity at the treeline. Oikos 123(6): 729-740.

Webb CO, Ackerly DD, Kembel SW. (2008) Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics 18: 2098-2100.

Williams MA, Rice CW. (2007) Seven years of enhanced water availability influences the physiological, structural, and functional attributes of a soil microbial community. App. Soil Ecol. 35: 535-545.

Chapter 4: Climate-mediated shifts in plant leachate chemistry: A novel driver of soil microbial community structure and function?

Plants are principal drivers of soil microbial community structure and function. We know that plant metabolites substantially contribute to soil function, through root exudation and litterfall, but less is known of the impact of the 16% of plant leaf carbon permeating the soil profile as water-soluble plant metabolites. Leaf chemistry varies among species, and can also vary within species due to certain abiotic cues. In light of forecasted climate shifts, understanding the link between plant litter leachate and soil function will help us better understand how terrestrial habitats will respond and function in the future. The main objective of this study was to assess inter- and intraspecific variability in aboveground plant litter leachate in a grassland system, and its consequential belowground impacts on microbial community structure and function. Overall, we found evidence that shifts in plant chemistry, both in terms of species and long-term exposure to different climate conditions, influenced belowground function and microbial community structure. In particular, we found that one of the invasive plants, Linaria dalmatica, possessed distinct leachate chemistry, impacted soil function, and imposed stronger influences on the bacterial community structure relative to other species' leachates. Separation in the chemistry of plant litter leachate was primarily driven by a handful of metabolites, including porphine, fumaric acid, shikimic acid and galacturonic acid. The primary driver of bacterial community structure in the incubation was harvest time, though L. dalmatica communities still separated from the other communities. This consistent shift in bacterial community is likely due to the drastic impact of altering soil physical structure and water availability. As the invasive L. dalmatica exhibited unique impact on soil microbial structure and

function, it may highlight the influence of plant invasions in general and how they can present substantial functional consequences under future climate. Broadly, our study assesses the importance of plant litter leachate in regulating the belowground response to forecasted climate.

Introduction

Plants are known to regulate their surrounding soil habitat (Hobbie 1992) through chemical contribution of plant metabolites via root exudation (Bais *et al* 2006, Bird *et al* 2011, Shi *et al* 2011, Phillips *et al* 2011) and litterfall (Freschet *et al* 2013, Kotroczó *et al* 2014, Fanin *et al* 2014). Less is known about the influence of plant litter leachate, defined as the water-soluble fraction of plant tissue, on belowground microbial community structure and function as it permeates the soil profile (Fanin *et al* 2014). Upwards of 16% of plant leaf tissue carbon can enter the soil as plant leachate (Soong *et al* 2014). Water-soluble plant metabolites have been shown to contribute substantially to belowground structure and function (Bowman *et al* 2004, Ushio *et al* 2013, Fanin *et al* 2013), due in part to the detailed differences in chemistry among plant species (Rinkes *et al* 2014). Plant leachate can serve as C-rich energy source for heterotrophic microorganisms (Waldrop & Firestone 2004), but some plant metabolites inhibit microbial function (Goldfarb *et al* 2011, Triebwasser *et al* 2012, Schmidt *et al* 2013). It is therefore essential to understand the influence this substantial portion of plant productivity entering the soil profile as leachate on soil microbial community structure and function. Additionally, it is important to understand how this influence may vary by plant species and/or by plant community composition (Fig. 4.1).

Intraspecific variability in plant tissue chemistry is well known (Coley 1987, Jamieson & Bowers 2010, Moore *et al* 2013). Species-specific niche conditions, including life history

strategies, photosynthetic pathway, and the plant's functional role in a habitat can greatly influence plant responses to environmental conditions, and the tissue chemistry often reflects this response. Stoichiometric constraints also limit what suite of metabolites can be produced by the plant, and to what extent a plant can physiologically respond to the proximate environmental conditions. These stoichiometric constraints could then result in shifts in the plant litter leachate chemistry.

Plant chemistry can also vary intraspecifically (Nicotra *et al* 2010, Coq *et al* 2011, Tharayil *et al* 2011, Suseela *et al* 2015a). For instance, sugar maple (*Acer rubrum* L.) shifts polyphenolic production to more reactive forms in response to drought stress (Tharayil *et al* 2011), which effects soil function by strongly binding with N-rich proteins, thereby inhibiting microbial extracellular enzyme activity (EEAs) and shifting the availability of N in the soil (Triebwasser *et al* 2012, Schmidt *et al* 2013, Halvorson *et al* 2011). Therefore, the indirect impacts of climate, through altered plant tissue chemistry, can have scaffolding effects on soil function and the ecosystem as a whole.

In the context of forecasted climate change, assessing the impacts of the primary drivers of belowground activity, like water availability and temperature, is key to understanding the soil response to climate shifts. However, in addition to the direct impacts of these abiotic drivers, the *indirect* impact of climate-mediated shifts in plant species chemistry and plant community structure may even be stronger than the direct impacts on soil function (Fig. 4.1). Plant communities and comprising individuals respond concurrently to altered climate conditions (Fig. 4.1), which may result in two-fold impacts on what leachate enters the soil. If inter- and intraspecific plant litter leachate differentially impact soil microorganisms, then there is a direct link between plant community and microbial community structure and function.

The primary objective of this study was to examine the importance of plant leaf litter leachate in driving the response of soil microbial community structure and function. The field site selected for this study was the Prairie Heating and Carbon Dioxide Enrichment (PHACE) site (Cheyenne, WY) to address the potential two-fold impacts of plant leachate in the context of forecasted climate change (Fig. 4.1).

The specific objectives of this study were:

1. Does plant tissue vary interspecifically? *What variability exists in aboveground plant tissue chemistry amongst the most abundant species present at PHACE?*
2. Do abiotic factors impose intraspecific difference in plant chemistry? *How does elevated CO2 and temperature influence intraspecific plasticity in aboveground tissue chemistry?*
3. Are soil function and bacterial community structure impacted by shifts in leachate chemistry?

Our study examines to what extent among species variability in aboveground plant litter leachate chemistry translates to changes in soil function (including assessment of microbial respiration, inorganic N, microbial biomass CN, and extracellular enzymes) and soil bacterial community structure. Because the plant tissue is from the PHACE field site, we can also test the influence of altered climate on species-specific plant chemistry and the consequences for belowground function. In total, this study allows us to examine the climate-mediated effects of altered plant leachate chemistry on soil microbial community structure and function in a grassland ecosystem.

Materials and Methods

Study site and sample collection

The soil and plant tissue used in this experiment are from the PHACE field experiment. PHACE was exposed to elevated CO_2 and warming for seven years in a complete two factor randomized block experimental design. The experimental site is located 15 km west of Cheyenne, WY (41.198700 N, -104.887048 W), and is characterized as a semi-arid northern mixed grassland prairie with a mean annual precipitation of 384 mm. The average soil at PHACE is pH is 7.0 and is characterized as mollisol (Morgan *et al* 2011).

The soils and plant tissue were collected from PHACE at the peak of the growing season in July 2013 and just prior to the end of the PHACE field treatments. One site level soil was collected from non-compacted areas immediately adjacent to the PHACE rings to ensure homogeneity of the soil and minimize any bias towards a given plant species. The soils were sieved to 2 mm on site, and stored in a cooler for transport and in a temperature controlled cold room at 4 ^{o}C until the laboratory incubation setup. Subsamples were taken to determine gravimetric water content (%GWC), pH, and saved for starting conditions for the bulk functional assays.

To assess the impact of leachate on this particular soil habitat, we collected aboveground plant litter from a subset of the most abundant species found at the PHACE site, capturing both functional breadth and native-invasive status. These plant species included: *Artemisia frigida* (Willd.), *Centaurea diffusa* (Lam.), *Linaria dalmatica* (L.), and *Hesperostipa comata* (Trin. & Rupr.). We also collected aboveground plant tissue from *Pinus sylvestris* (L.) to serve as the off-plot functional outgroup. Details on the plant species used, including broad functional group, family, and invasive status, are outlined in Table 4.1. In total, there were 17 plant litter samples,

from the four plant species exposed to each of the four PHACE treatments (ambient (ac); elev. CO_2, unwarmed (Ec); ambient CO_2, warmed (aH); elev. CO_2, warmed (EH);), and the ambient *Pinus sylvestris* outgroup. The timing of the plant litter harvest was ideal as it was late enough in the season for no substantial plant growth to continue, yet not after the bulk of water-soluble metabolites were lost to the soil as leachate.

Plant litter preparation and leachate

The aboveground plant litter was cut into 2 cm pieces to ensure similar size pieces among plant species. A subsample of each litter sample was also set aside for C:N analysis. Multiple room temperature water extraction events were imposed on 1.25g of each dried plant tissue to extract the water-soluble leachate. Each sample leachate was collected in a sterile modified filtration apparatus. In trial tests, the original 0.45 μm pore size was ideal for removing microorganisms. However, this small pore size severely clogged the filters after just one leaching, so modified filtration systems were developed by replacing the intact filter with sterile, acid washed/rinsed 20 μm nylon mesh and a layer of inert SiO_2 to prevent clogging by plant litter and leachate materials (Fig. 4.2). To minimize microbially processed plant-derived metabolites that develop soon after plant litter is wet (Bakker *et al* 2011), ten continuous leaching events were implemented over a 12-hour period. Each leaching event consisted of adding 20 mL of sterile pH-balanced deionized H_2O to the litter atop the filtration apparatus (Fig. 4.2). After two minutes, the vacuum was attached to pull the room temperature extractable metabolites from the intact plant litter. The leachate samples were concentrated 10:1 and stored frozen until use. Leachate subsamples were also shipped to Clemson University for small-molecule and polymeric metabolite analysis in the Tharayil laboratory at Clemson University. These

subsamples were lyophilized, re-dissolved in 100% methanol and transferred to low volume crimp top vials (VWR 97058-832), and re-dried over N gas for shipment. Small molecule metabolites identified using an Agilent 7980A gas chromatography-mass spectrometry (GC-MS). Twenty µL of arabitol was included with each leachate sample as an internal standard, and 20 µL of myristic acid as a retention lock. AMDIS (amdis.net, chemdata.nist.gov) identification was determined by retention index and mass spectral fragments using a curated metabolomics library (MUAL#203) and CAS# reported for each unique compound found within all the samples (Table 4.2).

Laboratory microcosm and respiration reading

To test the impact of both plant species and PHACE treatment of the plant litter on soil function, a laboratory incubation was implemented using 152 soil microcosms. The experimental design consisted of randomized three factor experimental design. One factor, plant species had four treatments (*A.frigida, C.diffusa, H.comata, L.dalmatica*), which allowed for examination of the interspecific differences in leachate impacts on soil microbial function. The second factor was the four PHACE field treatments (ac, aH, EH, Ec), which allowed for testing for any intraspecific differences in leachate impacts on soil, as influenced by the long-term PHACE field treatments. Finally, incubation time factor was implemented to test for any dissipation, or increased soil functional differences since time of leachate application. An off plot pine leachate sample served as the functional outgroup to this incubation. Finally, there were four experimental reps for each factor combination, along with four water-only and four no-water control soil microcosms.

Each soil microcosm consisted of ~60 g field fresh, homogenized soil in a sterilized, acid-washed Mason jar, equipped with two Wheaton natural rubber flange septa (Wheaton Scientific, Millville, NJ). The septa in each jar lid was sealed with 100% silicone and cured for five days. The efficacy of each jar's seal was tested by adding known amounts of CO_2 and testing for retention five days later. The soil in each jar was independently adjusted to 40% water holding capacity with the appropriate leachate and immediately sealed. Background CO_2 levels were recorded, and the incubation was placed in a dark, temperature controlled room at 25 ^oC.

Cumulative heterotrophic respiration (soil CO_2 production per g dry soil) was measured on each microcosm throughout the week and month incubation times. In the first week of the incubation experiment, daily measurements of CO_2 were collected using the EcoCore (http://ecocore.nrel.colostate.edu) in house LI-COR LI 6552 infrared gas analyzer (LI-COR Biosciences, Lincoln, NB). In subsequent weeks, CO_2 measurements were collected every other day. After each reading, the microcosms were flushed with CO_2-free air to prevent inhibitory effects of high CO_2 on microbial function (Debs-Louka *et al* 1999). The µmols of CO_2 produced by each jar was estimated using a standard curve (R^2>0.99) of a known CO_2 concentration, generated for each LI-COR run. Finally, the respiration values were adjusted to end units of µmols CO_2 per g dry soil per hr.

Functional Assays

In addition to the heterotrophic respiration readings collected during the incubation, destructive soil assays were also conducted at incubation harvest to assess the impact of plant species' leachates on soil function. These assays include: %GWC, microbial biomass carbon and nitrogen (MBC, MBN), inorganic nitrogen (NO_3^-, NH_4^+), and seven extracellular enzyme assays

(alpha-glucosidase (AG), beta-glucosidase (BG), cellobiohydrolase (CB), leucine aminopeptidase (LAP), N-acetylglutamate synthase (NAG), phosphatase (PHOS), and xylanase (XYL)).

Gravimetric water content (GWC) was estimated by weighing out 5 g of fresh soil, drying it overnight in a 60 ^{o}C oven and reweighing the cooled dry soil and subtracting off the pre-weighed tin. GWC was then calculated as the difference between the soil weights over the dry soil weight, multiplied times 100%. As the starting moisture conditions were experimentally controlled during incubation setup, GWC was primarily assessed as verification that moisture conditions did not significantly vary among the soil microcosms.

Microbial biomass carbon and nitrogen were determined using the standard chloroform fumigation protocol (Joergensen 1996). The general rationale of chloroform fumigation is to kill and lyse the soil microbial cells when exposed to chloroform for an extended period of time. The cell lysis releases their organic matter, which can then be isolated from the soil by dissolving in a 0.5M K_2SO_4 salt solution. Two separate 15 g experimental reps of each soil samples were subset from each microcosm at the week and month incubation lengths. One sample is saved at 4 ^{o}C while the other is fumigated for five days in the presence of 50 mL of 100% chloroform under vacuum conditions. After the fumigation was complete, the fresh soil from each sample was mixed with 75 mL 0.5M K_2SO_4 for an hour at 300 rpms, along with two 0.5M K_2SO_4 negative controls, and leached over pre-rinsed Whatman #1 filter paper. The K_2SO_4 extracts were diluted 1:10 and analyzed for MBC and MBN on the EcoCore (http://ecocore.nrel.colostate.edu) Shimadzu TOC-L (Shimadzu Scientific Instruments, Inc., Kyoto, Japan). Final measures of MBC and MBN were adjusted by soil moisture to assess microbial biomass per gram dry soil. Finally, these values were adjusted by conversion of chloroform-labile pool to microbial biomass

by dividing MBN by 0.54 and MBC by 0.45 (Joergensen 1996). The final reported units for microbial biomass were μg C (or N) per g dry soil.

Inorganic nitrogen (NO_3^- and NH_4^+) content of each soil microcosm was determined using a standard KCl salt extraction protocol (Keeney 1987). Fifteen grams of fresh soil was mixed with 75 mL 2M KCl for an hour at 300 rpms, along with two 2M KCl negative controls, and extracted over a pre-rinsed, N-free Whatman #1 filter paper. The KCl extracts were diluted 1:10 and analyzed for NO_3^- and NH_4^+ on the EcoCore (http://ecocore.nrel.colostate.edu) Alpkem Flow Solution IV Automated wet chemistry system (O.I. Analytical, College Station, TX). The final units for soil inorganic nitrogen were μg NO_3^--N (or NH_4^+-N) per g dry soil.

Finally, potential microbial extracellular enzyme activity was determined using seven common enzymes following a published high-throughput fluorometric assay protocol (Bell *et al* 2013). The enzymes include: alpha-glucosidase (AG), beta-glucosidase (BG), cellobiohydrolase (CB), leucine aminopeptidase (LAP), N-acetylglutamate synthase (NAG), phosphatase (PHOS), and xylanase (XYL). By adding non-limiting amounts of specific fluorescently labeled substrates (either with 7-animo-4-methylcoumarin (MUC) or 4-methylumbelliferone (MUB)), each as a target to one of the enzymes listed above, potential EEA activity was measured as the release of the fluorophore during enzymatic breakdown of the substrate. Standard curves were generated for each enzyme with a range of known substrate concentrations. Subsamples (2.75g) were taken from each soil microcosm at the week and month harvest and homogenized in a pH-adjusted 50 mM sodium acetate buffer. Soil slurries were then divided into deep well plates and subjected to the enzyme assays. Fluorescence was measured on the EcoCore (http://ecocore.nrel.colostate.edu) Tecan Infinite M200 Fluorescence Plate Reader at an excitation wavelength of 365 nm and emission wavelength of 450 nm (Tecan Group, Ltd., San

Jose, CA). The final calculated units for soil EEA activity were reported as μmols extracellular enzyme activity per gram dry soil.

Bacterial Community Structure

Bacterial community structure was determined using a modified standard protocol outlined in (Fierer & Jackson 2006). During the week and month incubation harvest, a 2 g subset from each microcosm was stored at (-)80 ^{o}C to preserve the intact microbial community until further processing. Total genomic DNA was isolated using the MoBio Power Soil kit (MoBio, Carlsbad, CA). The bacterial V4 hypervariable region of the conserved 16S rRNA gene was isolated from genDNA (515F and 806R with different 15-mer 454 barcodes, Caporaso *et al* 2012) using a Kapa2G Fast PCR kit (Kapa Biosystems, Wilmington, MA), and cleaned of PCR residuals using the MoBio UltraClean PCR Cleanup kit (MoBio, Carlsbad, CA). The V4 bacterial amplicons were quantified in each sample, using Quanti-iT PicoGreen Assay kit (Life Technologies, Grand Island, NY) and combined equimolarly among the samples into a single multiplexed sample.

The MSU Research Technology Support Facility sequenced the multiplex sample on an Illumina paired-end 2x250bp formatted run, using a 500 cycle Miseq V2 flow cell chemistry and an internal PhiX standard (Illumina, San Diego, CA). Nucleotide basepair calling was completed with Illumina Real Time Analysis (RTA) V1.18.54. The RTA output was then demultiplexed and converted to FASTQ format using Illumina Bcl2fastq V1.8.4.

FASTQ files were quality checked in the Qiime pipeline (V1.9.1, Caporaso *et al* 2010) by trimming the Illumina-specific barcodes, removing low quality reads (fastqual < 25, homopolymer acceptance of < 6 bp, no tolerance for 'N' base calls), and clustered to the standard 0.97 sequence similarity using USEARCH (Edgar 2010). After singleton and putative chimera

removal by USEARCH, representative sequences were identified in Qiime and used to create OTU biom-formatted tables. Ten rarefied biom table were generated from the master biom file to the sample with the lowest sequence depth (10,900 sequences). An alignment and phylogenetic tree were also inferred, against a curated Silva111 reference alignment using Practical Alignment using SATé and TrAnsitivity (PASTA, Mirarab *et al* 2014, SATé, Liu *et al* 2011). The phylogeny and rarefied biome tables were used to create abundance weighted and unweighted UniFrac distances among the soil bacterial communities, which were plotted into PCoA space (Fig. 4.12).

Statistical Analysis

A modified two-way ANOVA was implemented to test the effects of plant species, PHACE treatment origin, and the species * PHACE interaction on each of the soil functional assays. Unfortunately, during the leachate concentration process prior to the start of the soil incubation, one leachate sample was entirely lost (*Hesperostipa comata* of ac origin), so there were some *H. comata* leachate samples remaining, but not for those derived from ambient PHACE plots. Supplemental plant tissue was not available as some tissue was not consistently in excess for certain plots due to natural heterogeneity in plant community structure across the PHACE plots. As a result of this sample loss, the resultant ANOVA was unbalanced. Consequently, linear models with Type III SS (in Cran R, V3.2.0 as lm()) with a drop1 function were used instead of the standard ANOVA in R, aov(). The ANOVA tests were conducted on the calculated bulk measurements, and the graphical representation of the data was relative to the water-only control values.

Results

Overall, there were 74 unique extractable plant-derived metabolites in all the leachate samples (Table 4.2). Differences among the plant species' leachate chemistry were significant (Fig. 4.2, 4.3). All PHACE treatment origins of *L. dalmatica* leachate strongly separated from the other species and the *P. sylvestris* outgroup (Fig. 4.3). The NMDS ordination plot generated from abundance-weighted composite leachate chemistry (Fig. 4.3) shows the compound loading vectors that correspond strongly (pvalue < 0.0001) to separation of *L. dalmatica* and *P. sylvestris*. Specifically, 3-dehydroshikimic acid corresponds to the separation of *P. sylvestris* from the PHACE plant species' leachates. *A. frigida*, *C. diffusa*, and *H. comata* cluster together in ordination space, yet separate from *L. dalmatica* along the porphine compound vector. Finally, separation among the *L. dalmatica* PHACE treatment origins is substantially larger compared to collective separation from the other PHACE plant species. The overall separation of *L. dalmatica* is associated with the compound vectors of fumaric, galacturonic, and shikimic acids (p-value < 0.0001) (Fig. 4.3).

Overall, certain bulk indices of soil microbial function were significantly impacted by the species and/or the PHACE plot origin of the leachate treatment. MBC was significantly higher in *L. dalmatica* microcosms compared with other species' microcosms for the week incubation length, when adjusted to the water-only microcosm MBC (Fig. 4.4, Species: p-val < 0.0001). The elevated MBC dissipates for *L. dalmatica* after the month incubation length. MBN was also significantly impacted by the plant species of the leachate treatment in the week incubation length microcosms (Fig. 4.5, Species: p-val < 0.0001). MBC and MBN were not significantly impacted by the PHACE treatment origin of the leachate (Fig. 4.4, 4.5). There was also a marginally significant relationship between *L. dalmatica* and C:N of microbial biomass.

Nitrate levels were significantly impacted by PHACE treatment in both the week (p-val < 0.0001) and month (p-val = 0.05) incubation times (Fig. 4.6). For the week incubation length, microcosms treated with *A. frigida*, *C. diffusa*, and *H. comata* leachate had elevated NO_3^- levels relative to the water-only control, and *L. dalmatica* had significantly higher NO_3^- levels relative to the Ec and EH PHACE treatments. Consistently among the week and month incubation lengths and among plant species treatments, NO_3^- levels were higher in aH versus Ec and EH treated microcosms. Ammonium levels were not significantly driven by plant species or PHACE treatment origin in either the week or month harvest time points (Fig. 4.7).

Carbon- and nitrogen-degrading EEAs were also not significantly influenced by PHACE treatment origin or plant species leachate (Fig. 4.8, 4.9). However, plant species and the species*PHACE treatment origin interaction had marginal impacts on the activity of phosphorus-degrading EEAs for both week and month harvest time points (Fig. 4.10). Most of the P-EEA was elevated in the week harvest relative to water-only controls (Figure 4.10a), and the within plant species pattern held in the month harvest, but the values were lower. Overall there was lower P-EEA for aH treatment microcosms, and higher for the ambient treatments.

In all treatment combinations, the average cumulative heterotrophic respiration was higher than the respiration of the water-only control microcosms. Cumulative respiration was significantly influenced by plant species for both week and month incubation lengths (p-val < 0.0001, each), and the interaction term for the week harvest (p-val = 0.01) (Fig. 4.11). Microcosms treated with *L. dalmatica* leachate had significantly higher heterotrophic respiration relative to other plant species treatments in both the week and month harvests.

The bacterial community structure was primarily determined by incubation length (Fig. 4.12). As indicated by the PCoA ordination plot, 5.67% of variation for unweighted UniFrac

community scores corresponds primarily to incubation length (Fig. 4.12a). The second PCoA axis (2.4% variation explained) in the unweighted UniFrac ordination plot shows bacterial communities exposed to *L. dalmatica* leachate separating from the other treatments, which collectively cluster together and overlap by plant species leachate treatment (Fig. 4.12b). The PCoA ordinations for weighted UniFrac distances reveal a similar pattern (Axis 1 17.42% variation explained) corresponding to incubation length (Fig. 4.12c). There were no significant separation patterns in weighted UniFrac distances with respect to the plant species (Fig. 4.12d). Finally PHACE treatment origin had no significant impact on the beta diversity of either weighted or unweighted UniFrac distances (Suppl. Fig. 4.4).

Discussion

The chemical composition of leaf litter leachate varied by plant species and to a lesser extent among PHACE treatments. Differences in plant leachate chemistry also resulted in shifts in certain aspects of soil function, including microbial biomass C & N, nitrate levels, phosphorus-degrading enzyme activity, and heterotrophic respiration. There was also a correlation between the leachate of certain plant species and bacterial community structure.

Our findings support the concept that aboveground plant litter leachate can have strong impacts on the functional capacity of soil and the bacterial community structure. The invasive plant species' leachate, from *L. dalmatica*, had the most unique leachate chemistry (Fig. 4.3), which also resulted in the most exceptional soil microbial functional response.

The strong separation of the bacterial community structure, as measured by both weighted and unweighted UniFrac distances of beta-diversity, by incubation length was likely due to the extensive shift in microbial habitat conditions. Subjecting intact soil microbial communities to

the soil homogenization process by sieving to 2 mm, and subsequent treatment with a large water pulse are likely to restructure the soil microbial communities as well as their overall functional capacity. Because the water-only control microcosms separated by incubation length like those that received leachate indicates that it was the water pulse, and not the supplement of leachate-derived labile C, that drove the bacterial community structure.

In terms of bulk function, the invasive *L. dalmatica* has increased overall soil microbial functional, with increased MBC, MBN and cumulative heterotrophic respiration, though the relative increase in MBC/N is suppressed relative to the water-only control treatments by the month incubation. Curiously, the corresponding EEAs do not track these increase in microbial biomass. Potentially, there were enough existing EEAs already in the soil microcosms to compensate for the addition of the *L. dalmatica* leachate treatment. Regardless, the overall impact of this invasive is something to consider in the broader sense. The PHACE site has some indications that *L. dalmatica* may be positively influenced by the implemented PHACE treatments. Therefore, if *L. dalmatica* is more successful under future climate conditions, then the soil microbial functional impacts of altered plant community-level leachate might be substantial once the soil microbial functional data are scaled up to the plot level to account for differences in plant species abundance.

Finally, *L. dalmatica* soils separate from the other soils with unweighted UniFrac distances, but not with weighted UniFrac distances. This is likely due to the difference between community membership and community structure. The *L. dalmatica* microbial communities are separated from other communities because of distinct bacterial community membership, but this difference dissipates when abundance of bacterial taxa is incorporated into the UniFrac calculation, largely

driven by a few, abundant and ubiquitous taxa that are present in most soil microcosms and in large relative abundances.

The influence of the invasive *L. dalmatica* on soil microbial community structure and function, relative to the other plant species found at PHACE, highlights the unique influence of invasive plants on ecosystem function. *L. dalmatica* not only has the ability to alter the functional capacity of the surrounding soil, which could have nutrient cycling impacts, but this alteration may even benefit *L. dalmatica* invasion success. Specifically, I demonstrated in this laboratory incubation that soils exposed to *L. dalmatica* leachate tend to have elevated microbial function (respiration, MBC, and MBN) across all PHACE treatment origins. Though this increased activity dissipates a month after exposure to the *L. dalmatica* leachate, it is an ecologically important result. The initial pulse of *L. dalmatica* leachate imposes a strong pulse of microbial activity that then dissipates. From the broader scale of the plant community, if *L. dalmatica* are more abundant, the soil response is even greater, and could even result in enzymatic priming of existing organic carbon in the system. Additionally, previous studies have demonstrated that *L. dalmatica* possess iridoid glycosides (Jamieson & Bowers 2010), which may lend itself to their invasion success (Davidson *et al* 2011). In this study, I found that fumaric acid, shikimic acid and galacturonic acid clearly separated the *L. dalmatica* litter chemistry from the remaining plant species. These compounds may also be contributing to the unique elevated soil community functional and structural responses to *L. dalmatica* leachate (Fig. 4.3). Additionally, these plants had a waxy substance that passed through our leachate filtration system, which may have further influenced *L. dalmatica* unique impact on soil microbial structure and function.

In conclusion, this study has demonstrated that plant leachate chemistry of freshly senesced aboveground plant tissue varies among species, particularly between the invasive, *L. dalmatica*, and the other plant species. These chemical differences of *L. dalmatica* leachate resulted in distinct impacts on both soil microbial function and bacterial community structure. Additionally, incubation time strongly constrains the soil bacterial community structure, and even in the water only control microcosms. Despite being standard procedure for soil laboratory incubations, this clear shift in bacterial community structure does beg the question of whether soil ecologists should attempt to maintain intact soil structure for laboratory incubations. The practical downside is that the number of replicates would increase substantially as they are no longer simply experimental reps of a homogenized soil.

Broadly, this study highlights the importance of considering plant leachate as a key driver of soil microbial community structure and function. In the context of forecasted climate change, alternations to plant community structure may result in significant shifts in plant community structure, particularly those susceptible to invasion, with substantial impacts on soil microbial function. This chapter is important in that it examines a substantial suite of metabolites entering the soil by means other than the extensively studied inputs from rhizosphere and root exudation. Often the aboveground plant parts are studied in terms of tissue loss, and their rate of decomposition, but this study is unique in that is examines the impact of this leaf tissue degradation on soil microbial community structure and function.

Tables

Table 4.1. Description of plant species used in this experiment

Plant Species	Family	Native Status	Functional group	Common Name
Artemisia frigida (Willd)	Asteraceae	N	perennial forb	prairie sagewort
Centaurea diffusa (Lam)	Asteraceae	I	annual, perennial forb	white knapweed
Linaria dalmatica (L.)	Scrophulariaceae	I	perennial forb	dalmatian toadflax
Hesperostipa comata (Trin. & Rupr.)	Poaceae	N	annual, perennial graminoid	barkworth needle & thread
Pinus sylvestris (L.)	Pinaceae	N	perennial evergreen gymnosperm	scotch pine

Table 4.2. Unique extractable metabolites from the water-soluble fraction of the aboveground plant tissue samples used in the soil incubation, identified to CAS# and specific compound name.

CAS	Name	CAS	Name
8-3-36	ribonic acid-gamma-lactone 1	55-10-7	DL-4-hydroxy-3-methoxymandeli
10-8-63	citramalic acid	56-40-6	glycine
110-16-7	maleic acid	56-41-7	L-alanine 1
1114-34-7	D-lyxose 1	56-81-5	glycerol
1128-23-0	L-gulonic acid gamma-lactone	56-86-0	L-glutamic acid 3 (dehydrated)
114-63-6	4-hydroxybenzoic acid	57-13-6	urea
121-34-6	4-hydroxy-3-methoxybenzoic acid	57-48-7	fructose 1
127-17-3	pyruvic acid	579-36-2	D-allose 1
133-37-9	tartaric acid	585-88-6	palatinitol 1
137-08-6	pantothenic acid 2	59-23-4	D-glucose 1
13718-94-2	palatinose	5949-29-1	citric acid
138-59-0	shikimic acid	617-48-1	D-malic acid
141-82-2	malonic acid 1	63-42-3	lactose 2
144-62-7	oxalic acid	65-85-0	benzoic acid
149-91-7	gallic acid	66-22-8	uracil
14982-50-4	galacturonic acid 1	66009-10-7	melibiose 2
156-38-7	4-hydroxyphenylacetic acid	681295-24-	porphine 1
17013-01-3	fumaric acid	689-31-6	2-ketoadipic acid 2
24259-59-4	ribose	69-79-4	maltose 1
2595-98-4	talose 1	6968-16-7	D-threitol
2922-42-1	3-dehydroshikimic acid 1	7158-70-5	leucrose
29915-38-6	succinic acid	72-18-4	L-valine 2
342385-52-	2-keto-L-gulonic acid 1	7664-38-2	phosphoric acid
3715-29-5	3-methyl-2-oxobutanoic acid 1	77-95-2	quinic acid
39840-37-4	D-lyxosylamine 2	79-14-1	glycolic acid
443-79-8	DL-isoleucine 2	79-33-4	L-(+) lactic acid
4502-00-5	2-ketoisocaproic acid 2	80-69-3	tartronic acid
463-00-3	4-guanidinobutyric acid 2	828-01-3	3-phenyllactic acid
473-81-4	glyceric acid	87-78-5	D-mannitol
498-00-0	4-hydroxy-3-methoxybenzyl alcol	87-81-0	tagatose 1
50-70-4	D-sorbitol	87-89-8	allo-inositol
50-78-2	O-acetylsalicylic acid	87-99-0	xylitol
501-94-0	2-(4-hydroxyphenyl)ethanol	95-43-2	threose 2
503-49-1	3-hydroxy-3-methylglutaric acid	95-43-2	threose 1
526-95-4	gluconic acid 2	96-82-2	lactitol
526-99-8	mucic acid	98-79-3	L-pyroglutamic acid
534-46-3	sophorose 1	99-50-3	3,4-dihydroxybenzoic acid

Table 4.3. Soil functional activity by plant species and PHACE treatment (mean ± SE) in both the Week and Month harvest time points.

Week Harvest	**Microbial Biomass & Activity**				**Inorganic Nitrogen**		**Extracellular Enzyme Activity**		
	MBC	MBN	C:N	Respiration	NO3-	NH4+	Total C EEAs	Total N EEAs	Total P EEAs
Artemisia frigida									
Ambient	84 ± 1.6	11 ± 0.7	7 ± 1.6	33 ± 0.1	56 ± 1.9	0.13 ± 2	213 ± 13.7	187 ± 70.7	267 ± 5.5
Elev CO2, Unwarmed	78 ± 1.5	9 ± 1.2	8 ± 7.3	38 ± 0.1	38 ± 7.3	0.23 ± 2	199 ± 25.3	192 ± 63.9	226 ± 4.1
Ambient, Warmed	80 ± 1.6	11 ± 0.5	7 ± 1	31 ± 0.1	52 ± 4.7	0.14 ± 1	189 ± 7.6	171 ± 64.3	235 ± 5.6
Elev. CO2, Warmed	82 ± 1.2	11 ± 0.1	7 ± 1.5	35 ± 0.1	49 ± 7	0.13 ± 0.6	210 ± 11	190 ± 54.3	246 ± 3.7
Centaurea diffusa									
Ambient	84 ± 2.2	12 ± 0.3	7 ± 3.1	46 ± 0.1	48 ± 6.4	0.04 ± 3.1	215 ± 23.9	193 ± 86.7	267 ± 2.7
Elev CO2, Unwarmed	84 ± 0.8	11 ± 0.3	7 ± 1.5	43 ± 0	48 ± 3.6	0.02 ± 2.3	218 ± 12.5	162 ± 22.1	287 ± 0.8
Ambient, Warmed	86 ± 1.1	12 ± 0.4	7 ± 1.7	43 ± 0.1	52 ± 5.2	0.07 ± 2.4	202 ± 24.5	149 ± 17.2	260 ± 3
Elev. CO2, Warmed	84 ± 0.4	12 ± 1.3	7 ± 0.8	42 ± 0	44 ± 5.1	0 ± 3	217 ± 13.5	180 ± 35.7	267 ± 1
Linaria dalmatica									
Ambient	92 ± 1.5	12 ± 0.5	8 ± 0.5	72 ± 0.1	52 ± 6.1	0.05 ± 2.6	209 ± 24.8	236 ± 107.2	222 ± 2.8
Elev CO2, Unwarmed	94 ± 3.5	12 ± 0.7	8 ± 1.7	73 ± 0	43 ± 5.5	0.09 ± 3.3	186 ± 25.7	154 ± 30.1	272 ± 3
Ambient, Warmed	93 ± 1.1	14 ± 0.4	7 ± 1	69 ± 0.1	48 ± 13.4	0.06 ± 5.3	225 ± 43.6	245 ± 121.2	251 ± 1.5
Elev. CO2, Warmed	92 ± 2.7	13 ± 0.4	7 ± 3.1	60 ± 0.1	42 ± 14.9	0.04 ± 5.8	197 ± 55.4	162 ± 42.4	265 ± 2.9
Hesperostipa comata									
Ambient	--	--	--	--	--	--	--	--	--
Elev CO2, Unwarmed	83 ± 1.8	12 ± 0.4	7 ± 0.6	36 ± 0.1	54 ± 13.4	0.01 ± 4.1	216 ± 46.7	265 ± 139.9	241 ± 3.8
Ambient, Warmed	80 ± 2	12 ± 0.9	7 ± 2.3	38 ± 0.2	61 ± 17.9	0.04 ± 4	199 ± 52.4	198 ± 80.9	269 ± 2.1
Elev. CO2, Warmed	76 ± 2.4	11 ± 0.3	7 ± 1.5	33 ± 0.2	55 ± 11.4	0.38 ± 4.3	209 ± 50.4	154 ± 19.3	257 ± 7.2

Month Harvest	**Microbial Biomass & Activity**				**Inorganic Nitrogen**		**Extracellular Enzyme Activity**		
	MBC	MBN	C:N	Respiration	NO3-	NH4+	Total C EEAs	Total N EEAs	Total P EEAs
Artemisia frigida									
Ambient	61 ± 1.4	8 ± 0.5	8 ± 2.8	178 ± 0	116 ± 5	0.04 ± 3.2	245 ± 24.2	361 ± 37.6	277 ± 2.1
Elev CO2, Unwarmed	55 ± 1.3	7 ± 0.6	8 ± 5	159 ± 0	103 ± 17.5	0.11 ± 8.7	179 ± 75.8	308 ± 48.6	214 ± 5.9
Ambient, Warmed	52 ± 3.9	6 ± 0.4	9 ± 5.4	160 ± 0.1	114 ± 7.1	0.03 ± 7.8	209 ± 46.4	335 ± 26.4	231 ± 4.6
Elev. CO2, Warmed	54 ± 2.6	7 ± 0.9	8 ± 9.4	160 ± 0.1	104 ± 9.9	0.14 ± 2.2	251 ± 15.6	342 ± 27.4	282 ± 2.1
Centaurea diffusa									
Ambient	57 ± 2.9	7 ± 0.4	8 ± 4.3	182 ± 0	97 ± 2.2	0.09 ± 1.7	243 ± 10	372 ± 18.3	231 ± 1.8
Elev CO2, Unwarmed	56 ± 4.1	8 ± 1.1	7 ± 7	164 ± 0.1	100 ± 13.2	0.02 ± 6.2	201 ± 50.7	329 ± 45.9	233 ± 4.6
Ambient, Warmed	58 ± 2.7	7 ± 0.6	8 ± 2	178 ± 0.1	101 ± 18.7	0.03 ± 7.1	180 ± 69.3	310 ± 56.4	216 ± 4.8
Elev. CO2, Warmed	60 ± 2.6	8 ± 0.1	8 ± 4.3	178 ± 0.1	96 ± 19.2	0.05 ± 4.4	265 ± 36	383 ± 47	258 ± 2.3
Linaria dalmatica									
Ambient	52 ± 2.2	6 ± 0.9	9 ± 6.8	197 ± 0.1	121 ± 7.2	0.13 ± 1.6	214 ± 22.8	353 ± 40.2	257 ± 1.5
Elev CO2, Unwarmed	55 ± 6.3	8 ± 2.4	7 ± 6.6	217 ± 0	105 ± 12.4	0.06 ± 4.8	178 ± 45.8	334 ± 40.5	219 ± 3.5
Ambient, Warmed	52 ± 1.3	5 ± 0.4	10 ± 3.9	215 ± 0.1	118 ± 15.3	0.03 ± 6.9	178 ± 62.4	337 ± 59	217 ± 5.3
Elev. CO2, Warmed	54 ± 2.6	7 ± 0.7	8 ± 8	207 ± 0.1	91 ± 13	0.15 ± 4.2	183 ± 44.6	325 ± 23.1	204 ± 3.4
Hesperostipa comata									
Ambient	--	--	--	--	--	--	--	--	--
Elev CO2, Unwarmed	51 ± 3.5	6 ± 0.8	9 ± 7.5	165 ± 0.1	106 ± 8.7	0.6 ± 4.4	236 ± 35.1	364 ± 28.2	247 ± 3
Ambient, Warmed	56 ± 0.7	7 ± 0.2	9 ± 7.7	168 ± 0.1	116 ± 3.5	0.07 ± 3.9	206 ± 25.4	330 ± 31.3	230 ± 3.3
Elev. CO2, Warmed	55 ± 3.7	7 ± 0.6	8 ± 6	159 ± 0.1	108 ± 14.1	0.07 ± 8.5	216 ± 59.5	343 ± 34.2	234 ± 5.5

Figures

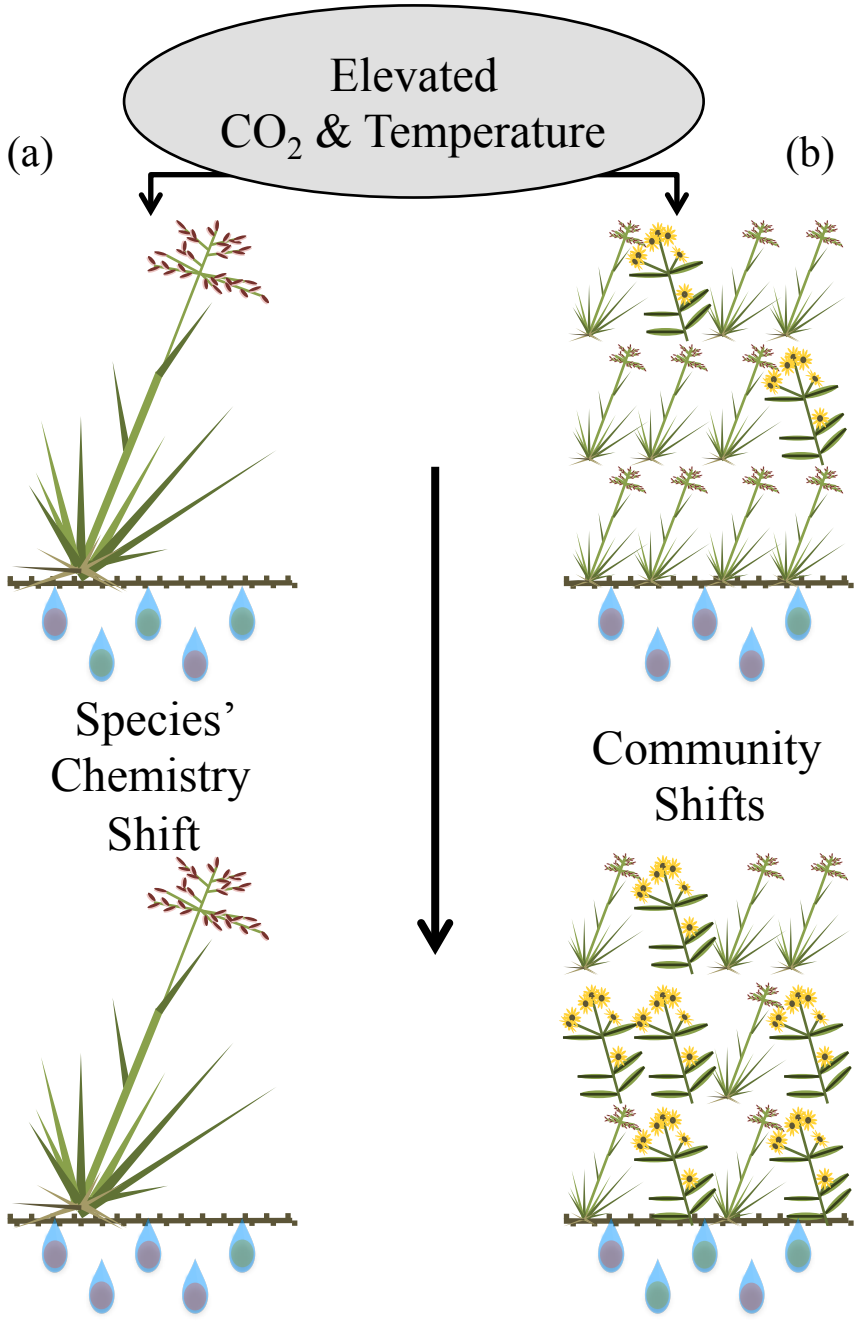


Figure 4.1. Hypothetical two-fold impact of forecasted climate on plant community chemistry. Elevated CO_2 and temperature may induce (a) intraspecific shift in plant tissue chemistry, and/or (b) changes in the plant community, potentially resulting in altered chemistry of the water-soluble leachate permeating the soil profile.

	ARFR	CEDI	LIDA	HECO	PISY
					
Species	*Artemisia frigida*	*Centaurea diffusa*	*Linaria dalmatica*	*Hesperostipa comata*	*Pinus sylvestris*
Family	Asteraceae	Asteraceae	Scroph.	Poaceae	Pinaceae
Fxn Group	Perennial Semi-shrub	Perennial Forb	Perennial Forb	C3 Perennial	Evergreen Gymnosperm
Native Status	N	I	I	N	N
Common Name	prairie sagewort	white knapweed	dalmatian toadflax	barkworth needle & thread	scotch pine

Figure 4.2. Visual differences among plant species' leachates. Leachate filtered through 20 µm mesh, and corresponding chopped plant tissue in the foreground.

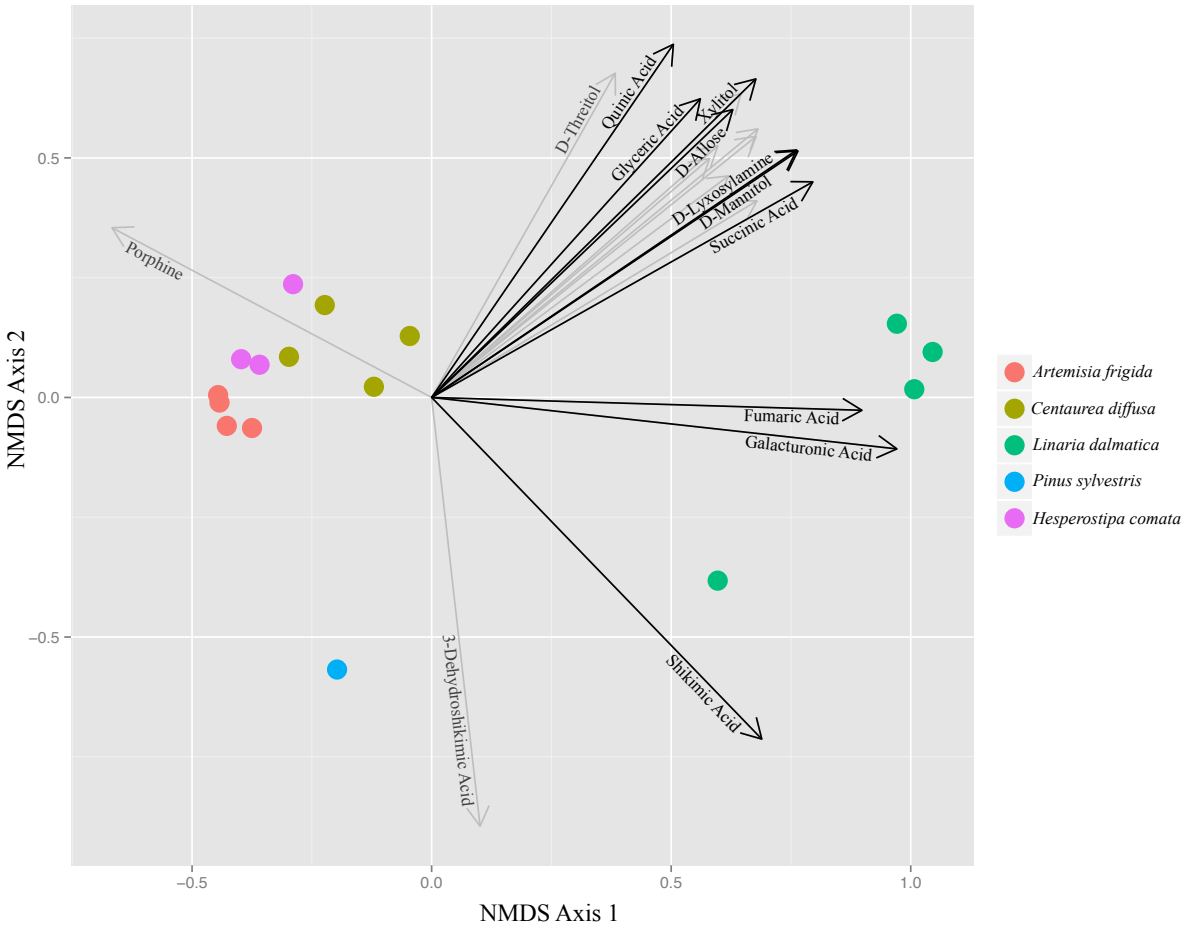


Figure 4.3. Nonmetric Multidimensional Scaling ordination plot of plant species' leachate chemistry. Points colored by plant species represent abundance-weighted composite leachate chemistry. Black (p-val < 0.0001) and gray (p-val < 0.001) arrows symbolize the compound-loading vectors for important chemical compounds in ordination space.

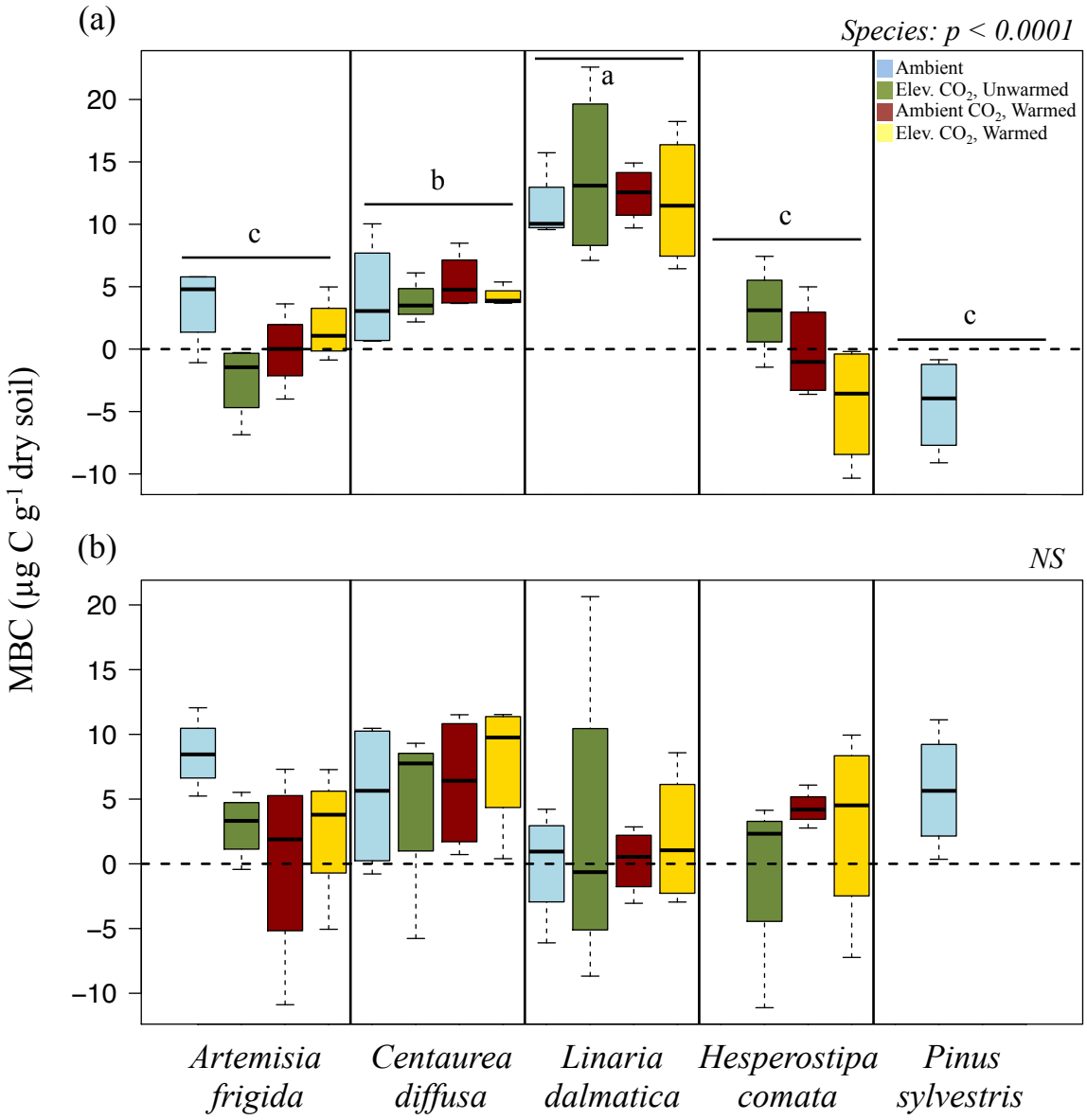


Figure 4.4. Microbial biomass carbon (MBC). Week (a) and month (b) harvest bulk measurements of MBC, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue. Letters above each boxplot represent Tukey HSD multiple comparisons for detecting significant differences in mean MBC (p-value: 0.05).

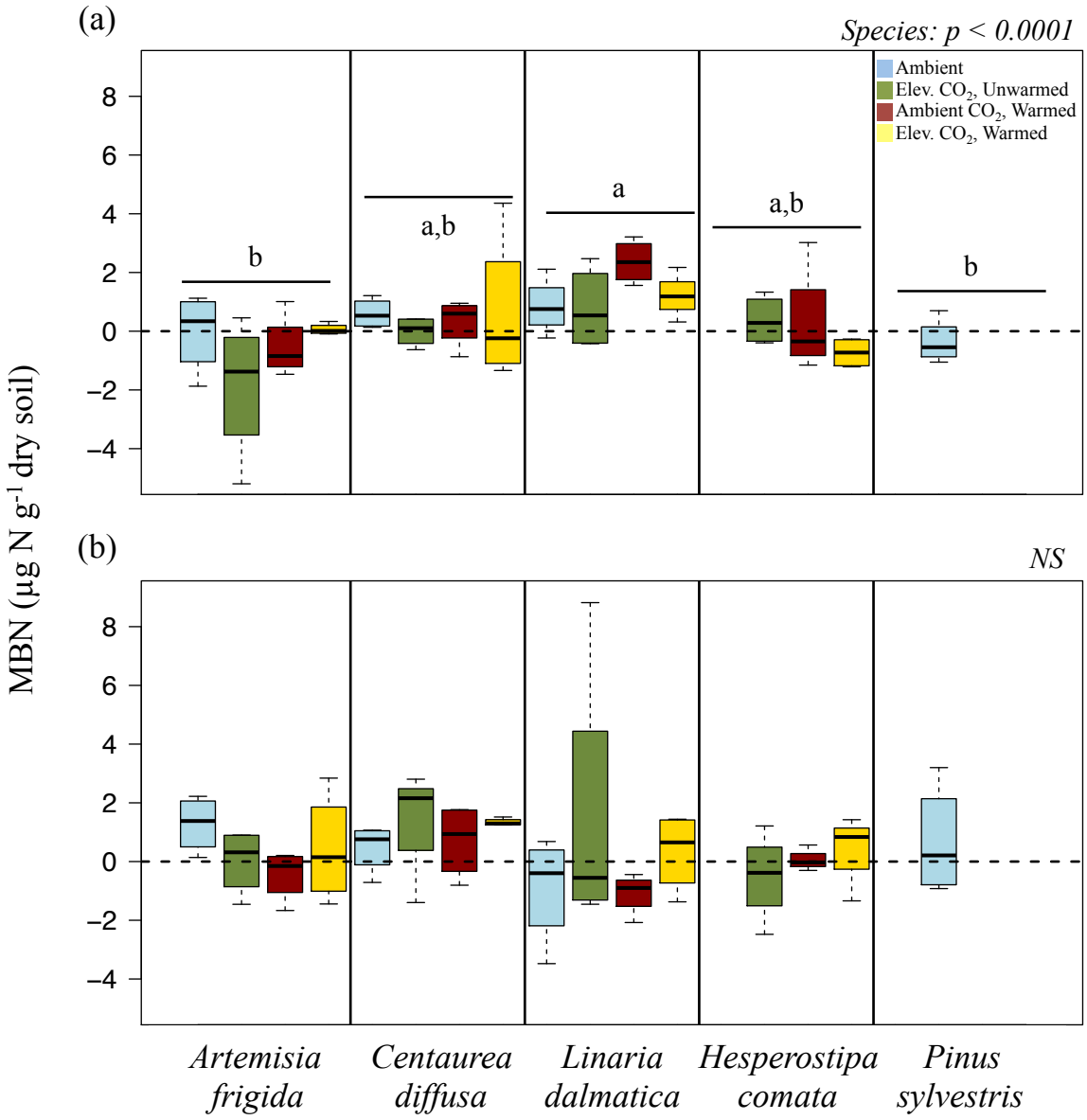


Figure 4.5. Microbial biomass nitrogen (MBN). Week (a) and month (b) harvest bulk measurements of MBN, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue. Letters above each boxplot represent Tukey HSD multiple comparisons for detecting significant differences in mean MBN (p-value: 0.05).

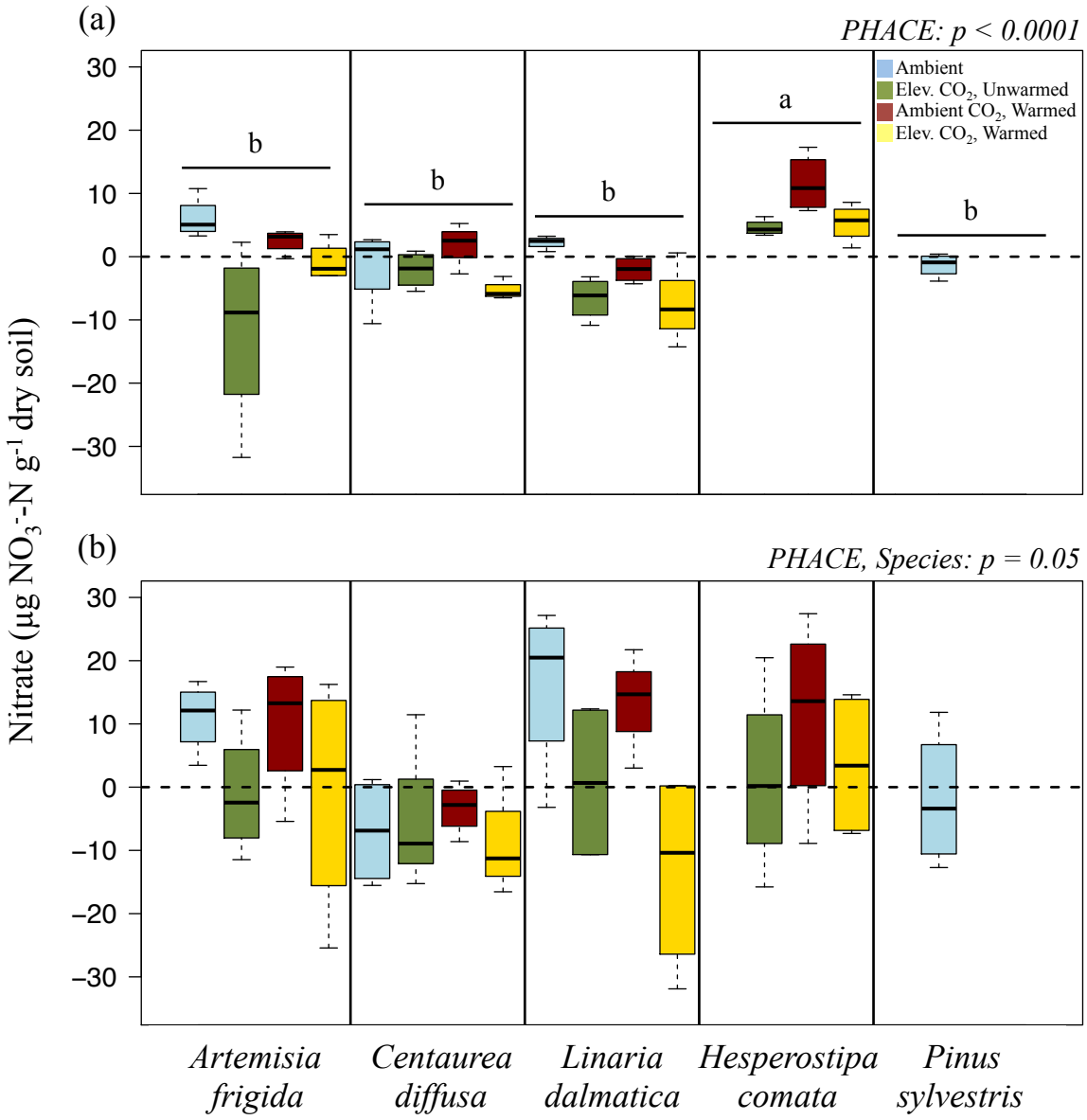


Figure 4.6. Inorganic nitrate (NO_3^--N). Week (a) and month (b) harvest bulk measurements of NO_3^-N, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue. Letters above each boxplot represent Tukey HSD multiple comparisons for detecting significant differences in mean nitrate values (p-value: 0.05).

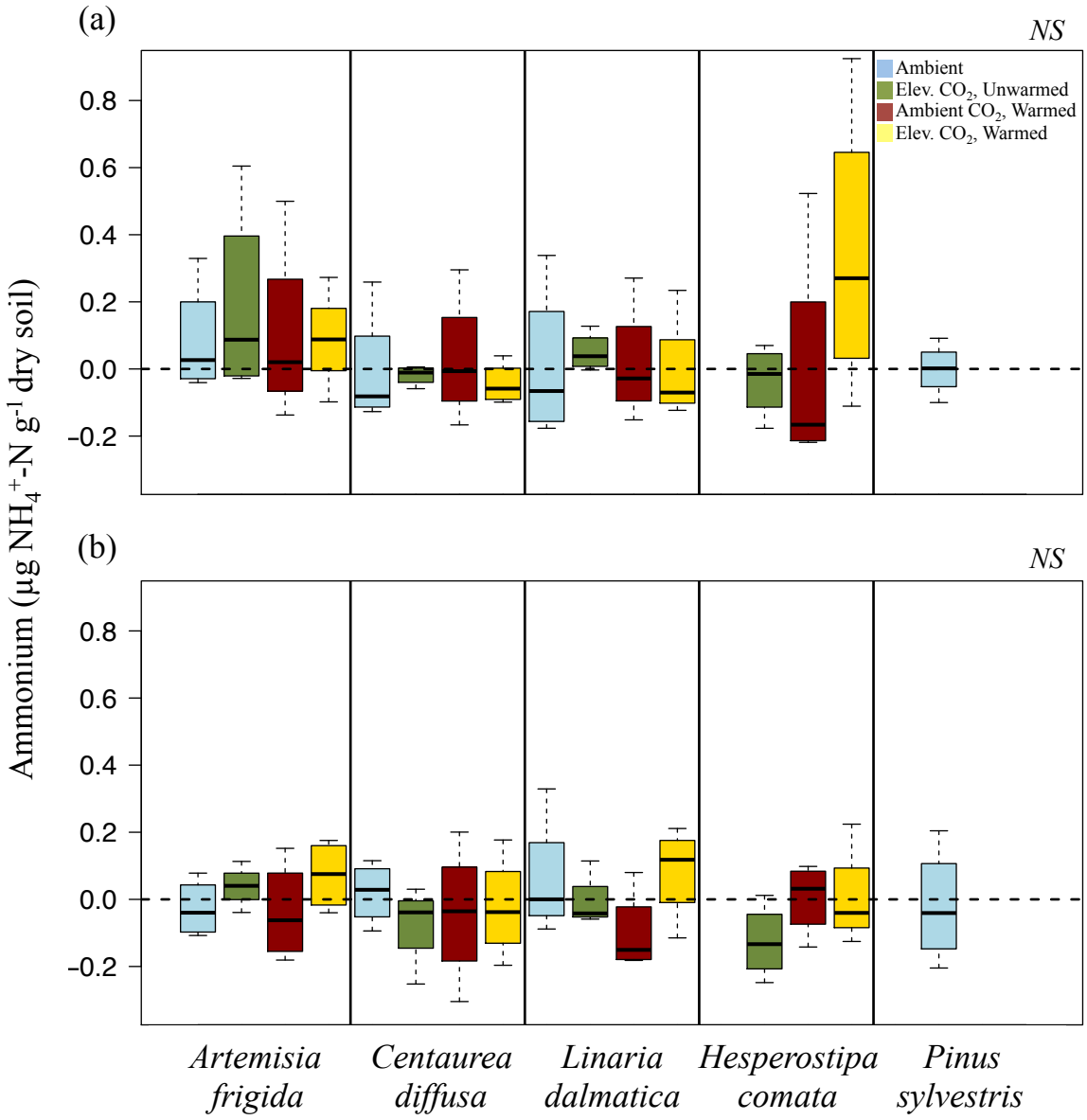


Figure 4.7. Inorganic ammonium (NH_4^+-N). Week (a) and month (b) harvest bulk measurements of NH_4^+-N, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue.

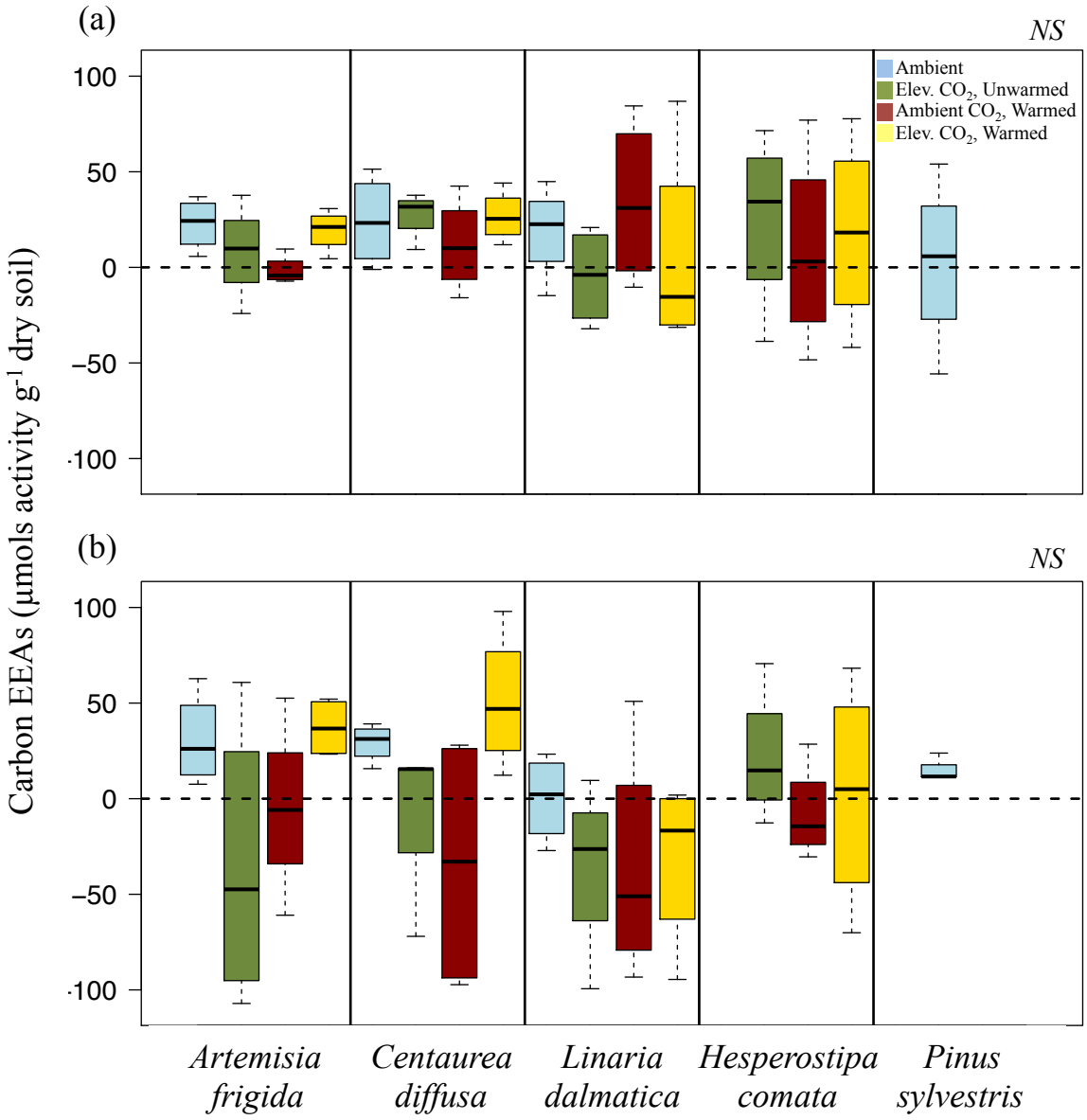


Figure 4.8. Total C-degrading extracellular enzyme activity (EEAs) (C EEAs). Week (a) and month (b) harvest bulk measurements of C EEAs, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue.

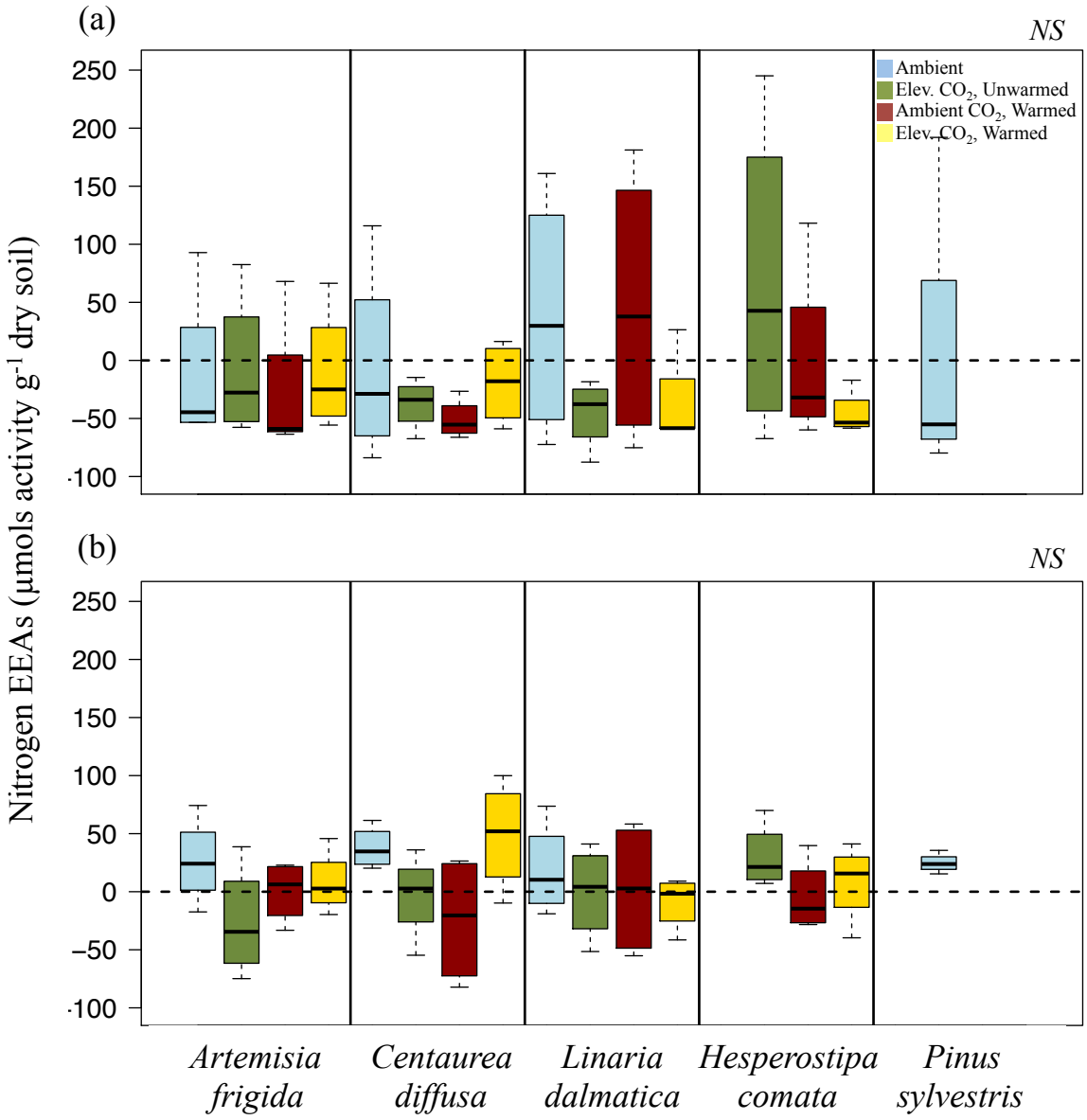


Figure 4.9. Total N-degrading extracellular enzyme activity (EEAs) (N EEAs). Week (a) and month (b) harvest bulk measurements of N EEAs, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue.

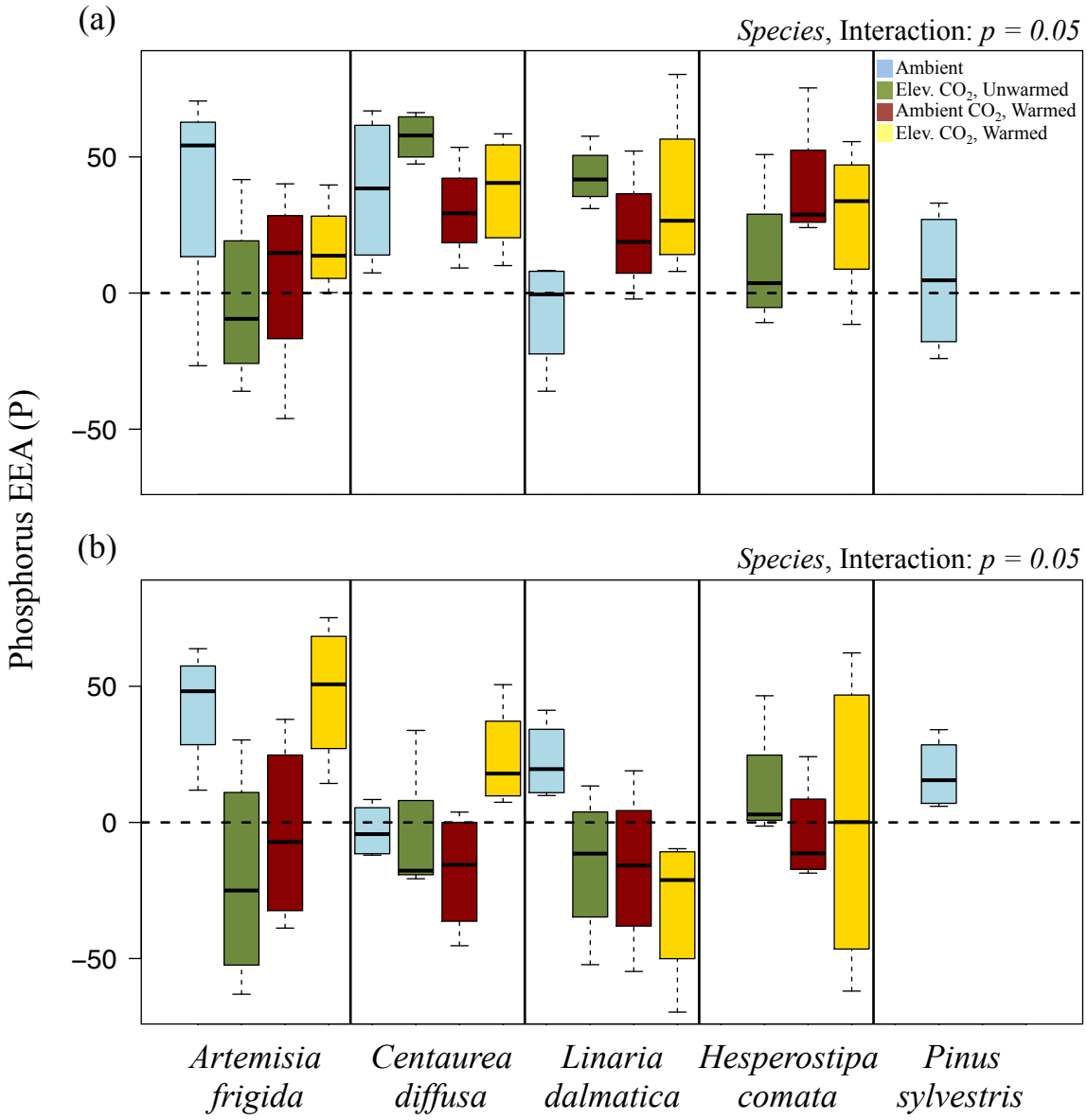


Figure 4.10. Total P-degrading extracellular enzyme activity (EEAs) (P EEAs). Week (a) and month (b) harvest bulk measurements of P EEAs, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue.

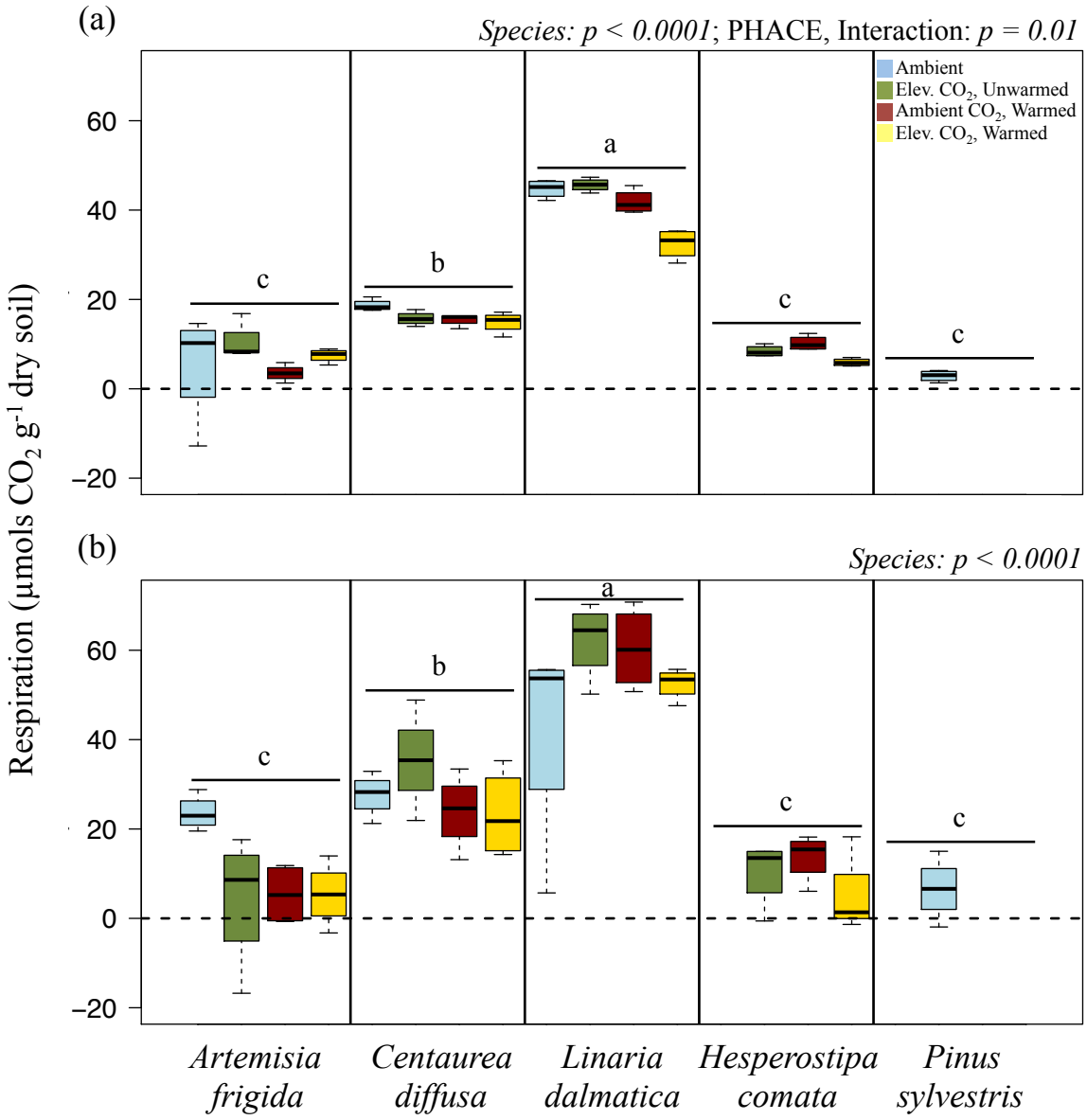


Figure 4.11. Microbial heterotrphic respiration (CO_2). Week (a) and month (b) harvest bulk measurements of CO_2, relative to the water-only control soil microcosms, sorted by plant species of leachate treatment, and colored by PHACE treatment. Pinus sylvestris is colored as an ambient outgroup plant tissue. Letters above each boxplot represent Tukey HSD multiple comparisons for detecting significant differences in mean respiration (p-value: 0.05).

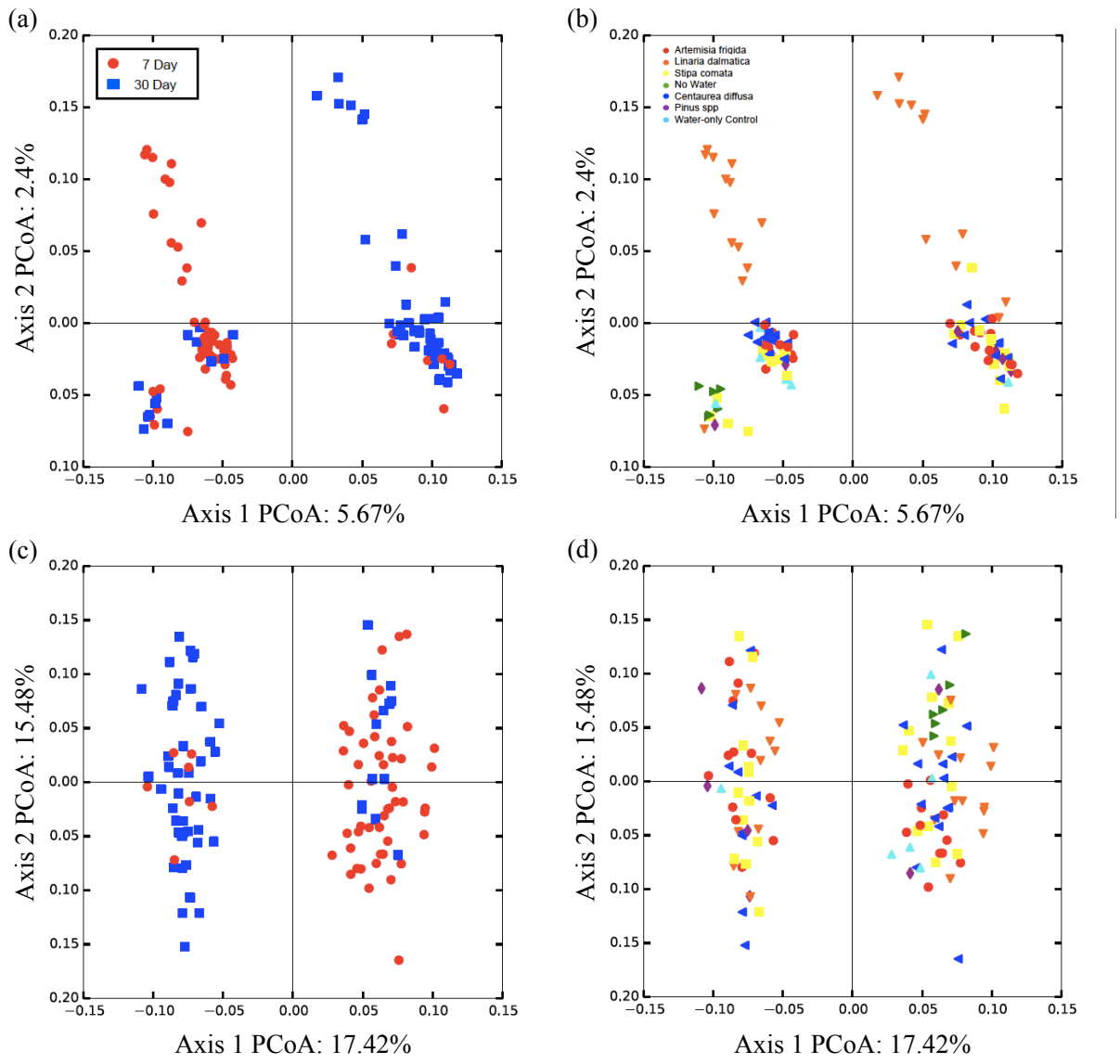


Figure 4.12. PCoA of differences in bacterial communities. Ordination plots of unweighted (a,b), and weighted (c, d) UniFrac distances, and colored by incubation harvest time (a, c), and plant species' leachate treatment (b, d).

References

Bais, H.P., T.L. Weir, L.G. Perry, S. Gilroy, J.M. Vivanco. (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. Annu Rev Plant Biol 57: 233-266.

Bakker MA, Carreno-Rocabado G, Poorter L. (2011) Leaf economics traits predict litter decomposition of tropical plants and differ among land use types. Fun. Ecol. 25: 473-483.

Bird, J.A., D.J. Herman, M.K. Firestone. (2011) Rhizosphere priming of soil organic matter by bacterial groups in a grassland soil. Soil Biol & Biochem. 43(4): 718-725.

Bowman, W.D., H. Steltzer, T.N. Rosenstiel, C.C. Cleveland, C.L. Meier. (2004) Litter effects of two cooccuring alpine species on plant growth, microbial activity and immobilization of nitrogen. Oikos 104(2): 336-344.

Caporaso, J.G., J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Pena, J.K. Goodrich, J.I. Gordon, G.A. Huttley, S.T. Kelley, D. Knights, J.E. Koenig, R.E. Ley, C.A. Lozupone, D. McDonald, B.D. Muegge, M. Pirrung, J. Reeder, J.R. Sevinsky, P.J. Turnbaugh, W.A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight. (2010) QIIME allows analysis of high-throughput community sequence data. Nature Methods 7(5): 335-336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6: 1621-1624.

Coley, P.D. (1987) Interspecific variation in plant anti-herbivore properties: the role of habitat quality and rate of disturbance. New Phyt. 106: 251-263.

Coq, S., J. Souquet, E. Meudec, V. Cheynier, S. Hattenschwiler. (2011) Interspecific variation in leaf litter tannins drives decomposition in a tropical rain forest of French Guiana. Ecology. 97(7): 2080-2091.

Davidson, A.M., M. Jennions, A.B. Nicotra. (2011) Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. Ecol Lett. 14(4): 419-431.

Debs-Louka, E., N. Louka, G. Abraham, V. Chabot, K. Allaf. (1999) Effect of compressed carbon dioxide on microbial cell viability. Appl. Environ. Microbiol. 65(2): 626-631.

Edgar RC. (2010) Search and clustering order of magnitude faster than BLAST. Bioinformatics 26(19): 2460-61.

Fanin N, Fromin N, Buatois B,Hattenschwiler S. (2013) An experimental test of the hypothesis of non-homeostatic consumer stoichiometry in a plant litter–microbe system. Ecol. Lett. doi: 10.1111/ele.12108

Fanin, N., S. Hattenschwiler, N. Fromin. (2014) Litter fingerprint on microbial biomass, activity, and community structure in the underlying soil. Plant Soil. 379: 79-91.

Fierer, N. and R.B. Jackson. (2006) The diversity and biogeography of soil bacterial communities. PNAS. 103(3): 626-631.

Freschet, G.T., W.K. Cornwell, D.A. Wardle, T.G. Elumeeva, W. Liu, B.G. Jackson, V.G. Onipchenko, N.A. Soudzilovskaia, J. Tao, J.H.C. Cornelissen. (2013) Linking litter

decomposition of above- and below-ground organs to plant-soil feedbacks worldwide. J Ecol 101(4): 943-952.
Goldfarb, K.C., U. Karaoz, C.A. Hanson, C. A. Santee, M.A. Bradford, K.K. Treseder, M.D. Wallenstein,E.L. Brodie. (2011) Differential growth responses of soil bacterial taxa to carbon substrates of varying chemical recalcitrance. Frontiers Microbiology. 2(94).
Halvorson, J.J., J.M. Gonzalez, A.E. Hagerman. (2011) Repeated applications of tannins and related phenolic compounds are retained by soil and affect cation exchange capacity. Soil Bio & Biochem 43: 1139.
Hobbie S. (1992) Effects of plant species on nutrient cycling. Trends in Ecol & Evol. 7:336–339.
Jamieson, M.A. and M.D. Bowers. (2010) Iridoid glycoside variation in the invasive plant Dalmation toadflax, Linaria dalmatica (Plantaginaceae), and sequesteration by the biological control agent, Calophasia lunula. Chem Ecol. 36: 70-79.
Joergensen, R.G. (1996) The fumigation-extraction method to estimate soil microbial biomass: calibration of the k(EC) factor. Soil Bio. Biochem. 28: 25-31.
Keeney, D. (1987) Nitrogen-availability indices. In: A. Page, R. Miller & D. Keeney (eds.) *Methods of soil analysis, part 2, chemical and microbiological properties*. Am. Soc. Agro. Madison, WI, pp: 711-734.
Kotroczo, Z., Z. Veres, I. Fekete, Z. Krakomperger, J.A. Toth, K. Lajtha, B. Tothmeresz. (2014) Soil enzyme activity in response to long-term organic matter manipulation. Soil Biol Biochem. 70: 237-243.
Liu, K., T.J. Warnow, M.T. Holder, S. Nelesen, J. Yu, A. Stamatakis, and C.R. Linder. (2012) SATe-II:Very Fast and Accurate Simultaneous Estimation of Multiple Sequence Alignments and Phylogenetic Trees. Systematic Biology 61(1): 90-106.
Mirarab, S., N. Nguyen, and T. Warnow. (2014) PASTA: ultra-large multiple sequence alignment. J Comp Biol.
Moore, B.D., R.L. Andrew, C. Kulheim, W.J. Foley. (2013) Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phyt Tansley Review. 201: 733-750.
Nicotra, A.B., O.K. Atkin, S.P. Bonser, A.M. Davidson, E.J. Finnegan, U. Mathesius, P.Poot, M.D. Purugganana, C.L. Richards, F. Valladares, M. van Kleunen. (2010) Plant phenotypic plasticity in a changing climate. Trends in Plant Sci. 15(2): 684.
Phillips, R.P., A.C. Finzi, E.S. Bernhardt. (2011) Enhanced root exudation induces microbial feedbacks to N cycling in a pine forest under long-term CO2 fumigation. Ecol Lett 14(2): 187-194.
Rinkes, Z.L., J.L. DeForest, A.S. Grandy, D.L. Moorhead, M.N. Weintraub. (2014) Interactions between leaf litter quality, particle size, and microbial community during the earliest stage of decay. Biogeochem 117(1): 153-168.
Schmidt, M.A., A.J. Kreinberg, J.M. Gonzalez, J.J. Halvorson, E. French, A. Bollmann, A.E. Hagerman. (2013) Soil microbial communities respond differently to three chemically defined polyphenols. Plant Phys & Biochem 72: 190-197.
Shi, S., A.E. Richardson, M. O'Callaghan, K.M. DeAngelis, E.E. Jones, A. Stewart, M.K. Firestone, L.M. Condron. (2011) Effects of selected root exudate components on soil bacterial communities. FEMS Microbiol Ecol 77(3): 600-610.
Soong, J.L., W. Parton, F.J. Calderon, E.E. Cambell, M.F. Cotrufo. Chemical traits of fresh or pyrolyzed aboveground plant inputs to soil determine the amount and quality of dissolved organic matter leaching during decomposition (In review).

Suseela V., Triebwasser-Freese D., Linscheid N., Morgan J.A., Tharayil N. (2014) Litters of photosynthetically divergent grasses exhibit differential metabolic responses to warming and elevated CO2. Ecosphere. 5(9): 106.

Suseela, V., N. Tharayil, B. Xing, J.S. Dukes. (2015) Warming and drought differentially influence the resorption of elemental and metabolic nitrogen pools in Quercus rubra. Global Change Biology. (In Press)

Tharayil, N. V. Suseela, D.J. Triebwasser, C.M. Preston, P.D. Gerard, J.S. Dukes. (2011) Changes in the structural composition and reactivity of Acer rubrum leaf litter tannins exposed to warming and altered precipitation: climatic-stress-induced tannins are more reactive. New Phyt. 191: 132-145.

Triebwasser , D.J., N. Tharayil, C.M. Preston, P.D. Gerard. (2012) The susceptibility of soil enzymes to inhibition by leaf litter tannins is dependent on the tannin chemistry, enzyme class and vegetation history. New Phyt. 196: 1122-1132.

Ushio, M., T.C. Balser, K. Kitayama. (2013) Effects of condensed tannins in conifer leaves on the composition and activity of the soil microbial community in a tropical montane forest. Plant Soil. 365: 157-170.

Waldrop, M.P. and M.K. Firestone. (2004) Microbial community utilization of recalcitrant and simple carbon compounds: impact of oak-woodland plant communities. Oecologia. 138(2): 275-284.

Chapter 5: Interspecific variability in senesced plant leachate chemistry and functional impacts on soils

To build on the results of Chapter 4, I conducted another laboratory soil incubation, to examine the resiliency of interspecific differences in plant chemistry long after leaf senescence. Soils were collected directly adjacent to the PHACE field site and exposed to leachate from an overlapping set of plant species found adjacent to PHACE. However, in this case, the plant tissue was senesced for several months prior to tissue harvest. The laboratory incubation was set up similarly to the incubation in Chapter 4, except that the PHACE treatment factor was not included in this study. The overall conclusion was that there were no differences in plant litter or leachate chemistry among plant species, and this translated to no interspecific impacts on soil microbial function. However, despite the overall negative result, the study does stress the importance of temporal release of plant leachate long into the dormant season. Despite being senesced for several months, and exposed to three significant rain events, the leaf tissue of all plant species still imposed significant increases in soil microbial activity with respect to water-only control soil microcosms.

Introduction

Plants are key drivers of soil microbial function (Hobbie 1992) through chemical contribution of metabolites via root exudation (Bais *et al* 2006, Bird *et al* 2011, Shi *et al* 2011, Phillips *et al* 2011) and senescence of plant tissue parts (Freschet *et al* 2013, Kotroczó *et al* 2014, Fanin *et al* 2014). Little is known about the influence of the water-soluble component of plant tissue on soil microbial function (Fanin *et al* 2014). It is estimated that upwards of 16% of

plant leaf tissue carbon permeates the soil as leachate (Soong *et al* 2014). Water-soluble plant metabolites are known to contribute substantially to soil microbial structure and function (Bowman *et al* 2004, Ushio *et al* 2013, Fanin *et al* 2013), since most of the metabolites contained in leachate are water-soluble metabolites that are easily assimilated by microbial biomass without necessarily requiring extracellular enzymatic breakdown of the compounds. Additionally, it is well known that plant species can vary substantially in the chemical makeup of their leaf tissue (Coley 1987, Wurst *et al* 2008), which in turn could result in chemical differences in their leachate (Rinkes *et al* 2014). Plant leachate can serve as C-rich energy source for heterotrophic microorganisms (Waldrop & Firestone 2004), but some plant metabolites also inhibit microbial function (Goldfarb *et al* 2011, Triebwasser *et al* 2012, Schmidt *et al* 2013). Distinctive leachate chemistries could therefore differentially impact soil microbial structure and function. In the broader context, this could mean that shifts in plant community structure could alter the functional of the soil as a result of the altered chemistry permeating the soil profile. Therefore, it is essential to understand how plant leachates vary among species, and how that influences soil microbial function.

Intraspecific variability in plant tissue chemistry is well known (Coley 1987, Jamieson & Bowers 2010, Moore *et al* 2013). Species-specific niche conditions, including life history strategies, photosynthetic pathway, and functional role in a habitat can greatly influence the way in which a plant responds to environmental conditions, and the chemistry often reflects this response. Stoichiometric constraints also limit what suite of metabolites can be produced by the plant, which translates to possible differences in the litter leachate portion of plant chemistry.

In the context of forecasted changes in climate, the plant community may respond to the habitat shifts. If there are interspecific differences in leachate chemistry that translate to changes

in soil microbial function, then there could be a direct connection between plant community shifts and changes in soil microbial function. The impacts of elevated CO_2, temperature, and altered precipitation regime may have substantial impacts on soil function, but climate-mediated shifts in plant community structure may be far more substantial to soil function. The goal of this study was to assess interspecific variability in plant leaf tissue and leachate, and the consequences for soil microbial function. If species-specific plant litter leachate differentially impacts soil microorganisms, then there is a direct link between plant community and soil microbial function.

The primary objective of this study was to examine the importance of aboveground plant litter leachate in driving the response of soil microbial community function. Our study examines to what extent among species variability in aboveground plant litter leachate chemistry translates to shifts in soil function.

Materials and Methods

Study site and sample collection

The soil and plant tissue used in this experiment are from undisturbed grassland (Cheyenne, WY; 41.198700 N, -104.887048 W). The site is characterized as a semi-arid northern mixed grassland with a mean annual precipitation of 384 mm, with an average soil pH of 7.0 that is characterized as mollisol (Morgan *et al* 2011).

To assess the variability of leaf and leachate chemistry among species at PHACE, senesced aboveground plant tissue was harvested from the seven most abundant plant species in December 2011 (Table 5.1). These plant species included: *Artemisia frigida* (Willd.), *Astragalus drummondii* (Douglas ex Hook), *Bouteloua dactyloides* (Nutt), *Chrysopsis villosa* (Pursh),

Koeleria pyramidata (Ledeb.), *Linaria dalmatica* (L.), and *Hesperostipa comata* (Trin. & Rupr.). These plant species encompass a broad range of phylogenetic breadth and both native and invasive statuses. Plants litter samples were collected as standing litter, i.e. no leaf tissue that was already in contact with the soil. The samples were stored in paper bags, and air-dried before storing long-term in a desiccator until further use.

The soil was harvested at the end of the growing season in July 2013. One site level soil sample was collected from non-compacted, undisturbed areas to ensure homogeneity of the soil and minimize any bias towards a given plant species. The soils were sieved to 2 mm on site, and stored in a cooler for transport and in a temperature controlled cold room at 4 ^{o}C until the laboratory incubation setup. Subsamples were taken to determine gravimetric water content (%GWC), pH, and saved for starting conditions for the bulk functional assays.

In November 2012, the plant litter samples were prepared for extraction of plant-derived leachate. The aboveground plant litter was cut into 2 cm pieces to ensure similar size pieces among plant species. A 0.5 g subsample of each litter sample was also set aside for C:N analysis. Multiple room temperature leaching events were executed on 1.25 g leaf tissue samples with 20 mL of deionized H_2O to extract the water-soluble, plant-derived metabolites. Each sample leachate was collected in a sterile modified filtration apparatus. Trial tests revealed that the original 0.45 μm pore size was ideal for removing microorganisms. However, this small pore size severely clogged the filters after just one leaching, so modified filtration systems were developed by replacing the intact filter with sterile, acid washed/rinsed 20 μm nylon mesh and a layer of inert SiO_2 to prevent clogging by plant litter and leachate materials. To minimize microbial processing of the plant-derived metabolites, which develop soon after plant litter is wet (Bakker *et al* 2011), ten continuous water extraction events were implemented over a 12-hour

period. Each leaching event consisted of adding 20 mL of sterile pH-balanced deionized H_2O to the litter atop the filtration apparatus (Fig. 5.2). After two minutes, the vacuum was attached to pull the room temperature extractable metabolites from the intact plant litter. The leachate samples were concentrated 10:1 and stored frozen until use. Leachate subsamples (250 µL) were also lyophilized in 0.5 g KBr for FT-IR analysis at the USDA in Akron, CO.

Leaf and leachate chemistry

The lyophilized leachate extracts and ground subsamples (0.5 g) of the original leaf litter tissue were analyzed with Fourier transformed mid- and near-infrared spectroscopy (MidIR and NearIR, respectively) on a Digilab FTS 7000 spectrometer (Varian, Inc., Palo Alto, CA) using a Pike AutoDIFF plate-like autosampler (Pike Technologies, Madison, WI) to rotate two analytical replicates per sample through the spectrometer. Each analytical replicate was a spectral average of 64 scans through separate absorbance ranges: 4000 to 400 cm^{-1} for MidIR; and for 10000 to 4000 cm^{-1} for NearIR.

Pyrolysis molecular beam mass spectrometry (py-MBMS) was also used to examine detailed chemistry for the senesced plant leaf tissue. Triplicate analytical replicates (0.02 g) of ground plant leaf tissue subsamples were also analyzed on a modified py-MBMS at the National Renewable Energy Laboratory (NREL, Boulder, CO). A categorical list for dividing the output dataset into general compound classes: alkyl aromatics, carbohydrates, lignin dimers, lipids (incl. alkanes, alkenes, fatty acids, n alkyl esters), nitrogenous compounds, peptides (incl. proteins, amino acids, enzymes, and nucleic acids), phenol and lignin monomers, sterols, and an aggregated group of unidentified compounds (Wallenstein *et al* 2013).

Soil Laboratory Incubation

To examine the influence of interspecific senesced plant litter leachate on soil microbial function, a laboratory incubation was implemented using 72 soil microcosms.

The experimental design consisted of randomized two factor experimental design with four experimental reps per factor combination. One factor, plant species had seven treatments comprised of the most common plant species at the field site [*Artemisia frigida* (Willd.), *Astragalus drummondii* (Douglas ex Hook), *Bouteloua dactyloides* (Nutt), *Chrysopsis villosa* (Pursh), *Koeleria pyramidata* (Ledeb.), *Linaria dalmatica* (L.), and *Hesperostipa comata* (Trin. & Rupr.)], which allowed for examination of the interspecific differences in leachate impacts on soil microbial function. The second factor was incubation time factor, which permitted examination of the soil microbial functional response to a leachate event over time. Finally, four water-only and four no-water soil microcosms were included as control soil microcosms.

Each soil microcosm consisted of ~60 g field fresh, homogenized soil in a sterilized, acid-washed Mason jar, equipped with two Wheaton natural rubber flange septa (Wheaton Scientific, Millville, NJ). The septa in each jar lid was sealed with 100% silicone and cured for five days. The efficacy of each jar's seal was tested by adding known amounts of CO_2 and testing for retention five days later. The soil in each jar was independently adjusted to 40% water holding capacity with the appropriate leachate and immediately sealed. Background CO_2 levels were recorded, and the incubation was placed in a dark, temperature controlled room at 25 ^{o}C.

Cumulative heterotrophic respiration (soil CO_2 production per g dry soil) was measured on each microcosm throughout the week and month incubation times. In the first week of the incubation experiment, daily measurements of CO_2 were collected using the LI-COR LI 6552 infrared gas analyzer (LI-COR Biosciences, Lincoln, NB). In subsequent weeks, CO_2

measurements were collected every other day. After each reading, the microcosms were flushed with CO_2-free air to prevent inhibitory effects of high CO_2 on microbial function (Debs-Louka *et al* 1999). The μmols of CO_2 produced by each jar was estimated using a standard curve (R^2>0.99) of a known CO_2 concentration, generated for each LI-COR run. Finally, the respiration values were adjusted to end units of μmols CO_2 per g dry soil per hr.

Bulk Functional Assays

In addition to the heterotrophic respiration readings collected during the incubation, destructive soil assays were also conducted at incubation harvest to assess the impact of plant species' leachates on soil function. These assays include: %GWC, microbial biomass C/N, inorganic nitrogen (NO_3^-, NH_4^+), and seven extracellular enzyme assays.

Gravimetric water content was estimated by weighing out 5 g of fresh soil, drying it overnight in a 60 ^oC oven and reweighing the cooled dry soil. GWC was then calculated as the difference between the soil weights over the dry soil weight, multiplied times 100%. As the starting moisture conditions were experimentally controlled during incubation setup, GWC was primarily assessed as verification that moisture conditions were not significantly different among the treatments.

Microbial biomass carbon and nitrogen were determined using the standard chloroform fumigation protocol (Joergensen 1996). The general rationale of chloroform fumigation is to kill and lyse the soil microbial cells when exposed to chloroform, releasing their organic matter for extraction in a 0.5M K_2SO_4 salt solution. Two separate 15 g soil samples were subset from each microcosm at the week and month incubation harvests. One sample is saved at 4 ^oC while the other is fumigated for five days with chloroform under vacuum conditions. After the fumigation

was complete, the fresh soil from each sample was mixed with 75 mL 0.5M K_2SO_4 for an hour at 300 rpms, along with two 0.5M K_2SO_4 negative controls, and leached over pre-rinsed Whatman #1 filter paper. The K_2SO_4 extracts were diluted 1:10 and analyzed for MBC and MBN on the EcoCore (http://ecocore.nrel.colostate.edu) Shimadzu TOC-L (Shimadzu Scientific Instruments, Inc., Kyoto, Japan). Final measures of MBC and MBN were adjusted by soil moisture to assess microbial biomass per gram dry soil. Finally, these values were adjusted by conversion of chloroform-labile pool to microbial biomass by dividing MBN by 0.54 and MBC by 0.45 (Joergensen 1996). The final reported units for microbial biomass were μg C (or N) per g dry soil.

Inorganic nitrogen (NO_3^- and NH_4^+) content of each soil microcosm was determined using a standard KCl salt extraction protocol (Keeney 1987). Fifteen grams of fresh soil was mixed with 75 mL 2M KCl for an hour at 300 rpms, along with two 2M KCl negative controls, and extracted over a pre-rinsed, N-free Whatman #1 filter paper. The KCl extracts were diluted 1:10 and analyzed for NO_3^- and NH_4^+ on the EcoCore (http://ecocore.nrel.colostate.edu) Alpkem Flow Solution IV Automated wet chemistry system (O.I. Analytical, College Station, TX). The final units for soil inorganic nitrogen were μg NO_3^--N (or NH_4^+-N) per g dry soil.

Finally, potential microbial extracellular enzyme activity was determined for seven common enzymes following a published high-throughput fluorometric assay protocol (Bell *et al* 2013). The enzymes include: alpha-glucosidase (AG), beta-glucosidase (BG), cellobiohydrolase (CB), leucine aminopeptidase (LAP), N-acetylglutamate synthase (NAG), phosphatase (PHOS), and xylanase (XYL). By adding non-limiting amounts of specific fluorescently labeled substrates (either with 7-animo-4-methylcoumarin (MUC) or 4-methylumbelliferone (MUB)), each as a target to one of the enzymes listed above, potential EEA activity was measured as the release of

the fluorophore during enzymatic breakdown of the substrate. Standard curves were generated for each enzyme with a range of known substrate concentrations. Subsamples (2.75g) were taken from each soil microcosm at the week and month harvest and homogenized in a pH-adjusted 50 mM sodium acetate buffer. Soil slurries were then divided into deep well plates and subjected to the enzyme assays. Fluorescence was measured on the EcoCore (http://ecocore.nrel.colostate.edu) Tecan Infinite M200 Fluorescence Plate Reader at an excitation wavelength of 365 nm and emission wavelength of 450 nm (Tecan Group, Ltd., San Jose, CA). The final calculated units for soil EEA activity were reported as μmols EEA Activity per g dry soil.

Statistical Analysis

Plant leaf and leachate chemistry was evaluated using a multivariate approach to estimate the principle components (%) and the corresponding variability in chemistry explained by the first two PC axes. Graphical representation of separation was done with nonmetric multidimensional scaling ordination plots. The subset of compounds (m/z) that contributed significantly to sample separation ($p < 0.01$) were overlain on the NMDS plots for the purposes of viewing what, if any, compounds matter to NMDS separation.

Two-way ANOVAs were used to test the effects of plant species, incubation time, and the species * incubation time interaction on each of the soil functional assays. The ANOVA tests were conducted on the calculated bulk measurements, and the graphical representation of the data was relative to the water-only control values.

Results

Plant leaf chemistry and corresponding leachate chemistry did not vary in composition among species. Overall variation in C:N did differ among plant species, from 23% C:N for *A. drumonndii* to 90% for *B. dactyloides* (Table 5.1). Implementation of two detailed chemical analytical technologies, FT-IR and py-MBMS, did not reveal any distinct compounds or compound classes that differentiate any one species from another (Table 5.2, Figure 5.1-5.4). The principle components of either MidIR or NearIR never exceeded 1% variance explained for either leaf tissue or leachate (Table 5.2). The ordination examination of the other detailed dataset, py-MBMS, also did not reveal any significant differentiation among species' leaf tissue (Fig. 5.1-5.4). Specifically, NMDS ordinations were displayed for the following compound classes: alkyl aromatics, carbohydrates, lignin dimers, lipids, peptides, phenols, sterols, and unidentified compounds. The separation of plant leaf chemistry in NMDS space never exceeded 0.5. Examination of the principal components of separation never exceeded 1% variability explained.

The impacts of the leachate treatments in the soil laboratory incubation were not influenced by plant species origin (Fig. 5.5-5.7). Indices of microbial biomass were not influenced by plant species treatment, though microbial biomass C (MBC) was significantly impacted by incubation length (Figure 5.5a, $p < 0.001$) with increased MBC in the month incubation relative to the week timeframe, which were not significantly different from the MBC of the water-only treatments. Microbial respiration (Figure 5.5c) was also strongly influenced by incubation length, with a consistent increase in the cumulative respiration with length of incubation. Inorganic N showed contrasting patterns, with no significant patterns for NO_3^-, and a significant influence of

incubation length on NH_4^+ ($p = 0.006$). Finally, there was a significant decrease in N-harvesting extracellular enzyme activity with incubation time ($p < 0.001$).

Overall, there were no significant differences in detailed chemistry among the plant species' senesced leaf tissue or the resultant leachate, determined through py-MBMS and FT-IR. Additionally these insignificant differences in leaf and leachate chemistry translated to no significant species-specific impact on soil function, though incubation length showed strong changes in microbial function relative to water-only controls in terms of MBC, microbial respiration, NH_4^+ levels, and N-harvesting EEAs.

Discussion

The plant litter leachate treatments imposed no interspecific differences on soil microbial function in this experiment. Interspecific chemical variation in the litter and leachate, derived from late season senesced plants, was minimal. Despite differences in bulk C:N of leaf tissue, these differences did not translate to significant changes in detailed litter or leachate chemistry. Not surprisingly, there was no impact of plant species identity on the soil microbial functional effects of leachate. No detectable difference in leachate chemistry could have still resulted in soil functional impacts. The reactivity of certain compounds could change with no change in the relative abundance of the compound class. For instance, maple trees exposed water stress alter the reactivity of their polyphenolic compounds, while maintaining more or less similar amounts (Tharayil *et al* 2011). In this case, the shift in the reactivity of polyphenolics could impose significant changes in soil function, while the overall level of polyphenolics remains the same. However, in my examination of interspecific differences in leachate chemistry derived from the same grassland yielded no soil microbial functional differences. Soil function treated with

leachate, regardless of plant species, only differed consistently from the water-only controls. This result was not surprising, as adding a water-soluble food source to soil microbial heterotrophs should impact soil function relative to just a pulse of water.

In addition to differences in leachate versus water-only control treatments, there was a significant impact of incubation length on MBC, microbial respiration, NH_4^+, and microbial extracellular activity targeting nitrogenous compounds. Despite a lack of differentiation in species-specific soil functional impacts, this study does track the temporal effects of a leachate pulse. There was a consistent increase in microbial biomass C and heterotrophic respiration with incubation length, consistent with an overall increase in microbial growth. The increase in soil microbial growth was likely prompted by the addition of labile plant metabolites from the leachate treatments. There is also an initial suppression in NH_4^+ relative to water-only control microcosms. This suppression dissipates after a month of incubation, when NH_4^+ levels become more or less even with the water-only control treatment. The initial decrease in NH_4^+ could be a stoichiometric uptake of available mineralized N by microorganisms to compensate for the new labile, C-rich metabolites. This is somewhat consistent with the initial increase in N-harvesting EEAs. Overall, the pulse of C-rich leachate resulted in a stoichiometric need for both uptake of mineralized N and enzymatic mining of more complex forms of N. Although this study did not reveal differences among plant species' leachate, or their impacts on soil microbial function, it does show the potential temporal influence of a leachate pulse, and the time scale of operation on soil microbial function. This is important from the scale of the plant community, as it highlights a potential mechanism for seasonal changes in nutrient availability, induced by leachate-mediated shifts in microbial function.

In addition to demonstrating the overall impact of a leachate pulse, relative to water-only controls on soil microbial function, this study also shows how plant-derived water-soluble metabolites continue to leach from plant litter well after senescence. Despite the incidence of several precipitation events between the onset of plant senescence and litter harvest at the PHACE field site, the plant tissue continues to lose metabolites to the leaching process far into the autumn. This is important as it highlights a potential mechanism for how soil microbial activity continues long after aboveground plant activity has largely abated for the season.

Broadly, this study shows a lack of variability in leaf and leachate chemistry among plant species from the northern mixed grassland prairie habitat adjacent to the PHACE field site. The lack of interspecific variability in leaf and leachate chemistry translated to no significant interspecific impacts on soil microbial function. If leaf tissue from other habitats, like *Quercus hemisphaerica*, which are known to have exceptionally high leaf polyphenolic content (Popović *et al* 2013) were added to this experiment, interspecific differences would likely be found. Tannins from *Q. hemisphaerica* could act as a microbial food source but also bind to microbial extracellular enzymes, which would retard soil microbial function. Inclusion of a broad set of plants from numerous terrestrial habitats would likely result in detection of interspecific differences in leaf and leachate chemistry, which in turn could also translate to altered soil microbial function. However, the purpose of this study was not to find plant tissue that would more than likely impose strong soil microbial functional effects. Instead, the purpose of this study was to examine the breadth of leaf and leachate chemistry within a given habitat. If upwards of 16% of plant leaf tissue carbon enters the soil profile as rapid pulses of water-soluble compounds in early autumn, and if there are differences in chemistry among plant species within a given habitat, then shifts in plant community structure would substantial impacts on soil

microbial function. I did not find significant separation in interspecific chemistry, and also no soil microbial functional impacts of leachate of similar chemistry. This negative result means that plant community shifts in this particular northern mixed grassland prairie comprising the species examined in this study will not translate to shifts in soil microbial function in terms of late season senesced leaf tissue. However, the incubation with fresh tissue did show interspecific differences in soil impacts, but here I demonstrate that these differences dissipate with time since senescence. Previous studies suggest that most of the water-soluble fraction of plant-derived metabolites leach out soon after programmed cell death. This initial burst of water-soluble metabolites may also be what differentiates the among species leachate chemistry. Temporally, we could hypothesize that any impacts of species-specific leachate would be most important at the time of leaf abscission, after which the leachate impacts converge interspecifically with time (Grandy & Neff 2008, Fierer *et al* 2009, Wallenstein *et al* 2013).

In conclusion, there are no interspecific differences in plant tissue leachate, and correspondingly no among-species impacts on soil microbial function. However, this study does emphasize the overall impacts of a leachate pulse on soil function, relative to water-only controls. In the context of the habitat, this highlights a possible mechanism for why microbial function continues long after aboveground plant activity has subsided for the season.

Tables

Table 5.1. Description of plant species used in this experiment Scientific name, family, native status, and bulk chemical makeup of each plant species examined in this study.

Species	Code	Family	Life History	Native Status	Carbon %	Nitrogen %	C:N
Artemisia frigidia (Willd.)	ARFR	Asteraceae	Perennial	Native	50.08	1.10	45.44
Astragalus drummondii (Douglas ex Hook.)	ASDR	Fabaceae	Perennial	Native	47.88	2.06	23.27
Bouteloua dactyloides (Nutt.) J.T. Columbus	BUDA	Poaceae	Perennial	Native	43.60	0.48	90.03
Chrysopsis villosa (Pursh) Shinners	CHVI	Asteraceae	Perennial	Native	48.00	0.69	69.79
Koeleria pyramidata (Ledeb.) Schult.	KOPY	Poaceae	Perennial	Native	42.70	0.67	63.32
Linaria dalmatica (L.) Mill.	LIDA	Scrophulariaceae	Perennial	Invasive	48.83	0.68	71.84
Hesperostipa comata (Trin. & Rupr.) Barkworth	STCO	Poaceae	Perennial	Native	48.75	0.75	64.89

Table 5.2. Principle components of multivariate examinations of the FT-IR datasets for the leaf tissue and leachate. The principle components (PC1, PC2), generated from MidIR and NearIR Fourier Transform Infrared Spectroscopy baseline adjusted absorbance data. The variables were scaled for unit variance prior to determining principle components. All values reported as percentage variance explained on given axis.

	PC-1 **% Var Explained**	**PC-2** **% Var Explained**
MidIR Leaf Tissue	0.3524	0.3261
MidIR Leachate	0.5262	0.2927
NearIR Leaf Tissue	0.773	0.1285
NearIR Leachate	0.5837	0.2236

Figures

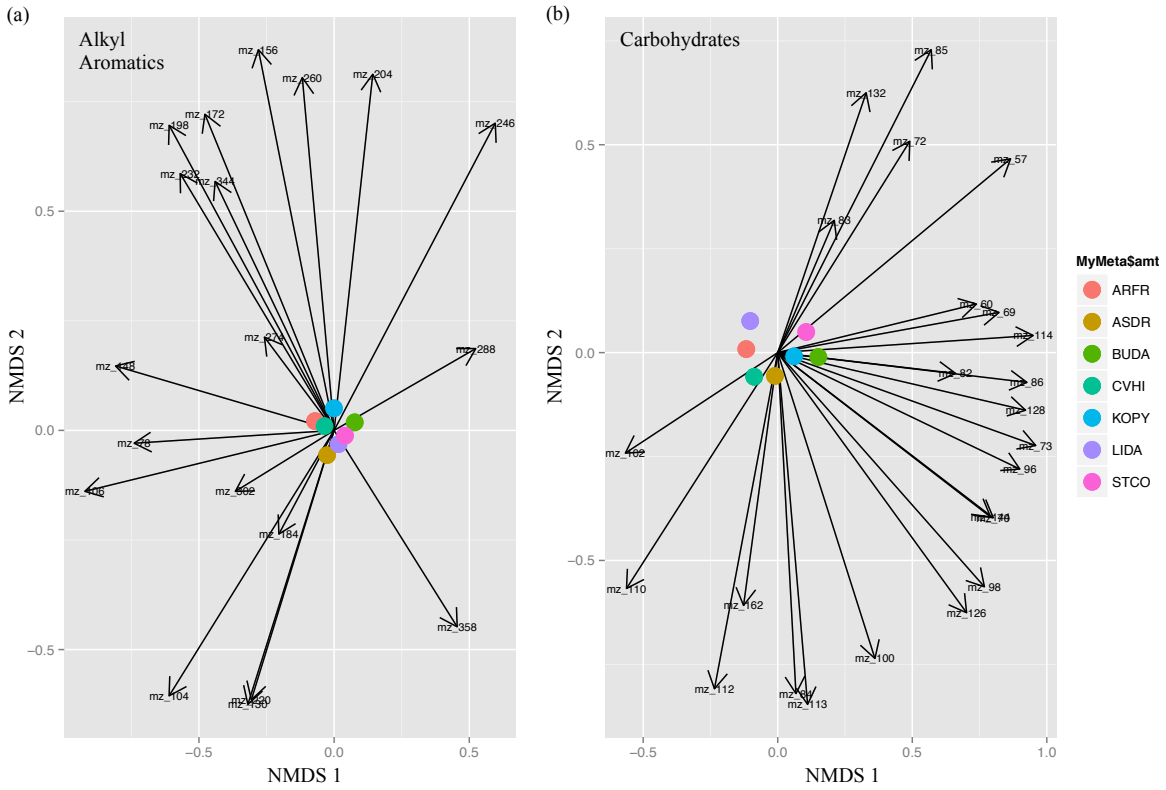


Figure 5.1. Ordination of alkyl-aromatics and carbohydrates of plant species' tissue. Non-metric multidimensional scaling (NMDS) ordination of the baseline-adjusted py-MBMS datasets categorized and colored by plant species. (a) alkyl aromatic compound class, and (b) carbohydrate compound class in py-MBMS. Black vectors represent compounds (m/z) that are significant contributors to separation (p-val = 0.01).

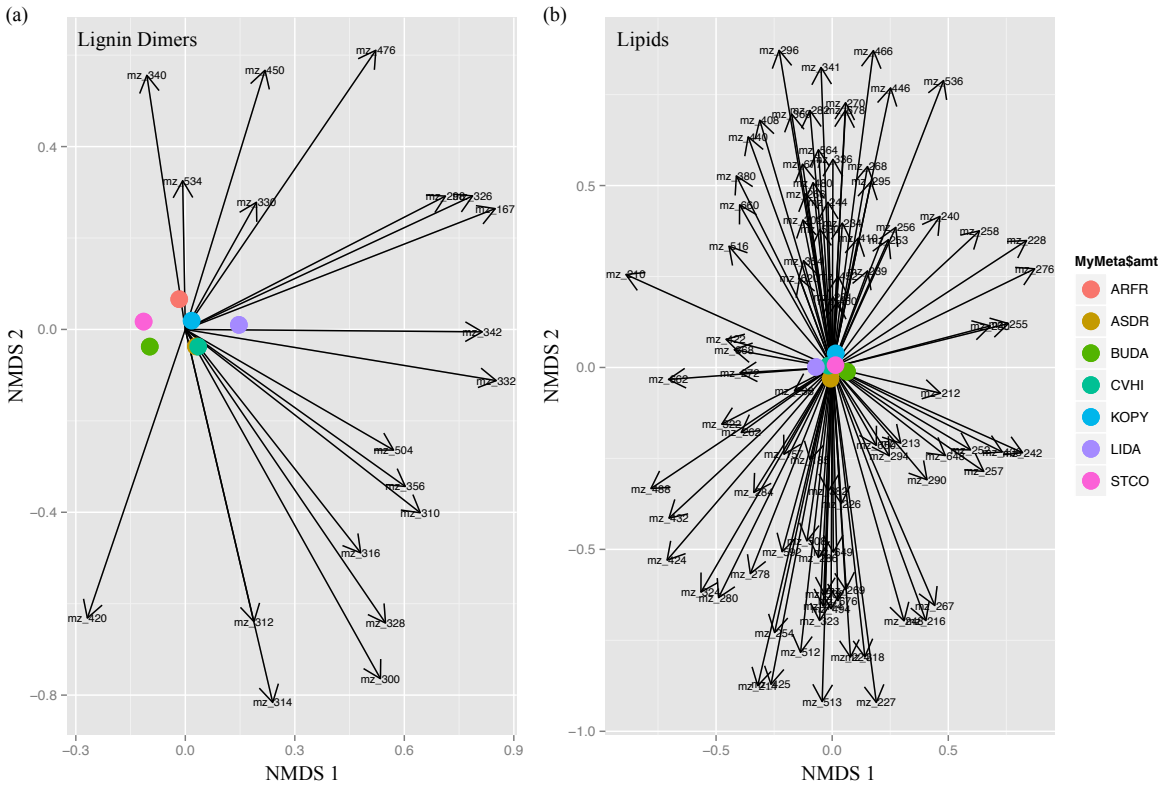


Figure 5.2. Ordination of lignin-dimers and lipids of plant species' tissue. Non-metric multidimensional scaling (NMDS) ordination of the baseline-adjusted py-MBMS datasets categorized and colored by plant species. (a) lignin dimers compound class, and (b) lipid compound class in py-MBMS. Black vectors represent compounds (m/z) that are significant contributors to separation (p-val = 0.01).

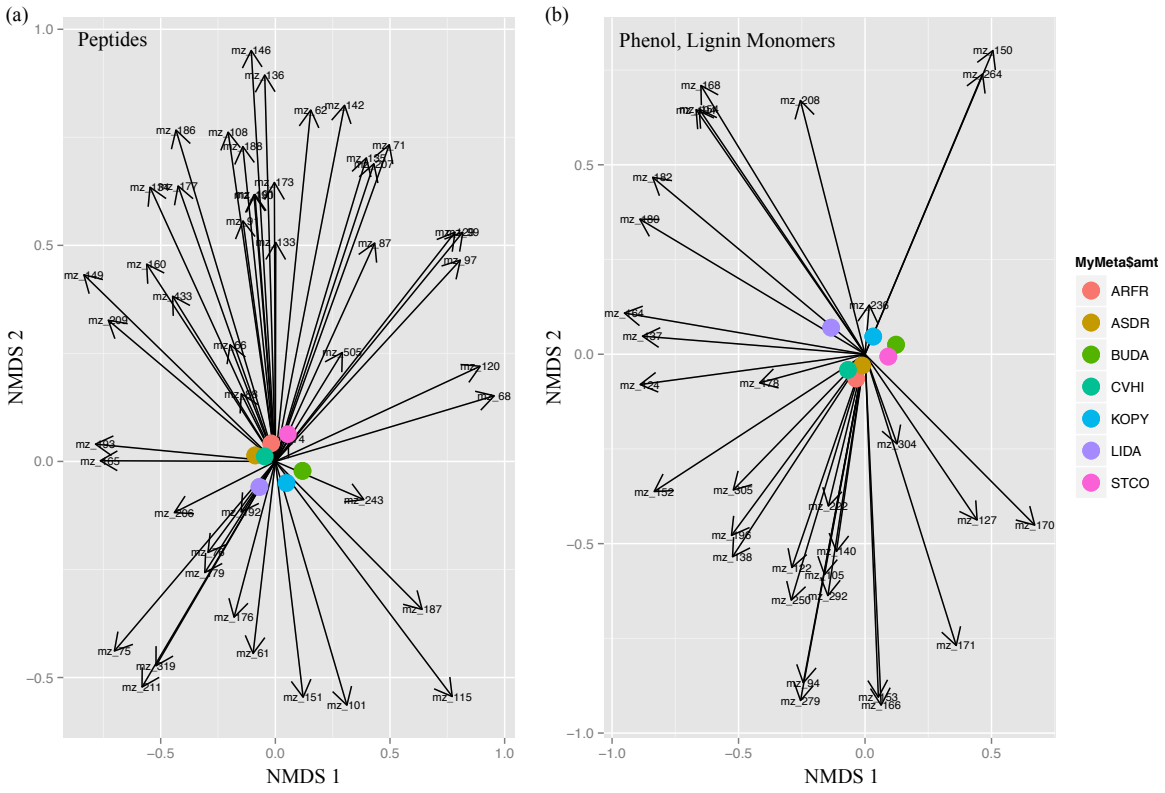


Figure 5.3. Ordination of peptides and lignin monomers of plant species' tissue. Non-metric multidimensional scaling (NMDS) ordination of the baseline-adjusted py-MBMS datasets categorized and colored by plant species. (a) peptide compound class, and (b) phenol, lignin monomer compound class in py-MBMS. Black vectors represent compounds (m/z) that are significant contributors to separation (p-val = 0.01).

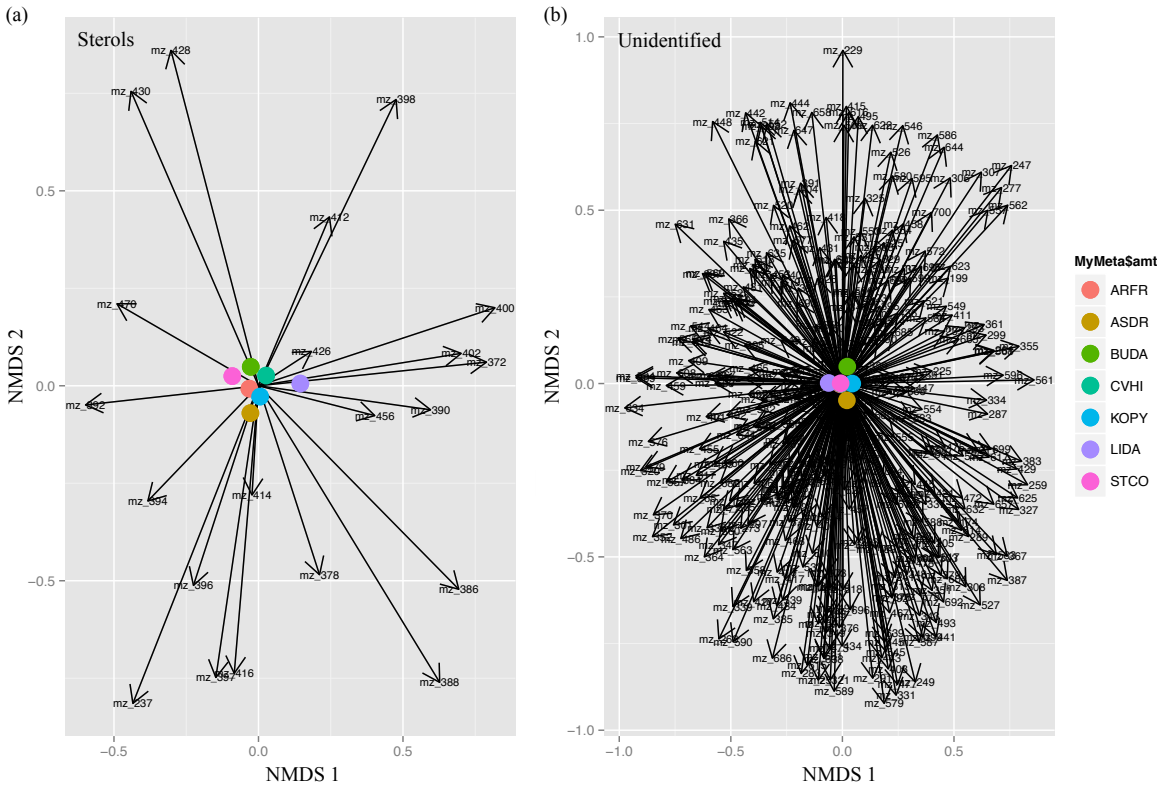


Figure 5.4. Ordination of sterols and unidentified compounds of plant species' tissue. Non-metric multidimensional scaling (NMDS) ordination of the baseline-adjusted py-MBMS datasets categorized and colored by plant species. (a) sterol compound class, and (b) unidentified compounds in py-MBMS. Black vectors represent compounds (m/z) that are significant contributors to separation (p-val = 0.01).

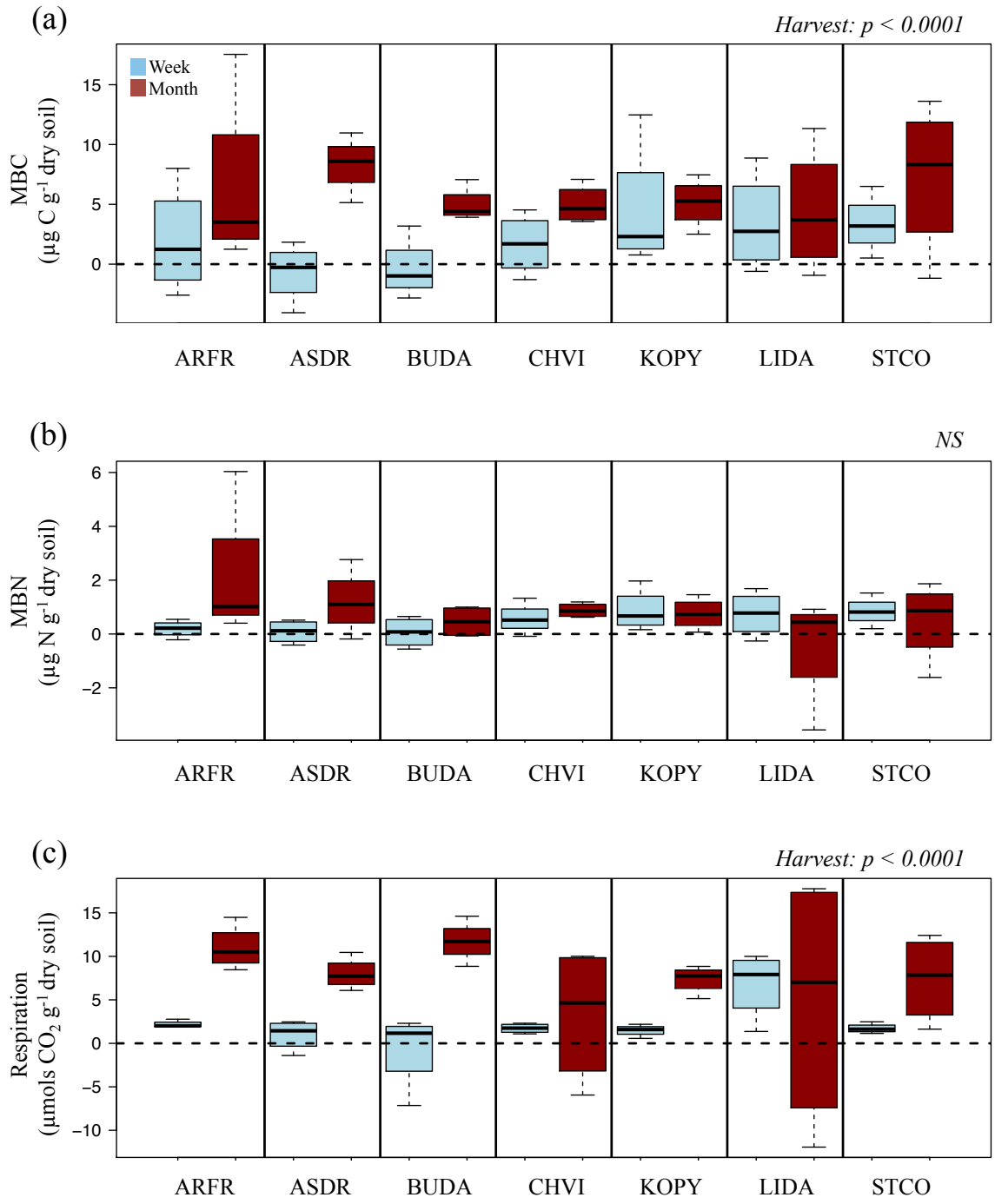


Figure 5.5. Microbial biomass and bulk activity by plant species and incubation length. MBC (a) MBN (b), and microbial respiration (c) bulk measurements per gram dry soil relative to the water-only control soil microcosms. The boxplots are colored by incubation length (blue: week, red: month). Letters above each boxplot represent Tukey HSD multiple comparisons for detecting significant differences in mean respiration (p-value: 0.05).

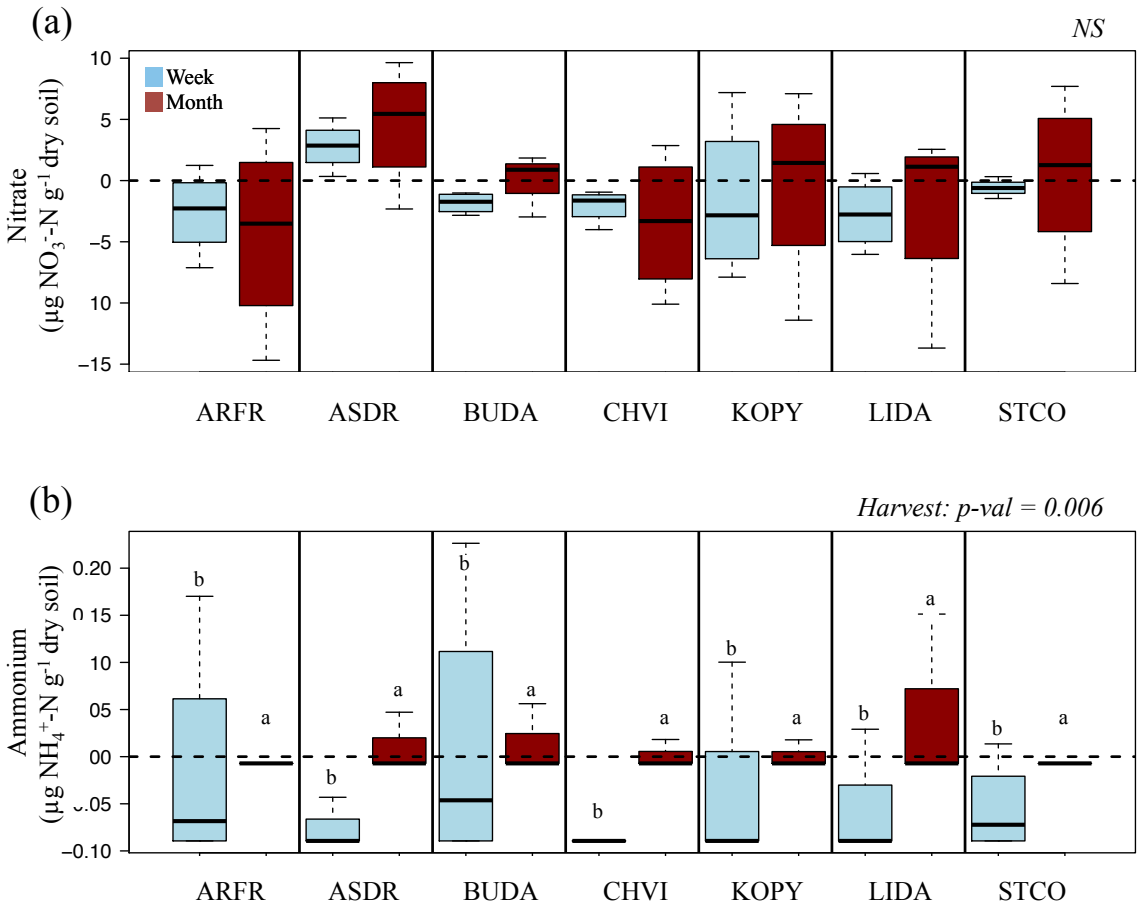


Figure 5.6. Soil inorganic N concentrations by plant species and incubation length. Nitrate (a) and ammonium (b) bulk measurements per gram dry soil relative to the water-only control soil microcosms. The boxplots are colored by incubation length (blue: week, red: month). Letters above each boxplot represent Tukey HSD multiple comparisons for detecting significant differences in mean respiration (p-value: 0.05).

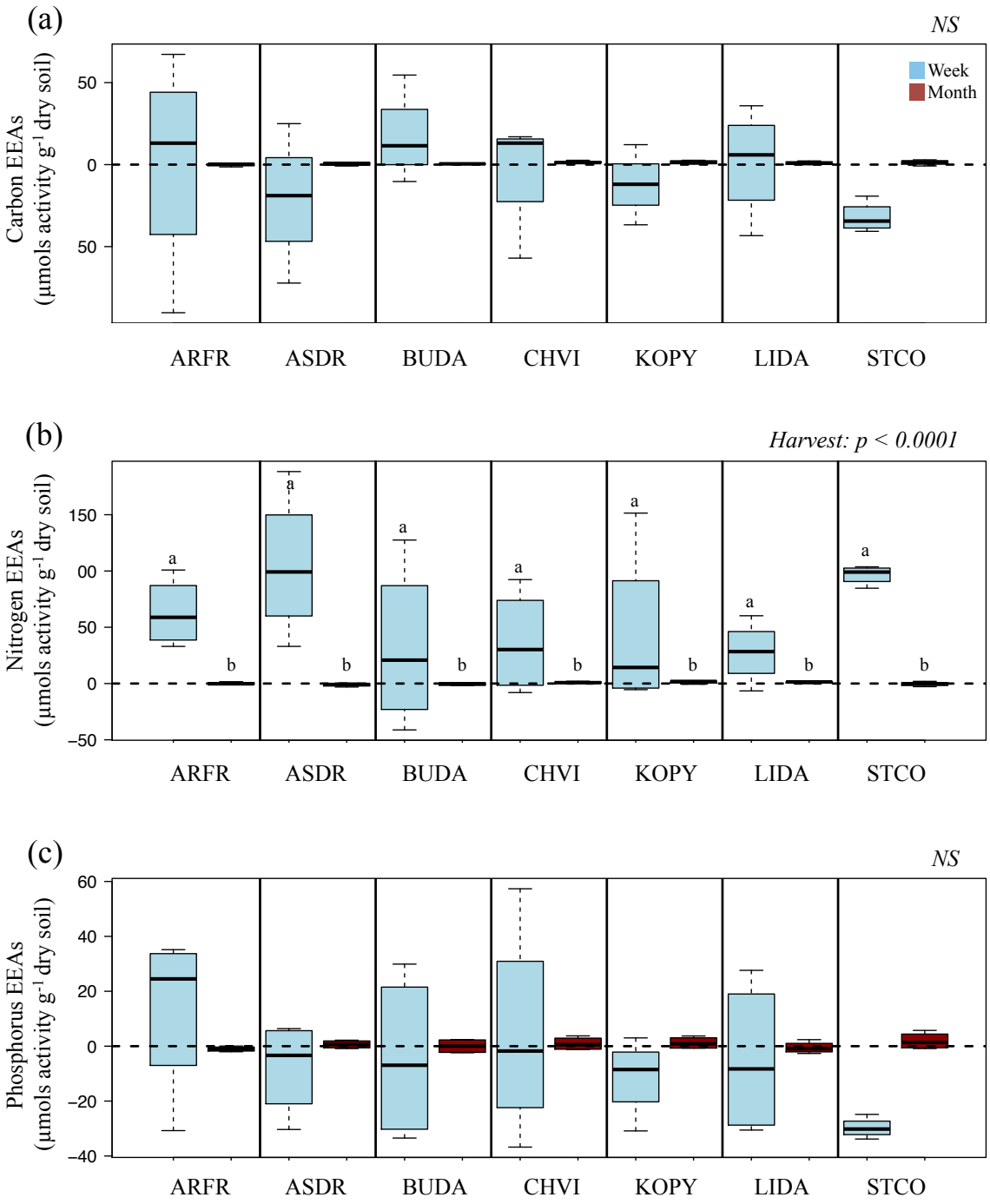


Figure 5.7. Extracellular enzyme activity. Total carbon EEAs (a), total nitrogen EEAs (b), total phosphorus EEAs bulk measurements per gram dry soil relative to the water-only control soil microcosms. The boxplots are colored by incubation length (blue: week, red: month). Letters above each boxplot represent Tukey HSD multiple comparisons for detecting significant differences in mean respiration (p-value: 0.05).

References

Bais, H.P., T.L. Weir, L.G. Perry, S. Gilroy, J.M. Vivanco. (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. Annu Rev Plant Biol 57: 233-266.

Bakker MA, Carreno-Rocabado G, Poorter L. (2011) Leaf economics traits predict litter decomposition of tropical plants and differ among land use types. Fun. Ecol. 25: 473-483.

Bird, J.A., D.J. Herman, M.K. Firestone. (2011) Rhizosphere priming of soil organic matter by bacterial groups in a grassland soil. Soil Biol & Biochem. 43(4): 718-725.

Bowman, W.D., H. Steltzer, T.N. Rosenstiel, C.C. Cleveland, C.L. Meier. (2004) Litter effects of two cooccuring alpine species on plant growth, microbial activity and immobilization of nitrogen. Oikos 104(2): 336-344.

Caporaso, J.G., J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Pena, J.K. Goodrich, J.I. Gordon, G.A. Huttley, S.T. Kelley, D. Knights, J.E. Koenig, R.E. Ley, C.A. Lozupone, D. McDonald, B.D. Muegge, M. Pirrung, J. Reeder, J.R. Sevinsky, P.J. Turnbaugh, W.A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight. (2010) QIIME allows analysis of high-throughput community sequence data. Nature Methods 7(5): 335-336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6: 1621-1624.

Coley, P.D. (1987) Interspecific variation in plant anti-herbivore properties: the role of habitat quality and rate of disturbance. New Phyt. 106: 251-263.

Coq, S., J. Souquet, E. Meudec, V. Cheynier, S. Hattenschwiler. (2011) Interspecific variation in leaf litter tannins drives decomposition in a tropical rain forest of French Guiana. Ecology. 97(7): 2080-2091.

Davidson, A.M., M. Jennions, A.B. Nicotra. (2011) Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. Ecol Lett. 14(4): 419-431.

Debs-Louka, E., N. Louka, G. Abraham, V. Chabot, K. Allaf. (1999) Effect of compressed carbon dioxide on microbial cell viability. Appl. Environ. Microbiol. 65(2): 626-631.

Edgar RC. (2010) Search and clustering order of magnitude faster than BLAST. Bioinformatics 26(19): 2460-61.

Fanin N, Fromin N, Buatois B,Hattenschwiler S. (2013) An experimental test of the hypothesis of non-homeostatic consumer stoichiometry in a plant litter–microbe system. Ecol. Lett. doi: 10.1111/ele.12108

Fanin, N., S. Hattenschwiler, N. Fromin. (2014) Litter fingerprint on microbial biomass, activity, and community structure in the underlying soil. Plant Soil. 379: 79-91.

Fierer, N. and R.B. Jackson. (2006) The diversity and biogeography of soil bacterial communities. PNAS. 103(3): 626-631.

Fierer, N., A.S. Grandy, J. Six, E.A. Paul. (2009) Searching for unifying principles in soil ecology. Soil Biol. & Biochem. 41: 2249-2256.

Freschet, G.T., W.K. Cornwell, D.A. Wardle, T.G. Elumeeva, W. Liu, B.G. Jackson, V.G. Onipchenko, N.A. Soudzilovskaia, J. Tao, J.H.C. Cornelissen. (2013) Linking litter decomposition of above- and below-ground organs to plant-soil feedbacks worldwide. J Ecol 101(4): 943-952.

Goldfarb, K.C., U. Karaoz, C.A. Hanson, C. A. Santee, M.A. Bradford, K.K. Treseder, M.D. Wallenstein,E.L. Brodie. (2011) Differential growth responses of soil bacterial taxa to carbon substrates of varying chemical recalcitrance. Frontiers Microbiology. 2(94).

Grandy, A.S. J.C. Neff. (2008) Molecular C Dynamics Downstream: the Biochemical Decomposition Sequenc and Its Impact on Soil Organic Matter Structure and Function. 297-307.

Halvorson, J.J., J.M. Gonzalez, A.E. Hagerman. (2011) Repeated applications of tannins and related phenolic compounds are retained by soil and affect cation exchange capacity. Soil Bio & Biochem 43: 1139.

Hobbie S. (1992) Effects of plant species on nutrient cycling. Trends in Ecol & Evol. 7:336–339.

Jamieson, M.A. and M.D. Bowers. (2010) Iridoid glycoside variation in the invasive plant Dalmation toadflax, Linaria dalmatica (Plantaginaceae), and sequesteration by the biological control agent, Calophasia lunula. Chem Ecol. 36: 70-79.

Joergensen, R.G. (1996) The fumigation-extraction method to estimate soil microbial biomass: calibration of the k(EC) factor. Soil Bio. Biochem. 28: 25-31.

Keeney, D. (1987) Nitrogen-availability indices. In: A. Page, R. Miller & D. Keeney (eds.) *Methods of soil analysis, part 2, chemical and microbiological properties*. Am. Soc. Agro. Madison, WI, pp: 711-734.

Kotroczo, Z., Z. Veres, I. Fekete, Z. Krakomperger, J.A. Toth, K. Lajtha, B. Tothmeresz. (2014) Soil enzyme activity in response to long-term organic matter manipulation. Soil Biol Biochem. 70: 237-243.

Liu, K., T.J. Warnow, M.T. Holder, S. Nelesen, J. Yu, A. Stamatakis, and C.R. Linder. (2012) SATe-II:Very Fast and Accurate Simultaneous Estimation of Multiple Sequence Alignments and Phylogenetic Trees. Systematic Biology 61(1): 90-106.

Mirarab, S., N. Nguyen, and T. Warnow. (2014) PASTA: ultra-large multiple sequence alignment. J Comp Biol.

Moore, B.D., R.L. Andrew, C. Kulheim, W.J. Foley. (2013) Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phyt Tansley Review. 201: 733-750.

Nicotra, A.B., O.K. Atkin, S.P. Bonser, A.M. Davidson, E.J. Finnegan, U. Mathesius, P.Poot, M.D. Purugganana, C.L. Richards, F. Valladares, M. van Kleunen. (2010) Plant phenotypic plasticity in a changing climate. Trends in Plant Sci. 15(2): 684.

Phillips, R.P., A.C. Finzi, E.S. Bernhardt. (2011) Enhanced root exudation induces microbial feedbacks to N cycling in a pine forest under long-term CO2 fumigation. Ecol Lett 14(2): 187-194.

Rinkes, Z.L., J.L. DeForest, A.S. Grandy, D.L. Moorhead, M.N. Weintraub. (2014) Interactions between leaf litter quality, particle size, and microbial community during the earliest stage of decay. Biogeochem 117(1): 153-168.

Schmidt, M.A., A.J. Kreinberg, J.M. Gonzalez, J.J. Halvorson, E. French, A. Bollmann, A.E. Hagerman. (2013) Soil microbial communities respond differently to three chemically defined polyphenols. Plant Phys & Biochem 72: 190-197.

Shi, S., A.E. Richardson, M. O'Callaghan, K.M. DeAngelis, E.E. Jones, A. Stewart, M.K. Firestone, L.M. Condron. (2011) Effects of selected root exudate components on soil bacterial communities. FEMS Microbiol Ecol 77(3): 600-610.

Soong, J.L., W. Parton, F.J. Calderon, E.E. Cambell, M.F. Cotrufo. Chemical traits of fresh or pyrolyzed aboveground plant inputs to soil determine the amount and quality of dissolved organic matter leaching during decomposition (In review).

Suseela V., Triebwasser-Freese D., Linscheid N., Morgan J.A., Tharayil N. (2014) Litters of photosynthetically divergent grasses exhibit differential metabolic responses to warming and elevated CO2. Ecosphere. 5(9): 106.

Suseela, V., N. Tharayil, B. Xing, J.S. Dukes. (2015) Warming and drought differentially influence the resorption of elemental and metabolic nitrogen pools in Quercus rubra. Global Change Biology. (In Press)

Tharayil, N. V. Suseela, D.J. Triebwasser, C.M. Preston, P.D. Gerard, J.S. Dukes. (2011) Changes in the structural composition and reactivity of Acer rubrum leaf litter tannins exposed to warming and altered precipitation: climatic-stress-induced tannins are more reactive. New Phyt. 191: 132-145.

Triebwasser , D.J., N. Tharayil, C.M. Preston, P.D. Gerard. (2012) The susceptibility of soil enzymes to inhibition by leaf litter tannins is dependent on the tannin chemistry, enzyme class and vegetation history. New Phyt. 196: 1122-1132.

Ushio, M., T.C. Balser, K. Kitayama. (2013) Effects of condensed tannins in conifer leaves on the composition and activity of the soil microbial community in a tropical montane forest. Plant Soil. 365: 157-170.

Waldrop, M.P. and M.K. Firestone. (2004) Microbial community utilization of recalcitrant and simple carbon compounds: impact of oak-woodland plant communities. Oecologia. 138(2): 275-284.

Wallenstein, M.D., M.L. Haddix, E. Ayres, H. Steltzer, K.A. Magrini-Bair, E.A. Paul. (2013) Litter chemistry changes more rapidly when decomposed at home but converges during decomposition – transformation. *Soil Bio. Biochem.* 57: 311-319.

Wurst, S., Van Dam N.M., Monroy F., Biere A., Van der Putten W.H. (2008) Intraspecific variation in plant defense alters effects of root herbivores on leaf chemistry and aboveground herbivore damage. J. Chem. Ecol. 34: 1360-1367.

Chapter 6: Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed[1]

For any enzyme-catalyzed reaction to occur, the corresponding protein-encoding genes and transcripts are necessary prerequisites. Thus, a positive relationship between the abundance of gene or transcripts and corresponding process rates is often assumed. To test this assumption, we conducted a meta-analysis of the relationships between gene and/or transcript abundances and corresponding process rates. We identified 415 studies that quantified the abundance of genes or transcripts for enzymes involved in carbon or nitrogen cycling. However, in only 59 of these manuscripts did the authors report both gene or transcript abundance and rates of the appropriate process. We found that within studies there was a significant but weak positive relationship between gene abundance and the corresponding process. Correlations were not strengthened by accounting for habitat type, differences among genes or reaction products versus reactants, suggesting that other ecological and methodological factors may affect the strength of this relationship. Our findings highlight the need for fundamental research on the factors that control transcription, translation and enzyme function in natural systems to better link genomic and transcriptomic data to ecosystem processes.

Introduction

Describing the relationship between microbial community structure and ecosystem function has emerged as an important yet elusive target in microbial ecology (Fuhrman, 2009). To evaluate these linkages, many studies use molecular techniques, including quantitative PCR

[1] Published: Rocca JD, Hall EK, Lennon JT, Evans SE, Waldrop MP, Cotner JB, Nemergut DR, Graham EB, Wallenstein MD (2015) *ISME J*. 9: 1693–1699; doi:10.1038/ismej.2014.252

(qPCR), quantitative reverse-transcription PCR (qRT-PCR) (Liu *et al*, 2010; Shannon *et al*, 2011), microarrays (Taroncher-Oldenburg *et al*, 2003; Kang *et al*, 2013) and most recently, metagenomics (Tyson *et al*, 2004; Mackelprang *et al*, 2011; Fierer *et al*, 2012; Yau *et al*, 2013). These techniques measure the relative abundances of nucleic acids encoding proteins that catalyze biogeochemical reactions, which we refer to as protein-encoding genes and transcripts, (or henceforth as gene or transcript). The use of these approaches to compare gene or transcript abundance to rates of biogeochemical processes stems from the 'sequence hypothesis' proposed in the 1950s and now a central tenet of molecular biology, where the transference of biological information is sequential, from gene(s) to transcript(s) to protein synthesis, finally leading to the resultant enzyme-catalyzed chemical reaction (Crick, 1958). On the most fundamental level, genes are prerequisites for any enzyme-catalyzed reaction and should therefore correlate. However, the relationship between gene and transcript abundances and the processes they facilitate can be obscured by various factors including habitat attributes, complexity of biogeochemical pathways, and differences in the turnover time of nucleic acids and reaction products. In addition, methodological constraints such as extraction efficiency, PCR biases and primer design can affect any apparent relationship between gene and function. Nonetheless, many studies assume that there is a quantitative relationship between functional enzymes and the processes that they catalyze. Our goal in this review was to evaluate whether this assumption is supported by data.

Materials and Methods

To address how well gene copy or transcript abundance correlates with process, we conducted a meta-analysis of microbial genes with known matching functions in environmental

habitats. Data were assembled from a wide range of studies to examine explanatory value of certain study-specific conditions, like specific gene type, habitat, DNA vs. RNA, and/or reaction states.

We obtained data for our meta-analysis by identifying studies that quantified genes or transcripts, which primarily involved C and N cycling genes using Google Scholar and ISI Web of Knowledge (for published datasets) and ProQuest Dissertation & Theses Database (to access the gray literature). We included unpublished literature in order to minimize potential publication biases against studies that found non-significant or weak correlations (Koricheva *et al*, 2013). Searches were limited to studies published after 2005, matching the time frame for when qPCR methods were optimized and routinely used. We analyzed only studies using qPCR, because it is the most well established and most quantitative method. We used the following search phrases to identify all potentially relevant studies for the initial literature search: '*functional gene expression and microb**', '*qpcr and microb**', '*quantitative pcr and microb**', and '*<gene name> and gene expression*'. Data were then extracted from figures using PlotDigitizer (http://plotdigitizer.sourceforge.net/), which we validated by inputting plots of known values to determine the accuracy and precision of the retrieved digitized data. Studies reporting log-transformed gene counts were antilog-converted, and those studies that reported gene counts per unit nucleic acid were excluded since the units must be comparable to that of the process in order to appropriately assess correlation strength.

We tested for bias in the final database in two ways. First, funnel plots of individual standard errors against residuals from the random effects model did not show asymmetry using an Egger regression test (Viechtbauer *et al*, 2010). We also assessed the database for publication bias using Rosenthal's estimation of the fail-safe n, which calculates the number of studies

yielding null outcomes that would need to be added to have a significant influence on the outcome of the existing set of studies in the database. The fail-safe n estimation must then be larger than five times n samples plus ten to assume a robust existing dataset (Rosenthal, 1991), which was true for our dataset.

Statistical analyses were performed with the 'metafor' package in R (version 3.0.2) (Viechtbauer *et al*, 2010), and a Fisher Z-transformed Pearson product-moment test of the correlations between genes and corresponding processes for the individual effect size ($Zr = 0.5 \times \ln\left[(1+r)/(1-r)\right]$). To account for the variable number of observations used to calculate a single effect size, we weighted the correlation by the inverse of the asymptotic variance ($V_{Zr} = 1/(n-3)$) (Aloe & Becker, 2009, Koricheva *et al*, 2013), which allowed us to treat effect sizes equally. To select the appropriate statistical model, we tested for homogeneity of all the effects, which was rejected, $Q = 3827, df = 223, p < 0.0001$, indicating that the distribution of true individual effect sizes were not uniformly distributed. Therefore, to assess the ability of the categorical moderators [i.e. nucleic acid form, process dynamics, biome type, functional reaction state, gene type (Table 6.1), and habitat type] to explain the heterogeneity in the relationship among nucleic acid and biogeochemical processes, we used random effects models (Viechtbauer *et al*, 2010, Koricheva *et al*, 2013), with an unbiased restricted maximum-likelihood estimator of aggregate effect sizes, as it is optimal for use in meta-analyses (Viechtbauer *et al*, 2010) to assess aggregate effect sizes and confidence intervals.

Results and Discussion

We identified 9 584 candidate studies that examined C and N genes and/or transcripts. Of these studies, 415 met the initial screening criteria for the meta-analysis (Figure 6.1). We then

removed studies that had mismatched gene and function, or those that had not quantified or not reported the corresponding function. This filtering step reduced our analysis to 59 studies containing 224 observations where qPCR was used to measure gene or transcript abundances as well as the reaction precursor immediate reaction product, or flux from one nutrient pool to the other on the same environmental sample (Figure 6.2, Suppl. Table 6.1). The 88% reduction (from 415 to 59 studies) in studies for our meta-analysis database does not suggest that qPCR is not widely used to gain inference about biogeochemical functions (Figure 6.1). Instead, most studies were excluded because (1) they only examined gene or transcript abundance (and not function), or (2) because the measured gene was not directly involved in the biogeochemical sub-process that was assessed. This suggests that gene abundances are often assumed to correlate with the related process without any evaluation of that relationship.

Contrary to the common assumption that gene abundances should be consistently correlated with process, we found extensive heterogeneity among the effect sizes of individual studies included in our meta-analysis. Overall, the abundances of genes were significantly positively correlated with process ($r = 0.26$, $p < 0.0001$, $n = 224$, Figure 6.2), though the distribution of effect sizes was centered at $Zr = 0.26$ (Figure 6.2), and resembled a normal distribution (Suppl. Figure 6.1). Only 38% of individual effects were significantly positively correlated, while nearly twice as many were either negatively correlated or showed no significant relationship. To evaluate whether gene or transcript copy abundance was more likely to be correlated with process rates, we separated the predictors by nucleic acid type (DNA versus RNA). We found a significant positive correlation between gene abundance (DNA) and process ($r = 0.30$, $p < 0.0001$, $n = 189$), while transcript abundances (RNA) were not significantly correlated ($r = 0.08$, $p = 0.60$, $n = 35$). Even though the sizes of these categories are substantially

different, the statistical analysis we used accounts for the difference in sample size when determining the aggregate effect size and confidence intervals (Nakagawa & Cuthill, 2007). The lack of correlation between transcript abundance and processes may be attributed to the rapid degradation of RNA in the environment, with the half-life of transcripts being reported to be as short as 30 seconds in some habitats (Moran *et al*, 2013). Additionally, the production rate of transcripts may not have a 1:1 relationship between enzyme activities and/or process rates, since multiple proteins can be translated from the same transcript molecule (Moran *et al*, 2013). Therefore, transcript abundance may track short-term processes (seconds to minutes), whereas gene abundance may be more reflective of microbial process performance potential integrated over longer time scales (hours to days or longer).

We also examined whether gene abundances correlated more strongly to the reaction product or reactant, as which side of the reaction is measured might influence the correlation strength. However, both reactants ($r = 0.23$, $p = 0.0051$, $n = 111$) and products ($r = 0.30$, $p = 0.0003$, $n = 113$) showed significant positive correlations with corresponding process and similar effect strengths.

When we compared gene-function relationships among terrestrial and aquatic ecosystems, we found that terrestrial studies exhibited significant positive correlations between the abundance of genes and processes ($r = 0.28$, $p < 0.0001$, $n = 172$), while those from aquatic studies did not ($r = 0.21$, $p = 0.09$, $n = 52$). This finding may reflect inherent differences between these habitats, due to a more rapid temporal decoupling of organisms and biogeochemical reactants in aquatic habitats where the matrix is more temporally static than soils. Within aquatic and terrestrial biomes we also examined more specific habitat types, and found that agricultural systems and coastal aquatic systems exhibited a significant positive gene-function relationship

[($r = 0.34$, $p < 0.0001$, $n = 110$), and ($r = 0.60$, $p = 0.02$, $n = 9$), respectively] while other habitat types (forest, grassland, meadow, tundra, estuary, and lake; Figure 6.2) did not show a significant relationship. The strength of the correlation between agricultural and coastal studies may be due to strong signal-to-noise ratios associated with fertilizer at agricultural sites, which due to well documented run-off and connectivity may also impact coastal ecosystems. As the number of studies in each habitat category is small, additional, well-designed experiments are needed to thoroughly examine the validity of these trends.

It has previously been proposed that biogeochemical processes that are narrowly distributed among bacterial or fungal taxa may be more likely to be constrained by microbial community composition than those biogeochemical processes conducted by a broader phylogenetic suite of organisms (Schimel, 1995; Schimel *et al*, 2005). We extended this concept to hypothesize that narrow processes should exhibit a stronger correlation between gene abundance and process rate than broad processes. To test this, we examined the data using gene type as the categorical variable, and found that only ammonia monooxygenases (both bacterial and archaeal *amoA*) and nitrite reductase genes (*nirK* and *nirS*), both considered to be narrowly regulated processes, yielded significant positive correlations with corresponding process rates [($r = 0.42$, $p = 0.03$, $n = 21$), ($r = 0.42$, $p = 0.03$, $n = 21$), ($r = 0.30$, $p = 0.03$, $n = 42$), and ($r = 0.28$, $p = 0.04$, $n = 39$), respectively]. Other genes, whether broadly (i.e. *pmoA*) or narrowly (i.e. *mcrA*) distributed across taxa, were not significantly correlated with process rates (Figure 6.2) The lack of correlation for certain genes may simply be due to the few available studies from our literature search which yielded only one study each for the following: *mcrA*, *nidA*, *nrfA*, and *PAHgn*, none of which showed significant correlations between gene and the corresponding process. However, it is also likely that biological factors may obscure the relationship. The potential impact of aggregated

processes likely obscures the gene-function correlation. For instance, *nifH* gene abundance may diverge from NH_4^+ concentrations because NH_4^+ is also generated by nitrogen mineralization, independent of microbial N fixation (Chapin *et al*, 2002), and gross rates of N-fixation are rarely measured. Alternatively, certain archaeal clades of methanogens also contain *nifH*-like genes (Dang *et al*, 2009), which if they are substantial contributors, may be unaccounted for in gene abundance estimates where the qPCR primers only target bacterial N-fixers.

Additional reasons for gene-process decoupling include methodological complications in quantifying rates of biogeochemical processes. For instance, the sequential reduction of nitrogen from NO_3^- to N_2 gas during denitrification is a complex pathway, where the additional contribution of NO from the anaerobic oxidation of NH_3, or potentially Annamox (anaerobic ammonium oxidation) may further obscure the relationship between *nar* genes and denitrification rates (Strous *et al*, 1999). Additionally, many studies measure NO_3^- concentrations in conjunction with *amoA* gene abundance. NO_2^- and NO_3^- gross fluxes are difficult to assess, since both molecules are mobile and are rapidly assimilated by other microbes. Finally, non-biological reduction of NO_2^- to NO, or 'chemodenitrification' (Chapin *et al*, 2002) may also obscure the gene-function relationship of the early denitrification steps.

Coupled biogeochemical reactions can make it difficult to detect relationships between a pool and a rate. For instance, the select clades of methanogens using *mcrA* genes to produce CH_4 are spatially separated from the broad range of microbes that consume CH_4 in soils (Brune *et al*, 2000). However, if both methanotrophy and methanogenesis rates are high, then the CH_4 pool might be smaller than what the *mcrA* counts would predict, and larger than what *pmoA* genes might suggest (Poret-Peterson *et al*, 2008). While the number of studies that examined CH_4 production and consumption in the meta-analysis was small (n= 8), neither the *mcrA* nor the

pmoA gene abundances were significantly correlated to CH_4 net fluxes, possibly due to the competing influences of methanogenesis and methanotrophy on changes in the CH_4 pool.

We examined studies that measured both *AOA* (archaeal *amoA*) and *AOB* (bacterial *amoA*) to determine whether correlations improved when aggregating gene analogs. Possession of the *amoA* gene was originally thought to be restricted to clades in the bacterial domain (*AOB*), but the discovery of *amoA* gene homologs in archaea (*AOA*) (Venter *et al*, 2004) and subsequent clarification of their potential importance in nitrification (Francis *et al*, 2005), indicated that neither *AOA* nor *AOB* abundances alone may accurately predict nitrification rates. However, based on our meta-analysis, *AOA+AOB* ($r = 0.12$, $p = 0.33$, $n = 28$) showed no enhancement of correlation strength over either *AOA* ($r = 0.09$, $p = 0.46$, $n = 28$) or *AOB* ($r = 0.43$, $p = 0.0001$, $n = 28$) alone, which may be due to the distinct niche requirements between the two domains of ammonia oxidizers. The archaeal and bacterial *amoA* functional clades are phylogenetically narrow, which may influence the strength of the correlation between the abundance of *AOA* and *AOB* in the environment and ammonia oxidation. Also, the two domains of ammonia oxidizers likely flourish under distinct environmental conditions, with *AOA* being more prevalent in low-pH habitats (Zhang *et al*, 2012), further highlighting the importance of measuring the abundance of both domains as a more accurate reflection of nitrifying potential.

Since genes are ultimately translated into enzymes catalyzing biogeochemical reactions, they may be more strongly correlated to fluxes than to pools of reactants or products. On the other hand, the abundance of protein-encoding genes themselves is a pool rather than a flux, and therefore it is likely that they correlate more with a pool than a flux. Despite a smaller sample size, the correlation between gene abundance and process flux showed a strong positive correlation ($r = 0.50$, $p < 0.0001$, $n = 76$), while correlation with pool size was not significant ($r =$

0.13, p = 0.08, n = 148). This is likely because pool sizes are affected not only by rates of production but also by transfer, consumption, and loss.

Our results suggest that while gene abundance may be an accurate reflection of a given microbial physiological pathway, comparing it with the corresponding biogeochemical process or pool is likely to reveal an even more complex story. For example, reactant pools may frequently be assimilated, leached or re-routed through other biochemical pathways independent of the pathway that corresponds to the gene of interest. The advent of relatively new technologies like gene microarrays (He *et al*, 2007) and metagenomics (Tyson *et al*, 2004; Mackelprang *et al*, 2011; Fierer *et al*, 2012; Yau *et al* 2013), has allowed researchers to rapidly assess more genes compared to qPCR techniques. However, the use of enticing new technologies for assessing gene abundance are also susceptible to many of the same biases that obscure the links between molecular information and biogeochemical processes illustrated by the meta-analysis presented here.

The quantification of genes through qPCR, gene microarrays, and now metagenomics, is widespread, but few studies have examined the ability of these metrics to predict process rates. At the broadest level we found a positive relationship between gene abundance and process among all studies, but the broad variation in correlation strengths shows that gene abundances cannot be used *a priori* as a proxy for biogeochemical processes. Rather, these relationships need to be examined for each study system before gene abundance can be used to infer microbial biogeochemical process. When multiple studies were examined, the concentration of products or reactants rarely correlated to gene abundance (38%). Even in the cases where the correlations are strongly positive, they should not be interpreted as demonstrating causation without further evidence. Spatial and temporal heterogeneity in environmental conditions, such as anoxic/oxic

and water and/or nutrient availability, may not track the variability of corresponding gene abundance in space and time due to different residence times of these indices. Monitoring gene abundance itself may be useful for determining the variability of that particular organism, but that may not explain fluctuations in the corresponding process. Therefore, studies whose main objective is to examine a particular biogeochemical process may not require, or even be hindered by inclusion gene abundance for determining and predicting process (Graham *et al*, 2014). However, state-of-the-art technologies, like metagenomic assessment of protein-encoding genes may more accurately reflect process, as they capture more gene diversity (Tyson *et al*, 2004; Mackelprang *et al*, 2011). Additionally, quantifying expressed proteins using metaproteomics may prove to be a robust indicator of corresponding function(s). Our findings point to a critical need for more fundamental studies on the factors controlling gene abundance, transcription, translation, and enzyme activity, so that we can better interpret the potential for patterns in gene or transcript abundance in a broad range of environments if genes will inform us about fundamental ecosystem processes.

Tables

Table 6.1. Description of the PEGs examined in this meta-analysis.

PEG	Encoded Enzyme	Function	Citation
Nitrogen Cycle			
nifH	Dinitrogenase reductase	NH_4^+ production (N-fixers)	*Yun & Szalay 1984*
AOA	Archaeal Ammonia monooxygenase, α subunit	NO2-, NO3- accumulation (Nitrifiers)	*Könneke et al, 2005*
AOB	Bacterial Ammonia monooxygenase, α subunit	NO2-, NO3- accumulation (Nitrifiers)	*Rotthauwe et al, 1997*
napA, narG	Nitrate reductase	NO_2^- <--> NO_3^-	*Graham et al, 2003, Flanagan et al, 1999*
nirK, nirS	Nitrite reductase	NO_2^-/NO_3^- --> NO (Denitrification)	*Braker et al, 1998*
nrfA	Nitrite reductase	NO_2^-/NO_3^- --> NO (Denitrification)	*Wang & Gunsalus, 2000*
cnorB	Cytochrome B nitric oxide reductase	NO --> N_2O (Denitrification)	*Braker & Tiedje 2003*
nosZ	Nitrous oxide reductase	N_2O --> N_2 (Denitrification)	*Kloos et al, 2001*
Methane			
mcrA	Methyl-coenzyme M reductase	CH_4 production	*Luton et al, 2002*
pmoA	Particulate methane monooxygenase	CH_4 consumption	*Heyer et al, 2002*
Other			
nidA	α subunit dioxygenase	Degradation of polycyclic aromatic hydrocarbons (PAH rings)	*Brezna et al, 2003*
PAHgn	Gram-negative bacterial PAH dioxygenase	PAH ring degradation	*Mueller et al, 1997*
tfdA	2,4 dichlorophenozyacetate monooxygenase	1st step of MCPA degradation	*Streber et al, 1987*

Figures

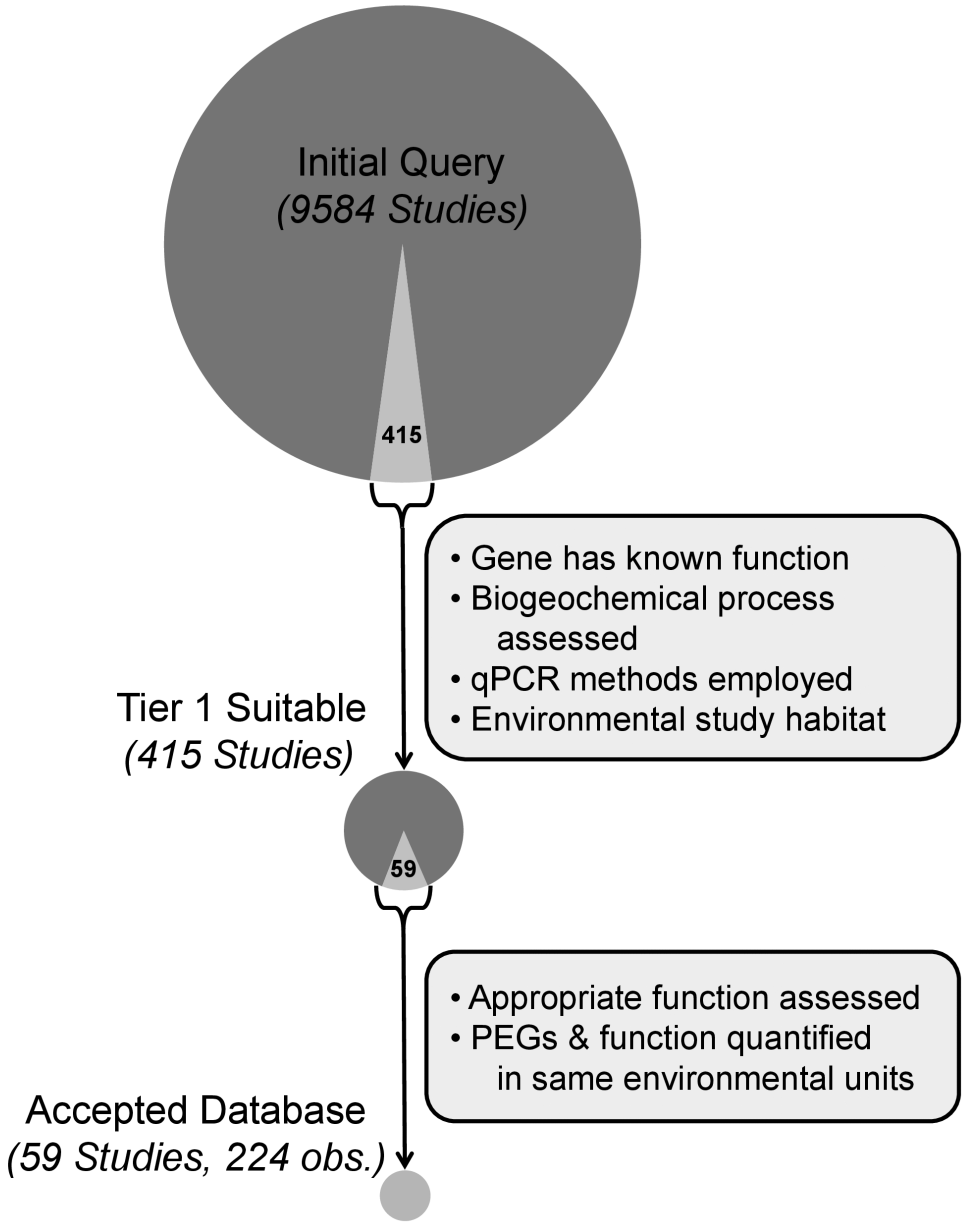


Figure 6.1. Pie charts illustrating the number of studies at each stage of evaluation criteria, the number of studies excluded by each set of criteria and the final number studies that were accepted and included in the meta-analysis.

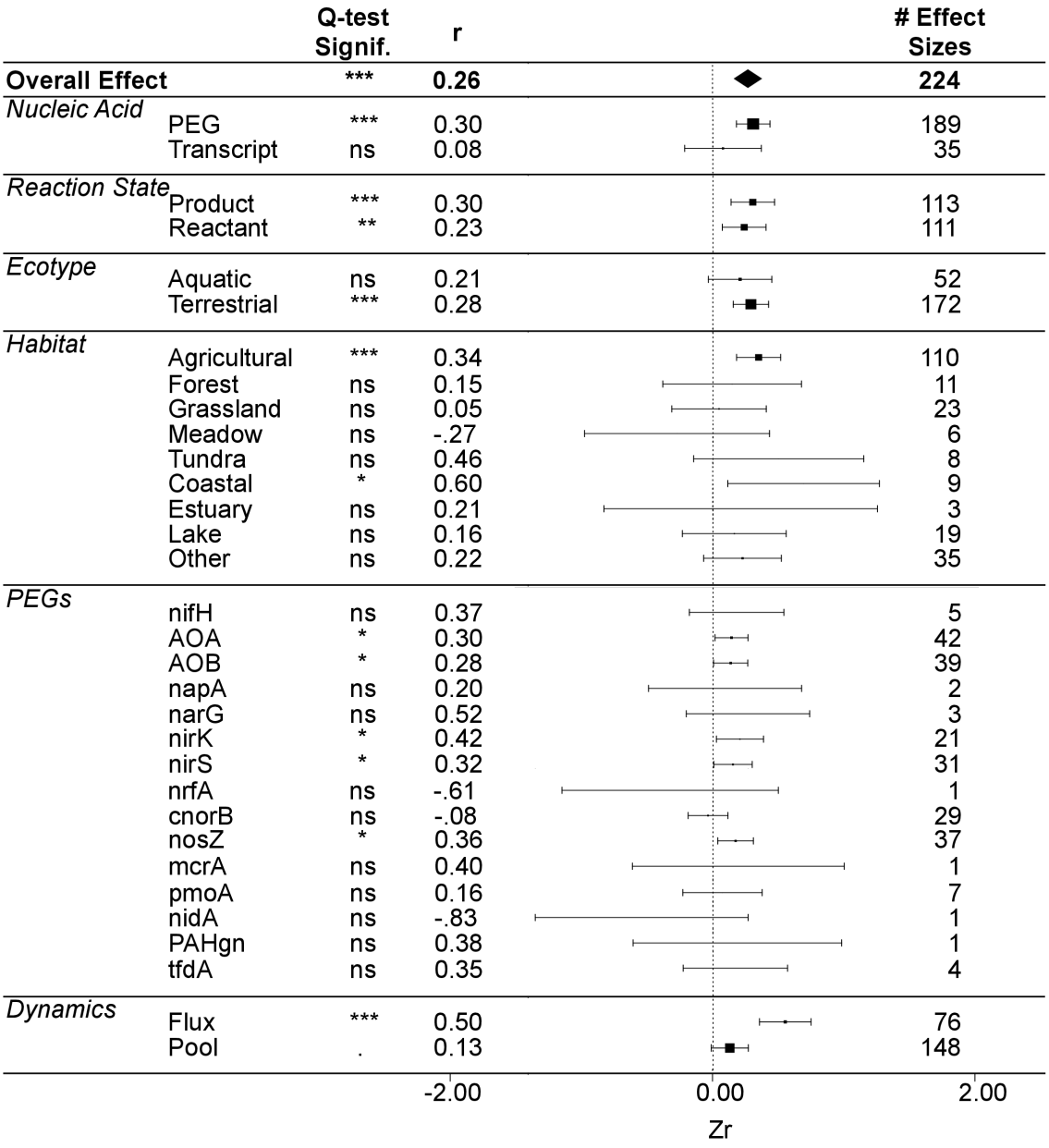


Figure 6.2. Overall effect sizes for the relationship between gene or transcript abundance and corresponding processes by categorical moderator. The Zr = 0 line indicates no correlation, with positive and negative values indicating the directionality of the relationship. Horizontal bars represent 95% confidence intervals of each effect, with bars not crossing Zr=0 indicative of a significant correlation. Q-test significance codes: '***' - p<0.001, '**' - p=0.001, '*' - p=0.05, '.' – p=0.1, ns - p>0.1.

References

Aloe AM, Becker BJ. (2009) Teacher Verbal Ability and School Outcomes: Where is the evidence? Educ Res 38: 612-624.

Braker G, Fesefeldt A, Witzel K. (1998) Development of PCR primer systems for amplification of nitrite reductase genes (nirK and nirS) to detect denitrifying bacteria in environmental samples. Appl Environ Microbiol 64: 3769-3775.

Braker G, Tiedje JM. (2003) Nitric oxide reductase (norB) genes from pure cultures and environmental samples. Appl Environ Microbiol 69: 3476-3483.

Brezna B, Khan AA, Cerniglia CE. (2003) Molecular characterization of dioxygenases from polycyclic aromatic hydrocarbon-degrading Mycobacterium spp. FEMS Microbiol Lett 223: 177-183.

Brune A, Frenzel P, Cypionka H. (2000) Life at the oxic-anoxic interface: microbial activities and adaptations. FEMS Microbiol Rev 24: 691-710.

Chapin FS, Matson PA, Mooney HA. (2002) Principles of Terrestrial Ecosystem Ecology. Springer-Verlag: New York.

Crick F. (1958) On Protein Synthesis. The Symposia of the Society for Experimental Biology 12: 138-163.

Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL *et al.* (2012) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proc Natl Acad Sci USA 109: 21390-21395.

Flanagan DA, Gregory LG, Carter JP, Karakas-Sen A, Richardson DJ, Spiro S. (1999) Detection of genes for periplasmic nitrate reductase in nitrate respiring bacteria and in community DNA. FEMS Microbiol Lett 177: 263-270.

Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB. (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. Proc Natl Acad Sci USA 102: 14683-14688.

Fuhrman JA. (2009) Microbial community structure and its functional implications. Nature 459: 193-199.

Graham LG, Bond PL, Richardson DJ, Spiro S. (2003) Characterization of a nitrate-respiring bacterial community using the nitrate reductase gene (narG) as a functional marker. Microbiology 149: 229-237.

Graham EB, Wieder WR, Leff JW, Weintraub SR, Townsend AR, Cleveland CC *et al.* (2014) Do we need to understand microbial communities to predict ecosystem function? A comparison of statistical models of nitrogen cycling processes. Soil Biol Biochem 68: 279-282.

He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC *et al.* (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J 1: 67-77.

Heyer J, Galchenko VF, Dunfield PF. (2002) Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. Microbiology 148: 2831-2846.

Kang S, Van Nostrand JD, Gough HL, He Z, Hazen TC, Stahl DA *et al.* (2013) Protein-encoding gene array-based analysis of microbial communities in heavy metals-contaminated lake sediment. FEMS Microbiol Ecol 86: 200-214.

Kloos K, Mergel A, Rösch C, Bothe H. (2001) Denitrification within the genus Azospirillum and other associative bacteria. Funct Plant Biol 28: 991-998.
Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature 437: 543-546.
Koricheva J, Gurevitch J, Mengersen K. (2013) Handbook of Meta-analysis in Ecology and Evolution. Princeton University Press: Princeton.
Liu B, Morkved PT, Frostegard A, Bakken LR. (2010) Denitrification gene pools, transcription and kinetics of NO, N_2O and N_2 production as affected by soil pH. FEMS Microbiol Ecol 72: 407-417.
Luton PE, Wayne JM, Sharp RJ, Riley PW. (2002) The mcrA gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. Microbiology 148: 3521-3520.
Moran MA, Satinsky B, Gifford SM, Luo H, Rivers A, Chan LK *et al.* (2013) Sizing up metatranscriptomics. ISME J 7: 237-243.
Mueller JG, Devereux R, Santavy DL, Lantz SE, Willis SG, Pritchard PH. (1997) Phylogenetic and physiological comparisons of PAH-degrading bacteria from geographically diverse soils. A van Leeuw 71: 329-343.
Nakagawa S, Cuthill IC. (2007). Effect size, confidence intervals and statistical significance: a practical guide for biologists. Biol Rev Camb Philos Soc 82: 591-605.
Poret-Peterson AT, Graham JE, Gulledge J, Klotz MG. (2008) Transcription of nitrification genes by the methane-oxidizing bacterium, Methylococcus capsulatus strain Bath. ISME J 2: 1213-1220.
Rosenthal R. (1991) Meta-analytic Procedures for Social Research. SAGE Publications: Newbury Park.
Rotthauwe JH, Witzel KP, Liesack W. (1997) The ammonia monooxygenase structural gene amoA as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. Appl Environ Microbiol 63: 4704-4712.
Schimel JP. (1995) Ecosystem consequences of microbial diversity and community structure. In: Chapin FS, Korner C (eds). Arctic and Alpine Biodiversity: patterns, causes, and ecosystem consequences. Springer-Verlag: Berlin. pp 239-254.
Schimel JP, Bennett J, Fierer N. (2005) Microbial community composition and soil N cycling: is there really a connection? In: Bardgett RD, Hopkins DW, Usher MB (eds). Biological diversity and function in soils. Cambridge University Press: Cambridge, pp 171-188.
Shannon EM, Saleh-Lakha S, Burton DL, Zebarth BJ, Goyer C, Trevors JT. (2011) Effect of nitrate and glucose addition on denitrification and nitric oxide reductase (cnorB) gene abundance and mRNA levels in Pseudomonas mandelii inoculated into anoxic soil. A van Leeuw 100: 183-195.
Streber WR, Timmis KN, Zenk MH. (1987) Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene tfdA of Alcaligenes eutrophus JMP134. J Bacteriol 169: 2950-2955.
Strous M, Kuenen JG, Jetten MS. (1999) Key physiology of anaerobic ammonium oxidation. Appl Environ Microbiol 65: 3248-3250.
Taroncher-Oldenburg G, Griner EM, Francis CA, Ward BB. (2003) Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. Appl Env Microbiol 69: 1159-1171.

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM *et al.* (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428: 37-43.
Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA *et al.* (2004) Environmental Genome Shotgun Sequencing of the Sargasso Sea. Science 304: 66-74.
Viechtbauer W. (2010) Conducting a Meta-Analysis in R with the metafor Package. J Stat Softw 36: 1-48.
Wang H, Gunsalus RP. (2000) The nrfA and nirB Nitrite Reductase Operons in Escherichia coli Are Expressed Differently in Response to Nitrate than to Nitrite. J Bacteriol 182: 5813-5822.
Yau S, Lauro FM, Williams TJ, DeMaere MZ, Brown MV, Rich J *et al.* (2013) Metagenomic insights into strategies of carbon conservation and unusual sulfur biogeochemistry in a hypersaline Antarctic lake. ISME J 7: 1944-1961.
Yun AC, Szalay AA. (1984) Structural genes of dinitrogenase and dinitrogenase reductase are transcribed from two separate promoters in the broad host range cowpea Rhizobium strain Irc78. Proc Natl Acad Sci USA 81: 7358-7362.
Zhang L, Offre PR, He J, Verhamme DT, Nicol GW, Prosser JI. (2010) Autotrophic ammonia oxidation by soil thaumarchaea. Proc Natl Acad Sci USA 107: 172-17245.

Chapter 7: Conclusions

Microorganisms are the chief regulators of biogeochemical cycling and ecosystem function. Therefore, understanding how their communities assemble and the factors controlling these assembly processes will help clarify the link between community structure and their resultant function that are so key to ecosystem processes. The primary objective of my research was to examine various aspects of scale (phylogenetic, plant community, and functional gene abundance) on microbial community structure and function. Specifically, I assessed the scalability of phylogenetic dispersion indices, the impact of environmental factors constraining soil microbial community structure and function, and to evaluate the usage of microbial functional gene abundance as a proxy for resultant function. With a collection of soil microbial communities, two laboratory incubations, and assembling many studies into a comprehensive meta-analysis, I was able to address my main research questions, which were:

1. Does taxonomic scale influence estimations of phylogenetic dispersion in soil bacterial communities?
2. What are the impacts of long term elevated CO_2 and warming, and interannual variability in water availability on soil bacterial community structure?
3. What are the impacts of altered abiotic conditions on plant leachate, and how do these changes translate to shifts in soil microbial community structure and function?
4. Does late season senesced plant leachate chemistry vary interspecifically, and does this impact soil function?
5. Is microbial functional gene abundance a valid proxy for corresponding function?

In Chapter 2, I examined the influence of taxonomic scale on estimations of phylogenetic community dispersion. Microbial ecologists are moving beyond site descriptions of what microbial taxa are present at a given site to unraveling assembly mechanisms that determine microbial community structure. Phylogenetic community dispersion begins to shed light on what assembly processes might be primarily structuring microbial communities: environmental filtering and/or competition. However, the bacterial domain is the typical taxonomic scale used to estimate phylogenetic dispersion. I set out to explore the influence of taxonomic scale, because I hypothesized that the domain-level was far too broad for valid interpretation. Other studies have tested this impact of taxonomic scale with macroorganism, and it has been tested with microorganisms but still at broad scales. In examining all internal phylogenetic nodes of a bacterial tree with a real soil bacterial community dataset, I determined that the bacterial domain is too broad for anything other than clustering signals to occur. Overdispersion, indicative of potential competitive interactions, was rare in my dataset and only occurred in taxonomic scales finer than the bacterial phylum. This chapter contributes to the field of microbial ecology in that if we want to use phylogenetic dispersion indices, we should consider the scale of assessment and what it means for the interpretation.

In Chapter 3, I used the information generated from Chapter 2 to assess the same soil bacterial community dataset in the context of the abiotic factors, elevated CO_2 and temperature, and interannual variability in soil moisture availability in determining bacterial community structure. I found that elevated CO_2 and temperature had insignificant influence on bacterial community structure. Surprisingly, soil moisture availability also had no significant impact on bacterial community structure. Many previous studies have demonstrated the strong influence of temporal heterogeneity of water availability on soil microbial communities, but the PHACE

bacterial dataset did not support this idea. Season was a strong driver of bacterial community structure, and I conjecture that this has something to do with overall seasonality of plant function. Because PHACE is situated in a strong seasonal northern mixed grassland, the plants undergo drastic changes throughout the year. In spring, the plants are likely 'waking up', and increasing photosynthesis, and increasing uptake of available soil nutrients. By late summer, photosynthesis has subsided and root exudation increases. These changes in plant functionality are likely the drivers of season that structure bacterial communities that separate strongly between April and July collection dates.

In Chapters 4 and 5, I examined two scales of plant community structure and their impact on soil function. In the context of forecasted climate change, there could be direct impacts of altered climate on plant physiology, which in turn may result in altered leaf and leachate chemistry. Climate change may also induce shifts in plant community structure. If there are among plant species differences in plant leaf and leachate chemistry, then shifts in plant community structure could impose strong changes on soil function. Therefore, I examined both inter- and intraspecific differences in plant litter leachate chemistry. In the first incubation, I selected fresh aboveground plant tissue from the four most abundant plants present in the PHACE plots, which were exposed to the long-term elevated CO_2 and warming treatments, as well as a off plot pine functional outgroup. The plant species selection included a suite of functional and native status. Overall, I found evidence that shifts in plant chemistry, both in terms of plant species and long-term exposure to different climate conditions, influenced soil microbial function and bacterial community structure. Specifically, one of the invasive plants, *Linaria dalmatica*, possessed distinct leachate chemistry, impacted soil function, and imposed stronger influences on the bacterial community structure relative to other species' leachates.

Separation in the chemistry of plant litter leachate, which primarily separated *L. dalmatica* from the rest of the species, was primarily driven by a handful of metabolites, including porphine, fumaric acid, shikimic acid and galacturonic acid. The primary driver of bacterial community structure in the incubation was harvest time, though soil communities exposed to *L. dalmatica* leachate still separated from the other soil communities. This consistent shift in bacterial community is likely due to the drastic impact of altering soil physical structure and water availability. As the invasive *L. dalmatica* exhibited unique impact on soil microbial structure and function, it may highlight the influence of plant invasions in general and how they can present substantial functional consequences under future climate. In the second experiment, I assessed the among species differences in leachate from late season senesced plant tissue. The rationale was to see if interspecific differences in plant tissue chemistry persisted late after the initial leachate compounds were lost from the plant tissue. I harvest late season senesced leaf tissue from the most common plant species at the PHACE, from adjacent untreated grassland. I found very little interspecific differences in either plant litter or leachate chemistry derived from senesced plants. Not surprisingly, the consequences of this leachate were minimal in terms of soil microbial function. Broadly, these incubations highlight the importance of considering climate-mediated shifts in plant community structure as potential drivers of soil function through altered fresh litter leachate permeating the soil profile, but that these interspecific differences likely dissipate with late senescence.

Finally, in Chapter 6, I assess the scalability of microbial functional gene abundance to corresponding functional capacity. I conducted a meta-analysis where I examined the link between microbial protein-encoding gene abundance and actual functional capacity, flux, or pool as part of the USGS Powell Center working group ‘The next generation of ecological indicators:

defining which microbial properties matter most to ecosystem function and how to measure them'. For any enzyme-catalyzed reaction to occur, the matching gene and transcript are necessary prerequisites. It is therefore reasonable to assume a positive relationship between the abundance of protein-encoding genes and/or transcripts for enzymes catalyzing biogeochemical processes and the process itself. The assumption was tested by conducting a meta-analysis and found that of 415 studies that quantified protein-encoding genes or transcripts abundance, only 59 reported both gene or transcript abundance and the corresponding biogeochemical process. This result suggests that the relationship between molecular information and corresponding biogeochemical rate is commonly assumed yet rarely evaluated. Of the 224 individual relationships between gene or transcript abundance and process that were evaluated (reported in the 59 studies), we found a significant positive correlation, yet extensive variability amongst individual effect sizes were not explained by study site characteristics, suggesting that other ecological and methodological factors may affect this relationship, or that this type of molecular data does not reflect functional capacity. Since our findings contradict how gene and transcript abundance information are currently being interpreted, we feel that our analysis has the potential to make an important contribution to advancing the field of microbial ecology.

In conclusion, the overall objective of my research was to examine various aspects of scale on the structure and function of microbial communities. Taxonomic scale impacts our interpretation of what community assembly process primarily regulates microbial communities. The influence of climate-mediate shifts in plant community structure may scale down to changes in soil function with altered leachate permeating the soil profile. Finally, the utility of microbial protein-encoding gene abundance does not consistently scale to the corresponding function.