BIOGEOGRAPHY OF PLANT AND SOIL MICROBIAL COMMUNITIES

by

JONATHAN WINSTON LEFF

B.A., University of Colorado, 2007

M.S., University of Montana, 2011

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Ecology and Evolutionary Biology

2016

This thesis entitled:
Biogeography of plant and soil microbial communities
written by Jonathan Winston Leff
has been approved for the Department of Ecology and Evolutionary Biology

Dr. Noah Fierer

Dr. William D. Bowman

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Leff, Jonathan Winston (Ph.D., Ecology and Evolutionary Biology)
Biogeography of Plant and Soil Microbial Communities
Thesis directed by Associate Professor Noah Fierer

ABSTRACT

Plants and soil are fundamental components of terrestrial ecosystems, and microorganisms play key roles in the health of plants and in the ecosystem processes that take place in soils. Thus, in order to understand the functioning of terrestrial ecosystems, it is critical that plant and soil associated microbial communities are studied. Yet, remarkably little is known about the general distributions of plant and soil communities and the factors determining their structure. Here, I present an exploration of the biogeography of plant and soil microbial communities through four independent studies. In the first, I show how bacterial community structure varies predictably throughout the leaves and bark in individual *Ginkgo biloba* trees. Next, I investigate whether domestication of the sunflower plant (*Helianthus annuus*) has led to differences in root and rhizosphere fungal and bacterial communities. The results demonstrate that domestication has affected fungal but not bacterial communities. Third, using two complementary experiments set in a grassland, I show that plant species identity causally affects soil fungal, bacterial, protistan, and metazoan community composition and that the composition of soil communities in field samples were predictable based on plant community composition. Fourth, I show that increased inputs of nitrogen and phosphorus had consistent effects on soil fungal and bacterial communities in grasslands located around the globe. I describe how the observed effects relate to the function and ecology of the belowground community members and concomitant shifts in the composition of plant communities. Together, these studies reveal general patterns in plant and soil associated microbial communities and demonstrate important factors determining the structure of these communities. These results will help enable a predictive understanding of the biogeography of plant and soil microbial communities and will lead to a better understanding of terrestrial ecosystems.

ACKNOWLEDGEMENTS

First and foremost, I express my sincere gratitude for the tireless assistance given to me by my advisor, Noah Fierer. He is a true role model, and I am very lucky to have him as a mentor. I am also grateful to my dissertation committee. Jessica Henley provided general laboratory support. There were many people and funding sources that have been essential to individual chapters. Their contributions are indicated below:

Chapter II. Samantha Weintraub provided extensive help collecting hundreds of leaf and bark samples as well as maintaining the sanity of the author. Chris Steenbock assisted with the laboratory work, and Susan Leff was responsible for ensuring the samples made it safely to the laboratory in Colorado. Funding for the project was generously provided by Elisabeth Dudley and Janine Luke with additional funding to Noah Fierer from the National Science Foundation.

Chapter III. Xavier Rojas provided integral laboratory assistance, and Indigo Agriculture generously provided funding.

Chapter IV. Richard Bardgett, Ben Jackson, James Pritchard, and many others in the Bardgett lab established and maintained the experiments used in this project. Emily Morgan performed essential laboratory work.

Chapter V. I am grateful to The Nutrient Network (http://nutnet.org) and its funders for making this work possible. Monte Lunacek of University of Colorado Research Computing provided valuable assistance for using CU computational resources. Elizabeth DeLorenze and Ryan Williams provided valuable feedback on earlier drafts of the chapter. Xavier Rojas helped with the laboratory sample preparation. The shotgun metagenomic analyses were made possible with support from the U.S. Department of Energy's Joint Genome Institute and their Community Sequencing Program. This work was supported by a grant to Noah Fierer from the National Science Foundation.

CONTENTS

FIGURES

CHAPTER I

INTRODUCTION AND SUMMARY

The microbial biodiversity of planet Earth is immense. There are thought to be upwards of one trillion microbial species (Locey and Lennon 2016), which is several orders of magnitude times the number of plant (~380,000) and animal (~1.5 million) species combined (Roskov et al. 2016). These organisms are engaged in myriad activities, and they are relevant to many of the ecosystem processes that take place on our planet. Yet, in contrast with plants and animals, our knowledge of geographic patterns in microbial diversity is comparatively trivial (Decaëns 2010). In particular, the geographic distributions of microbial communities vary throughout the world, but the details of how and why they vary are relatively unknown. In this thesis, I present four independent studies which investigate the spatial distribution, or biogeography, of microbial communities associated with plants and soil. These studies aim to develop our nascent understanding of plant and soil microbial biogeography so that we can enhance our understanding of how they function and influence the world around us.

I chose to study microbial communities associated with plants and soil because they are key components of ecosystem functioning and human well-being. In plants, bacteria and fungi can cause disease or protect plants from disease, and they can impact plant growth by facilitating their access to nutrients and water and protecting them from heavy metals (Berendsen et al. 2012, Bulgarelli et al. 2013, Turner et al. 2013). Still, many of the microorganisms that closely associate with plants are not known, and many more are not well understood. Likewise, soil health and functioning underlie most terrestrial ecosystem processes, agricultural productivity, and human health (Schlesinger and Bernhardt 2013, Wall et al. 2015). Microbial communities regulate soil processes (Bardgett and van der Putten 2014). Yet, soil microbial communities are extremely diverse (Roesch et al. 2007), and many questions

regarding soil microbial communities remain unanswered. Therefore, it is important that we work towards improving our understanding of plant and soil microbial communities. Moreover, it is appropriate for plant and soil microbial communities to be studied in concert since plants and soil interact with one another, and intricate feedbacks between plants, soil, and microorganisms can drive the biogeography of plant and microbial communities (Wardle et al. 2004).

In Chapter II of this thesis, I present an investigation of the biogeography of bacterial communities within individual trees. Despite the growing appreciation of the importance of bacteria for the health of plants, the spatial distribution of bacterial communities within individual plants is not well resolved. This work sought to address this knowledge gap by examining leaf and bark bacterial communities from >300 samples within individual ginkgo trees (Fig. 1.1). This investigation demonstrated that there can be predictable patterns of bacterial community structure within individual trees. The findings have implications for the way plant-associated bacterial communities are studied since they show that samples collected from one region of a plant are not necessarily representative of

Figure 1.1. Photograph of ginkgo leaves from one of the trees sampled in Chapter II.

other regions. Furthermore, it establishes a baseline set of results to guide future hypotheses regarding the assembly of plant associated communities. For example, bacterial communities in the exterior leaves and bark of the tree had higher relative abundances of *Proteobacteria* while the interior tissues harbored higher relative abundances of *Acidobacteria*. These patterns are likely caused by differences in environmental conditions in different regions of the tree and differences in the life history strategies of the dominant members of these bacterial phyla, but future research is needed to corroborate this speculation.

It is possible that there are certain factors controlling the assembly of microbial communities across closely related plant genotypes, and in Chapter III, I present a study where we investigated whether the selective breeding associated with domestication is one of those factors. For this study, we

Figure 1.2. Photograph of a sunflower plant (*Helianthus annuus* L.) from the common garden experiment described in Chapter III.

grew >30 sunflower (*Helianthus annuus* L.) strains (Fig. 1.2) in a common environment and assessed the relationship between root and rhizosphere fungal and bacterial community structure and domestication level. Fungal community composition was associated with the domestication of its host while bacterial community composition was not, suggesting that fungal communities are more strongly affected by domestication than bacterial communities. Moreover, these results suggest that different portions of plant-associated microbial communities can be assembled in different ways and that a nuanced approach is necessary for understanding and predicting microbial linkages with different plant strains.

The study presented in Chapter IV addresses the research question: Can plant community attributes, including composition, phylogeny, and traits, be used to predict spatial variability in soil community composition? In an English grassland study system (Fig. 1.3), I found that plant species

Figure 1.3. An old-process gum bichromate print created from a digital photograph of the grassland featured in Chapter IV.

identity and differences in plant community composition were predictive of soil fungal, bacterial, protistan, and metazoan communities. However, differences in plant phylogeny and traits were not strong predictors of soil communities. These results show that different plant species affect soil communities differently. This information can be used to shape our understanding of the determinants of soil biogeography.

Chapter V focuses on a study investigating the effect of elevated nutrient inputs on soil fungal, bacterial, and archaeal communities. While previous work has identified effects of elevated nitrogen (N) inputs on soil microbial communities (e.g., Ramirez et al. 2012), this study sought to determine whether there were consistent effects across grasslands worldwide and to investigate the impact of elevated phosphorus inputs on soil microbial communities. The chapter describes the observed consistent effects in detail and links them to shifts in the ecology and function of the belowground communities as well as shifts in plant community composition. Given the vast soil microbial diversity described above, it is striking that there are consistencies in the responses of these communities in such disparate locations. This knowledge can be applied to enhance predictions of shifts in soil communities with altered levels of nutrient inputs, which are predicted to change due to human activity. In addition, since nutrient inputs fluctuate through time and space under natural conditions, this information will improve our understanding of the biogeography of these communities.

Taken together, this thesis identifies and describes patterns that exist in the biogeography of plant and soil microbial communities. It demonstrates that plant and soil microbial biogeography can be explained by specific factors and is therefore predictable under certain circumstances and with particular information. For example, bacterial communities are, to some degree, predictable based on their location within a ginkgo tree, and sunflower fungal communities can be predicted based on their host's level of domestication. However, the results of this thesis also demonstrate that plant and soil microbial communities are still difficult to predict accurately under other circumstances. For example,

information on the level of domestication of sunflowers does little to improve our ability to predict bacterial communities. Thus, it is unclear what factors, if any, drive variation in bacterial community composition across sunflower strains. Similarly, Chapter IV shows that variation in soil community composition is driven, in part, by differences across plant species. However, while valuable, this understanding will not enable us to predict variation in soil communities without a priori knowledge of plant-microbe associations. Unfortunately, knowledge of plant phylogeny and commonly-measured traits do not appear to be predictive of soil communities. Therefore, future research must continue to investigate the factors determining the biogeography of plant and soil microbial communities in order improve our ability to predict their composition and function across space and time.

CHAPTER II

SPATIAL STRUCTURING OF BACTERIAL COMMUNITIES WITHIN INDIVIDUAL *GINKGO BILOBA* TREES

(Leff, J.W., Del Tredici, P., Friedman, W.E. & Fierer, N. (2014). Spatial structuring of bacterial communities within individual *Ginkgo biloba* trees. *Environ. Microbiol.*)

Abstract

Plant-associated microorganisms affect the health of their hosts in diverse ways, yet the distribution of these organisms within individual plants remains poorly understood. To address this knowledge gap, we assessed the spatial variability in bacterial community diversity and composition found on and in aboveground tissues of individual *Ginkgo biloba* trees. We sampled bacterial communities from > 100 locations per tree, including leaf, branch, and trunk samples, and used high-throughput sequencing of the 16S rRNA gene to determine the diversity and composition of these communities. Bacterial community structure differed strongly between bark and leaf samples, with bark samples harboring much greater bacterial diversity and a community composition distinct from leaves. Within sample types, we observed clear spatial patterns in bacterial diversity and community composition that corresponded to the samples' proximity to the exterior of the tree. The composition of the bacterial communities found on trees is highly variable, but this variability is predictable and dependent on sampling location. Moreover, this work highlights the importance of carefully considering plant spatial structure when characterizing the microbial communities associated with plants and their impacts on plant hosts.

Introduction

Large numbers of bacteria live inside plant tissues and on plant surfaces (Lindow and Brandl 2003). These microbes have many potential effects on their host's growth and survival through a wide array of mechanisms (Lindow and Leveau 2002, Gnanamanickam 2006, Bulgarelli et al. 2013). Still, we have a limited understanding of the diversity and spatial distribution of bacterial communities associated with plants, especially those taxa living in and on aboveground plant tissues. This knowledge gap persists in part because plants harbor a large diversity of bacteria that cannot be readily cultivated, and therefore, many of the bacterial taxa associated with plants have not been captured in the long history of culture-based surveys (Yang et al. 2001, Yashiro et al. 2011). Recent work on plant-associated bacteria has yielded a more comprehensive understanding of the diversity of bacterial communities and their spatial distributions while focusing on herbaceous model plants, such as *Arabidopsis thaliana,* using culture-independent techniques (Lundberg et al. 2012, Bulgarelli et al. 2012). In addition, a few studies have begun to apply similar techniques to characterize bacterial diversity associated with other plant species, including long-lived woody species (Redford et al. 2010). Together, this work has highlighted that the microbial diversity associated with plants is vast and variable, and that the identities and functional attributes of many of these taxa remain unknown (Chelius and Triplett 2001, Rastogi et al. 2013).

Since plants do not represent a homogeneous microbial habitat, understanding the spatial heterogeneity in microbial distributions across different anatomical tissues and morphological regions of plants is important for building a more complete picture of the bacterial communities associated with plants. For example, the biotic and abiotic characteristics of leaves differ from those of stems. Moreover, branch surfaces change over time from first-year epidermal primary tissues to phellogen-derived bark, that itself will vary significantly over the course of many years. Likewise, shade and sun leaves represent distinct microbial environments even though they exist in close proximity.

Heterogeneous environmental conditions in tree canopies have been shown to affect the distribution of arboreal animals (Stork and Hammond 2001, Scheffers et al. 2013). Thus, the spatial distribution of arboreal microorganisms within a single canopy is also likely to covary with environmental characteristics. At the scale of an individual microbe, trees may represent heterogeneous environments where colonization, community assembly, and succession patterns may shape microbial community structure.

A more detailed understanding of the spatial patterns exhibited by plant-associated bacteria will not only provide insight into the factors shaping the diversity and composition of plant-associated bacterial communities; such studies can also yield important information on the natural history of these microorganisms, establishing a foundation for using plant-associated microbial communities as systems for testing ecological theories (Meyer and Leveau 2012, Peñuelas and Terradas 2014). Nutrient availability, for instance, may represent a particularly important broad-scale determinant of microbial communities across individual plants since there are ecological trade-offs between traits that contribute to an affinity for high nutrient conditions (copiotrophy) and low nutrient conditions (oligotrophy; Andrews and Harris 2000). Similarly, other environmental conditions, including ultraviolet (UV) radiation and water availability, are likely important drivers of variation in microbial community composition both within and between plants (Lindow and Brandl 2003, Whipps et al. 2008, Vorholt 2012). Taken together, these environmental factors should lead to predictable spatial patterns in microbial community structure within and between individual plants. Correspondingly, there is some evidence that this appears to be the case. For example, studies using culture-based techniques have linked microbial community differences in plants to microbial adaptations to specific environments (Kinkel 1997, Andrews and Harris 2000, Hirano and Upper 2000). Likewise, culture-independent work has demonstrated that plant interspecific variation is an important driver of differences in phyllosphere bacterial communities, patterns that are likely driven by a wide range of biotic and abiotic factors that

differ among plant species (Baily et al. 2006, Redford et al. 2010, Kim et al. 2012, Kembel et al. 2014). One recent study found that different tree species living in close proximity harbored distinct communities while different individuals of the same plant species growing thousands of kilometers apart harbored strikingly similar communities (Redford et al. 2010). However, the majority of these studies have focused on bacterial community differences between plant species or between individual plants of the same species (Redford et al. 2010, Hunter et al. 2010, Finkel et al. 2012, Rastogi et al. 2013), with few studies examining the spatial variability in bacterial communities across different tissue types and locations within individual plants (Ottesen et al. 2013, Lambais et al. 2014). As a result, we do not know how the diversity and composition of bacterial communities vary as a function of plant anatomy, position within the architecture of the organism, and ontogeny (e.g., one year old bark vs. 10 year old bark; young leaf surface vs. older leaf surface), and whether there are predictable spatial patterns driven by heterogeneity in environmental conditions within plants. Given that plant-associated bacteria can have myriad impacts on plant health, documenting the spatial variability in these bacterial communities is critical to improving our understanding of plant-microbe interactions.

We analyzed the bacterial communities found in about 100 individual aboveground locations, including trunk, branch, and leaf-associated communities, on each of three *Ginkgo biloba* trees using high-throughput DNA sequencing of 16S rRNA genes. We focused on *Ginkgo biloba*, since it is an exceedingly common tree in urban environments (Crane 2013), it has a relatively simple morphology and architecture that made it feasible to design a sampling scheme that was consistent across replicate trees, and estimating the relative ages of various tree segments on *Ginkgo biloba* from its growth scars and its branching patterns is relatively straightforward. We hypothesized that bacterial community structure would exhibit predictable spatial heterogeneity across tissue types and across different locations within individual tissue types.

Results

Variation in bacterial diversity within individuals

In total, we collected 314 aboveground samples from the three trees in the Arnold Arboretum in Boston, MA, USA (42.297°N, 71.129°W; Fig. A2.1 and A2.2) and assessed microbial community diversity and composition in the samples using amplicon sequencing of a region of the 16S rRNA gene (see Methods). Bacterial sequences were clustered into groups that share ≥97% sequence identity and are hereafter referred to as phylotypes. We observed 2,345 unique bacterial phylotypes on average per tree at the rarefied sequencing depth of 575 sequences per sample. Bacterial phylotype richness was significantly different among the tissue types sampled, which included trunk bark, branch bark, new branch growth, and leaves ($P < 0.001$; Fig. 2.1A). This was also true for Shannon diversity, phylotype evenness, and phylogenetic diversity, which displayed similar patterns (Fig. A2.3), and therefore, we use the term diversity to refer to phylotype richness hereafter. Trunk samples had the greatest diversity (156 unique phylotypes per sample on average), and new branch growth and leaves had the lowest diversity (54 and 71, respectively), reflecting generally lower diversity on new plant tissue compared to branch and trunk tissue >1 year old (Fig. 2.1A).

Bacterial diversity also varied spatially within tissue types. We categorized samples according to the age of the tissue they were collected from using bud scale scars resulting from annual terminal buds. Older branch and lower trunk surfaces and the undersides of branches tended to harbor more diverse bacterial communities than young branch and trunk surfaces further from the ground (Fig. 2.1B). Leaf-associated bacterial communities exhibited a similar pattern with leaves growing from older branch segments generally harboring more diverse communities than leaves growing from younger branch segments (Fig. 2.1B).

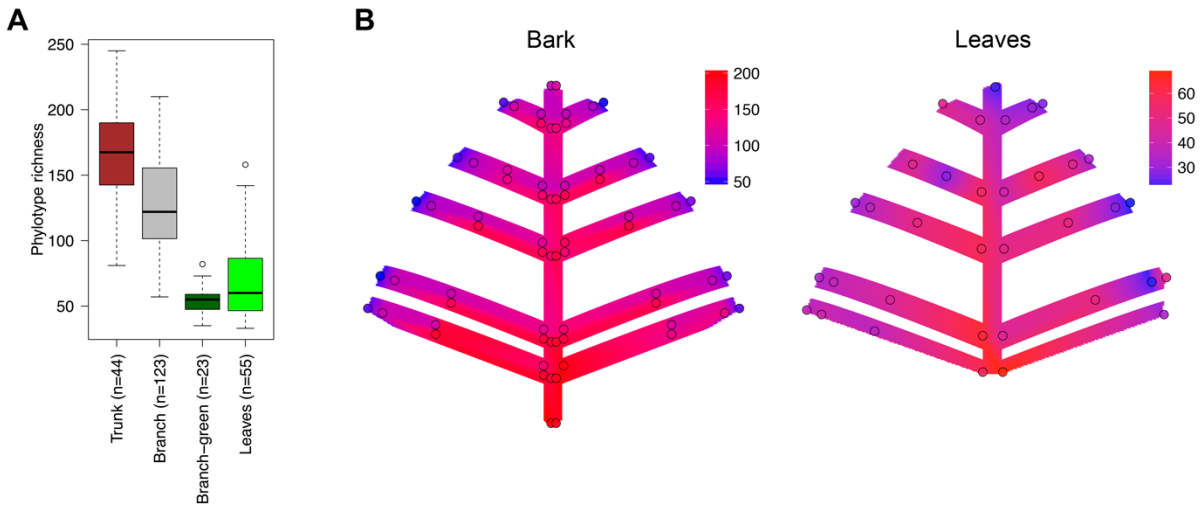

Figure 2.1. (A) Boxplots showing the distribution of phylotype richness across the sampled ginkgo tissue types. Richness was estimated following rarefying to 575 sequences per sample. (B) Spatial variation of bacterial phylotype richness across the trees for bark and leaf samples. Shading is based on linear interpolation between mean sample values (represented by circles) with red indicating greater diversity and blue indicating lower diversity. Graphical representation of sample locations were adjusted to accommodate slight variations among the three replicate trees prior to computing mean richness values for a given location across the three trees. Branches on left and right extend approximately south and north, respectively. Results based on samples rarefied to 575 and 200 sequences per sample for bark and leaves, respectively.

Variation in bacterial community composition within individuals

Across all samples, the bacterial taxa observed were largely members of the Acidobacteria, Actinobacteria, Bacteroidetes, and Proteobacteria phyla (Fig. 2.2A). However, bacterial community composition was significantly different across the samples taken from the different organ and tissue types ($P = 0.001$; Fig. 2.3). Additionally, pairwise comparisons revealed that each sample type had a significantly distinct community composition from one another ($P < 0.02$ in all cases). Differences among the three individual trees also contributed to variation among samples ($P = 0.001$). However, the differences between the trees for a given tissue type were considerably less than the differences between tissue types (Fig. A2.4).

Even at a coarse level of taxonomic resolution, it is apparent that bacterial community composition varied across the different organ and tissue types (Fig. 2.2A). For example, Acidobacteria had a median relative abundance of 13% on trunk and branch samples > 1 year old but were < 1% on

B

Phylum	Median relative abundances: trunk	branch	branch-green	leaves
Acidobacteria*	0.13	0.13	0.00	0.01
Actinobacteria*	0.06	0.03	0.02	0.03
Armatimonadetes*	0.01	0.01	0.00	0.00
Bacteroidetes*	0.38	0.31	0.19	0.06
Firmicutes*	0.00	0.00	0.00	0.01
Proteobacteria*	0.36	0.42	0.80	0.82
WYO*	0.02	0.01	0.00	0.00

B

Phylum	Family	Median relative abundances: trunk	branch	branch-green	leaves
Acidobacteria	Acidobacteriaceae*	0.13	0.13	0.00	0.00
Actinobacteria	Microbacteriaceae*	0.01	0.01	0.00	0.00
Actinobacteria	Nocardioidaceae*	0.01	0.01	0.00	0.00
Actinobacteria	Pseudonocardiaceae*	0.01	0.00	0.00	0.00
Bacteroidetes	Flexibacteraceae*	0.27	0.24	0.19	0.05
Bacteroidetes	Sphingobacteriaceae*	0.05	0.03	0.00	0.00
Proteobacteria	Acetobacteraceae*	0.05	0.06	0.00	0.02
Proteobacteria	Beijerinckiaceae*	0.02	0.03	0.00	0.01
Proteobacteria	Caulobacteraceae*	0.01	0.01	0.00	0.02
Proteobacteria	Comamonadaceae*	0.00	0.00	0.01	0.02
Proteobacteria	Methylobacteriaceae	0.03	0.02	0.02	0.02
Proteobacteria	Oxalobacteraceae*	0.00	0.01	0.59	0.07
Proteobacteria	Sphingomonadaceae*	0.08	0.09	0.04	0.05
Proteobacteria	Unclass. Alphaproteobacteria*	0.01	0.01	0.00	0.00
Proteobacteria	Unclass. Rhizobiales*	0.04	0.06	0.02	0.04
Proteobacteria	Unclass. Rhizobiales*	0.06	0.06	0.01	0.02
WYO	Unclassified*	0.02	0.01	0.00	0.00

Figure 2.2. Heat maps showing the median relative abundances (proportions) of those taxa represented by ≥ 1% of the sequence reads in any sample type. Phylum level taxa (A), and family level taxa (B) that were significantly different in relative abundances across sample types (Bonferroni-corrected $P \leq 0.001$) are indicated with an asterisk. Colors are scaled to the range of values within each row.

leaf samples and green branch samples. Proteobacteria exhibited the opposite pattern with a median relative abundance of 38% and 31% on trunk and branch samples, respectively, with mean relative abundances of approximately 80% on the leaf and green branch samples (Fig. 2.2A). Relative abundances of Actinobacteria, Armatimonadetes (formerly OP10), Bacteroidetes, Firmicutes, and the WYO candidate phylum were also significantly different across the sample types (Fig. 2.2A). The organ and tissue types also significantly differed in their community compositions when viewed at the family

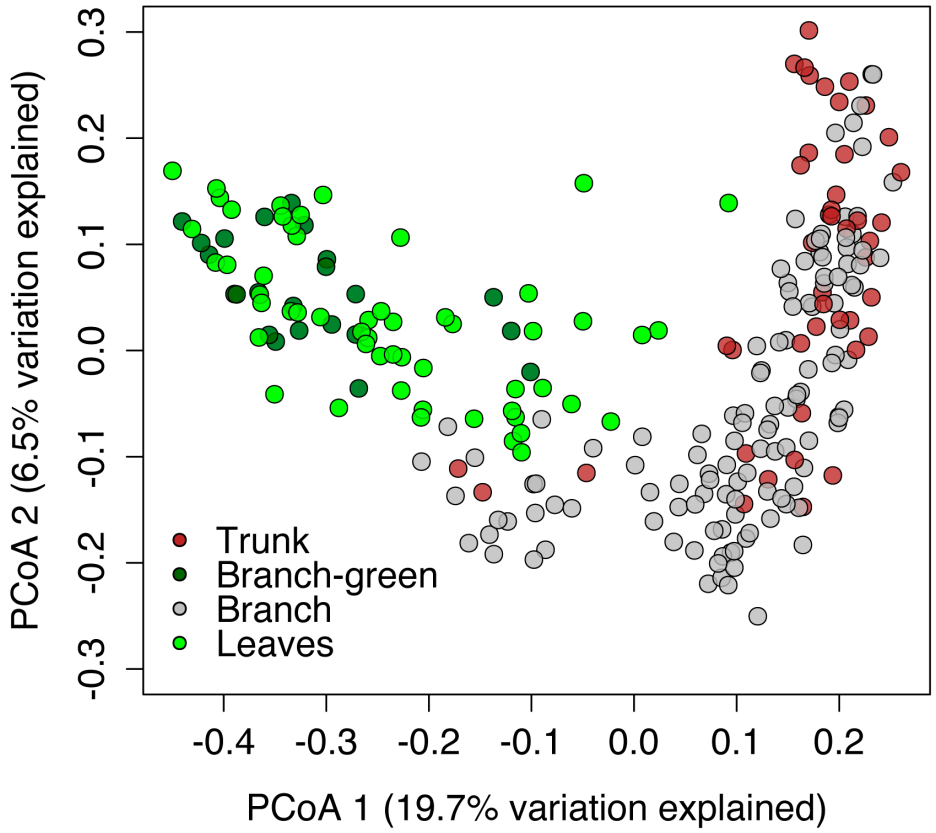

Figure 2.3. Principal coordinates analysis plot showing the relative similarities of samples taken from different tissue types. This ordination was created from unweighted UniFrac distances.

level; younger plant tissues (leaves and green branches) were largely dominated by members of the Flexibacteraceae, Oxalobacteraceae, and Sphingomonadaceae, whereas trunk and older branch communities were composed of several distinct family-level groups (Fig. 2.2B). The individual bacterial phylotypes with the greatest relative abundances across the sample types were identified as *Hymenobacter* sp. and one unclassified member of the Rhizobiales on the older branch and trunk samples, while an unclassified Oxalobacteraceae phylotype dominated on the green branches and leaves (Table A2.1).

To visualize spatial patterns in bacterial community composition for a given tissue type, we first investigated patterns within individual trees as the sampling locations differed slightly from tree to tree. These visualizations revealed that the general spatial patterns were similar across the trees (Fig. A2.5), a

finding confirmed by the statistical analyses described above. Therefore, in order to simplify our analyses and visualizations, we averaged sample values that came from similar locations across the three individual trees that were sampled. These sampling locations included five vertical branch levels and seven vertical trunk levels that were shared across all trees. Similarly, four distances from the trunk were sampled along each branch (in most cases), which corresponded to four distinct branch ages as younger portions of the branch occur further from the tree. We found that the undersides of branches and lower parts of the trunk had bacterial communities that were relatively similar in composition and distinct from those found on the tops of branches and the upper parts of the trunk (Fig. A2.6). These differences were apparent when we compared the relative abundances of the four dominant phyla in these samples. Acidobacteria and Proteobacteria were relatively more abundant on the tops of branches and the upper trunk and lower in abundance on the undersides of branches and the lower trunk with Actinobacteria exhibiting the opposite pattern. The relative abundance of Bacteriodetes was typically highest on branches of an intermediate age (Fig. 2.4).

Leaf-associated bacterial communities also exhibited predictable spatial structuring with leaves growing from younger branch segments harboring communities that were distinct from those found on older branch segments in the interior of the tree (Fig. A2.7). The Acidobacteria were relatively more abundant on the north inner leaves with Gammaproteobacteria being relatively more abundant on the outer leaves (Fig. A2.8).

Discussion

The *Ginkgo* tissues harbored diverse bacterial communities containing representatives of seven bacterial phyla and seventeen families, confirming previous reports that plant leaves can harbor a wide range of microbial taxa (Whipps et al. 2008, Redford et al. 2010, Vorholt 2012). In addition, we found that this bacterial diversity was not evenly or randomly distributed across tissue and organ types (Fig. 2.1A). In particular, the trunk and branch bark (i.e. periderm) samples harbored more bacterial diversity

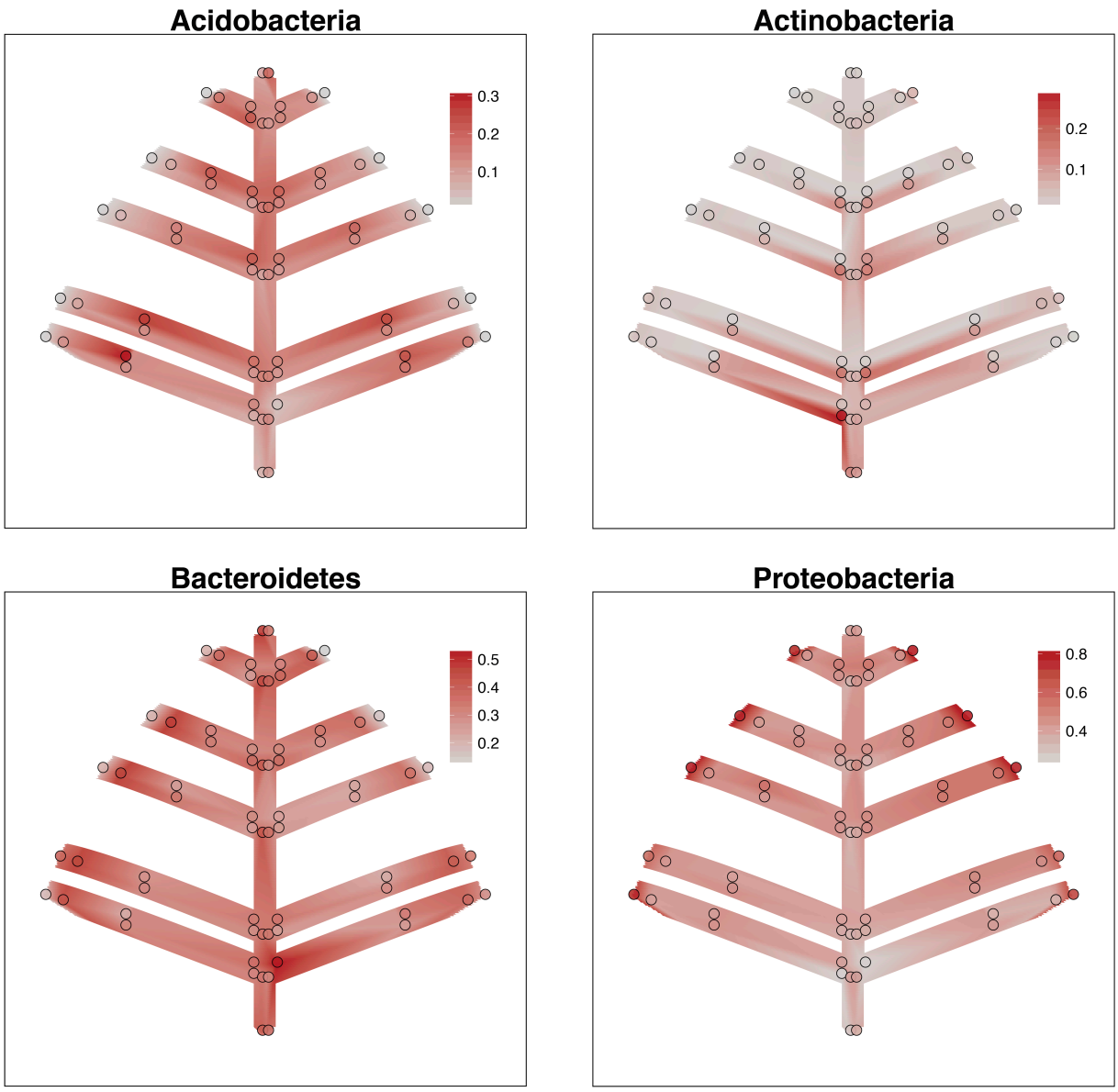

Figure 2.4. Schematic of tree branch and trunk sampling sites with the shading indicating the relative abundances of the four dominant phyla on those surfaces. Branches on left and right extend approximately south and north, respectively.

than younger tissues such as leaves and new lateral stem segments, which potentially reflects broad differences between dead and living tissue. These differences in diversity were also accompanied by differences in the bacterial community composition among tissue and organ types (Fig. 2.2 and Fig. 2.3).

Few studies have directly compared leaf microbial community structure to that of bark or characterized differences in bacterial community structure across various plant organs within the same species. However, one recent study showed differences in microbial communities across various organs within the tomato plant (Ottesen et al. 2013), and another demonstrated differences in bacterial community structure across various grape vine tissues including bark and leaves (Martins et al. 2013). Additionally, differences between rhizosphere and endosphere or root and aboveground bacterial communities have been reported (Garbeva et al. 2001, Idris et al. 2004, Gottel et al. 2011, Lundberg et al. 2012, Bulgarelli et al. 2012, Bodenhausen et al. 2013). Our finding that the diversity and composition of these bacterial communities found on trunk and branch surfaces were different across leaf and young branch tissues corroborates these previous findings and extends them to a tree species, supporting our hypothesis that different tree tissue and organ types support microbial communities that differ in their structure.

The leaf-associated bacterial taxa we observed were similar to those described by previous culture-based and culture-independent work, in that these communities were largely comprised of members of the Proteobacteria, Bacteroidetes, and Actinobacteria phyla (Vorholt 2012, Bulgarelli et al. 2013, Turner et al. 2013). At the family level, the common leaf-associated taxa we observed were generally similar to those found in previous studies. For example, Sphinogomonadaceae, Methylobacteriaceae, and Flexibacteraceae were relatively abundant across the leaves as has been observed previously with other plant species (Delmotte and Knief 2009). We also observed a high relative abundance of Oxalobacteraceae, which were not particularly abundant in a previous survey of other tree leaf communities (Redford et al. 2010), but may be more abundant on individual species of plants as they have been associated with the leaves of lettuce (Rastogi et al. 2012, Leff and Fierer 2013). The trunk and older branch surfaces had high relative abundances of taxa belonging to the phylum Acidobacteria, a phylum that was relatively rare on the leaves. Since there are few cultured

representatives in the Acidobacteria phylum, and there are few other studies using culture-independent techniques to study bark surface bacteria, it is difficult to compare this result with previous work. For example, Martins *et al*. (2013) did not observe members of this phylum on grape vine bark using a culture-based survey. Acidobacteria are dominant members of soil microbial communities (Lauber et al. 2009), and these arboreal microbes may have arrived on the tree via dispersal from the soil beneath the tree.

Alternatively, Acidobacteria may be more specialized inhabitants of bark surfaces. They are thought to be largely comprised of slow growing and oligotrophic taxa (Ward et al. 2009), which could mean they are adapted to inhabit older plant tissues such as bark rather than young tissues such as leaves or may take sufficiently long to colonize their plant substrate as to be detectable only on older surfaces. Besides the Acidobacteria, the trunk and branch samples contained members classified as belonging to the phylum Armatimonadetes and the candidate division, WYO. Representatives of Armatimonadetes [formerly candidate division OP10 (Tamaki et al. 2011)] have been previously observed in plant-associated communities (Chelius and Triplett 2001), but generally, members of these lineages are quite rare, and this finding suggests that tree bark can harbor many novel bacterial taxa that are far less abundant on more commonly studied plant tissues. The observed community differences between bark and leaf tissues could have been partially due to differences in the sampling methods (swabbing versus whole-leaf extractions). However, we highlight multiple reasons it is unlikely that the sampling methods alone account for the large differences in community composition observed here: (1) Green branch and leaf tissue tended to have more similar communities (both in terms of diversity and community composition) than green branch tissue had with the other bark samples that were sampled in an identical manner, (2) The taxa dominating the bark communities were quite rare in the leaf communities, and (3) The quantities of tissue sampled for the DNA extractions were similar for the bark swabs and leaf samples.

In addition to differences across plant tissue types, we observed spatial structuring of bacterial diversity and community composition within tissue types. Overall, there were similar patterns with leaf and bark samples, but the patterns were stronger on bark surfaces than on leaves. Across leaf and bark samples, there was generally a greater diversity in samples from the inner, less exposed parts of the tree when compared to the more distal (outer) parts (Fig. 2.1B). Likewise, there were similar patterns with differences in community composition (Fig. A2.6). These patterns were expected given that bacterial richness has been tied to differences in environmental conditions such as moisture (Turner et al. 2013). Although we did not measure moisture availability within the trees we sampled, there are likely to be differences in humidity and exposure to precipitation between the samples taken from the central and distal portions of the tree where we saw the strongest differences in phylotype richness. Other factors such as the age of the tissue or the disturbance regime (including disturbances from UV radiation and precipitation) may have also contributed to this pattern. In addition to the observed differences in community structure between interior and exterior portions of the tree, we observed distinct differences in structure between the upper and lower sides of branches (Fig. 2.1B; Fig. A2.6). This pattern could be due to similar factors as the interior/exterior differences since the upper sides of the branches receive greater precipitation and UV radiation than the lower sides, which could lower diversity. However, the upper portions of the branches were less similar to the exterior portions of the tree with regard to their community composition (Fig. A2.6B), suggesting that characteristics of these tissues support bacterial taxa specific to their individual environments.

Differences in the relative abundance of Acidobacteria partially drove the overall differences in community composition within the bark samples. Members of this phylum tended to have greater relative abundances on interior portions of the tree and lower relative abundances on new branch growth (Fig. 2.4), a pattern that could be a product of the life history strategies of these taxa as discussed above, with these slower-growing taxa preferring older tissue types. In addition, Acidobacteria

tended to be more abundant on northern interior leaves when compared to leaves found on other locations within a tree, which could suggest that greater UV radiation or lower moisture on southern leaves inhibits the growth of Acidobacteria taxa.

Proteobacteria generally displayed an opposite pattern to the Acidobacteria as Proteobacteria were relatively more abundant on the new branch growth. This pattern may be due to the tendency of many Proteobacteria to grow fast and proliferate in environments with fewer existing microorganisms or in environments where organic carbon resources are more available (Fierer et al. 2007a). The relative abundances of the other dominant phyla also showed spatial patterning, highlighting that there are often predictable patterns in the spatial distributions of a wide array of microbial taxa associated with different locations of a given tissue type. These patterns could have been driven by a number of mechanisms, which have been proposed in other studies to explain the occurrence of specific taxa in the phyllosphere: UV radiation, moisture, nutrient availability, and the amounts and types of available organic C (Andrews and Harris 2000, Lindow and Brandl 2003, Vorholt 2012) as well as other leaf characteristics or microbe-microbe interactions (Hunter et al. 2010).

Along with previous work demonstrating variation in phyllosphere communities over time (Redford and Fierer 2009), between plant species (Redford et al. 2010, Kembel et al. 2014), and with the geographic location of plants (Finkel et al. 2012), this study demonstrates that plant-associated microbial communities also exhibit a high degree of variation within individual plants. For example, the differences we observed in the relative abundances of the dominant phyla across different locations within individual ginkgo trees were on par with the differences in phylum-level abundances observed on leaves from diverse tree species (Redford et al. 2010). Our results also suggest that the distribution of these diverse bacterial taxa across individual trees is predictable, with community composition differing between plant tissues and exhibiting strong spatial patterns within individual trees that relate to their anatomical structure. Thus, just as single forehead skin sample would not be adequate to characterize

the bacterial communities found on and in an individual human (Costello et al. 2009), a single sample from a tree does not adequately characterize that entire tree's bacterial community. This predictable spatial variation should be considered when studying plant-microbe relationships and trying to understand the impacts of these plant-associated microbes on plant health.

Methods

Sample collection

Bacterial communities were sampled from three ginkgo (*Ginkgo biloba*) trees growing in a ~500 m^2 area at the Arnold Arboretum in Boston, MA, USA (42.297°N, 71.129°W; Fig. A2.1 and A2.2). All trees were raised from cuttings collected in 1989 from separate individual trees growing wild on Tian Mu Mountain in Zhejiang Province, China. At the time of sampling, tree #1 (AA 1073-89-B, a female tree) was 7.3 m tall with an average spread of 5.4 m; tree #2 (AA 1223-89-A, a female tree) was 9.7 m tall with an average spread of 7.0 m; and tree #3 (AA 1072-89-F, male tree) was 9.2 m tall with an average spread of 6.2 m. All three trees were growing with supplemental irrigation and fertilization in the same general area and experienced the same growing conditions. All samples were collected between June 26 and 28, 2012 just after a short period of light rain. Four types of samples were collected: trunk (bark tissue), branches (> 1 year old; bark tissue), new branch growth (< 1 year old epidermal primary tissue; green branches with intact epidermis), and leaves. Trunk and branch bacterial communities were collected by swabbing an approximately 5 cm^2 area on the surfaces with sterile swabs. Bacterial communities from leaves were collected by directly sampling four whole leaves collected in sterile plastic bags due to the inability to recover sufficient microbial biomass by swabbing. In all cases, sampling was done aseptically while wearing nitrile gloves. Bacterial communities were sampled at sites in relation to branches at five vertical levels on each tree and separately on north and south sides of the trunk. When possible, samples were collected from branches at four distances away from the main trunk corresponding to distinct ages of the branch segments, and the tops and bottoms of branches were sampled separately

on the older segments. The ages of the branch segments were calculated by identifying compact areas of bud scale scars left in the bark from annual terminal buds. The leaves from the youngest branch segments were associated with long shoots, while other leaves were associated with the short shoots characteristic of the ginkgo morphology in older branch segments.

Leaf samples were collected from each of the sampled branch segments. Swabs and leaves were stored at -20°C prior to molecular analysis. Leaf samples were homogenized prior to DNA extraction by freezing at -80°C and crushing them in the bags. DNA was extracted from swabs and the homogenized leaf material (each containing a similar amount and less than 50 mg of sample material) using the PowerPlant Pro kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA). While only surface-associated bacteria were extracted from the trunk and branch samples, the leaf communities include those bacteria found on the leaf surfaces and those found inside the individual leaves.

Determination of bacterial community composition and diversity

To assess bacterial diversity and community composition in each of the 314 samples collected, a region of the 16S rRNA gene was amplified and analyzed via barcoded sequencing following a procedure described in Redford *et al.* (2010). The DNA was PCR amplified in triplicate using a primer pair (799f/1115r), which was designed to avoid amplification of chloroplast DNA (Chelius and Triplett, 2001; Redford et al., 2010), and reduced the percentage of chloroplast sequences to minimal levels in this study (Table A2.2). To enable simultaneous sequencing of all samples, each sample was amplified with a primer set containing a unique 12-bp barcode. Following amplification, triplicate reactions were combined and visualized on an agarose gel along with positive and negative controls. Amplicons were combined in equimolar ratios, cleaned using the UltraClean PCR Clean-Up Kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA), and sequenced at the Engencore facility at the University of South Carolina on the Roche 454 sequencing platform.

Raw 16S rRNA gene sequences were processed using the QIIME v1.6.0 pipeline (Caporaso et al. 2010) in order to perform initial steps of bacterial community characterization. These steps included: demultiplexing, quality filtering, phylotype clustering, taxonomy assignments, sequence alignments, and computing pair-wise community dissimilarities. Default parameters were used except for the following: only sequences between 270 and 370 bp were retained to remove poor quality sequences as the expected length was ~316 bp, both forward and reverse primers were trimmed when contained in the sequence, phylotype clustering was performed using the open reference (reference-based + *de novo*) implementation of UCLUST (Edgar 2010) and relied on the Greengenes October 2012 16S rRNA database clustered at 97% similarity (McDonald et al. 2012a), and taxonomic identities were assigned to phylotypes using the RDP classifier (Wang et al. 2007b) trained on the aforementioned Greengenes database with a confidence threshold of 0.5. As an additional quality control measure, phylotypes classified as mitochondria or chloroplasts and those that could not be assigned to a specific phylum were removed prior to further analyses. Following initial processing, the sequence data were rarefied to 575 sequences per sample to account for the variable sequencing depth obtained (Table A2.2), which resulted in 191 of the 201 collected trunk and branch samples having sufficient sequencing depth to be retained for downstream analyses. For spatial (within sample type) analyses of leaf-associated bacterial communities, samples were rarefied to 200 sequences per sample in order to retain more samples (with 89 of the 113 collected leaf samples included in downstream analyses). Diversity metrics and unweighted UniFrac distances (Lozupone et al. 2011a) were calculated within QIIME. Although we could not assess the full extent of bacterial diversity with this sequencing depth, previous work demonstrates that this sequence depth is sufficient for accurate assessments of patterns in bacterial diversity and community composition on leaf surfaces (Redford et al. 2010) and in other microbial habitats (Kuczynski et al. 2010). Amplicon sequences were deposited in the EMBL-EBI European Nucleotide Archive (http://www.ebi.ac.uk/ena/) and can be accessed using the accession number, ERP005154.

Statistical analyses and visualizations

Differences in diversity among sample types were assessed using boxplots and ANOVA, while differences in community composition were assessed using principal coordinate analysis and permutational multivariate ANOVA (PERMANOVA). In ANOVA and PERMANOVA models, sample type was included as a fixed factor and tree individual was included as a random effect. For pairwise comparisons among sample types, post-hoc tests were used within the PERMANOVA function. Two-dimensional linear interpolation was used to visualize spatial patterns across an individual model tree using the 'akima' package in R (R Core Team 2013). Significant differences in the relative abundances of individual bacterial taxa across sample types were determined using Kruskal-Wallis tests and Bonferroni corrections. Univariate analyses and principal coordinate analysis were performed using R (R Core Team 2013), and PERMANOVA was performed using PRIMER 6 (Clarke and Gorley 2006).

CHAPTER III

PLANT DOMESTICATION AND THE ASSEMBLY OF BACTERIAL AND FUNGAL COMMUNITIES ASSOCIATED WITH STRAINS OF THE COMMON SUNFLOWER, *HELIANTHUS ANNUUS* L.

(Leff, J.W., Lynch, R.C., Kane, N.C. & Fierer, N. (Accepted). Plant domestication and the assembly of bacterial and fungal communities associated with strains of the common sunflower, *Helianthus annuus* L. *New Phytol.*)

Abstract

Root and rhizosphere microbial communities can affect plant health, but it remains undetermined how plant domestication may influence these bacterial and fungal communities. We grew 33 sunflower (*Helianthus annuus* L.) strains (*n* = 5) that varied in their extent of domestication and assessed rhizosphere and root endosphere bacterial and fungal communities. We also assessed fungal communities in the sunflower seeds to investigate the degree to which root and rhizosphere communities were influenced by vertical transmission of the microbiome through seeds. Neither root nor rhizosphere bacterial communities were affected by the extent of sunflower domestication, but domestication did affect the composition of rhizosphere fungal communities. In particular, more modern sunflower strains had lower relative abundances of putative fungal pathogens. Seed-associated fungal communities strongly differed across strains, but several lines of evidence suggest that there is minimal vertical transmission of fungi from seeds to the adult plants. Our results indicate that plant-associated fungal communities are more strongly influenced by host genetic factors and plant breeding than bacterial communities, a finding that could influence strategies for optimizing microbial communities to improve crop yields.

Introduction

Root and rhizosphere microbial communities play key roles in determining plant health and productivity (Berendsen et al. 2012, Chaparro et al. 2012, Bulgarelli et al. 2013, Verbon and Liberman 2016), yet our understanding of how these communities are assembled remains rudimentary. For example, although different strains of *Arabidopsis* are known to harbor distinct root microbiomes (Lundberg et al. 2012, Bulgarelli et al. 2012), we currently lack a predictive understanding of the biotic and abiotic factors responsible for these observed differences. A better understanding of root and rhizosphere microbial community assembly will improve our ability to predict the structure of plant-associated microbial communities and their effects on plant health. Ultimately, such knowledge could be used to directly or indirectly manipulate these microbial communities to enhance the health and productivity of agricultural crops (Berg 2009, Compant et al. 2010, Singh et al. 2011, Turner et al. 2013, Bender et al. 2016).

Previous work has demonstrated that soil properties are the dominant factors structuring root and rhizosphere microbial communities (Bulgarelli et al. 2012, Philippot et al. 2013). However, other factors such as plant species identity or genotype can also have measurable influence on their composition (Hardoim et al. 2011, Bulgarelli et al. 2012, Philippot et al. 2013, Coleman-Derr et al. 2015). Therefore, a critical next step is to determine how and when differences in plant genotypes matter for root and rhizosphere microbial community assembly (Lareen et al. 2016). For example, root and rhizosphere bacterial communities have been shown to differ across related strains of *Arabidopsis* (Lundberg et al. 2012, Schlaeppi et al. 2014), maize (Peiffer et al. 2013), and rice (Hardoim et al. 2011, Edwards et al. 2015), but the specific nature of these interactions and the factors driving these host genotype-microbial associations remain largely undetermined.

In agriculture, knowing the factors that favor the assembly of beneficial bacterial and fungal associations with crop plants could be leveraged to enhance crop yields given the potential importance

of these belowground microbial associations in mediating nutrient acquisition, environmental tolerances, and disease resistance (Rodriguez et al. 2008, Mei and Flinn 2010, Farrar et al. 2014). In particular, we need to know whether crop-associated microbial communities have shifted due to plant domestication and whether potentially beneficial interactions between plants and their microbial symbionts have been lost during the domestication process. If so, it could be possible to modify agricultural practices to account for this loss and improve crop yields. More generally, domestication of plants has resulted in a number of changes to their interactions with other organisms and their effects on agroecosystems – processes that are still poorly understood (García-Palacios et al. 2013, Milla et al. 2015, Turcotte et al. 2015). Likewise, there has been some speculation that the selective breeding involved with domestication and the conditions in which domesticated plants are typically grown can influence their microbial communities (Pérez-Jaramillo et al. 2015, Pieterse et al. 2016, Schmidt et al. 2016). For instance, differences in the necessity for stronger stress tolerance or the ability to grow under systematic disturbances such as plowing could promote differences between microbial communities associated with wild and domesticated plants. Previous work suggests that differences in rhizosphere bacterial communities associated with maize might be linked to domestication (Bouffaud et al. 2014). Other studies have proposed that modern crops do not support beneficial microbes in their rhizosphere as readily as their wild ancestors due to modification of plant traits (Philippot et al. 2013). For example, mutually beneficial associations between arbuscular mycorrhizal fungi (AMF) and wild relatives of crops could be less prevalent in modern crops since those crops might be less dependent on AMF in an agricultural setting (Sawers et al. 2008). However, there is mixed empirical evidence for whether AMF more commonly form associations with wild plants compared to their domesticated relatives (Zhu et al. 2001, An et al. 2010, Lehmann et al. 2012, Xing et al. 2012, Leiser et al. 2016, Turrini et al. 2016). More broadly, relationships between crop evolutionary history and associated microbial communities are not always easily detectable (Bouffaud et al. 2012), or the effects of domestication are sufficiently subtle

that they can be difficult to quantify in highly diverse belowground microbial communities (Bulgarelli et al. 2015).

Differences in the microbes contained in seeds and the passage of these microbes to offspring (*i.e.* vertical transmission) is one possible mechanism that could promote genotype-specific differences in plant-associated microbial communities (Nelson 2004, Truyens et al. 2015). Plants could be under selective pressure to package microbes in seeds to ensure progeny are able to form associations with their most beneficial microbes, thus promoting plant genotype-specific seed communities (Ewald 1987, Rudgers et al. 2009, Truyens et al. 2015). Previous work has shown that seeds could serve as an important vector for the transmission of microbes from one generation of plant to another for certain species (Kaga et al. 2009, Johnston-Monje and Raizada 2011, Hardoim et al. 2012, Cope-Selby et al. 2016, Pitzschke 2016). Still, it is unclear whether vertical transmission of microbes through seeds can contribute to differences in microbial community composition across plant varieties and whether such vertical transmission can influence plant performance. Alternatively, given that the types of microbes found in soil appear to have a strong influence on what types of microbes associate with plant roots (Bulgarelli et al. 2012), any potential influence of vertical transmission via seeds could be negligible in ultimately structuring the belowground associations adult plants form with microbes and have little effect on plant development and growth.

Here we sought to determine whether the structure of bacterial and fungal communities in roots and rhizosphere was predictable based on host genotype across 33 sunflower (*Helianthus annuus*) strains grown in the same soil type. We chose sunflower as a model species since it is a globally important crop, and we were able to obtain seeds from strains which spanned a wide gradient of wild, landrace (*i.e.* early domesticated strains), and modern domesticated cultivars. This enabled us to investigate whether domestication through selective breeding affected microbial community structure. We hypothesized that root and rhizosphere microbial community structure differs in consistent ways

across sunflower strains and that these differences are related to the extent of plant domestication (*i.e.* whether the sunflower strain was classified as a wild, landrace, or modern strain). In addition, we characterized the bacterial and fungal communities in seeds from the same batches used to grow the sunflowers to assess the extent to which bacteria and fungi were transmitted from the seeds to the adult plant root and/or rhizosphere communities. Given that previous work has shown evidence in favor of vertical transmission in other plant species, we hypothesized that vertical transmission of seed-associated microbial communities may contribute to the observed strain-specific differences in the microbial communities associated with adult plants.

Methods

Plant material used

A total of 33 common sunflower (*Helianthus annuus* L.) strains were selected to span the full range of genetic variation in the species (Harter et al. 2004) including 11 wild populations from across the native range in North America, 14 Native American landraces representing the diversity of pre-modern cultivated lineages that occurred following a single domestication event, approximately 4,000 – 5,000 years ago (Harter et al. 2004, Blackman et al. 2011, Kane et al. 2013, Smith 2014), and 8 modern domestic varieties that are the product of the last few hundred years of more intensive breeding efforts (Table S1). This sample set thus allows us to evaluate the role of the original domestication event in North America, the result of thousands of years of selection by Native Americans in what is now the south-eastern US, as well as the more recent, scientific breeding for modern agriculture. All seeds were obtained from the USDA National Plant Germplasm System (https://npgsweb.ars-grin.gov/gringlobal/; see Table S1 for seed accession information).

Plant growth and sample collection

Seeds of each strain (n = 5 per strain) were started in sterile petri dishes containing moist paper towels and transferred to potting soil (type) once germinated. After two weeks, seedlings were

transplanted to a 350 m^2 field outside Boulder, Colorado, USA (40°02'24"N 105°07'48"W). Plants were watered as needed. The soil in the field was a Mollisol classified as a Manter sandy loam. This soil type is characterized as being well drained and having a circumneutral acidity (http://websoilsurvey.nrcs.usda.gov/). Individuals were planted in random locations across the field since small variations in soil and environmental conditions that might have existed in the field. There was minimal pre-existing plant cover in the experimental plot, and small weeds were removed prior to planting.

Plants were grown for 53 days and harvested when all plants were expected to be at or near peak height. Immediately prior to harvest, plant height, stem diameter, most recent fully expanded leaf length and width, the number of nodes, and the number of branches were recorded for each individual. Plants were manually uprooted from the soil and roots were aggressively shaken in order to remove loose soil. The ends of multiple representative roots were cut from each plant and transferred to sterile 50 mL conical tubes (filling approximately half the volume of the tube). The samples in tubes were immediately transferred to the laboratory on ice. To remove rhizosphere soil, 10 mL of DNA-free water was added to each tube and vortexed for 10 seconds, with the rhizosphere soil collected after the slurry was allowed to settle for 24 hours and the supernatant was decanted, following the general approach described previously (minus the centrifugation step; Lundberg *et al.*, 2012). The washed roots were then transferred to new tubes and the roots were further cleaned by adding 10 mL of DNA-free water to tubes, vortexing for 10 seconds, pouring out the water, and repeating. Subsamples (0.2 g) of the cleaned roots were then transferred to 1.7 mL tubes where they were processed with 100% ethanol, rinsed with water and then treated with propidium monoazide (PMA) as in (Nocker et al. 2007) to remove superficial, dead bacterial and fungal cells. Roots were macerated in their tubes with sterile pestles prior to DNA extraction.

Seeds that were from the same batch used to grow the plants were included in our bacterial and fungal community analyses (n = 4 per strain). Seeds were prepared by soaking the seeds in DNA-free water for 24 h, briefly submerging in 95% ethanol, and rinsing with water. This procedure was intended to soften seeds and remove superficial microbial cells, but microbial cells integrated in the seed coat were purposefully retained since they could influence the adult plant's microbial community. Each seed was then macerated separately with a sterile and DNA-free glass mortar and pestle.

Bacterial and fungal community analysis

Subsamples (100 μl) of rhizosphere slurries, macerated roots, and macerated seeds were transferred to 96-well plates for DNA extraction by mixing 150 μl DNA-free water with each sample and transferring 50 μl to an individual well. DNA extraction was performed using the PowerPlant Pro DNA Isolation Kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA) according to the manufacturer's instructions, with appropriate negative control ('blanks') included in all steps of the process. The V4 region of the 16S rRNA gene and the ribosomal internal transcribed spacer 1 (ITS) were PCR amplified to assess bacterial and fungal diversity, respectively. PCR protocols followed those used previously (McGuire et al. 2013, Lundberg et al. 2013) and included primers with barcodes unique to each sample to permit sample multiplexing. The primer set with linkers and adapters used for 16S rRNA gene amplification was 515-F (AATGATACGGCGACCACCGAG ACGTACGTACG GT GTGCCAGCMGCCGCGGTAA) and 806-R (CAAGCAGAAGACGGCATACGAGAT XXXXXXXXXXXX AGTCAGTCAG CC GGACTACHVGGGTWTCTAAT), where "X" characters represent the 12-bp barcodes (Caporaso et al. 2011, Fierer et al. 2012b). The primer set with linkers and adapters used for fungal ITS amplification was ITS1-F (AATGATACGGCGACCACCGAGATCTACAC GG CTTGGTCATTTAGAGGAAGTAA) and ITS2 (CAAGCAGAAGACGGCATACGAGAT XXXXXXXXXXXX AGTCAGTCAG AT GCTGCGTTCTTCATCGATGC; White *et al.*, 1990; McGuire *et al.*, 2013; Smith & Peay, 2014). Peptide nucleic acid PCR clamps were used when targeting 16S rRNA genes to inhibit the amplification of chloroplast and mitochondria genes as

described in (Lundberg et al. 2013). PCR products from triplicate reactions per sample were cleaned and pooled in equimolar concentrations using the SequalPrep kit (Thermo Fisher Scientific, Waltham, MA, USA). The cleaned amplicons were sequenced in three runs (two for bacteria and one for fungi) on an Illumina MiSeq instrument at the University of Colorado BioFrontiers Institute Next Generation Sequencing Facility using a paired-end 2x151 bp kit for the two bacterial sequencing runs and a 2x251 bp kit for the fungal sequencing run. All raw sequence data have been deposited in the National Center for Biotechnology Information Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under the accession number, SRP075934.

Initial sequence processing was conducted similarly to (Prober et al. 2015) following the UPARSE pipeline implemented in USEARCH v.8 (Edgar 2013). Briefly, sequencing adapters were removed from fungal ITS sequences using cutadapt (Martin 2011), both bacterial and fungal sequences were assigned to individual samples (*i.e.* demultiplexed), and a *de novo* database was created by merging paired-end reads, quality filtering, dereplicating, removing unique (*i.e.* singleton) sequences, and clustering sequences into phylotypes at the 97% similarity threshold. Representative sequences from those phylotypes that were not ≥ 75% similar to any sequence in the Greengenes or UNITE databases were removed as they were assumed to be low quality, chimeric, or a product of non-specific amplification. Raw merged reads were then mapped to the *de novo* database in order to determine the number of sequences representing each phylotype for each sample. Taxonomy was determined for each phylotype using the RDP classifier (Wang et al. 2007a) trained on the Greengenes (McDonald et al. 2012b) and UNITE (Abarenkov et al. 2010) databases for bacterial and fungal sequences. 16S rRNA sequences from chloroplasts, mitochondria, or archaea were removed prior to downstream processing as were bacterial or fungal sequences that were not classified to at least the phylum level of resolution. Due to the high relative abundance of chloroplasts and sparsity of robust information on the seed bacterial communities, we did not use those samples for downstream analyses investigating the strain-specific

differences in plant-associated bacterial communities. Bacterial data were rarefied to 2,000 sequences per sample, and fungal data were rarefied to 1,000 sequences per sample prior to all downstream analyses, rarefaction depths that were chosen to balance the number of samples that could be included while maximizing the available number of sequences per sample. FunGuild (Nguyen et al. 2015) was used to assign fungal phylotypes from the rarefied data to one of three trophic modes (saprotroph, symbiont, or pathogen) when possible.

Statistical analyses

We used the Shannon diversity metric to quantify bacterial and fungal diversity. Diversity across sample types was compared using linear mixed effect models with sunflower strain as a random factor. Overall differences in bacterial or fungal community composition were assessed by calculating pair-wise Bray-Curtis dissimilarities from square-root transformed phylotype relative abundances. Differences in community composition across sample types were tested using permutational analysis of variance (PERMANOVA), and pairwise comparisons between sample types were tested by comparing two factor levels at a time using PERMANOVA and adjusting the resulting p-values for multiple comparisons with false discovery rate corrections. This was done using the function, 'calc_pairwise_permanovas' in mctoolsr (Leff 2016), which implements the 'adonis' function in the vegan package (Oksanen et al. 2016) in R (R Core Team 2016). Differences in community composition across sunflower strains and across domestication levels were also tested using PERMANOVA. We verified that the data met the assumption of multivariate homogeneity of dispersions using the 'betadisper' function in the vegan package prior to running these tests. When testing for differences in community composition across domestication levels, dissimilarities were first averaged across replicates of each sunflower strain to avoid pseudoreplication. Differences in the relative abundances of fungal pathogens among domestication levels were compared using a linear mixed effects model with sample type and domestication level as fixed effects and sunflower strain as a random effect. Analysis of variance (ANOVA) was used to test for

differences in the proportion of samples in each strain that had detectable fungal symbionts across domestication levels. Only sunflower strains with data from at least three replicate samples were used when making comparisons across strains or levels of domestication. Relationships between plant characteristics and microbial diversity were assessed with Spearman correlations, and relationships with microbial community composition were assessed with permutational Mantel tests. Differences in the relative abundance of fungal pathogens across sample types were tested using a linear mixed model with sample type as a fixed effect and plant strain as a random effect. We used R (R Core Team 2016) for all statistical analysis.

Results

Microbial communities differ across plant compartments

Across all samples, fungal and bacterial community structure differed strongly among root, rhizosphere, and seed samples, regardless of sunflower strain (Fig. 3.1). Fungal and bacterial diversity differed significantly across compartments ($P < 0.001$ in both cases), and fungal and bacterial diversity in seeds was lower than in root and rhizosphere communities (Fig. 3.1A). Root bacterial communities were significantly more diverse than rhizosphere communities ($P < 0.001$), but fungal rhizosphere and root communities had equivalent levels of diversity ($P > 0.1$; Fig. 3.1A). The higher bacterial diversity in root endosphere compared to rhizosphere samples was related to the dominance of *Pseudomonas* in the rhizosphere samples.

For both bacteria and fungi, the rhizosphere, root, and seed communities were each significantly different in composition from one another ($P < 0.001$ in all pairwise comparisons; Fig. 3.1B). Root and rhizosphere bacterial and fungal communities were generally more closely related to one another than to seed communities. Seeds tended to have high relative abundances of the bacterial families, *Nocardiopsaceae, Enterobacteriaceae,* and *Sphingomonadaceae* compared to root and rhizosphere samples, which harbored high relative abundances of a number of families not commonly

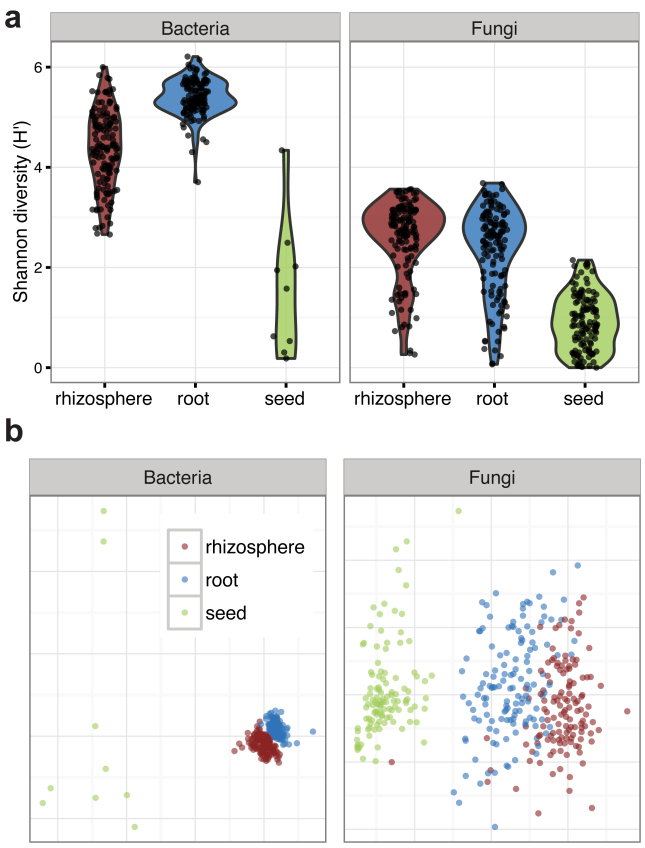

Figure 3.1. Shannon diversity distributions (a) and nonmetric multidimensional scaling ordinations showing differences in community composition (b) of bacterial and fungal communities across rhizosphere, root, and seed samples. The ordinations were based on Bray-Curtis dissimilarities calculated from square-root transformed data.

observed in seeds (Fig. A3.1). The composition of fungal communities in seeds were distinguished from those in roots and rhizospheres due to high relative abundances of *Pleosporaceae*, while root and rhizosphere communities had high relative abundances of *Nectriaceae*, *Olpidiaceae*, and *Mortierellaceae* (Fig. A3.2).

Differences in microbial communities across sunflower strains

Neither bacterial nor fungal community diversity significantly differed across the root or rhizosphere communities from different sunflower strains ($P > 0.05$ in all cases). Bacterial root and rhizosphere and fungal root community composition also did not significantly differ across the strains ($P > 0.05$). However, fungal rhizosphere community composition did differ across the strains, albeit

somewhat weakly (R^2 = 0.25, *P* = 0.01). No single sunflower strain appeared to drive the overall differences; instead the level of dissimilarity in fungal community composition was similar across all sunflower strains (Fig. A3.3).

Differences in microbial communities across levels of domestication

Neither bacterial nor fungal diversity differed across domestication levels in rhizospheres or roots (*P* > 0.5). Domestication level did significantly affect fungal rhizosphere community composition (R^2 = 0.08, *P* = 0.03). Within rhizospheres, unclassified *Pleosporales*, *Preussia* spp., unclassified *Thelebolaceae*, *Fusarium* spp., and *Conocybe* spp. tended to have higher relative abundances in modern strains, while unclassified *Chaetomiaceae* and *Mortierella* spp. had higher relative abundances in wild strains, and the *Chytridiomycota* genus *Olpidium* had higher relative abundances in Native American and wild strains than in modern strains (Fig. A3.4).

In addition to individual taxa, putative fungal pathogens had a lower relative abundance in the root and rhizosphere microbial communities of modern strains when compared to wild or native American strains (*P* = 0.04; Fig. 3.2A). Conversely, fungal symbionts (primarily *Glomeromycota*; Fig. A3.5) were only detected in a subset of the samples, and while not statistically significant, they were detected in a greater proportion of the roots of modern strains compared to wild strains (*P* > 0.1; Fig. 3.2B).

Are differences in microbial communities explained by plant attributes?

More domesticated strains tended to be taller, have broader leaves, wider stems and fewer branches after growing for the same length of time (Fig. 3.3). However, differences in root and rhizosphere bacterial and fungal communities were not strongly related to differences in plant attributes (Table A3.2 & A3.3). The diversity of rhizosphere bacterial communities and root fungal communities was weakly and inversely related to measured plant characteristics indicative of growth rate (larger plants tended to have lower bacterial and fungal diversity; Table A3.2). Neither bacterial nor fungal

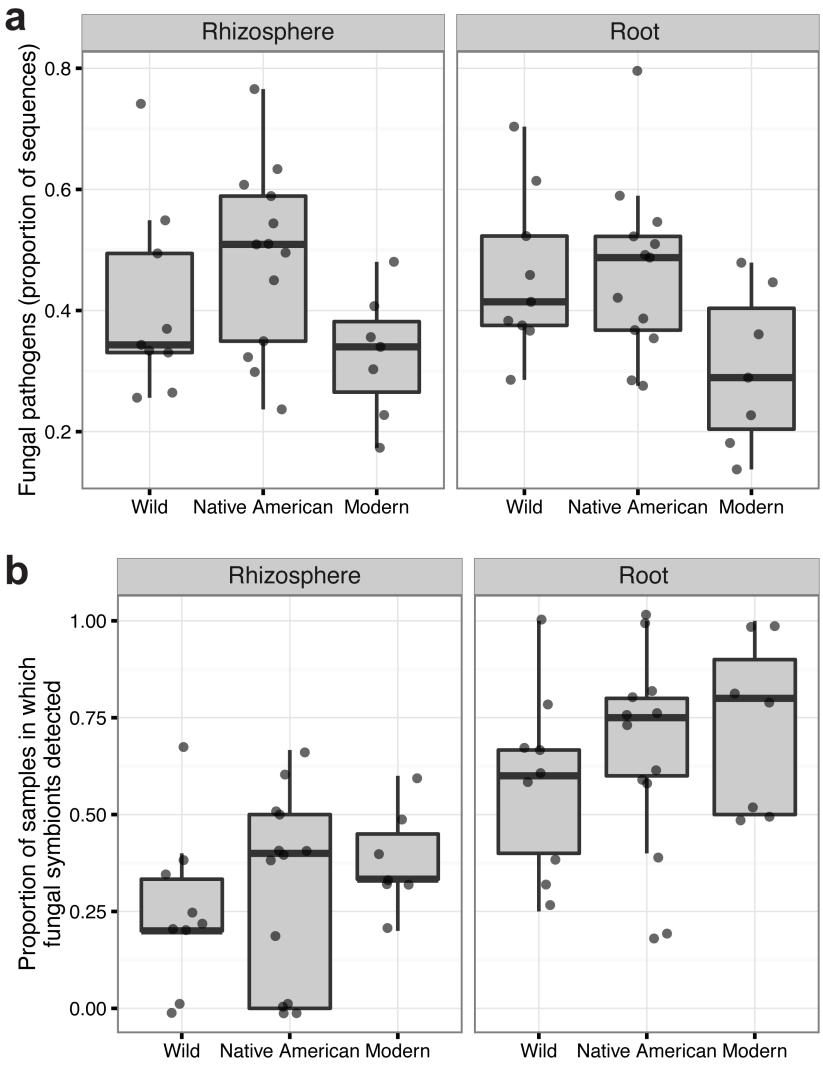

Figure 3.2. The relative abundance of putative fungal pathogens (a) and the presence of putative fungal symbionts (b) in root and rhizosphere communities across wild, Native American, and modern sunflower (*Helianthus annuus* L.) strains. Each point represents the mean value for each strain, and individual points are presented over boxplots.

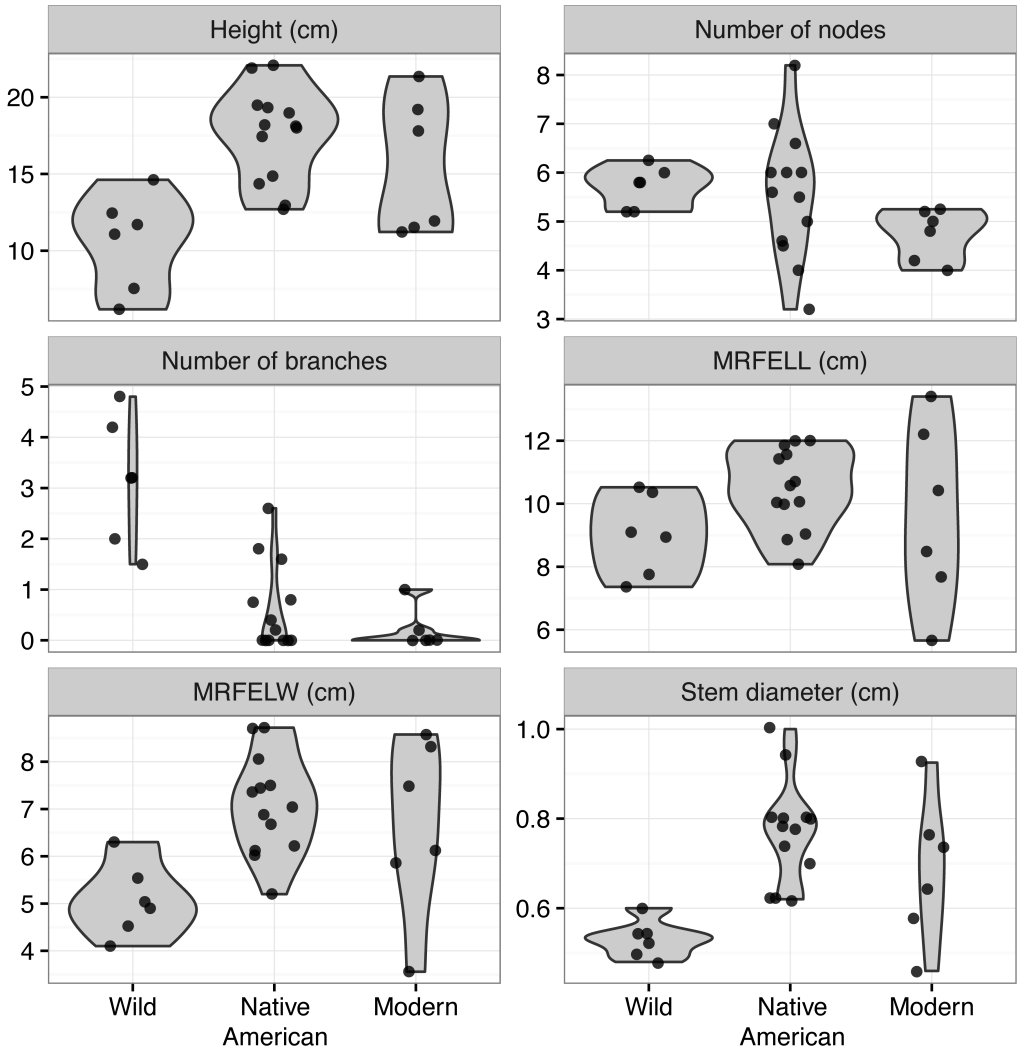

Figure 3.3. Violin plots showing differences in phenotypes across wild, Native American, and modern sunflower (*Helianthus annuus* L.) strains at the time of sample collection. Each overlying point represents the mean value for sunflower strain. MRFELL = most recent fully expanded leaf length. MRFELW = most recent fully expanded leaf width.

community composition were strongly related to any of the measured phenotypic characteristics (Table A3.3).

Seed fungal communities and their relationship with root and rhizosphere communities

The various sunflower strains harbored different seed-associated fungal communities (Fig. 3.4). On average, 7% of the fungal phylotypes in a given root and 5% of the phylotypes in a given rhizosphere sample were also observed in a given seed. These proportions were the same (7% and 5% for root and

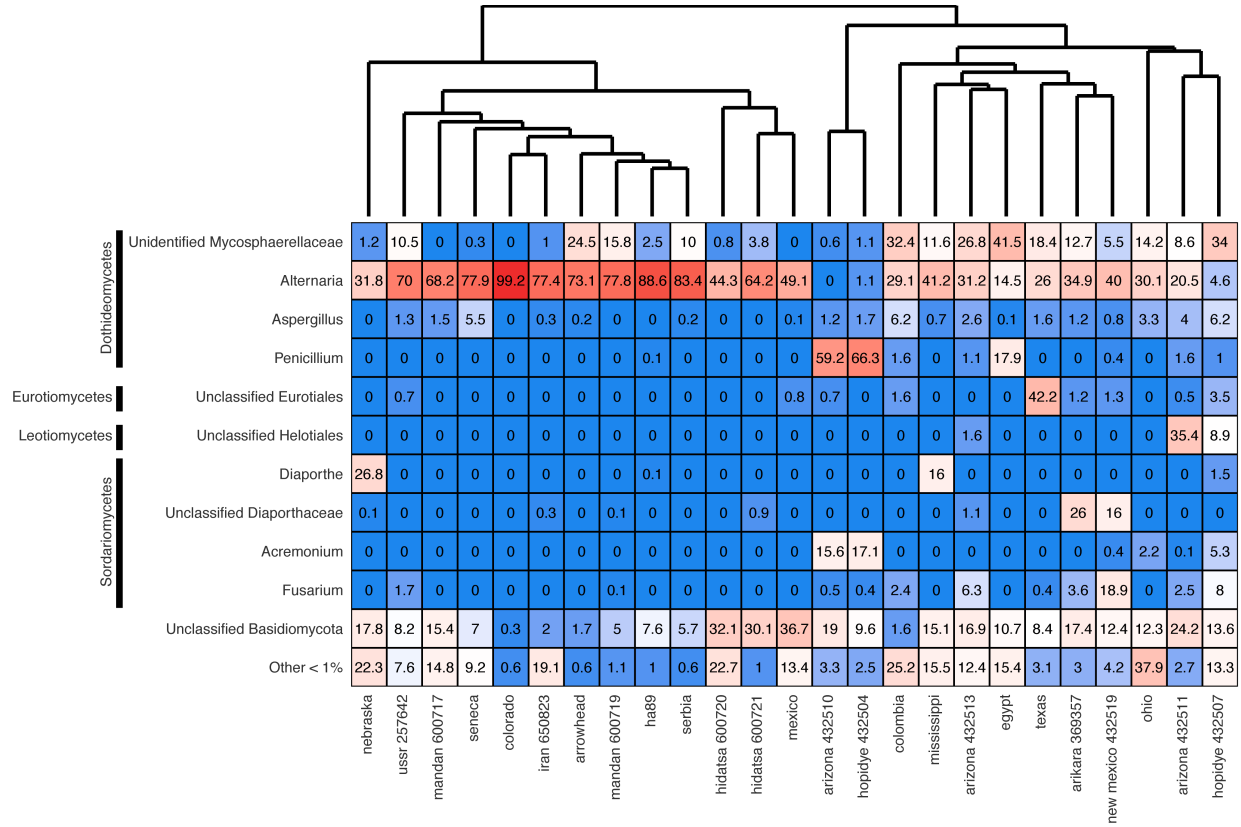

Figure 3.4. Composition of fungal communities in seeds across sunflower (*Helianthus annuus* L.) strains. The heat map represents the relative abundance of genera ≥1%. Values represent mean relative abundances within individual strains (%), and colors indicate lower relative abundances (blue) and higher relative abundances (red). The cluster diagram shows clustering of sunflower strains by fungal community composition based on Bray-Curtis dissimilarities of square-root transformed phylotypes relative abundances. See Table A3.1 for details on all the sunflower strains shown in this plot.

rhizosphere samples, respectively) whether calculated within individual strains or across all strains. There was no significant relationship between pairwise dissimilarities across seed communities and pairwise dissimilarities across root or rhizosphere communities ($P > 0.4$ in both cases).

Most seed, root, and rhizosphere communities contained relatively few putatively symbiotic taxa, but we found that roots (63%) and rhizospheres (37%) had the largest proportion of samples that contained detectable fungal symbionts (primarily members of the phylum, *Glomeromycota*; Fig. A3.5) while seeds had the lowest proportion (3.3%; Fig. 3.5A). Although seeds rarely contained detectable levels of known symbiotic fungi, 55% of fungal sequences from seeds were from putative pathogens. In

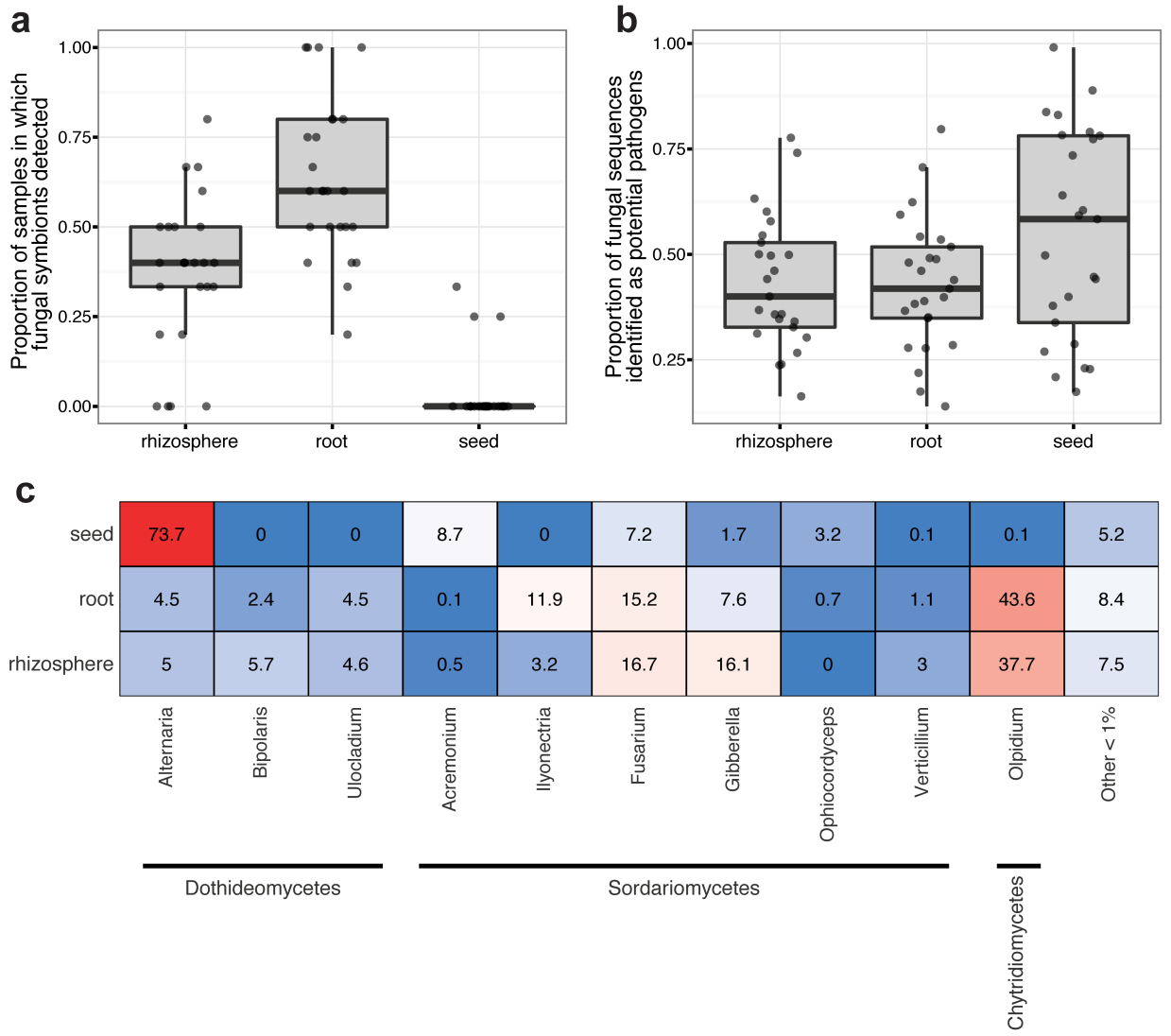

Figure 3.5. The presence of putative fungal symbionts (a) and mean relative abundance of putative fungal pathogens (b) across rhizosphere, root, and seed samples. Each point represents the value for each sunflower (*Helianthus annuus* L.) strain, and individual points are presented over boxplots. Genus-level differences in putative pathogen fungal community composition across rhizosphere, root, and seed samples are shown as a heat map (c). Values represent mean relative abundances within individual strains (%), and colors indicate lower relative abundances (blue) and higher relative abundances (red).

comparison, 43% of fungal sequences from both root and rhizosphere samples were from putative pathogens, a significant difference (P = 0.005; Fig. 3.5B). There were multiple taxa identified as being potentially pathogenic, including *Alternaria* spp. and *Acremonium* spp., that had high relative abundances in seeds and were nearly absent in roots and rhizospheres (Fig. 3.5C). Relative abundances

of putative pathogens in seeds also did not correspond to relative abundances in roots (P = 0.3), and they had a weak inverse relationship with relative abundances in rhizospheres (r = -0.46, P = 0.02).

Discussion

The root and rhizosphere bacterial and fungal communities were more similar to one another than to seed communities (Fig. 3.1B), which suggests that many root endophytes are derived from the rhizosphere, a finding in line with results from other plant species (Rosenblueth and Martínez-Romero 2006, Bulgarelli et al. 2012, Bai et al. 2015). Likewise, our results corroborate previous work showing that taxa found in rhizospheres (Philippot et al. 2013) tend to be distinct from those found in seeds (Truyens et al. 2015). Since the different plant compartments clearly have distinct bacterial and fungal communities, we investigated the factors influencing community assembly separately for each of these habitats.

One of the central goals of this study was to assess whether root and rhizosphere microbial communities differ in consistent ways across closely related strains of *H. annuus*. While previous work has shown that different plant hosts can harbor distinct rhizosphere and root bacterial communities (Wieland et al. 2001, Ofek et al. 2014), the magnitude of genotype effects on bacterial communities within individual host plant species are typically quite small (Inceoglu et al. 2010, Weinert et al. 2011, Lundberg et al. 2012, Peiffer et al. 2013, Marques et al. 2014, Schlaeppi et al. 2014, Wagner et al. 2016). Fewer studies have investigated the effect of host genotype on overall fungal communities in rhizosphere or on root endophyte microbial communities (Bacilio-Jiménez et al. 2003). Our observation that bacterial communities were not structured by sunflower genotype does not necessarily conflict with previous studies that have shown subtle genotype effects on bacterial communities in other plant species since it is possible that belowground bacterial communities respond more strongly to differences in certain host traits [*e.g.*, starch content (Marques et al. 2014)] that may not differ across the sunflower strains examined here. It is noteworthy that we found a significant relationship between

sunflower strain identity and fungal community composition since this suggests that fungal taxa are more sensitive to host traits and are more likely to exhibit strain-specificity than bacteria (Cassman et al. 2016). At a more basic level, the significant association between host genotype and fungal community suggests an important effect of host genotype on the rhizosphere community.

Although there was generally a weak effect of sunflower strain identity on the microbial communities found in roots and rhizospheres, we investigated whether there were broad effects of domestication on these communities across all strains. Our results indicated that the degree of domestication has little effect on overall bacterial communities in the rhizosphere and root, but there are potentially important effects of domestication on fungal communities. Domesticated crops likely interact with rhizosphere microbial communities in different ways than their wild counterparts (Wissuwa et al. 2009, Pérez-Jaramillo et al. 2015). The fact that the sunflower rhizosphere fungal community composition was related to the level of domestication could be driven by indirect effects of domestication on soil characteristics moderated by differences in root traits. Domesticated strains have been bred to grow more quickly and develop at different rates, and they likely exude different quantities and types of organic compounds in their roots, which could have important effects on belowground microbial communities (Haichar et al. 2008, Pérez-Jaramillo et al. 2015). Differences in organic compound production across levels of domestication may be due to known trade-offs between growth rates and defense against biotic and abiotic stressors (Mayrose et al. 2011). Some of these interactions are mediated by important secondary metabolites or defense compounds such as sesquiterpene lactones (Dempewolf et al. 2008, Prasifka et al. 2015), which could stimulate or hinder growth of different fungal taxa. Additionally, the conditions under which crops have been selectively bred may have contributed to the proliferation and demise of particular crop-associated microbial taxa (Wissuwa et al. 2009). Few other studies have directly assessed effects of domestication on microbial communities, but there is some evidence that bacterial communities from the roots and rhizosphere of landrace wheat

cultivars are more diverse than their modern counterparts (Germida and Siciliano 2001). Likewise, Szoboszlay *et al.* (2015) demonstrated that rhizosphere microbial community structure varies between a domesticated maize cultivar and more ancestral relatives. More generally, our results demonstrate that the effects of domestication on plant-fungal associations are in line with the numerous described effects of domestication on plant interactions with insects (Chen and Welter 2005, Chen et al. 2015).

It has been hypothesized that the domestication of crops might affect the prevalence of pathogens and symbionts naturally occurring with those crops (Pérez-Jaramillo et al. 2015). Our results suggest that domestication of sunflowers actually decreased the prevalence of pathogens associated with the plants and might have even increased the prevalence of symbionts, but these patterns need to be investigated further as it is difficult to infer whether fungi (aside from *Glomeromycota*) are symbiotic or pathogenic from taxonomy alone (Nguyen et al. 2015).

As expected, domesticated strains had phenotypic differences compared to more ancestral relatives (Fig. 3.3; Purugganan & Fuller, 2009), but these differences were largely unrelated to differences in root and rhizosphere bacterial and fungal communities. The one exception was that faster growing plants tended to have lower bacterial and fungal diversity, suggesting that plant growth rate can directly or indirectly control root and rhizosphere microbial diversity or that diverse microbial associations hinder plant growth. It is possible that slower and faster growing plants alter soil conditions in particular ways that promote more or less microbial diversity. For instance, faster growing plants could excrete compounds into the rhizosphere that promote certain taxa and thus lower diversity (Grayston et al. 1998, Oger et al. 2004, Haichar et al. 2008, Shi et al. 2011). Variation in community composition across the root and rhizosphere samples was unrelated to the measured phenotypes, and thus, the observed patterns were likely driven by other unmeasured traits that may have varied across the sunflower strains (*e.g.*, root exudates, root physiology, or nutrient concentrations). We only measured aboveground plant traits in this study, and it is possible that microbial community

composition is more strongly related to unmeasured belowground plant traits (e.g. root exudate production). Additionally, it is possible that relationships between microbial community composition and plant traits were obscured by variation in soil properties across the field. Future greenhouse experiments could help elucidate subtler relationships by more carefully controlling variation in soil and environmental factors.

We next sought to determine if seed microbial communities were important in structuring the microbial communities found in the roots and in the rhizosphere of the adult plants (*i.e.* if there was evidence for vertical transmission via seeds). We did observe strong differences in seed fungal community composition across sunflower strains, which could have been responsible for the observed differences in adult plant fungal communities. Our results are consistent with previous work showing differences in microbial communities across seeds from different plants (Barret et al. 2015, Truyens et al. 2015). However, contrary to expectations, the seed communities were not very similar to root and rhizosphere communities, and the diversity in seeds was much lower than in root and rhizosphere communities (Fig. 3.1). Moreover, only a small proportion of the root or rhizosphere fungal phylotypes were represented in the seeds. It is possible that this small fraction was derived from the seed, but it is equally likely that those phylotypes could have come from other environmental sources, such as the surrounding soil. If those seed-associated microbes served as the inocula for root and rhizosphere communities, we would expect the proportion of phylotypes observed in adult plants that were also observed in seeds to be greater within individual sunflower strains than between strains. However, we found that the proportion was very consistent (5-7%, on average across all strains) regardless of whether it was calculated within or between strains, suggesting that root and rhizosphere fungal communities are not predictable based on differences in seed communities across sunflower strains.

We also investigated whether those sunflower strains with more distinct root and rhizosphere communities also had more distinct seed communities. If seed-associated microbes served as important

inocula to root and rhizosphere communities, we would expect seeds with more dissimilar communities to also have more dissimilar root and/or rhizosphere communities. Yet, there was no significant relationship in either case, again highlighting that the strain-specific differences in seed-associated fungal communities were likely not responsible for the observed differences in root and rhizosphere communities across the sunflower strains.

Although we found minimal evidence that seed fungi contribute significantly to the assembly of root and rhizosphere fungal communities, certain fungal taxa might be transmitted from seeds to adult plants. For instance, there could be a selective pressure for plants to have symbiotic fungal taxa residing in their seeds (Ewald 1987, Rudgers et al. 2009). While there were relatively few putatively symbiotic taxa, likely due to symbiotic fungi being rare and/or not well represented in the database we used, our results suggest that symbiotic fungi are much more readily promoted in roots and rhizosphere than in seeds. In addition, our results indicate lower fungal pathogen prevalence in root and rhizosphere samples compared to seeds (Fig. 5), suggesting that the sunflower seeds may in fact be more susceptible to fungal pathogens than roots or rhizosphere and that those pathogens are filtered out of roots as the plant grows (Vaidehi et al. 2002, Afzal et al. 2010).

Given our findings and previous work showing that soil is an important inoculum governing the assembly of the plant microbiome (Turner et al. 2013, Souza et al. 2016), microorganisms external to seeds and seedlings are likely to be the principal inocula for newly developing plants. If plant-associated microbial communities are largely assembled from external sources, it may be possible to manipulate these sources and thus the plant microbiome. Together, these results suggest that this may be done by altering the environment or potentially by selecting for plant varieties with different effects on microbial communities (e.g. Panke-buisse et al. 2014). This could provide important opportunities for agricultural improvements where specific plant-associated microbial communities can increase crop yield, improve crop characteristics, and potentially decrease the reliance on irrigation and fertilizer inputs.

Conclusions

Since plants and their associated microbial communities depend on one another (Vandenkoornhuyse et al. 2015), a more comprehensive understanding of these relationships is critical for efforts to manage or manipulate the plant microbiome to improve crop yields (Farrar et al. 2014). Our results demonstrate that intraspecific differences in bacterial communities across sunflower strains spanning different levels of domestication are relatively minor but that there is likely an important effect of plant genotype on the assembly of rhizosphere fungal communities. Moreover, intrinsic differences in microbial communities across the seeds of different strains are unlikely to be important contributors to root and rhizosphere community assembly in healthy adult plants. This information helps us better understand plant-microbe relationships and could be used to improve crop yields.

CHAPTER IV

PREDICTING THE STRUCTURE OF SOIL COMMUNITIES FROM PLANT COMMUNITY ATTRIBUTES

Abstract

There are myriad ways in which soil organisms can impact plants and plants can alter the diversity and composition of belowground communities. However, it remains unclear whether we can use information on plant community attributes (their taxonomic, phylogenetic, or trait-based composition) to predict the overall structure of belowground communities. In this study, we grew 21 grassland plant species individually to assess their effects on soil communities (fungi, bacteria, protists, and metazoa), and we linked these monoculture results to the relationships between grassland plant communities and soil communities in a nearby field. We found that the composition of each soil community was affected by plant species identity regardless of whether plants were grown individually in mesocosms or when examining mixed plant communities in the field. Moreover, plant community composition was predictive of additional variation in soil community composition beyond what could be predicted from commonly measured soil characteristics. However, neither plant phylogeny nor the plant traits we measured were strong predictors of the composition of the soil communities in either case. In the field, plant community composition, whether quantified by measuring aboveground biomass directly or by sequencing plant DNA recovered from soil, was predictive of belowground community composition, but root-based plant community composition was not predictive, suggesting that plant shoot and leaves are more important in structuring soil communities than the belowground distributions of plants. Together, our results suggest that information on aboveground plant distributions can improve our ability to predict soil community composition predictions, but the strength of these associations depend on the soil community members of interest. Our results also highlight the need to identify and quantify those plant traits driving differences in soil communities in order to

generalize plant-soil community relationships across landscapes and predict plant effects on belowground communities a priori.

Introduction

Soil biological communities are a key component of terrestrial ecosystems, and they play an integral role in many ecosystem processes (Schlesinger and Bernhardt 2013, Wagg et al. 2014). While the majority of belowground taxa remain undescribed, it is increasingly evident that belowground communities, including fungal, bacterial, protistan, and faunal communities, can have important effects on ecosystem function and soil health (Waldrop et al. 2000, van der Heijden et al. 2008, Schimel and Schaeffer 2012). For example, belowground community composition can affect soil carbon (C) cycling rates (Strickland et al. 2009, Nielsen et al. 2011, Clemmensen et al. 2013), nitrogen (N) cycling dynamics (Balser and Firestone 2005), and soil fertility (Neher 2001, Jeffries et al. 2003). Thus, a central goal of soil ecology is to understand how these belowground communities vary across space. We are now able to predict differences in the composition of soil communities across large geographic scales and identify certain site and soil conditions that explain some of this spatial variation (Fierer et al. 2009, Griffiths et al. 2011, Bates et al. 2013, Tedersoo et al. 2014b, Kaiser et al. 2016). However, there is still a lot of unexplained variation in these predictions, and we often cannot accurately predict soil communities across smaller spatial scales where those environmental factors that are important drivers of soil communities (e.g. soil pH, climate) are more consistent.

Under natural settings, there is accumulating evidence that differences in overall plant community composition and soil bacterial and fungal community composition are associated at the landscape scale (de Vries et al. 2012, Barberán et al. 2015) and at the global scale (Prober et al. 2015). However, it is unclear whether these relationships are driven by shared environmental preferences or by direct species effects. Thus, we currently lack a predictive understanding of how plant species shape the composition of soil communities. The conceptual basis for plant species driving differences in soil

microbiota is well established, and certain soil organisms are known to form close associations with particular plant species (Wardle et al. 2004, Garbeva et al. 2004, Wardle 2006, Bardgett and Wardle 2010). Mycorrhizal relationships, for instance, involve a direct exchange of nutrients between plants and symbiotic soil fungi, and these relationships can influence plant-soil diversity linkages (van der Heijden et al. 1998, Hiiesalu et al. 2014). Less direct mechanisms, such as the release of root exudates and microbial attraction to those exudates, can drive associations between specific microbes and plant species (Singh et al. 2004), and plants can recruit specific microbes from soil to help combat pathogens (Berendsen et al. 2012). Due to these direct and indirect interactions, previous studies have observed associations between individual plant species and soil fungal (Peay et al. 2013, Lekberg and Waller 2016), bacterial (Kuske et al. 2002, Berg and Smalla 2009), nematode (De Deyn et al. 2004, Viketoft et al. 2005, Bezemer et al. 2010), and arthropod (St. John et al. 2006) communities. However, the effects of plant species identity of the overall composition of belowground communities are often weak or difficult to quantify, with numerous studies having failed to identify strong links between changes in plant assemblages and corresponding changes in soil communities (Porazinska et al. 2003, Bezemer et al. 2006, Tedersoo et al. 2015, Lekberg and Waller 2016). The general relationships between plants and soil communities remain uncertain or at least difficult to predict a priori.

There are multiple plant community attributes that could potentially be used to predict variation in soil communities. At the most basic level, plant species identity alone could be used to predict variation in soil communities, for reasons described above. Likewise, we might assume that the evolutionary history (i.e. the phylogeny) of plants could be associated with differences in soil communities and more closely related plants would be associated with more similar belowground communities (Barberán et al. 2015). Such patterns could arise if coevolution has taken place between plants and soil microbes or if phylogenetic relatedness corresponds to other plant attributes affecting soil organisms (De Deyn and Van Der Putten 2005, Burns et al. 2015). If plant phylogenetic information

can be used to predict differences in soil communities, we should be able to predict plant effects on belowground communities a priori for individual plant species that have not already been studied.

Given that plant species' distributions and community diversity are generally predictable based on their traits (Uriarte et al. 2010, Ben-Hur et al. 2012, Adler et al. 2013), and soil communities can form associations with plants based on these traits (Wardle et al. 2004), plant traits are a potentially useful lens with which to investigate plant-microbe associations. Plant species are prone to tradeoffs in their physiologies (Grime 1977, McGill et al. 2006, Reich 2014), which affect life history strategies and traits such as leaf carbon quality and growth rates. These traits could influence soil communities since the members likely have life history strategies of their own and thus different affinities for specific soil conditions such as C availability (Fierer et al. 2007b). There has been some work suggesting that plant traits are linked to belowground microbial processes, including both C and N dynamics (Meier and Bowman 2008, Orwin et al. 2010), suggesting that soil community composition would also be directly or indirectly influenced by plant traits. Additionally, one study found a relationship between plant traits and soil microbial communities in grasslands at the landscape scale (de Vries et al. 2012). Still, it is unknown whether variation in belowground communities is directly caused by, or merely associated with, differences in plant traits, and links between plant traits and the composition of soil communities are not always observed (Barberán et al. 2015).

Here we sought to address the overarching question: Can plant community information be used to predict spatial variability in soil community composition? To address this question, we sampled soils from both a field and a mesocosm experiment in a grassland in northern England, using DNA sequencing-based approaches to target soil fungal, bacterial, protistan, and metazoan (faunal) communities. We first assessed whether different plant species' identities, phylogenetic history, and/or traits drove differences in soil communities using replicated mesocosms, where each of 21 individual plant species were grown in isolation. Next, we augmented field plots with plants from the site to

produce mixed plant communities that spanned a compositional gradient. We used soil samples collected from these sites to assess whether differences in soil communities could be predicted based on plant community attributes. Through these two experiments, a mono-specific mesocosm study and a field study with mixed plant communities, we evaluated the predictive power of plant community attributes for understanding spatial variation in soil communities at the landscape scale.

Results and Discussion

The effect of plant species identity on soil communities

We evaluated whether plant species identity affected soil community composition by comparing the soil communities from the mesocosms, where each plant species (n = 4 per species) were grown over three years in the same homogenized soil collected from the site (see Methods). Overall, the soils contained diverse communities (Fig. A4.1A). Soil fungal communities were primarily composed of *Ascomycota* (43% of ITS reads, on average), *Basidiomycota* (31%), and *Zygomycota* (21%), bacterial communities were primarily composed of *Acidobacteria* (31% of 16S rRNA reads, on average), *Proteobacteria* (20%), and *Verrucomicrobia* (16%), protistan communities were primarily composed of *Rhizaria* (26%), *Amoebozoa* (25%), *Alveolata* (22%), and *Stramenopiles* (16%), and metazoan communities were primarily composed of *Nemotoda* (33%), *Arthropoda* (28%), and *Annelida* (15%; Fig. A4.1B). The general structure of these communities is similar to what has been observed in comparable surveys of other temperate grasslands (Wu et al. 2011, Bates et al. 2013, Leff et al. 2015).

Control mesocosms, where plants were actively removed, harbored fungal, protistan, and metazoan communities with significantly different compositions compared to those communities found in mesocosms containing plants ($P < 0.005$ in all cases), but bacterial community composition did not significantly differ between control and planted mesocosms ($P = 0.15$). For the communities that differed between control and planted mesocosms, certain taxa were indicative of soils containing plants (Fig. A4.2). For example, planted mesocosms had, on average, an 11-fold greater relative abundance of

fungi belonging to the phylum *Glomeromycota*, which is expected given that this phylum is almost entirely composed of arbuscular mycorrhizal fungi (Redecker and Raab 2006) but still demonstrates that these plants drive recruitment of mycorrhizal fungi in soil. Protistan communities from planted mesocosms had a 63% lower relative abundance of *Archaeplastida* (algae) compared to control mesocosms, likely due to lower light availability, and the metazoan phylum *Tardigrada* had a 14-fold greater relative abundance in planted compared to control mesocosms. These results demonstrate the important role plants play in altering soil communities on a relatively short time scale. While it is widely thought that plants can influence soil biota (Bardgett and Wardle 2010), few studies have comprehensively described the direct influence of plants on a wide array of organisms from natural soils. In addition, there was a notably small effect of plants on bacterial community composition, suggesting that bacterial communities are not strongly influenced by plants (or lack of plants) on yearly time scales. A large body of work has shown that individual bacterial taxa can form close associations with plants (Bardgett and Wardle 2010), but our results indicate that the overall composition of soil bacterial communities is more strongly affected by other soil factors than the overall presence or absence of plants over three years. Another potentially complementary explanation for the weak effect of plants on bacterial community composition is that soil bacteria are prone to legacy effects (Bartelt-Ryser et al. 2005, Van der Putten et al. 2013) whereby the observed communities, including those from control plots, were reflective of the plant community from which the soil was collected when establishing the mesocosms. Thus, it is possible that additional time would reveal plant effects on bacterial community composition as the lingering effects of those plants previously grown in those soils would diminish.

We also found that plant species identities drove differences in the overall composition of soil fungal (R^2 = 0.33; P = 0.001), bacterial (R^2 = 0.27; P = 0.02), protistan (R^2 = 0.32; P = 0.001), and metazoan (R^2 = 0.31; P = 0.001) communities (Fig. 4.1A). These plant species effects were not driven by a small

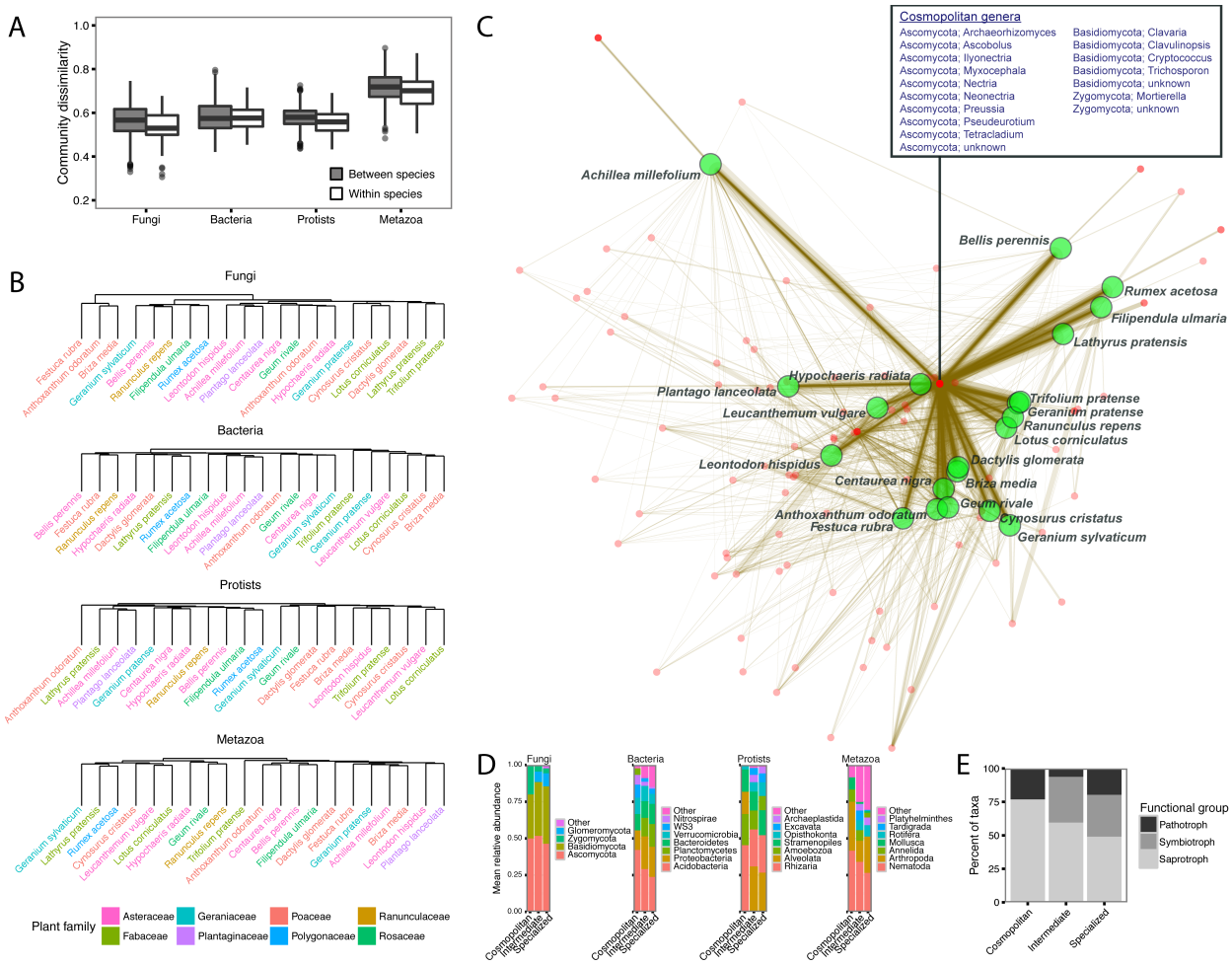

Figure 4.1. The effects of plant species identity on the composition of soil communities. Pairwise Bray-Curtis dissimilarities in community composition between vs. within soils from the same plant species (A). Hierarchical clustering diagrams based on mean dissimilarities across the plant species (B). Bipartite network diagram, where edges (lines) connect plant species (green circles) to fungal taxa (red points) that were present in their mesocosms. Taxa were considered present if their mean relative abundance was ≥0.1%, and only taxa with a relative abundance of >0.5% in ≥1 species are shown (C). The composition of cosmopolitan (present with all plant species), intermediate (present with 2 to 20 plant species), and specialized (present with only one plant species) soil taxa (D). The composition of functional groups in of cosmopolitan, intermediate, and specialized fungal taxa. Only those taxa that could be identified were included (E).

number of plant species having belowground communities distinct from the others (Fig. 4.1B). Certain fungal, protistan, and metazoan taxa tended to be strongly associated with individual plant species, while others tended to be more cosmopolitan (Fig. 4.1C, Fig. A4.3). For example, the fungal taxa identified as *Olpidium brassicae* and *Phoma sp*. associated with *Achillea millefolium* in particular, while several Ascomycota, Basidiomycota, and Zygomycota taxa were associated with all plant species (Fig.

4.1C). We used an indicator analysis approach to identify those taxonomic groups that were most strongly associated with each of the individual plant species and found that many of the plant species formed specific associations (Fig. A4.4). Since there are likely to be different traits associated with soil taxa that are more specialized or more cosmopolitan, we investigated whether taxa unique to individual plant species tended to represent different taxonomic groups when compared to taxa that were more ubiquitous across plant species. Cosmopolitan taxa were represented by a higher proportion of *Zygomycota*, *Acidobacteria*, *Rhizaria*, and *Nematotoda*, while more specialized taxa were represented by a greater proportion of *Glomeromycota*, *Planctomycetes*, *Alveolata*, and *Rotifera* (Fig. 4.1D). Additionally, we found that cosmopolitan fungal taxa were represented by a greater proportion of putative saprotrophs compared to more specialized taxa, which were represented by a greater proportion of pathogens and symbionts (Fig. 4.1E). This suggests that pathogens and symbionts tend to be more strongly limited to individual plant species, while saprotrophs are more cosmopolitan and less influenced by plant species identity. This finding is in concordance with a previous study conducted in an Amazon rainforest showing stronger plant-soil linkages for fungal pathogens and mycorrhizae compared to saprotrophs (Peay et al. 2013).

Our results corroborate findings from a number of previous experiments, which have shown that plant species identity can affect soil communities (e.g., St. John et al. 2006, Berg and Smalla 2009, Bezemer et al. 2010, Lekberg and Waller 2016). It is possible that certain previous studies did not observe species identity effects because they included fewer species and/or used soil community assessment methods that had lower power to resolve individual soil taxa (Porazinska et al. 2003, Nunan et al. 2005, Bezemer et al. 2006, Singh et al. 2007). Some of the soil taxa that we observed to associate with individual plant species belonged to the same genus or family as indicator taxa for other plant species (Fig. A4.4), demonstrating that sufficient taxonomic resolution is important for identifying differences across plant species.

Can the effect of plant species identity be explained by plant phylogeny or traits?

Simply documenting plant species identity effects in one ecosystem would allow for predictions in that ecosystem, but such predictions would not extend to other ecosystems with different plant species that had unknown associations with belowground communities. Therefore, we sought to assess whether plant species identity effects could be explained by plant phylogeny or traits, two attributes that could potentially be used to predict plant associations with belowground communities a priori.

The mesocosm plant species represented eight families including grasses, asters, and legumes, providing an opportunity to evaluate the influence of a wide ranging phylogeny on the composition of soil communities. Plant phylogenetic distances were not significantly related to differences in fungal, bacterial, or metazoan community composition ($P > 0.1$ in all cases; Fig. 4.2A). Differences in protistan community composition were related to plant phylogenetic distance, but this relationship was fairly weak (rho = 0.29, $P = 0.002$; Fig. 4.2A). Nonetheless, the relative abundance of *Stramenopiles* was significantly related to the phylogeny of the plant species ($K = 0.51$, $P = 0.004$; Fig. A4.5). We might expect plant phylogenetic differences to be associated with the structure of belowground communities due to coevolution with mutualists or pathogens (De Deyn and Van Der Putten 2005), but this did not appear to be the case across the majority of soil organisms. Our results parallel those of a recent study which found that negative plant-soil feedbacks, presumably due to the accumulation of pathogens following the growth of one plant species, were not related to plant phylogeny of the species tested in the preconditioned soil (Mehrabi and Tuck 2015).

The measured plant traits were highly variable across the mesocosm species. Grassland plants vary in their ecological strategies, with certain species growing fast under high nutrient conditions and having characteristically high specific leaf areas and N contents, and others are selected for surviving under lower nutrient conditions and have opposite traits (Lavorel and Garnier 2002, Reich 2014, Roumet et al. 2016). For each plant in the mesocosms, we measured the major traits that are known to be

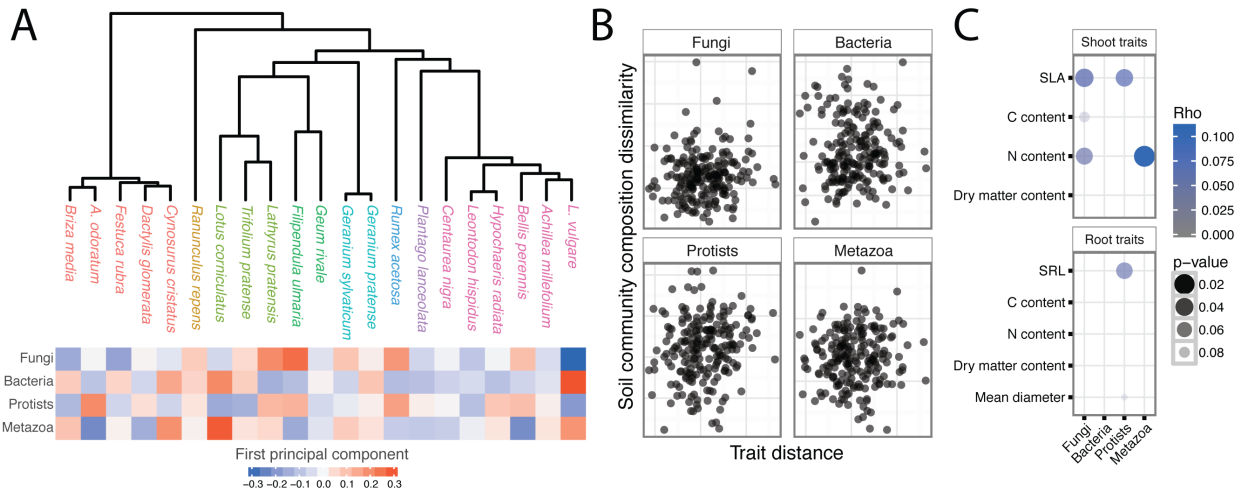

Figure 4.2. Relationships between plant species' relatedness and differences in the composition of soil communities (A). Plant phylogenetic tree with species names colored by family (key shown in Fig. 1). Heatmap representing the dissimilarities in the composition of each soil community. Colors represent the first principal coordinate analysis axis calculated from Bray-Curtis dissimilarities. The relationship between differences in the composition of soil communities and plant trait distances (B). Euclidean trait distances were calculated using all the traits shown in panel C. The relationship between differences in the composition of soil communities and individual plant traits. Points represent Spearman correlation coefficients (rho) and Mantel test results (*P* value).

indicative of the tradeoffs in these life history strategies (Fig. A4.6A). For example, the *Fabaceae* species tended to have a greater shoot and root N and C content while *Poaceae* species tended to have high leaf dry matter contents (Fig. A4.6B). Yet, multivariate dissimilarity in the traits of the plant species were not predictive of differences in soil communities ($P > 0.1$ in all cases; Fig. 4.2B). Furthermore, there were no strong or significant (i.e. Bonferroni corrected $P < 0.05$) relationships between belowground community compositions and individual traits (Fig. 4.2C). These results suggest that the traits we measured are not good indicators of the specific relationships plants form with communities belowground. Some previous work corroborates this finding. For example, Porazinska *et al*. (2003) found that certain soil communities were linked to individual plant species, but they were unable to identify traits that could predict soil communities. Likewise, other studies have shown that plant species identity is more predictive of soil communities than traits (Viketoft et al. 2005, Barberán et al. 2015). Nonetheless, the plant-soil organism associations we observed could have been driven by plant traits we did not measure. For example, root exudates (Haichar et al. 2008) and leaf litter chemistry (Bardgett and Wardle 2010, Cline and Zak 2015)

can influence soil community composition, and these traits are not necessarily strongly related to the plant traits we measured.

Are soil communities in the field predictable based on plant community attributes?

The results from the mesocosm study demonstrated that plant species identity can causally affect soil community compositions. However, there are reasons why these effects might not translate to more realistic conditions with multi-species plant communities. For example, plants can excrete organic molecules into the soil when interacting with other plants (Bertin et al. 2003, Semchenko et al. 2014), and these excretions could alter the associated soil communities. Additionally, the more complex conditions existing in situ could affect soil communities in the field differently than in the mesocosms. Therefore, we set out to assess whether results observed in the mesocosms corresponded to similar trends in the field, where we analyzed 80 individual samples collected from a 0.5 ha site. Plant community composition was evaluated in one of 3 ways: using aboveground biomass DNA, root biomass DNA, and plant DNA in the sieved soil (see Methods). DNA sequencing approaches were used because roots are difficult to identify visually, and assessing plant communities via soil DNA provides an alternate approach to determine plant species which have occupied a given location currently or in the past (Yoccoz et al. 2012). Our molecular approach was verified for efficacy by comparing it to visual inspection assessments of aboveground biomass (Fig. A4.7).

Differences in the compositions of each of the soil communities were related to differences in aboveground plant community composition and plant community composition assessed using the soil DNA ($P < 0.05$ in all cases). However, the differences in the composition of soil communities were not significantly related to differences in plant community composition assessed using roots ($P > 0.1$ in all cases; Fig. 4.3A). These results indicate that there is a relationship between plant and soil community compositions but that the root biomass we collected was not an appropriate indicator of the plant species the soil communities relate to. It is possible that much of the root biomass consisted of dormant

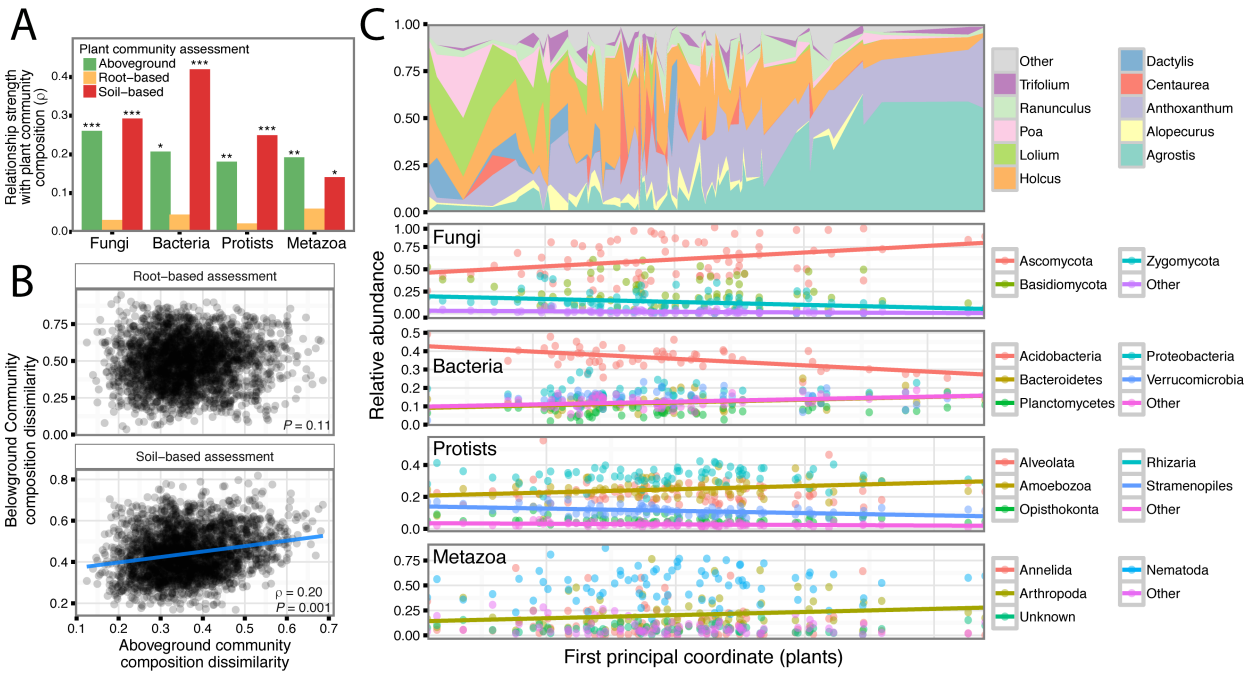

Figure 4.3. Soil community composition is related to plant community composition in the field. Pairwise Bray-Curtis dissimilarities in plant community composition as assessed using aboveground tissue are not related to dissimilarities in plant community composition as assessed using root tissue but are related to dissimilarities in plant community composition as assessed using plant DNA in soil (A). Relationship strength between dissimilarities in soil communities and dissimilarities in plant communities (* = $P < 0.05$, ** = $P < 0.01$, *** = $P = 0.001$; B). Relationships were assessed using Mantel tests in panels (A – B). Variation in plant community composition across the field samples ordered by the first principal coordinate score, and relationships between soil taxonomic group relative abundance and the plant first principal coordinate score (C). Linear trend lines were only plotted for groups that had a Pearson correlation $P \leq 0.05$.

plants or dead tissue (Tesařová et al. 1982), and thus, it is more likely that soil organisms would interact more heavily with the aboveground tissues as we observed. Nonetheless, our results demonstrate that we can use plant community composition to predict differences in overall soil community composition across a field.

Differences in aboveground plant community composition were unrelated to differences in root plant community composition ($P = 0.11$), but they were related to differences in the plant community composition as assessed using plant DNA in soil (rho = 0.2; $P = 0.001$; Fig. 4.3B). This shows that shoot and root biomass in a given location do not represent the same plant community. This is consistent with observations made in a tropical rainforest (Barberán et al. 2015). Additionally, these results suggest that

plant DNA in soil can be used as a proxy for the community composition of the aboveground biomass as observed previously (Yoccoz et al. 2012, Taberlet et al. 2012). This has implications for future research since it is often logistically simpler to obtain a representative sample of surface soils rather than sampling and homogenizing aboveground plant biomass.

By comparing the compositions of the plant communities (using the first principal coordinate score based on aboveground assessments), we were able to identify specific plant genera that drove variation in community composition across the samples (Fig. 4.3C). For instance, some samples had relatively high relative abundances of *Lolium spp*. while other samples had high relative abundances of *Agrostis spp*. These differences in plant community composition were related to the relative abundance of certain coarse taxonomic groups of soil taxa, including the *Ascomycota, Zygomycota*, *Acidobacteria*, *Amoebozoa, Stramenopiles*, and *Arthropoda* (Fig. 4.3C). These specific associations between plant taxa and belowground taxa can ultimately be used to predict the composition of soil communities from plant species abundances. For example, our results suggest that communities dominated by *Agrostis spp*. are likely to correspond to greater relative abundances of *Ascomycota* and lower relative abundances of *Acidobacteria* in the soils in which they grow. However, future research is needed to determine whether these patterns would also be observed in other site and soil types.

We evaluated whether the phylogenetic structure or aggregated traits of plant communities could explain relationships with soil communities. In other words, we evaluated whether plant communities containing genera with a more shared phylogenetic history or those with more similar trait values tended to be associated with more similar soil communities. Plant community phylogenetic structure was not significantly related to the composition of any of the soil communities ($P > 0.3$ in all cases) suggesting that plant community phylogenetic relatedness is not predictive of soil community composition. This finding is in agreement with another recent study that did not find a strong effect of tree species phylogenetic relationships on soil communities in a tropical rainforest (Barberán et al.

2015). Differences in aggregated trait values, including root and shoot N and C content, also did not significantly relate to the composition of any of the soil communities ($P > 0.1$ in all cases). Like phylogenetic structure, this suggests that the trait values we measured were not predictive of soil community composition. These results are consistent with our findings in the mesocosm experiment where phylogeny of the plants and the measured plant traits were not related to soil community composition.

Are relationships between plant community compositions and soil community compositions driven by soil characteristics?

In order to assess whether relationships between soil communities and plant communities in the field plots were attributable to the direct effects of the plants rather than mutual environmental drivers or intermediary effects of the plants on soil properties, we evaluated whether information on plant community composition contributed additional explanatory power to variation in soil community composition given differences in edaphic characteristics. Differences in the composition of soil communities were significantly correlated with multiple, individual edaphic properties (Table A4.1), and combinations of these properties explained 13 – 29% of the variation in soil community composition pairwise dissimilarities ($P = 0.001$ in all cases; Fig. A4.8A). For example, soil N content and soil pH were typically correlated with the composition of the four soil communities. Only differences in fungal community composition could be predicted significantly more accurately when information on aboveground plant community composition was added to the models containing only soil characteristics as predictor variables ($P = 0.01$; Fig. A4.8). When soil DNA-based plant community composition information was used instead of aboveground plant community composition, fungal, bacterial, and protistan community composition could all be predicted more accurately with the addition of plant community composition as a predictor ($P < 0.02$ in all cases; Table A4.8). These results suggest that the aboveground composition of plant communities likely influence soil communities in ways not accounted

for in commonly measured soil properties. Previous work has found that rhizosphere communities are affected by a combination of plant species identity and soil characteristics (Berg and Smalla 2009), which support our findings.

Future directions

Our results show that knowledge of differences in plant community composition across locations in a field are informative for predicting soil communities, and combining this information with information on other soil variables allowed us to predict up to 13 - 36% of variation in the overall soil community composition at the landscape scale, depending on the community of interest. However, while we were able to identify associations between certain plant species and soil organisms in this study, we were largely unable to generalize these findings across plant species based on knowledge of their phylogeny or commonly measured traits. Thus, it will be difficult to transfer the knowledge of specific plant-microbe associations we gained here to the ability to predict differences in soil communities across landscapes or ecosystems with different plant species. Therefore, a critical next step is to identify whether there are other plant traits that are more directly responsible for driving differences in belowground communities.

Information on plant community composition will be especially useful when predicting the distributions of particular soil organisms compared to others. For instance, we found that certain soil organisms, such as putative fungal pathogens, are more closely associated with particular plant species, and therefore their distributions will be easier to predict based on plant community information compared to more cosmopolitan taxa. Finally, we note that much of the variation in soil community composition across even the relatively small field used in this study could not be explained by the soil characteristics or plant community attributes we measured. This highlights that soil community structure is difficult to predict in general and that much more research is needed to identify additional

variables which influence soil communities and to identify the level of variation which can be explained by stochastic processes.

Methods

Mesocosms experiment

The mesocosms used to evaluate effects of individual plant species and their phylogeny and traits on microbial communities were established at a field station within the Ingleborough National Nature Preserve in northern England (54°11'38.7"N 2°20'54.4”W). Mesocosms were constructed in May 2012 from large planter boxes (~40 L) filled with 5 cm of rinsed gravel and the remainder with soil from the site. Soil was homogenized prior to distributing to the boxes. Grassland plant species were grown in a greenhouse with seeds collected from the site and transplanted to the mesocosms. 4 replicate mesocosms were constructed for each of the 21 plant species used. In addition, 4 replicate controls where no plants were planted were constructed. The planted and control mesocosms were arranged in a randomized block design with four blocks and one of each species per block. Mesocosms were actively weeded throughout the growing seasons. Plants were harvested and soils were collected (using a 2.5 cm core the height of the planter box) in July 2015. Soils were sieved to 4 mm, frozen, and shipped on dry ice to the University of Colorado for soil community analysis. Aboveground and root plant tissue was saved for trait measurements.

Field plots design and sampling

Field plots were established ~1 km from the mesocosm field site. Experimental plots were constructed by planting species belonging to one of three treatments (legumes, forbs, and grasses) or controls with no added plants. Plants were grown from seed in a greenhouse and planted in plots every spring over 3 years prior to samples being collected in July 2015. To sample vegetation and soil, 30 cm diameter collars were placed at representative locations within plots (n = 4 per plot with 5 plots per treatment; i.e. N = 20 per treatment), and aboveground plant biomass was harvested from within the

collar. 6.8 cm x 10 cm soil cores were collected from within the collars and processed identically to the mesocosm soil samples. Roots collected while sieving the soil were retained for the root-based assessment of plant community composition.

Soil community composition

Fungal, bacterial, protistan, and metazoan communities were assessed in soil samples following molecular marker gene sequencing protocols as described in Prober *et al*. (2015) and Ramirez *et al*. (2014b). Briefly, DNA was extracted from each sample, and ribosomal marker genes were amplified using PCR with barcoded primers unique to each sample. For fungi, we used the ITS1F/ITS2 primer pair (White et al. 1990, Smith and Peay 2014), for bacteria, we used the 515f/926r primer pair (Walters et al. 2015), and for protists and metazoa, we used the 1391f/EukBr primer set (Ramirez et al. 2014b). For each primer pair, equal quantities of amplicons from each sample were pooled and then sequenced on an Illumina MiSeq instrument using 2x251 bp sequencing kits at the BioFrontiers sequencing facility at the University of Colorado. Appropriate controls were used throughout the laboratory process to screen for contaminants.

Raw sequences were processed using the DADA2 pipeline (Callahan et al. 2016), which is designed to resolve exact biological sequences from Illumina sequence data and does not involve sequence clustering. Raw sequences were first demultiplexed by comparing index reads (barcodes) to a key, and paired sequences were trimmed to uniform lengths within marker genes. Sequences were then dereplicated, and the unique sequence pairs were denoised using the 'dada' function with 'err=NULL' and 'selfConsist = TRUE'. Potential primers and adapters were then screened and removed using a custom script (https://github.com/leffj/dada2helper). Next, paired end sequences were merged, chimeras were determined and removed, and a taxa table was produced, which contained the sequence counts per unique sequence and sample. Taxonomy assignments were determined using the RDP classifier trained on the UNITE (Abarenkov et al. 2010), Greengenes (McDonald et al. 2012a), or PR2

databases (Guillou et al. 2013) for fungi, bacteria, and protists and metazoa, respectively. 16S rRNA sequences identified as chloroplasts, mitochondria, or *Archaea* were removed. Protists and metazoa were separated into different taxa tables based on taxonomic classifications. To account for differences in sequencing depths, samples were rarefied to 5,300, 1,300, 2,400, and 1,250 sequences per sample for fungi, bacteria, protists, and metazoa. Putative fungal functional groups were identified using FUNGuild (Nguyen et al. 2015).

Plant community composition

Plant community composition in the field plot samples was assessed in three ways: (1) by visual inspection and sorting of the aboveground biomass, (2) by molecular analysis of the aboveground biomass, (3) by molecular analysis of the roots contained in the soil cores, and (4) by molecular analysis of DNA extracted from the soil samples. For visual inspection, harvested aboveground biomass was identified the same day as collection, and tissue from each species was dried and weighed. For molecular assessments, aboveground and root biomass samples were freeze-dried, ground, and homogenized prior to DNA extraction. We prepared DNA for sequencing following a protocol similar to (Kartzinel et al. 2015). We identified the genus-level plant community composition by targeting both the P6 loop of the *trnL* gene and the rRNA ITS region. We extracted DNA using the PowerSoil DNA Isolation Kit (Mo Bio Laboratories, Inc.), and soil samples were diluted 1:10 prior to amplification. The primer set trnL(UAA)c/trnL(UAA) with included Illumina sequencing adapters (Taberlet et al. 2007) was used to amplify the *trnL*-P6 marker following a PCR protocol of: denaturing at 94°C for 2 min followed by 36 cycles of 94°C for 1 min, 55°C for 30 s, and 72°C for 30 s, with a 5-min final extension at 72°C. To amplify the ITS region, we used the forward primer, ITS1-F, and included two reverse primers, ITS1Ast-R and ITS1Poa-R (Ait Baamrane et al. 2012, Kartzinel et al. 2015), to specifically target *Asteraceae* and *Poaceae* species. All primers included appropriate Illumina adapters, and PCR reactions were carried out as for *trnL* amplification. Each PCR was done in duplicate and the amplification product was combined. All

products for each sample were combined in equal volumes and cleaned using the UltraClean PCR Clean-Up Kit (Mo Bio Laboratories, Inc.). Illumina Nextera barcodes were added to the amplicons using an 8-cycle PCR, amplicons were cleaned and pooled using the SequalPrep kit (Invitrogen), and sequenced on an Illumina MiSeq instrument with a 2x151 bp kit at the University of Colorado BioFrontiers sequencing facility.

We processed raw plant sequences in a similar manner as for soil community sequences described above. We used the DADA2 pipeline (Callahan et al. 2016) to trim forward and reverse paired reads to 145 and 130 bp, respectively. Following the denoising step, Illumina adapters were removed, paired end reads were merged, and chimeras were filtered. We assigned taxonomy to each sequence using BLAST searches against the GenBank NR database. Sequences were assigned taxonomy only if ≥80% of the sequence aligned to a reference sequence and they matched the reference sequence with ≥95% identity. If a sequence had multiple best matches to reference sequences, a common genus and/or family name was assigned if one existed. Otherwise, sequences were assigned as 'unknown'. Taxonomy assignments were manually checked and verified in reference to species known to exist at the site. Separate taxa tables were created based on *trnL* amplicons and each of the *Asteraceae* and *Poaceae* ITS amplicons. Samples with fewer than 550, 1000, and 100 sequences were removed from taxa tables based on *trnL*, *Asteraceae* ITS, and *Poaceae* ITS amplicons. We calculated the relative abundance of individual plant genera in each sample using the *trnL* sequence counts. Because there is poor taxonomic resolution for the *Asteraceae* and *Poaceae* within the *trnL* gene, we replaced the total relative abundances of taxa (mostly unknown genera) within these two families with normalized relative abundances of genera determined using the ITS sequences.

Plant traits

Plant leaf traits were determined from fresh leaves and roots from the mesocosms as in Orwin et al. (2010) and Legay et al. (2015). Briefly, we measured specific leaf area, specific root length, leaf dry

matter content, root dry matter content, and root mean diameter using standard methods. Shoot and root N and C contents from the mesocosm-grown plants and the field sample plant communities were measured on an elemental analyzer. In both cases, plant material was freeze-dried and thoroughly homogenized prior to measurement.

Soil characteristics

Soil characteristics sere measured as in Orwin *et al.* (2010). pH was measured using a ratio of 1 g soil: 2.5 ml dH_2O. Total C and N were measured using an elemental analyzer. Dissolved inorganic N, individual ions, and net N mineralization were assessed using 0.5M KCl extracts, and dissolved organic N was assessed using water extracts as in Bardgett *et al*. (2003). Total soluble N was determined following oxidation of these extracts using potassium persulphate (Bardgett et al. 2003).

Statistical analysis

All statistical analyses were performed in R (R Core Team 2016) using specific packages where noted, and the package 'mctoolsr' (Leff 2016) was used to facilitate data manipulation and analyses. For mesocosms, only plant species with ≥ 3 replicates (21 species) were included in downstream analyses. To represent differences in community composition, we calculated Bray-Curtis dissimilarities using square-root transformed relative abundances. Permutational analysis of variance (PERMANOVA), as implemented in the 'adonis' function from the 'vegan' package (Oksanen et al. 2016), was used to test for differences in soil community compositions across factors. We compared the relative abundances of taxa from control mesocosm communities to the relative abundances of taxa from planted mesocosms using linear mixed effects models based on rank-transformed data with block included as a random effect. *P* values were corrected for multiple comparisons using false discovery rate corrections, and zeros were replaced with an estimate of the lower detection limit (1×10^{-5}) when creating Fig. S2 to avoid infinite fold changes. To test for differences in soil community composition across mesocosm plant

species, we used PERMANOVA and included block identity as a factor in the model. We displayed differences in community compositions across the plant species using hierarchical clustering based on mean dissimilarities. Network analysis plots were created using the 'igraph' package (Csardi and Nepusz 2006). We identified particular soil taxa that associated with specific plant species using indicator analysis (Dufrêne and Legendre 1997).

To test the relationship between the composition of soil communities and plant species relatedness in the mesocosms, we used the phylogeny from (Durka and Michalski 2012). Relationships between difference in soil community composition and plant phylogenetic distances were evaluated using Mantel tests with Spearman correlations. We tested for a phylogenetic signal in the relative abundance of individual protist taxa using the phylosig function in the 'phytools' package (Revell 2012), where the statistic, *K*, represents the strength of the signal (Blomberg et al. 2003). We calculated multivariate dissimilarities in trait values by normalizing and standardizing individual trait values and calculating Euclidian distances. We tested the relationship between Euclidian trait distances and community composition dissimilarities using Mantel tests.

For the field samples, we calculated differences in the phylogenetic structure of plant communities (i.e. phylogenetic dissimilarity) using UniFrac (Lozupone et al. 2011b) as implemented in the package, 'picante' (Kembel et al. 2010). We used the plant phylogenetic tree as reported in Durka and Michalski (2012), and plants not identified to the genus level were removed. We assessed the relationship between phylogenetic dissimilarity and the Bray-Curtis dissimilarities in soil community composition using Mantel tests with Spearman correlations.

To assess whether differences in plant community composition predicted variation in soil community composition beyond the explanatory power of soil characteristics, we built models of soil community composition dissimilarity using multiple regression on distance matrices (MRM; Lichstein 2007) as implemented in the 'ecodist' package (Goslee and Urban 2007) and compared the explanatory

power of the model with and without the addition of plant community dissimilarity as a predictor variable. In these models, each soil variable was transformed using log or inverse transformations where necessary to approximate a normal distribution, and they were standardized prior to calculating Euclidian distances. MRM was implemented with rank (i.e. Spearman) correlations, and the "best" models containing only soil variables were derived by first including all soil variables and using backwards elimination until all predictors explained significant levels of variation in the response dissimilarities.

CHAPTER V

CONSISTENT RESPONSES OF SOIL MICROBIAL COMMUNITIES TO ELEVATED NUTRIENT INPUTS IN GRASSLANDS ACROSS THE GLOBE

(Leff, J.W., Jones, S.E., Prober, S.M., Barberan, A., Borer, E.T., Firn, J.L., *et al.* (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proc. Natl. Acad. Sci. U.S.A.*, 112, 10967–10972)

Abstract

Soil microorganisms are critical to ecosystem functioning and the maintenance of soil fertility. However, despite global increases in the inputs of nitrogen (N) and phosphorus (P) to ecosystems due to human activities, we lack a predictive understanding of how microbial communities respond to elevated nutrient inputs across environmental gradients. Here we used high-throughput sequencing of marker genes to elucidate the responses of soil fungal, archaeal, and bacterial communities using an N and P addition experiment replicated at 25 globally distributed grassland sites. We also sequenced metagenomes from a subset of the sites to determine how the functional attributes of bacterial communities change in response to elevated nutrients. Despite strong compositional differences across sites, microbial communities shifted in a consistent manner with N or P additions, and the magnitude of these shifts was related to the magnitude of plant community responses to nutrient inputs. Mycorrhizal fungi and methanogenic archaea decreased in relative abundance with nutrient additions, as did the relative abundances of oligotrophic bacterial taxa. The metagenomic data provided additional evidence for this shift in bacterial life history strategies since nutrient additions decreased the average genome sizes of the bacterial community members and elicited changes in the relative abundances of representative functional genes. Our results suggest that elevated N and P inputs lead to predictable

shifts in the taxonomic and functional traits of soil microbial communities, including increases in the relative abundances of faster growing, copiotrophic bacterial taxa, with these shifts likely to impact belowground ecosystems worldwide.

Introduction

Human activities associated with fossil fuel combustion, agricultural fertilization, and dust or ash production have greatly increased nitrogen (N) and phosphorus (P) inputs to ecosystems around the globe relative to their pre-industrial levels (Galloway et al. 2004, Wang et al. 2015). The impacts of elevated N and P inputs on grassland ecosystems, which cover 26% of the global land surface (Foley et al. 2011), are expected to occur on relatively short time scales, with potentially important effects on plant biodiversity and terrestrial carbon (C) dynamics (Suding et al. 2005, LeBauer and Treseder 2008, Craine et al. 2008, Clark and Tilman 2008). A large body of research focusing on plant community responses has demonstrated consistent loss of grassland plant diversity with nutrient additions (Suding et al. 2005, Borer et al. 2014). In many cases, nutrient additions also shift the composition of plant communities with faster-growing plants that are good competitors for light being favored under conditions where nutrients are less limiting to growth (Grime 1977, Tilman and Wedin 1991). The associated belowground microbial responses to nutrient additions remain poorly understood, even though soil microbes represent a large fraction of the living biomass in grassland systems (Fierer et al. 2009) and can have important effects on terrestrial C dynamics, soil fertility, and plant diversity (van der Heijden et al. 2008). In particular, integrated, cross-site, experimental investigations of both plant and soil microbial responses to nutrient additions are needed to inform understanding of how the structure and functional attributes of soil microbial communities shift in response to anthropogenic inputs of N and P and whether these shifts are consistent across sites.

Soil microbial communities are often sensitive to nutrient inputs. For instance, N fertilization typically reduces microbial biomass and respiration rates (Treseder 2008, Janssens et al. 2010, Ramirez

et al. 2012), with specific functional groups of microbes, including ammonia oxidizers and mycorrhizal fungi, often being very sensitive to N additions (Treseder 2004, Egerton-Warburton et al. 2007, Wessén et al. 2010). A few studies conducted at individual sites also have shown that elevated N inputs can alter the overall composition of bacterial or fungal communities (Egerton-Warburton et al. 2007, Allison et al. 2007, Campbell et al. 2010, Ramirez et al. 2010, Coolon et al. 2013). Understanding of soil microbial community responses to elevated P inputs remains more limited even though many regions experience elevated inputs of both N and P (Wang et al. 2015), and anthropogenic activities can alter N:P ratios in soil (Galloway et al. 2004, Peñuelas et al. 2013). We are not aware of any studies that have used standardized nutrient treatments to evaluate the generality and local context dependence of soil bacterial, archaeal, and fungal communities to N and P amendments across a wide range of soil types. Individual studies conducted at specific sites are useful, but inconsistencies in methods and site characteristics limit the ability to make robust generalizations of how belowground microbial communities will respond to elevated nutrient inputs across sites.

While previous studies have shown that soil microbial communities can shift in response to nutrient additions at individual grassland sites (Treseder 2004, Ramirez et al. 2010, Coolon et al. 2013, Pan et al. 2014), relating these taxonomic or phylogenetic shifts to changes in the functional attributes of these communities is not trivial. Simply documenting how communities shift in composition might not tell us how the aggregated traits of these communities change in response to nutrient additions because soil bacteria are incredibly diverse and most soil bacterial taxa remain uncharacterized (Ramirez et al. 2014a). Such trait-level information is arguably more important for linking changes in soil bacterial communities to changes in belowground processes than simply documenting how nutrients increase or decrease the relative abundances of community members (Fierer et al. 2014). Just as the aggregated traits of plant communities can shift in predictable directions with nutrient additions (Grime 1977, Tilman and Wedin 1991), we expect that the aggregated traits of soil bacterial communities will also

shift in a predictable manner with fertilization. Specifically, we expect that increases in nutrient availability will tend to favor copiotrophic (i.e. fast growing, low C use efficiency) bacterial taxa and reduce the abundances of more oligotrophic (i.e. slow growing, high C use efficiency) taxa (Ramirez et al. 2010, Fierer et al. 2011). Although there is some evidence that we can use taxonomic information to place soil bacteria along this continuum in life history strategies (Fierer et al. 2007b), we can use shotgun metagenomic information to more accurately infer the aggregated traits of soil bacterial communities and determine whether copiotrophic traits are actually favored under conditions of elevated nutrient availability.

For this study we sought to build a predictive understanding of the responses of diverse soil microbes to elevated nutrient inputs that is generalizable across grasslands. We collected soils from an N and P addition experiment replicated at 25 grassland sites spanning four continents and quantified shifts in bacterial, archaeal, and fungal community structure in response to experimentally increased soil nutrients using high-throughput sequencing of marker genes. In addition, we investigated potential shifts in community-level traits by analyzing functional gene metagenomic sequences from a subset of those sites. We hypothesized that N and P additions would: induce shifts in fungal communities with mycorrhizal fungi decreasing in relative abundance, alter archaeal community composition by increasing the abundances of those taxa presumed to be capable of ammonia oxidation (Leininger et al. 2006), and shift bacterial communities to favor copiotrophic over more oligotrophic taxa. Further, we hypothesized that the degree to which microbial communities shifted in response to nutrient additions would be positively correlated with the magnitude of the shifts in plant community composition. Those sites where nutrient additions have the largest effects on plant communities are also those sites where we would expect to see the largest responses in belowground microbial communities due to the direct associations between plants and microbes or their shared responses to fertilization.

Results and Discussion

Effect of nutrient additions on soil fungal communities

Fungal diversity and community composition differed strongly across the 25 globally distributed grassland sites regardless of nutrient treatment ($P < 0.001$ in all cases; Fig. A5.1). Mean fungal phylotype (i.e. species) richness ranged 1.7-fold across the sites, and there were large variations in the relative abundances of major taxonomic groups (Table A5.1). The strong site effects are not surprising given the range in environmental conditions and soil characteristics found across sites spanning four continents and elevations from 50 to 2320 m (Table A5.2). In particular, the sites represented a broad range in soil acidity, climate, and plant community composition, factors that have previously been associated with differences in soil fungal community structure at these sites and others (Tedersoo et al. 2014a, Prober et al. 2015).

We investigated the within-site effects of nutrient additions on fungal community structure by statistically controlling for the strong cross-site differences by including site as a random effect in our models. Fungal Shannon diversity responded weakly to nutrient additions, decreasing by only 2.7% on average when N and P were added together ($P = 0.05$), a response consistent with the weak response observed for plants (Borer et al. 2014).

In contrast to the weak effects of nutrients on fungal diversity, we observed significant effects of N ($R^2 = 0.003$; $P < 0.001$) and/or P ($R^2 = 0.002$; $P = 0.04$) additions on fungal community composition, with the same taxa generally responding to nutrient additions across sites despite the large cross-site variation in fungal community types (Fig. 5.1). With combined addition of N and P, there were increases in *Ascomycota* and significant decreases in the relative abundances of *Glomeromycota* (Fig. 5.2A). The *Glomeromycota* phylum is composed almost entirely of arbuscular mycorrhizal fungi (Redecker and Raab 2006), and we expected these fungi to decrease in relative abundance with nutrient additions since they would be less valuable to their hosts and thus provided with less plant C under conditions of

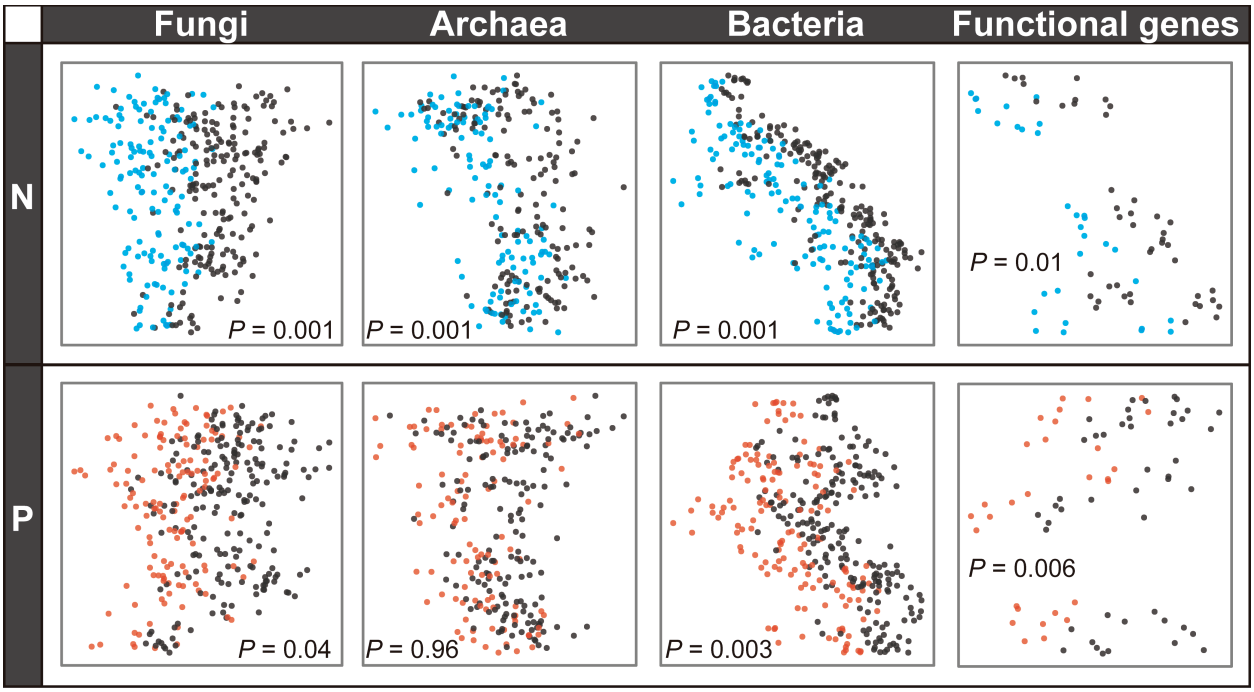

Figure 5.1. Constrained ordinations showing differences between microbial communities from plots that did not receive the indicated nutrient (gray points) and from plots receiving N (blue) or P (red) additions (colored points). Colored points include samples receiving both nutrients. P-values refer to PERMANOVA results.

increased N and P availability (Van Diepen et al. 2007, Johnson et al. 2010, Wei et al. 2013). We further investigated nutrient effects on mycorrhizal fungi by assessing the collective responses of mycorrhizal fungi, including those taxa outside the *Glomeromycota* phylum that are reported in the literature as being mycorrhizal. These taxa also consistently decreased in plots receiving N and P relative to the control plots ($P = 0.016$), corroborating results from a meta-analysis demonstrating declines in mycorrhizal fungi with N additions (Treseder 2004). Interestingly, adding N and P together led to far larger decreases in the relative abundances of *Glomeromycota* than when these nutrients were added individually ($P > 0.1$; Table A5.3), suggesting a role for both of these nutrients in shaping arbuscular mycorrhizal communities.

The overall decrease in the proportion of mycorrhizal fungi with N and P additions, and shifts in fungal community composition more broadly, could be caused by plant community shifts, changes in plant biomass, and/or the direct effects of added nutrients. The magnitudes of the responses of major

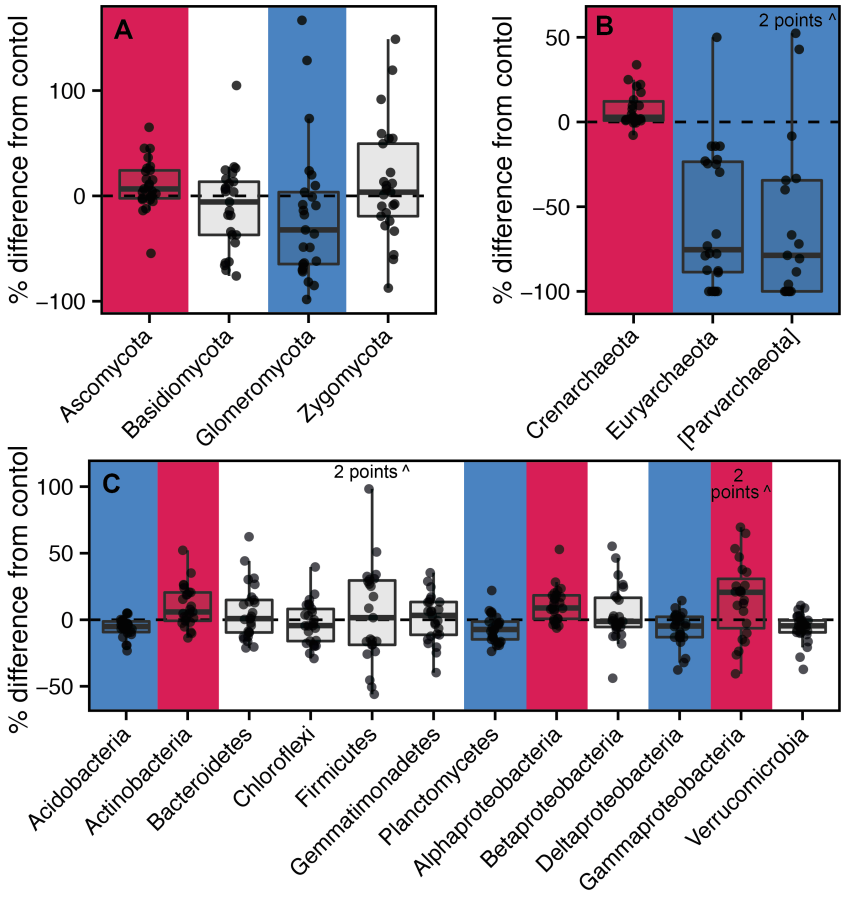

Figure 5.2. Differences in the relative abundance of higher-level taxa between control and nutrient addition plots. Fungal (A) and bacterial (C) taxa differences are comparisons to +N,+P plots, and archaeal taxa differences (B) are comparisons to +N differences since P additions did not significantly affect the relative abundance of archaeal taxa, nor was there an interaction between N and P additions. Points represent site means, and boxplots show quartile values for each taxon. Red and blue backgrounds show significant increases and decreases in the relative abundances of specific taxa, respectively (FDR-corrected $P < 0.05$). Only taxa with relative abundances >1% in any of the treatments are shown. Points with values greater than the plot axis maximum are indicated.

fungal taxonomic groups were not significantly correlated with changes in key soil characteristics (Table A5.4). However, the magnitude of fungal community composition response (i.e. the mean community dissimilarity between samples with added N and P and control samples) was significantly correlated with the magnitude of the response of plant community composition to added N and P (r = 0.44; $P = 0.03$; Fig. 5.3), helping to explain site-to-site variability in shifts in belowground communities. Those sites where nutrients had the largest impacts on plant communities were also the sites that had the strongest nutrient effects on fungal communities. This suggests either that shifts in plant community composition

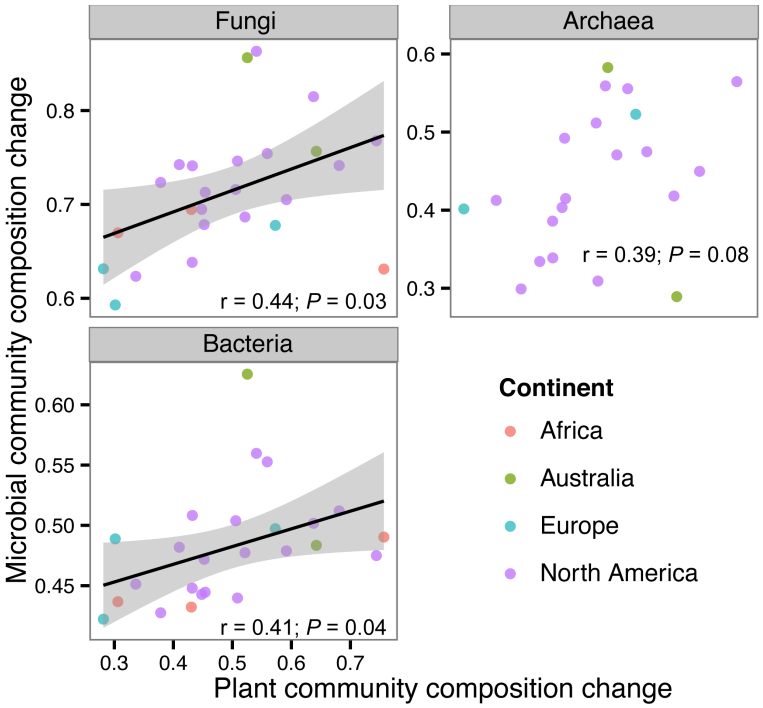

Figure 5.3. Correlations between changes in microbial and plant community composition with N and P additions across the sites for fungal, archaeal, and bacterial communities. Change in community composition was calculated as the mean Bray-Curtis dissimilarity between control plots and those plots amended with nutrients. Relationships were assessed using Pearson correlations.

drive shifts in fungal community composition, or that both plant and fungal communities respond similarly to changes in edaphic factors. Although overall fungal compositional shifts correlated with plant community composition shifts, changes in the relative abundance of *Glomeromycota* were not related to changes in live plant biomass with fertilization ($P > 0.1$), nor were they related to changes in surface soil nitrogen concentrations ($P > 0.1$; Table A5.4), suggesting that plant nutrient limitation was not a good predictor of the differential responses observed across the sites.

Effect of nutrient additions on soil archaeal communities

Archaea were rare at most sites, and archaeal diversity (Fig. A5.1A) and community composition (Fig. A5.1B) were highly variable across sites regardless of nutrient additions ($P < 0.001$). Archaeal phylotype richness ranged 3.7-fold across the sites, and the archaeal communities were dominated by *Crenarchaeota* (92% on average) and *Euryarchaeota* (4.3% on average; Table A5.1). The proportion of

16S rRNA reads that were of archaeal origin was also highly variable across the sites (Fig. A5.2A), ranging from 0 to 0.16. This variability in archaeal communities was likely due to the large cross-site differences in environmental conditions mentioned above. For instance, previous work has shown a correlation between archaeal relative abundances and soil nutrient content (Bates et al. 2011), we know that soil N concentrations varied 33-fold across the control plots, and archaea relative abundances were inversely related to soil C:N ratios ($r = -0.67$; $P < 0.001$).

We next assessed whether there were consistent shifts in archaeal relative abundance and community structure with nutrient additions by statistically controlling for the strong cross-site differences. Archaeal relative abundances generally increased with N additions ($P < 0.001$; Fig. A5.2B), and there was a mean 4.8% decrease in archaeal diversity with N additions when compared to control plots ($P = 0.01$). This decrease in diversity was possibly related to an N-induced growth of specific archaeal taxa. Specifically, the phylum *Crenarchaeota*, which was primarily comprised of members of the family *Nitrososphaeraceae*, consistently increased in relative abundance with N additions across the majority of sites while *Euryarchaeota*, and the candidate division *Parvarchaeota* consistently decreased (Fig. 5.2B). These shifts are likely related to *Archaea* being active drivers of the soil N cycle. For example, *Nitrososphaeraceae* can oxidize ammonia (Leininger et al. 2006, Gubry-Rangin and Hai 2011), a metabolism that is expected to be advantageous with elevated ammonium supply, which should have been elevated in the N addition plots, as urea is readily hydrolyzed to ammonium. Abundances of soil *Crenarchaeota* also are positively correlated with soil N content (Bates et al. 2011). Conversely, several reports have shown the potential for members of the *Euryarchaeota*, which are predominately methanogens, to fix atmospheric N_2 (Leigh 2000, Offre et al. 2013). This could place them at a competitive disadvantage under conditions of elevated N availability and explain their strong proportional decrease with N fertilization. While it has been shown that N can inhibit methanogenesis *in vitro* (Klüber and Conrad 1998), this is, to our knowledge, the first direct evidence that N additions may

also decrease methanogen populations in non-wetland soils. Still, it is important to note that these shifts in the relative abundances of archaeal phyla are not independent of one another, and decreased methanogen relative abundances could simply be the result of increased relative abundances of *Crenarchaeota*. Nonetheless, these results highlight that soil archaeal communities are sensitive to N additions, but additional research is required to determine if these community responses are associated with changes in methane fluxes or soil N cycling rates.

Effect of nutrient additions on soil bacterial communities

As with fungal and archaeal communities, bacterial diversity and community composition differed strongly across the 25 grassland sites (Fig. A5.1). These differences were likely due to factors such as acidity, climate, and plant community composition as has been previously observed (Lauber et al. 2009, Fierer et al. 2012a, Prober et al. 2015). Mean phylotype richness ranged 1.7-fold, and the abundant phyla, including *Proteobacteria*, *Acidobacteria*, *Verrucomicrobia*, *Actinobacteria*, and *Bacteroidetes*, all varied considerably in their relative abundances across the sites (Table A5.1).

Nutrient additions did not strongly alter bacterial diversity; P additions caused marginal (0.5%) increases in bacterial diversity ($P = 0.06$), and N had no significant effect. Our results stand in contrast to negative relationships between bacterial diversity and N additions reported from previous studies conducted at individual sites (Campbell et al. 2010, Koyama et al. 2014). This points to the importance of local context and highlights the pitfalls associated with extrapolating results obtained from individual sites to other ecosystems or soil types.

Bacterial community composition was significantly affected by N ($R^2 = 0.002$; $P < 0.001$) and P additions ($R^2 = 0.002$; $P = 0.003$; Fig. 5.1). The community shifts corresponded to changes in the relative abundances of numerous major taxa. For example, the relative abundances of *Actinobacteria*, *Alphaproteobacteria*, and *Gammaproteobacteria* consistently increased with nutrient additions across sites, while those of *Acidobacteria*, *Planctomycetes*, and *Deltaproteobacteria* consistently decreased

(Fig. 5.2C). However, these taxonomic shifts were not always in the same direction or magnitude when N or P was added alone (Table A5.3). Overall, the taxonomic patterns in our cross-site study were in agreement with previous work conducted at individual grassland sites (Ramirez et al. 2010), and they corroborate laboratory studies which have noted similar shifts in the relative abundances of these major bacterial groups with nutrient additions (Ramirez et al. 2012). Our findings are generally consistent with our hypothesized shifts in general life history strategies with bacterial taxa that are faster growing and more copiotrophic (Fierer et al. 2007b) being favored under conditions of elevated nutrient availability (Fierer et al. 2011). In particular, soil bacterial groups that are generally considered to be more copiotrophic, including *Actinobacteria* and *Alphaproteobacteria,* increased in relative abundance with nutrient additions, and the largely oligotrophic *Acidobacteria* phylum decreased in relative abundance. While original evidence for generalizations of these life history strategies across broad bacterial taxonomic groups was based on responses to labile carbon inputs (Fierer et al. 2007b, Bastian et al. 2009, Eilers et al. 2010), our results extend evidence for these ecological classifications to the direct or indirect bacterial responses to nutrient additions.

Genomic and metagenomic evidence for shifts in bacterial life history strategy with nutrient additions

We recognize that it is difficult to confidently assign bacterial clades into groups with copiotrophic and oligotrophic life history strategies, especially given the overwhelming amount of undescribed bacterial diversity found in soil (Ramirez et al. 2014a). Thus, we used a combination of genomic and metagenomic approaches to provide independent assessments of how copiotroph:oligotroph ratios shifted in response to added nutrients. First, we estimated aggregate community growth rates since we expected increases in the relative abundance of copiotrophic taxa to be reflected by faster growth rates (Pianka 1970, Fierer et al. 2007b). Thus, an increase in the estimated growth rate [i.e. a decrease in mean minimum generation time (MGT)] would suggest an increase in the relative abundance of copiotrophs. Mean MGTs were calculated for all samples from a combination of

our bacterial marker gene data and published genomes; 757 of the 46,534 phylotypes could be matched to genomes. As with other attributes of community structure, estimates of MGT strongly varied across sites (Fig. A5.3A). Within-site differences between nutrient-amended and control samples showed that adding nutrients tended to decrease MGTs (Fig. A5.3B), but this trend was not significant for N additions (P = 0.57) or P additions (P = 0.34) individually. However, this analysis has important limitations in that only a small proportion (~10%) of the 16S rRNA gene sequences from our samples could be mapped to genomes for which we had MGT estimates, and this proportion differed across nutrient treatments (Fig. A5.3C). Thus, this analysis likely provides a conservative estimate of potential differences in MGTs associated with nutrient additions and is only weakly supportive of the hypothesis that soil bacterial MGT decreases with nutrient additions.

To further confirm the putative shifts in life history strategies in bacterial communities, we assessed functional attributes directly from functional gene (i.e. shotgun metagenomic) data collected from six of the sites used in the taxonomic analyses (Table A5.2). These sites were selected because they spanned a wide geographic range, encapsulated a variety of environmental conditions, and the marker gene analyses suggested the N and P effects on microbial community composition were particularly strong. The shotgun metagenomic data (hereafter referred to as "metagenomic data") were found to be almost entirely derived from bacterial genomes – 94.8 ± 2.3% (mean ± SD) of the metagenomic small subunit (SSU) rRNA gene reads were identified as bacterial. Just as the marker gene data revealed that bacterial diversity and community composition differed strongly across sites, the metagenomic data revealed that functional gene diversity and composition also varied strongly across sites (Fig. A5.1). In addition, the diversity of annotated genes identified from the metagenomic data was significantly correlated with the diversity of bacterial phylotypes across the samples (r^2 = 0.27, P < 0.001; Fig. A5.4A), and the dissimilarity in functional gene composition was strongly related to the dissimilarity in bacterial community composition across the six sites (ρ = 0.87, P < 0.001; Fig. A5.4B). These findings suggest that

bacterial communities that are distinct in composition tend to have distinct functional attributes, and bacterial communities that are taxonomically more diverse also have more diverse metagenomes with a broader array of annotated genes. Correspondingly, the diversity of functional genes did not change with nutrient additions ($P > 0.1$), but there were significant shifts in overall functional gene composition with N ($P = 0.01$) and P additions ($P = 0.006$; Fig. 5.1) as was observed for bacterial taxa. These results are supported by previous work showing a relationship between the taxonomic structure of soil bacteria and functional genes across ecosystems (Fierer et al. 2012a) and significant N effects on functional gene composition at two North American sites (Fierer et al. 2011).

The metagenomic data yielded additional lines of evidence to support our hypothesis that nutrient additions favor copiotrophic bacterial taxa. Previous work has suggested that soil microorganisms with larger genomes should be more successful in resource-poor environments (Konstantinidis and Tiedje 2004), and thus, we expect copiotrophic taxa to have smaller genomes. To assess this, we calculated mean effective genome size, the estimated mean size of a genome in a given sample, and found that it significantly decreased with added N or P ($P < 0.03$ in both cases; Fig. 5.4A). More generally, this result highlights that genome size can be considered an important ecological trait, just as bacterial genome size is correlated with range size (Barberán et al. 2014) and plant genome size is an important predictor of species' ability to invade (Suda et al. 2014).

We investigated the specific gene categories that changed in proportion with nutrient additions by analyzing the quality-filtered metagenomic sequences that could be annotated. First, it is important to note that only 28.7 - 32.7% of sequences could be annotated, and soils receiving N or P had a 0.3% higher annotation rate on average ($P \leq 0.01$ in both cases; Fig. 5.4B), a pattern likely driven by the over-representation of copiotrophic bacteria, which are easier to culture, and are thus more commonly found in genome databases. Similarly, soils receiving N amendments tended to have a lower relative abundance of annotated, but unclassified, metabolic genes compared to control samples, likely also

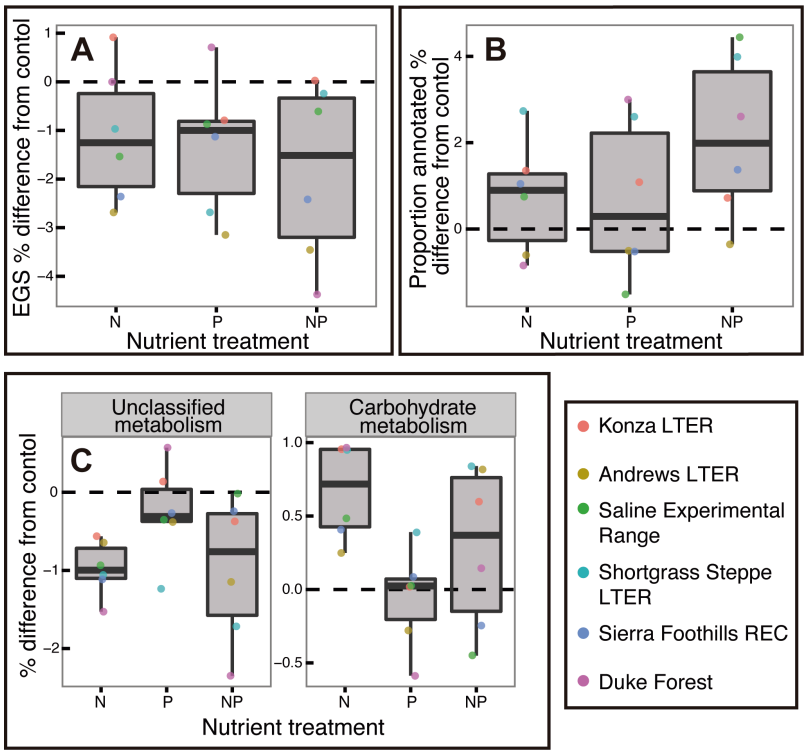

Figure 5.4. Shifts in metagenomic characteristics with the addition of nutrients. Differences in the proportion of annotated genes (A), effective genome size (B), and the relative abundance of metabolic genes (C) are shown with boxplots and mean responses for each site (points). Gene categories in (C) were chosen by selecting those that most greatly differed between control and treatment plots ($P < 0.02$ for each; Table A5.5).

reflecting the better representation of copiotrophs in genome databases (Fig. 5.4C; Table A5.5). We also observed a significant increase in the relative abundances of genes associated with carbohydrate metabolism (Fig. 5.4C) in fertilized plots. This is consistent with the added nutrients increasing copiotroph:oligotroph ratios and potentially increasing plant carbon inputs to soil. Although <33% of the sequence reads could be annotated, a percentage that is similar to that reported in other metagenomic analyses of diverse bacterial communities e.g., (Fierer et al. 2011), our results highlight that the annotated reads can be used to infer shifts in the functional capabilities of communities, shifts that are consistent with nutrient additions increasing the proportional abundance of bacteria with copiotrophic life history strategies.

Nutrients can have both direct and indirect effects on belowground bacterial communities making it difficult to unravel the mechanisms underlying the community responses described above.

Potential mechanisms include direct effects of the nutrients themselves, nutrient effects on soil characteristics (e.g., pH), nutrient inputs increasing plant productivity and organic matter inputs to soils (Ramirez et al. 2010), and nutrient inputs mediating microbial shifts through changes in plant community composition. With N addition, soil pH decreased by an average of 0.16 units across the sites ($P < 0.001$), and pH has been shown to strongly drive shifts in soil bacterial communities (Fierer and Jackson 2006, Lauber et al. 2009, Rousk et al. 2011). However, pH alone is not likely to have been a major driver of community shifts observed here, as the pH change was relatively small, it did not change with P additions ($P = 0.36$), and the magnitude of change in pH was unrelated to the change in the relative abundance of any of the major bacterial taxa with N and P additions across the sites (Table A5.4). Proportional changes in plant productivity were also unrelated to changes in the relative abundance of bacterial taxa, suggesting that elevated plant productivity in fertilized plots was not responsible for the bacterial community responses. On the other hand, the magnitude of shifts in plant community composition was directly related to the magnitude of shifts in bacterial community composition ($r = 0.41$, $P = 0.04$; Fig. 5.3), a pattern that mirrored that observed for fungi (Fig. 5.3). These findings suggest that changes in plant community composition may be more important for mediating bacterial community responses to elevated nutrient inputs than changes in edaphic characteristics or plant growth.

Conclusions

Taken together, our results demonstrate that while microbial community composition varied considerably across the diverse grassland sites examined, nutrient availability controls the composition of microbial communities in consistent ways across sites by selecting for microbial groups that have certain functional traits. Understanding the responses of soil microbial communities to changes in nutrient availability is critical given that ecosystems across the globe are receiving increasing inputs of N and P. Our analyses represent one of the first attempts to empirically assess whether there are

generalizable patterns in these responses across a wide range of climatic and edaphic environments and confirm their existence despite large cross-site differences in microbial community structure. The observed patterns correspond to broader ecological theory, and set the stage for more targeted hypothesis testing. For example, nutrient-induced shifts in copiotrophic versus oligotrophic traits could have important implications for soil C cycling (Wieder et al. 2013) if their traits elicit effects rather than solely reflect responses (Lavorel and Garnier 2002). Likewise, decreases in mycorrhizae and methanogens could have important impacts on ecosystem-level processes (van der Heijden et al. 1998, Offre et al. 2013). This work moves us towards a more mechanistic understanding of how shifts in microbial community composition mediate and reflect the effects of anthropogenically elevated nutrient inputs on terrestrial ecosystems.

Methods

Complete documentation of the experimental design, sample collection, and analytical methods are provided in Appendix A5 Methods.

Identical full factorial N and P addition experiments were established at each of the 25 sites used in this study, which included temperate-zone grasslands in Africa, Australia, Europe, and North America (Table A5.2). Nutrients were added annually in 10 g N or P m^{-2} yr^{-1}. Plant communities and soil characteristics were assessed as in (Prober et al. 2015). Fungal, archaeal, and bacterial community structure were characterized using barcoded Illumina sequencing of the internal transcribed spacer region of the ribosomal operon and the 16S rRNA gene for fungi and bacteria, respectively, using an approach described previously (Prober et al. 2015). These raw sequence data are available in the Sequence Read Archive at the National Center for Biotechnology Information (accession: SRP052716). The shotgun metagenomic sequences were collected and processed using an approach similar to (Fierer et al. 2013) with annotation performed using the KEGG hierarchy (Kanehisa et al. 2012). These data are available at the Integrated Microbial Genomes and Metagenomes website (http://img.jgi.doe.gov) and

referenced in the Genomes Online Database (GOLD Study ID: Gs0053063). We estimated MGTs for bacterial communities by calculating MGTs in available whole bacterial genomes using the method described in (Vieira-Silva and Rocha 2010) and mapping the 16S rRNA sequences we collected to these genomes.

REFERENCES

Abarenkov, K., R. Henrik Nilsson, K.-H. Larsson, I. J. Alexander, U. Eberhardt, S. Erland, K. Høiland, R. Kjøller, E. Larsson, T. Pennanen, R. Sen, A. F. S. Taylor, L. Tedersoo, B. M. Ursing, T. Vrålstad, K. Liimatainen, U. Peintner, and U. Kõljalg. 2010. The UNITE database for molecular identification of fungi – recent updates and future perspectives. New Phytologist 186:281–285.

Adler, P. B., A. Fajardo, A. R. Kleinhesselink, and N. J. B. Kraft. 2013. Trait-based tests of coexistence mechanisms. Ecology Letters 16:1294–1306.

Afzal, R., S. M. Mughal, M. Munir, K. Sultana, R. Qureshi, M. Arshad, and M. K. Laghari. 2010. Mycoflora associated with seeds of different sunflower cultivars and its management. Pakistan Journal of Botany 42:435–445.

Ait Baamrane, M. A., W. Shehzad, A. Ouhammou, A. Abbad, M. Naimi, E. Coissac, P. Taberlet, and M. Znari. 2012. Assessment of the food habits of the Moroccan dorcas gazelle in M'Sabih Talaa, west central Morocco, using the trnL approach. PLoS ONE 7.

Allison, S. D., C. A. Hanson, and K. K. Treseder. 2007. Nitrogen fertilization reduces diversity and alters community structure of active fungi in boreal ecosystems. Soil Biology and Biochemistry 39:1878–1887.

An, G. H., S. Kobayashi, H. Enoki, K. Sonobe, M. Muraki, T. Karasawa, and T. Ezawa. 2010. How does arbuscular mycorrhizal colonization vary with host plant genotype? An example based on maize (Zea mays) germplasms. Plant and Soil 327:441–453.

Andrews, J., and R. Harris. 2000. The ecology and biogeography of microorganisms on plant surfaces. Annual review of phytopathology 38:145–80.

Bacilio-Jiménez, M., S. Aguilar-Flores, E. Ventura-Zapata, P.-C. Eduardo, S. Bouquelet, and E. Zenteno. 2003. Chemical characterization of root exudates from rice (Oryza sativa) and their effects on the chematactic response of endophytic bacteria. Plant and Soil 249:271–277.

Bai, Y., D. B. Müller, G. Srinivas, R. Garrido-Oter, E. Potthoff, M. Rott, N. Dombrowski, P. C. Münch, S. Spaepen, M. Remus-Emsermann, B. Hüttel, A. C. McHardy, J. A. Vorholt, and P. Schulze-Lefert. 2015. Functional overlap of the Arabidopsis leaf and root microbiota. Nature 528:364–369.

Baily, M., A. Lilley, T. Timms-Wilson, and P. Spencer-Phillips, editors. 2006. Microbial Ecology of Aerial Plant Surfaces. CABI, Wallingford.

Balser, T. C., and M. K. Firestone. 2005. Linking microbial community composition and soil processes in a California annual grassland and mixed-conifer forest. Biogeochemistry 73:395–415.

Barberán, A., K. L. Mcguire, J. A. Wolf, F. A. Jones, S. J. Wright, B. L. Turner, A. Essene, S. P. Hubbell, B. C. Faircloth, and N. Fierer. 2015. Relating belowground microbial composition to the taxonomic, phylogenetic, and functional trait distributions of trees in a tropical forest. Ecology Letters 18:1397–1405.

Barberán, A., K. S. Ramirez, J. W. Leff, M. A. Bradford, D. H. Wall, and N. Fierer. 2014. Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. Ecology Letters 17:794–802.

Bardgett, R. D., and W. H. van der Putten. 2014. Belowground biodiversity and ecosystem functioning. Nature 515:505–511.

Bardgett, R. D., T. C. Streeter, and R. Bol. 2003. Soil microbes compete effectively with plants for organic-nitrogen inputs to temperate grasslands. Ecology 84:1277–1287.

Bardgett, R. D., and D. A. Wardle. 2010. Aboveground-belowground linkages: biotic interactions, ecosystem processes, and global change. Oxford University Press Oxford, New York, NY, USA.

Barret, M., M. Briand, S. Bonneau, A. Préveaux, S. Valière, O. Bouchez, G. Hunault, P. Simoneau, and M.-A. Jacques. 2015. Emergence shapes the structure of the seed microbiota. Applied and Environmental Microbiology 81:1257–1266.

Bartelt-Ryser, J., J. Joshi, B. Schmid, H. Brandl, and T. Balser. 2005. Soil feedbacks of plant diversity on

soil microbial communities and subsequent plant growth. Perspectives in Plant Ecology Evolution and Systematics 7:27–49.

Bastian, F., L. Bouziri, B. Nicolardot, and L. Ranjard. 2009. Impact of wheat straw decomposition on successional patterns of soil microbial community structure. Soil Biology and Biochemistry 41:262–275.

Bates, S. T., D. Berg-Lyons, J. G. Caporaso, W. a Walters, R. Knight, and N. Fierer. 2011. Examining the global distribution of dominant archaeal populations in soil. The ISME journal 5:908–17.

Bates, S. T., J. C. Clemente, G. E. Flores, W. A. Walters, L. W. Parfrey, R. Knight, and N. Fierer. 2013. Global biogeography of highly diverse protistan communities in soil. The ISME Journal 7:652–659.

Ben-Hur, E., O. Fragman-Sapir, R. Hadas, A. Singer, and R. Kadmon. 2012. Functional trade-offs increase species diversity in experimental plant communities. Ecology Letters 15:1276–1282.

Bender, S. F., C. Wagg, and M. G. A. van der Heijden. 2016. An Underground Revolution: Biodiversity and Soil Ecological Engineering for Agricultural Sustainability. Trends in Ecology and Evolution 31:440–452.

Berendsen, R. L., C. M. J. Pieterse, and P. A. H. M. Bakker. 2012. The rhizosphere microbiome and plant health. Trends in Plant Science 17:478–486.

Berg, G. 2009. Plant-microbe interactions promoting plant growth and health: Perspectives for controlled use of microorganisms in agriculture. Applied Microbiology and Biotechnology 84:11–18.

Berg, G., and K. Smalla. 2009. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. FEMS Microbiology Ecology 68:1–13.

Bertin, C., X. Yang, and L. A. Weston. 2003. The role of root exudates and allelochemicals in the rhizosphere. Plant and Soil 256:67–83.

Bezemer, T. M., M. T. Fountain, J. M. Barea, S. Christensen, S. C. Dekker, H. Duyts, R. van Hal, J. A.

Harvey, M. Maraun, J. Mikola, A. G. Mladenov, C. Robin, P. C. de Ruiter, S. Scheu, H. Setälä, P. Smilauer, and W. H. van der Putten. 2010. Divergent composition but similar function of soil food webs beneath individual plants: plant species and community effects. Ecology 91:3027–3036.

Bezemer, T. M., C. S. Lawson, K. Hedlund, A. R. Edwards, A. J. Brook, J. M. Igual, S. R. Mortimer, and W. H. Van Der Putten. 2006. Plant species and functional group effects on abiotic and microbial soil properties and plant-soil feedback responses in two grasslands. Journal of Ecology 94:893–904.

Blackman, B. K., M. Scascitelli, N. C. Kane, H. H. Luton, D. a Rasmussen, R. a Bye, D. L. Lentz, and L. H. Rieseberg. 2011. Sunflower domestication alleles support single domestication center in eastern North America. Proceedings of the National Academy of Sciences of the United States of America 108:14360–14365.

Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–745.

Bodenhausen, N., M. W. Horton, and J. Bergelson. 2013. Bacterial communities associated with the leaves and the roots of Arabidopsis thaliana. PloS one 8:e56329.

Borer, E. T., E. W. Seabloom, D. S. Gruner, W. S. Harpole, H. Hillebrand, E. M. Lind, P. B. Adler, J. Alberti, T. M. Anderson, J. D. Bakker, L. Biederman, D. Blumenthal, C. S. Brown, L. a Brudvig, Y. M. Buckley, M. Cadotte, C. Chu, E. E. Cleland, M. J. Crawley, P. Daleo, E. I. Damschen, K. F. Davies, N. M. DeCrappeo, G. Du, J. Firn, Y. Hautier, R. W. Heckman, A. Hector, J. HilleRisLambers, O. Iribarne, J. a Klein, J. M. H. Knops, K. J. La Pierre, A. D. B. Leakey, W. Li, A. S. MacDougall, R. L. McCulley, B. a Melbourne, C. E. Mitchell, J. L. Moore, B. Mortensen, L. R. O'Halloran, J. L. Orrock, J. Pascual, S. M. Prober, D. a Pyke, A. C. Risch, M. Schuetz, M. D. Smith, C. J. Stevens, L. L. Sullivan, R. J. Williams, P. D. Wragg, J. P. Wright, and L. H. Yang. 2014. Herbivores and nutrients control grassland plant diversity via light limitation. Nature 508:517–20.

Bouffaud, M. L., M. Kyselková, B. Gouesnard, G. Grundmann, D. Muller, and Y. Moënne-Loccoz. 2012. Is

diversification history of maize influencing selection of soil bacteria by roots? Molecular Ecology 21:195–206.

Bouffaud, M. L., M. A. Poirier, D. Muller, and Y. Moënne-Loccoz. 2014. Root microbiome relates to plant host evolution in maize and other Poaceae. Environmental Microbiology 16:2804–2814.

Bulgarelli, D., R. Garrido-Oter, P. C. Münch, A. Weiman, J. Dröge, Y. Pan, A. C. McHardy, and P. Schulze-Lefert. 2015. Structure and function of the bacterial root microbiota in wild and domesticated barley. Cell Host & Microbe 17:1–12.

Bulgarelli, D., M. Rott, K. Schlaeppi, E. Ver Loren van Themaat, N. Ahmadinejad, F. Assenza, P. Rauf, B. Huettel, R. Reinhardt, E. Schmelzer, J. Peplies, F. O. Gloeckner, R. Amann, T. Eickhorst, and P. Schulze-Lefert. 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 488:91–5.

Bulgarelli, D., K. Schlaeppi, S. Spaepen, E. Ver Loren van Themaat, and P. Schulze-Lefert. 2013. Structure and functions of the bacterial microbiota of plants. Annual review of plant biology 64:807–38.

Burns, J. H., B. L. Anacker, S. Y. Strauss, and D. J. Burke. 2015. Soil microbial community variation correlates most strongly with plant species identity, followed by soil chemistry, spatial location and plant genus. AoB PLANTS 7:1–10.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. Johnson, and S. P. Holmes. 2016. DADA2 : High resolution sample inference from amplicon data. Nature Methods 13:581–583.

Campbell, B. J., S. W. Polson, T. E. Hanson, M. C. Mack, and E. A. G. Schuur. 2010. The effect of nutrient deposition on bacterial communities in Arctic tundra soil. Environmental microbiology 12:1842–54.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, 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, and R. Knight. 2010. QIIME allows analysis of

high-throughput community sequencing data. Nature Methods 7:335–336.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer, and R. Knight. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl:4516–4522.

Cassman, N. A., M. F. A. Leite, Y. Pan, M. de Hollander, J. A. van Veen, and E. E. Kuramae. 2016. Plant and soil fungal but not soil bacterial communities are linked in long-term fertilized grassland. Scientific Reports 6:23680.

Chaparro, J. M., A. M. Sheflin, D. K. Manter, and J. M. Vivanco. 2012. Manipulating the soil microbiome to increase soil health and plant fertility. Biology and Fertility of Soils 48:489–499.

Chelius, M. K., and E. W. Triplett. 2001. The Diversity of Archaea and Bacteria in Association with the Roots of Zea mays L. Microbial ecology 41:252–263.

Chen, Y. H., R. Gols, and B. Benrey. 2015. Crop domestication and naturally selected species trophic interactions. Annual Review of Entomology 60:35–58.

Chen, Y. H., and S. C. Welter. 2005. Crop domestication disrupts a native tritrophic interaction associated with the sunflower, Helianthus annuus (Asterales: Asteraceae). Ecological Entomology 30:673–683.

Clark, C. M., and D. Tilman. 2008. Loss of plant species after chronic low-level nitrogen deposition to prairie grasslands. Nature 451:712–5.

Clarke, K., and R. Gorley. 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.

Clemmensen, K. E., A. Bahr, O. Ovaskainen, A. Dahlberg, A. Ekblad, H. Wallander, J. Stenlid, R. D. Finlay, D. A. Wardle, and B. D. Lindahl. 2013. Roots and Associated Fungi Drive Long-Term Carbon Sequestration in Boreal Forest. Science 339:1615–1618.

Cline, L. C., and D. R. Zak. 2015. Soil microbial communities are shaped by plant-driven changes in resource availability during secondary succession. Ecology 96:3374–3385.

Coleman-Derr, D., D. Desgarennes, C. Fonseca-Garcia, S. Gross, S. Clingenpeel, T. Woyke, G. North, A. Visel, L. P. Partida-Martinez, and S. Tringe. 2015. Plant compartment and biogeography affect microbiome composition in cultivated and native Agave species. New Phytologist 209:798–811.

Compant, S., C. Clément, and A. Sessitsch. 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. Soil Biology and Biochemistry 42:669–678.

Coolon, J. D., K. L. Jones, T. C. Todd, J. M. Blair, and M. A. Herman. 2013. Long-term nitrogen amendment alters the diversity and assemblage of soil bacterial communities in tallgrass prairie. PLoS ONE 8.

Cope-Selby, N., A. Cookson, M. Squance, I. Donnison, R. Flavell, and K. Farrar. 2016. Endophytic bacteria in Miscanthus seed: implications for germination, vertical inheritance of endophytes, plant evolution and breeding. GCB Bioenergy.

Costello, E. K., C. L. Lauber, M. Hamady, N. Fierer, J. I. Gordon, and R. Knight. 2009. Bacterial community variation in human body habitats across space and time. Science 326:1694–7.

Craine, J. M., C. Morrow, and W. D. Stock. 2008. Nutrient concentration ratios and co-limitation in South African grasslands. The New phytologist 179:829–36.

Crane, P. R. 2013. Ginkgo: the tree that time forgot. Yale University Press.

Csardi, G., and T. Nepusz. 2006. The igraph software package for complex network research. InterJournal Complex Sy:1695.

Decaëns, T. 2010, May. Macroecological patterns in soil communities. Blackwell Publishing Ltd.

Delmotte, N., and C. Knief. 2009. Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proceedings of the National Academy of Sciences of the United States of America 106:16428–16433.

Dempewolf, H., L. H. Rieseberg, and Q. C. Cronk. 2008. Crop domestication in the Compositae: A family-

wide trait assessment. Genetic Resources and Crop Evolution 55:1141–1157.

De Deyn, G. B., and W. H. Van Der Putten. 2005. Linking aboveground and belowground diversity. Trends in Ecology and Evolution 20:625–633.

De Deyn, G. B., C. E. Raaijmakers, J. van Ruijven, F. Berendse, and W. H. van der Putten. 2004. Plant species identity and diversity effects on different trophic levels of nematodes in the soil food web. Oikos 106:576–586.

Van Diepen, L. T. a, E. a. Lilleskov, K. S. Pregitzer, and R. M. Miller. 2007. Decline of arbuscular mycorrhizal fungi in northern hardwood forests exposed to chronic nitrogen additions. New Phytologist 176:175–183.

Dufrêne, M., and P. Legendre. 1997. Species assemblages and indicator species: The need for a flexible asymmetrical approach. Ecological Monographs 67:345–366.

Durka, W., and S. G. Michalski. 2012. Daphne: a dated phylogeny of a large European flora for phylogenetically informed ecological analyses. Ecology 93:2297.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461.

Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature methods 10:996–8.

Edwards, J., C. Johnson, C. Santos-Medellín, E. Lurie, N. K. Podishetty, S. Bhatnagar, J. A. Eisen, and V. Sundaresan. 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. Proceedings of the National Academy of Sciences of the United States of America 112:E911-20.

Egerton-Warburton, L., N. Johnson, and E. Allen. 2007. Mycorrhizal community dynamics following nitrogen fertilization: A cross-site test in five grasslands. Ecological Monographs 77:527–544.

Eilers, K. G., C. L. Lauber, R. Knight, and N. Fierer. 2010. Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. Soil Biology and

Biochemistry 42:896–903.

Ewald, P. W. 1987. Transmission modes and evolution of the parasitism-mutualism continuum. Annals of the New York Academy of Sciences 503:295–306.

Farrar, K., D. Bryant, and N. Cope-Selby. 2014. Understanding and engineering beneficial plant-microbe interactions: plant growth promotion in energy crops. Plant Biotechnology Journal 12:1193–1206.

Fierer, N., A. Barberán, and D. C. Laughlin. 2014. Seeing the forest for the genes: using metagenomics to infer the aggregated traits of microbial communities. Frontiers in Microbiology 5:1–6.

Fierer, N., M. A. Bradford, and R. B. Jackson. 2007a. Toward an ecological classification of soil bacteria. Ecology 88:1354–1364.

Fierer, N., M. A. Bradford, and R. B. Jackson. 2007b. Toward an ecological classification of soil bacteria. Ecology 88:1354–64.

Fierer, N., and R. B. Jackson. 2006. The diversity and biogeography of soil bacterial communities. Proceedings of the National Academy of Sciences of the United States of America 103:626–31.

Fierer, N., J. Ladau, J. C. Clemente, J. W. Leff, S. M. Owens, K. S. Pollard, R. Knight, J. A. Gilbert, and R. L. McCulley. 2013. Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. Science 342:621–4.

Fierer, N., C. L. Lauber, K. S. Ramirez, J. Zaneveld, M. a Bradford, and R. Knight. 2011. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. The ISME journal 6:1–11.

Fierer, N., J. W. Leff, B. J. Adams, U. N. Nielsen, S. T. Bates, C. L. Lauber, S. Owens, J. A. Gilbert, D. H. Wall, and J. G. Caporaso. 2012a. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences 109:21390–21395.

Fierer, N., J. W. Leff, B. J. Adams, U. N. Nielsen, S. T. Bates, C. L. Lauber, S. Owens, J. A. Gilbert, D. H.

Wall, and J. G. Caporaso. 2012b. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences 109:21390–21395.

Fierer, N., M. S. Strickland, D. Liptzin, M. a. Bradford, and C. C. Cleveland. 2009. Global patterns in belowground communities. Ecology Letters 12:1238–1249.

Finkel, O. M., A. Y. Burch, T. Elad, S. M. Huse, S. E. Lindow, A. F. Post, and S. Belkin. 2012. Distance-decay relationships partially determine diversity patterns of phyllosphere bacteria on Tamarix trees across the Sonoran Desert. Applied and environmental microbiology 78:6187–93.

Foley, J. A., N. Ramankutty, K. a. Brauman, E. S. Cassidy, J. S. Gerber, M. Johnston, N. D. Mueller, C. O'Connell, D. K. Ray, P. C. West, C. Balzer, E. M. Bennett, S. R. Carpenter, J. Hill, C. Monfreda, S. Polasky, J. Rockström, J. Sheehan, S. Siebert, D. Tilman, and D. P. M. Zaks. 2011. Solutions for a cultivated planet. Nature 478:337–342.

Galloway, J. N., F. J. Dentener, D. G. Capone, E. W. Boyer, R. W. Howarth, S. P. Seitzinger, G. P. Asner, C. C. Cleveland, P. A. Green, E. A. Holland, D. M. Karl, A. F. Michaels, J. H. Porter, A. R. Townsend, and C. J. Vörösmarty. 2004. Nitrogen cycles: past, present, and future. Biogeochemistry 70:153–226.

Garbeva, P., L. S. Overbeek, J. W. L. Vuurde, and J. D. Elsas. 2001. Analysis of Endophytic Bacterial Communities of Potato by Plating and Denaturing Gradient Gel Electrophoresis (DGGE) of 16S rDNA Based PCR Fragments. Microbial ecology 41:369–383.

Garbeva, P., J. A. van Veen, and J. D. van Elsas. 2004. Microbial Diversity in Soil: Selection of Microbial Populations by Plant and Soil Type and Implications for Disease Suppressiveness. Annual Review of Phytopathology 42:243–270.

García-Palacios, P., R. Milla, M. Delgado-Baquerizo, N. Martín-Robles, M. Álvaro-Sánchez, and D. H. Wall. 2013. Side-effects of plant domestication: Ecosystem impacts of changes in litter quality. New Phytologist 198:504–513.

Germida, J. J., and S. D. Siciliano. 2001. Taxonomic diversity of bacteria associated with the roots of modern, recent and ancient wheat cultivars. Biology and Fertility of Soils 33:410–415.

Gnanamanickam, S., editor. 2006. Plant-Associated Bacteria. Springer, Dordrecht.

Goslee, S. C., and D. L. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22:1–19.

Gottel, N. R., H. F. Castro, M. Kerley, Z. Yang, D. a Pelletier, M. Podar, T. Karpinets, E. Uberbacher, G. a Tuskan, R. Vilgalys, M. J. Doktycz, and C. W. Schadt. 2011. Distinct microbial communities within the endosphere and rhizosphere of Populus deltoides roots across contrasting soil types. Applied and environmental microbiology 77:5934–44.

Grayston, S. J., S. Wang, C. D. Campbell, and A. C. Edwards. 1998. Selective influence of plant species on microbial diversity in the rhizosphere. Soil Biology and Biochemistry 30:369–378.

Griffiths, R. I., B. C. Thomson, P. James, T. Bell, M. Bailey, and A. S. Whiteley. 2011. The bacterial biogeography of British soils. Environmental Microbiology 13:1642–1654.

Grime, J. P. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory.

Gubry-Rangin, C., and B. Hai. 2011. Niche specialization of terrestrial archaeal ammonia oxidizers. Proceedings of National Academy of Science 108:21206–21211.

Guillou, L., D. Bachar, S. Audic, D. Bass, C. Berney, L. Bittner, C. Boutte, G. Burgaud, C. De Vargas, J. Decelle, J. Del Campo, J. R. Dolan, M. Dunthorn, B. Edvardsen, M. Holzmann, W. H. C. F. Kooistra, E. Lara, N. Le Bescot, R. Logares, F. Mahé, R. Massana, M. Montresor, R. Morard, F. Not, J. Pawlowski, I. Probert, A. L. Sauvadet, R. Siano, T. Stoeck, D. Vaulot, P. Zimmermann, and R. Christen. 2013. The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Research 41:597–604.

Haegeman, B., J. Hamelin, J. Moriarty, P. Neal, J. Dushoff, and J. S. Weitz. 2013. Robust estimation of

microbial diversity in theory and in practice. The ISME journal 7:1092–101.

Haichar, F. E. Z., C. Marol, O. Berge, J. I. Rangel-Castro, J. I. Prosser, J. Balesdent, T. Heulin, and W. Achouak. 2008. Plant host habitat and root exudates shape soil bacterial community structure. The ISME journal 2:1221–1230.

Hardoim, P. R., F. D. Andreote, B. Reinhold-Hurek, A. Sessitsch, L. S. van Overbeek, and J. D. van Elsas. 2011. Rice root-associated bacteria: Insights into community structures across 10 cultivars. FEMS Microbiology Ecology 77:154–164.

Hardoim, P. R., C. C. P. Hardoim, L. S. van Overbeek, and J. D. van Elsas. 2012. Dynamics of seed-borne rice endophytes on early plant growth stages. PloS one 7:e30438.

Harter, A. V, K. A. Gardner, D. Falush, D. L. Lentz, R. A. Bye, and L. H. Rieseberg. 2004. Origin of extant domesticated sunflowers in eastern North America. Nature 430:201–5.

van der Heijden, M. G. A., R. D. Bardgett, and N. M. van Straalen. 2008. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecology letters 11:296–310.

van der Heijden, M. G. A., J. N. Klironomos, M. Ursic, P. Moutoglis, R. Streitwolf-Engel, T. Boller, A. Wiemken, and I. R. Sanders. 1998. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. Nature 396:69–72.

Hiiesalu, I., M. Pärtel, J. Davison, P. Gerhold, M. Metsis, M. Moora, M. Öpik, M. Vasar, M. Zobel, and S. D. Wilson. 2014. Species richness of arbuscular mycorrhizal fungi: Associations with grassland plant richness and biomass. New Phytologist 203:233–244.

Hirano, S., and C. Upper. 2000. Bacteria in the leaf ecosystem with emphasis on Pseudomonas syringae—a pathogen, ice nucleus, and epiphyte. Microbiology and Molecular Biology Reviews 64:624–653.

Hunter, P. J., P. Hand, D. Pink, J. M. Whipps, and G. D. Bending. 2010. Both leaf properties and microbe-

microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (Lactuca Species) phyllosphere. Applied and environmental microbiology 76:8117–25.

Idris, R., R. Trifonova, M. Puschenreiter, W. W. Wenzel, and A. Sessitsch. 2004. Bacterial Communities Associated with Flowering Plants of the Ni Hyperaccumulator Thlaspi goesingense. Applied and Environmental Microbiology 70:2667–2677.

Inceoglu, Ö., J. F. Salles, L. Van Overbeek, and J. D. Van Elsas. 2010. Effects of plant genotype and growth stage on the betaproteobacterial communities associated with different potato cultivars in two fields. Applied and Environmental Microbiology 76:3675–3684.

Janssens, I. A., W. Dieleman, S. Luyssaert, J.-A. Subke, M. Reichstein, R. Ceulemans, P. Ciais, A. J. Dolman, J. Grace, G. Matteucci, D. Papale, S. L. Piao, E.-D. Schulze, J. Tang, and B. E. Law. 2010. Reduction of forest soil respiration in response to nitrogen deposition. Nature Geoscience 3:315–322.

Jeffries, P., S. Gianinazzi, S. Perotto, K. Turnau, and J. M. Barea. 2003. The contribution of arbuscular mycorrhizal fungi in sustainable maintenance of plant health and soil fertility. Biology and Fertility of Soils 37:1–16.

St. John, M. G., D. H. Wall, and V. M. Behan-Pelletier. 2006. Does plant species co-occurrence influence soil mite diversity? Ecology 87:625–633.

Johnson, N. C., G. W. T. Wilson, M. a Bowker, J. a Wilson, and R. M. Miller. 2010. Resource limitation is a driver of local adaptation in mycorrhizal symbioses. Proceedings of the National Academy of Sciences of the United States of America 107:2093–2098.

Johnston-Monje, D., and M. N. Raizada. 2011. Conservation and diversity of seed associated endophytes in Zea across boundaries of evolution, ethnography and ecology. PloS one 6:e20396.

Kaga, H., H. Mano, F. Tanaka, A. Watanabe, S. Kaneko, and H. Morisaki. 2009. Rice seeds as sources of endophytic bacteria. Microbes and environments 24:154–162.

Kaiser, K., B. Wemheuer, V. Korolkow, F. Wemheuer, H. Nacke, I. Schöning, M. Schrumpf, and R. Daniel. 2016. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. Scientific Reports 6:33696.

Kane, N. C., J. M. Burke, L. Marek, G. Seiler, F. Vear, G. Baute, S. J. Knapp, P. Vincourt, and L. H. Rieseberg. 2013. Sunflower genetic, genomic and ecological resources. Molecular Ecology Resources 13:10–20.

Kanehisa, M., S. Goto, Y. Sato, M. Furumichi, and M. Tanabe. 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic acids research 40:D109-14.

Kartzinel, T. R., P. a. Chen, T. C. Coverdale, D. L. Erickson, W. J. Kress, M. L. Kuzmina, D. I. Rubenstein, W. Wang, and R. M. Pringle. 2015. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. Proceedings of the National Academy of Sciences 112:8019–8024.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

Kembel, S. W., T. K. O'Connor, H. K. Arnold, S. P. Hubbell, S. J. Wright, and J. L. Green. 2014. Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. Proceedings of the National Academy of Sciences 111:13715–13720.

Kent, W. J. 2002. BLAT---The BLAST-Like Alignment Tool. Genome Research 12:656–664.

Kim, M., D. Singh, A. Lai-Hoe, R. Go, R. A. Rahim, A. A.N., J. Chun, and J. M. Adams. 2012. Distinctive phyllosphere bacterial communities in tropical trees. Microbial ecology 63:674–81.

Kinkel, L. L. 1997. Microbial population dynamics on leaves. Annual Review of Phytopathology 35:327–347.

Klüber, H. D., and R. Conrad. 1998. Inhibitory effects of nitrate, nitrite, NO and N2O on methanogenesis by Methanosarcina barkeri and Methanobacterium bryantii. FEMS Microbiology Ecology 25:331–

339.

Konstantinidis, K. T., and J. M. Tiedje. 2004. Trends between gene content and genome size in prokaryotic species with larger genomes. Proceedings of the National Academy of Sciences of the United States of America 101:3160–5.

Koyama, A., M. D. Wallenstein, R. T. Simpson, and J. C. Moore. 2014. Soil bacterial community composition altered by increased nutrient availability in Arctic tundra soils. Frontiers in Microbiology 5:1–16.

Kuczynski, J., Z. Liu, C. Lozupone, and D. McDonald. 2010. Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. Nature methods 7:813–819.

Kuske, C. R., L. O. Ticknor, M. E. Miller, M. Dunbar, J. a Davis, S. M. Barns, J. M. Dunbar, and J. Belnap. 2002. Comparison of Soil Bacterial Communities in Rhizospheres of Three Plant Species and the Interspaces in an Arid Grassland Comparison of Soil Bacterial Communities in Rhizospheres of Three Plant Species and the Interspaces in an Arid Grassland. Applied and Environmental Microbiology 68:1854–1863.

Lambais, M., A. Lucheta, and D. Crowley. 2014. Community Assemblages Associated with the Phyllosphere, Dermosphere, and Rhizosphere of Tree Species of the Atlantic Forest are Host Taxon Dependent. Microbial Ecology 68:567–574.

Lareen, A., F. Burton, and P. Schäfer. 2016. Plant root-microbe communication in shaping root microbiomes. Plant Molecular Biology 90:575–587.

Lauber, C. L., M. Hamady, R. Knight, and N. Fierer. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Applied and environmental microbiology 75:5111–20.

Lavorel, S., and E. Garnier. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Functional Ecology 16:545–556.

LeBauer, D., and K. Treseder. 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecology 89:371–379.

Leff, J. W. 2016. mctoolsr: Microbial Community Data Analysis Tools. https://github.com/leffj/mctoolsr.

Leff, J. W., and N. Fierer. 2013. Bacterial communities associated with the surfaces of fresh fruits and vegetables. PLoS ONE 8:e59310.

Leff, J. W., S. E. Jones, S. M. Prober, A. Barberan, E. T. Borer, J. L. Firn, W. S. Harpole, S. E. Hobbie, K. S. Hofmockel, J. M. H. Knops, R. L. McCulley, K. La Pierre, A. C. Risch, E. W. Seabloom, M. Schuetz, C. Steenbock, C. J. Stevens, and N. Fierer. 2015. Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. Proceedings of the National Academy of Sciences of the United States of America 112:10967–10972.

Lehmann, A., E. K. Barto, J. R. Powell, and M. C. Rillig. 2012. Mycorrhizal responsiveness trends in annual crop plants and their wild relatives-a meta-analysis on studies from 1981 to 2010. Plant and Soil 355:231–250.

Leigh, J. A. 2000. Nitrogen fixation in methanogens: the archaeal perspective. Current issues in molecular biology 2:125–31.

Leininger, S., T. Urich, M. Schloter, L. Schwark, J. Qi, G. W. Nicol, J. I. Prosser, S. C. Schuster, and C. Schleper. 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442:806–9.

Leiser, W. L., M. O. Olatoye, H. F. W. Rattunde, G. Neumann, E. Weltzien, and B. I. G. Haussmann. 2016. No need to breed for enhanced colonization by arbuscular mycorrhizal fungi to improve low-P adaptation of West African sorghums. Plant and Soil 401:51–64.

Lekberg, Y., and L. P. Waller. 2016. What drives differences in arbuscular mycorrhizal fungal communities among plant species? Fungal Ecology:10–13.

Lichstein, J. W. 2007. Multiple regression on distance matrices: A multivariate spatial analysis tool. Plant

Ecology 188:117–131.

Lindow, S. E., and M. T. Brandl. 2003. Microbiology of the Phyllosphere. Applied and environmental microbiology 69:1875–1883.

Lindow, S. E., and J. H. . Leveau. 2002. Phyllosphere microbiology. Current Opinion in Biotechnology 13:238–243.

Locey, K. J., and J. T. Lennon. 2016. Scaling laws predict global microbial diversity. Proceedings of the National Academy of Sciences 113:5970–5975.

Lozupone, C., M. E. Lladser, D. Knights, J. Stombaugh, and R. Knight. 2011a. UniFrac: an effective distance metric for microbial community comparison. The ISME journal 5:169–172.

Lozupone, C., M. E. Lladser, D. Knights, J. Stombaugh, and R. Knight. 2011b. UniFrac: an effective distance metric for microbial community comparison. The ISME journal 5:169–72.

Lundberg, D. S., S. L. Lebeis, S. H. Paredes, S. Yourstone, J. Gehring, S. Malfatti, J. Tremblay, A. Engelbrektson, V. Kunin, T. G. Del Rio, R. C. Edgar, T. Eickhorst, R. E. Ley, P. Hugenholtz, S. G. Tringe, and J. L. Dangl. 2012. Defining the core Arabidopsis thaliana root microbiome. Nature 488:86–90.

Lundberg, D. S., S. Yourstone, P. Mieczkowski, C. D. Jones, and J. L. Dangl. 2013. Practical innovations for high-throughput amplicon sequencing. Nature methods 10:999–1002.

Magoč, T., and S. L. Salzberg. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics (Oxford, England) 27:2957–63.

Markowitz, V. M., I.-M. a Chen, K. Palaniappan, K. Chu, E. Szeto, Y. Grechkin, A. Ratner, B. Jacob, J. Huang, P. Williams, M. Huntemann, I. Anderson, K. Mavromatis, N. N. Ivanova, and N. C. Kyrpides. 2012. IMG: the Integrated Microbial Genomes database and comparative analysis system. Nucleic acids research 40:D115-22.

Marques, J. M., T. F. da Silva, R. E. Vollu, A. F. Blank, G. C. Ding, L. Seldin, and K. Smalla. 2014. Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown

sweet potato plants. FEMS Microbiology Ecology 88:424–435.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10.

Martins, G., B. Lauga, C. Miot-Sertier, A. Mercier, A. Lonvaud, M.-L. Soulas, G. Soulas, and I. Masneuf-Pomarède. 2013. Characterization of epiphytic bacterial communities from grapes, leaves, bark and soil of grapevine plants grown, and their relations. PloS ONE 8:e73013.

Mayrose, M., N. C. Kane, I. Mayrose, K. M. Dlugosch, and L. H. Rieseberg. 2011. Increased growth in sunflower correlates with reduced defences and altered gene expression in response to biotic and abiotic stress. Molecular Ecology 20:4683–4694.

McDonald, D., M. N. Price, J. Goodrich, E. P. Nawrocki, T. Z. DeSantis, A. Probst, G. L. Andersen, R. Knight, and P. Hugenholtz. 2012a. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME Journal 6:610–618.

McDonald, D., M. N. Price, J. Goodrich, E. P. Nawrocki, T. Z. DeSantis, A. Probst, G. L. Andersen, R. Knight, and P. Hugenholtz. 2012b. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME journal 6:610–8.

McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology from functional traits. Trends in ecology & evolution (Personal edition) 21:178–85.

McGuire, K. L., S. G. Payne, M. I. Palmer, C. M. Gillikin, D. Keefe, S. J. Kim, S. M. Gedallovich, J. Discenza, R. Rangamannar, J. A. Koshner, A. L. Massmann, G. Orazi, A. Essene, J. W. Leff, and N. Fierer. 2013. Digging the New York City Skyline: soil fungal communities in green roofs and city parks. PloS one 8:e58020.

Mehrabi, Z., and S. L. Tuck. 2015. Relatedness is a poor predictor of negative plant-soil feedbacks. New Phytologist 205:1071–1075.

Mei, C., and B. S. Flinn. 2010. The use of beneficial microbial endophytes for plant biomass and stress

tolerance improvement. Recent patents on biotechnology 4:81–95.

Meier, C. L., and W. D. Bowman. 2008. Links between plant litter chemistry, species diversity, and below-ground ecosystem function. Proceedings of the National Academy of Sciences of the United States of America 105:19780–5.

Meyer, K. M., and J. H. J. Leveau. 2012. Microbiology of the phyllosphere: a playground for testing ecological concepts. Oecologia 168:621–9.

Milla, R., C. P. Osborne, M. M. Turcotte, and C. Violle. 2015. Plant domestication through an ecological lens. Trends in Ecology and Evolution 30:463–469.

Neher, D. A. 2001. Role of nematodes in soil health and their use as indicators. Journal of nematology 33:161–168.

Nelson, E. B. 2004. Microbial Dynamics and Interactions in the Spermosphere. Annual Review of Phytopathology 42:271–309.

Nguyen, N. H., Z. Song, S. T. Bates, S. Branco, L. Tedersoo, J. Menke, J. S. Schilling, and P. G. Kennedy. 2015. FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. Fungal Ecology 20:241–248.

Nielsen, U. N., E. Ayres, D. H. Wall, and R. D. Bardgett. 2011, February. Soil biodiversity and carbon cycling: A review and synthesis of studies examining diversity-function relationships. Blackwell Publishing Ltd.

Nocker, A., P. Sossa-Fernandez, M. D. Burr, and A. K. Camper. 2007. Use of propidium monoazide for live/dead distinction in microbial ecology. Applied and Environmental Microbiology 73:5111–5117.

Nunan, N., T. J. Daniell, B. K. Singh, A. Papert, J. W. McNicol, and J. I. Prosser. 2005. Links between plant and rhizoplane bacterial communities in grassland soils, characterized using molecular techniques. Applied and Environmental Microbiology 71:6784–6792.

Ofek, M., M. Voronov-Goldman, Y. Hadar, and D. Minz. 2014. Host signature effect on plant root-

associated microbiomes revealed through analyses of resident vs. active communities. Environmental Microbiology 16:2157–2167.

Offre, P., A. Spang, and C. Schleper. 2013. Archaea in biogeochemical cycles. Annual review of microbiology 67:437–57.

Oger, P. M., H. Mansouri, X. Nesme, and Y. Dessaux. 2004. Engineering root exudation of Lotus toward the production of two novel carbon compounds leads to the selection of distinct microbial populations in the rhizosphere. Microbial Ecology 47:96–103.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2016. vegan: Community Ecology Package. https://cran.r-project.org/package=vegan.

Orwin, K. H., S. M. Buckland, D. Johnson, B. L. Turner, S. Smart, S. Oakley, and R. D. Bardgett. 2010. Linkages of plant traits to soil properties and the functioning of temperate grassland. Journal of Ecology 98:1074–1083.

Ottesen, A. R., A. González Peña, J. R. White, J. B. Pettengill, C. Li, S. Allard, S. Rideout, M. Allard, T. Hill, P. Evans, E. Strain, S. Musser, R. Knight, and E. Brown. 2013. Baseline survey of the anatomical microbial ecology of an important food plant: Solanum lycopersicum (tomato). BMC microbiology 13:114.

Pan, Y., N. Cassman, M. de Hollander, L. Mendes, H. Korevaar, R. Geerts, J. van Veen, and E. Kuramae. 2014. Impact of long term N, P, K and NPK fertilization on the composition and potential functions of the bacterial community in grassland soil. FEMS microbiology ecology.

Panke-buisse, K., A. C. Poole, J. K. Goodrich, R. E. Ley, and J. Kao-kniffin. 2014. Selection on soil microbiomes reveals reproducible impacts on plant function. The ISME Journal 9:980–989.

Peay, K. G., C. Baraloto, and P. V. A. Fine. 2013. Strong coupling of plant and fungal community structure across western Amazonian rainforests. The ISME journal 7:1852–61.

Peiffer, J. A., A. Spor, O. Koren, Z. Jin, S. G. Tringe, J. L. Dangl, E. S. Buckler, and R. E. Ley. 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. Proceedings of the National Academy of Sciences of the United States of America 110:6548–53.

Peñuelas, J., B. Poulter, J. Sardans, P. Ciais, M. van der Velde, L. Bopp, O. Boucher, Y. Godderis, P. Hinsinger, J. Llusia, E. Nardin, S. Vicca, M. Obersteiner, and I. a Janssens. 2013. Human-induced nitrogen-phosphorus imbalances alter natural and managed ecosystems across the globe. Nature communications 4:2934.

Peñuelas, J., and J. Terradas. 2014. The foliar microbiome. Trends in plant science 19:278–80.

Pérez-Jaramillo, J. E., R. Mendes, and J. M. Raaijmakers. 2015. Impact of plant domestication on rhizosphere microbiome assembly and functions. Plant molecular biology 90:635–644.

Philippot, L., J. M. Raaijmakers, P. Lemanceau, and W. H. van der Putten. 2013. Going back to the roots: the microbial ecology of the rhizosphere. Nature reviews. Microbiology 11:789–99.

Pianka, E. 1970. On r-and K-selection. American naturalist 104:592–597.

Pieterse, C. M. J., R. de Jonge, and R. L. Berendsen. 2016. The soil-borne supremacy. Trends in Plant Science 21:171–173.

Pitzschke, A. 2016. Developmental peculiarities and seed-borne endophytes in Quinoa: Omnipresent, robust bacilli contribute to plant fitness. Frontiers in Microbiology 7:1–15.

Porazinska, D. L., R. D. Bardgett, M. B. Blaauw, H. W. Hunt, A. N. Parsons, T. R. Seastedt, and D. H. Wall. 2003. Relationships at the Aboveground-Belowground Interface: Plants, Soil Biota, and Soil Processes. Ecological Monographs 73:377–395.

Prasifka, J. R., O. Spring, J. Conrad, L. W. Cook, D. E. Palmquist, and M. E. Foley. 2015. Sesquiterpene lactone composition of wild and cultivated sunflowers and biological activity against an insect pest. Journal of Agricultural and Food Chemistry 63:4042–4049.

Prober, S. M., J. W. Leff, S. T. Bates, E. T. Borer, J. Firn, W. S. Harpole, E. M. Lind, E. W. Seabloom, P. B.

Adler, J. D. Bakker, and others. 2015. Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. Ecology letters 18:85–95.

Purugganan, M. D., and D. Q. Fuller. 2009. The nature of selection during plant domestication. Nature 457:843–848.

Van der Putten, W. H., R. D. Bardgett, J. D. Bever, T. M. Bezemer, B. B. Casper, T. Fukami, P. Kardol, J. N. Klironomos, A. Kulmatiski, J. A. Schweitzer, K. N. Suding, T. F. J. Van de Voorde, and D. A. Wardle. 2013. Plant-soil feedbacks: The past, the present and future challenges. Journal of Ecology 101:265–276.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Research 41:590–596.

R Core Team. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramirez, K., J. W. Leff, A. Barberán, S. Bates, J. Betley, T. Crowther, E. Kelly, E. Oldfield, E. Shaw, C. Steenbock, M. Bradford, D. Wall, and N. Fierer. 2014a. Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. Proceedings of the Royal Society B 281:20141988.

Ramirez, K. S., J. M. Craine, and N. Fierer. 2012. Consistent effects of nitrogen amendments on soil microbial communities and processes across biomes. Global Change Biology 18:1918–1927.

Ramirez, K. S., C. L. Lauber, R. Knight, M. A. Bradford, and N. Fierer. 2010. Consistent effects of nitrogen fertilization on soil bacterial communities in contrasting systems. Ecology 91:3463–3470.

Ramirez, K. S., J. W. Leff, A. Barberán, S. T. Bates, J. Betley, T. W. Crowther, E. F. Kelly, E. E. Oldfield, E. A.

Shaw, C. Steenbock, M. A. Bradford, D. H. Wall, and N. Fierer. 2014b. Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. Proceedings of the Royal Society B: Biological Sciences 281:1–9.

Rastogi, G., G. L. Coaker, and J. H. J. Leveau. 2013. New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. FEMS microbiology letters 348:1–10.

Rastogi, G., A. Sbodio, J. J. Tech, T. V Suslow, G. L. Coaker, and J. H. J. Leveau. 2012. Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. The ISME journal:1–11.

Redecker, D., and P. Raab. 2006. Phylogeny of the glomeromycota (arbuscular mycorrhizal fungi): recent developments and new gene markers. Mycologia 98:885–95.

Redford, A. J., R. M. Bowers, R. Knight, Y. Linhart, and N. Fierer. 2010. The ecology of the phyllosphere : geographic and phylogenetic variability in the distribution of bacteria. Environmental microbiology 12:2885–2893.

Redford, A. J., and N. Fierer. 2009. Bacterial succession on the leaf surface: a novel system for studying successional dynamics. Microbial ecology 58:189–98.

Reich, P. B. 2014. The world-wide "fast–slow" plant economics spectrum: a traits manifesto. Journal of Ecology 102:275–301.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.

Rodriguez, R. J., J. Henson, E. Van Volkenburgh, M. Hoy, L. Wright, F. Beckwith, Y.-O. Kim, and R. S. Redman. 2008. Stress tolerance in plants via habitat-adapted symbiosis. The ISME journal 2:404–416.

Roesch, L., R. Fulthorpe, A. Riva, G. Casella, A. Hadwin, A. Kent, S. Daroub, F. Camargo, W. Farmerie, and

E. Triplett. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. The ISME Journal 1:283–290.

Rosenblueth, M., and E. Martínez-Romero. 2006. Bacterial endophytes and their interactions with hosts. Molecular plant-microbe interactions 19:827–837.

Roskov, Y., L. Abucay, T. Orrell, D. Nicolson, C. Flann, N. Bailly, P. Kirk, T. Bourgoin, R. E. DeWalt, W. Decock, and A. De Wever. 2016. 2016 Annual Checklist. http://www.catalogueoflife.org/annual-checklist/2016/info/totals.

Roumet, C., M. Birouste, C. Picon-Cochard, M. Ghestem, N. Osman, S. Vrignon-Brenas, K. fang Cao, and A. Stokes. 2016. Root structure-function relationships in 74 species: Evidence of a root economics spectrum related to carbon economy. New Phytologist.

Rousk, J., P. C. Brookes, H. C. Glanville, and D. L. Jones. 2011. Lack of correlation between turnover of low-molecular-weight dissolved organic carbon and differences in microbial community composition or growth across a soil pH gradient. Applied and environmental microbiology 77:2791–5.

Rudgers, J. A., M. E. Afkhami, M. A. Rúa, A. J. Davitt, S. Hammer, and V. M. Huguet. 2009. A fungus among us: Broad patterns of endophyte distribution in the grasses. Ecology 90:1531–1539.

Sawers, R. J. H., C. Gutjahr, and U. Paszkowski. 2008. Cereal mycorrhiza: an ancient symbiosis in modern agriculture. Trends in Plant Science 13:93–97.

Scheffers, B. R., B. L. Phillips, W. F. Laurance, N. S. Sodhi, A. Diesmos, and S. E. Williams. 2013. Increasing arboreality with altitude: a novel biogeographic dimension. Proceedings. Biological sciences / The Royal Society 280:20131581.

Schimel, J. P., and S. M. Schaeffer. 2012. Microbial control over carbon cycling in soil. Frontiers.

Schlaeppi, K., N. Dombrowski, R. G. Oter, E. Ver Loren van Themaat, and P. Schulze-Lefert. 2014. Quantitative divergence of the bacterial root microbiota in Arabidopsis thaliana relatives.

Proceedings of the National Academy of Sciences of the United States of America 111:585–92.

Schlesinger, W. H., and E. S. Bernhardt. 2013. Biogeochemistry: an analysis of global change. Academic press.

Schmidt, J. E., T. M. Bowles, and A. C. M. Gaudin. 2016. Using ancient traits to convert soil health into crop yield: impact of selection on maize root and rhizosphere function. Frontiers in Plant Science 7:1–11.

Schmieder, R., and R. Edwards. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics (Oxford, England) 27:863–4.

Semchenko, M., S. Saar, and A. Lepik. 2014. Plant root exudates mediate neighbour recognition and trigger complex behavioural changes. The New phytologist 204:631–7.

Shi, S., A. E. Richardson, M. O'Callaghan, K. M. Deangelis, E. E. Jones, A. Stewart, M. K. Firestone, and L. M. Condron. 2011. Effects of selected root exudate components on soil bacterial communities. FEMS Microbiology Ecology 77:600–610.

Singh, B. K., P. Millard, A. S. Whiteley, and J. C. Murrell. 2004. Unravelling rhizosphere–microbial interactions: opportunities and limitations. Trends in Microbiology 12:386–393.

Singh, B. K., S. Munro, J. M. Potts, and P. Millard. 2007. Influence of grass species and soil type on rhizosphere microbial community structure in grassland soils. Applied Soil Ecology 36:147–155.

Singh, L. P., S. S. Gill, and N. Tuteja. 2011. Unraveling the role of fungal symbionts in plant abiotic stress tolerance. Plant signaling & behavior 6:175–191.

Smith, B. D. 2014. The domestication of Helianthus annuus L. (sunflower). Vegetation History and Archaeobotany 23:57–74.

Smith, D. P., and K. G. Peay. 2014. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. PloS one 9:e90234.

Souza, R. S. C. de, V. K. Okura, J. S. L. Armanhi, B. Jorrín, N. Lozano, M. J. da Silva, M. González-Guerrero,

L. M. de Araújo, N. V. Ferreira, H. C. Bagheri, J. Imperial, and P. Arruda. 2016. Unlocking the bacterial and fungal communities assemblages of sugarcane microbiome. Scientific Reports:28774.

Stork, N., and P. Hammond. 2001. The spatial distribution of beetles within the canopies of oak trees in Richmond Park , U.K. Ecological Entomology 26:302–311.

Strickland, M., C. Lauber, N. Fierer, and M. Bradford. 2009. Testing the functional significance of microbial community composition. Ecology 90:441–451.

Suda, J., L. a Meyerson, I. J. Leitch, and P. Py. 2014. The hidden side of plant invasions : the role of genome size. New Phytologist.

Suding, K. N., S. L. Collins, L. Gough, C. Clark, E. E. Cleland, K. L. Gross, D. G. Milchunas, and S. Pennings. 2005. Functional- and abundance-based mechanisms explain diversity loss due to N fertilization. Proceedings of the National Academy of Sciences of the United States of America 102:4387–92.

Szoboszlay, M., J. Lambers, J. Chappell, J. V. Kupper, L. A. Moe, and D. H. McNear. 2015. Comparison of root system architecture and rhizosphere microbial communities of Balsas teosinte and domesticated corn cultivars. Soil Biology and Biochemistry 80:34–44.

Taberlet, P., E. Coissac, F. Pompanon, L. Gielly, C. Miquel, A. Valentini, T. Vermat, G. Corthier, C. Brochmann, and E. Willerslev. 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Research 35:e14–e14.

Taberlet, P., S. M. Prud'Homme, E. Campione, J. Roy, C. Miquel, W. Shehzad, L. Gielly, D. Rioux, P. Choler, J. C. Clément, C. Melodelima, F. Pompanon, and E. Coissac. 2012. Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. Molecular Ecology 21:1816–1820.

Tamaki, H., Y. Tanaka, H. Matsuzawa, M. Muramatsu, X.-Y. Meng, S. Hanada, K. Mori, and Y. Kamagata. 2011. Armatimonas rosea gen. nov., sp. nov., of a novel bacterial phylum, Armatimonadetes phyl. nov., formally called the candidate phylum OP10. International journal of systematic and

evolutionary microbiology 61:1442–7.

Tedersoo, L., M. Bahram, T. Cajthaml, S. Põlme, I. Hiiesalu, S. Anslan, H. Harend, F. Buegger, K. Pritsch, J. Koricheva, and K. Abarenkov. 2015. Tree diversity and species identity effects on soil fungi, protists and animals are context dependent. The ISME Journal:1–17.

Tedersoo, L., M. Bahram, S. Polme, U. Koljalg, N. S. Yorou, R. Wijesundera, L. V. Ruiz, a. M. Vasco-Palacios, P. Q. Thu, A. Suija, M. E. Smith, C. Sharp, E. Saluveer, A. Saitta, M. Rosas, T. Riit, D. Ratkowsky, K. Pritsch, K. Poldmaa, M. Piepenbring, C. Phosri, M. Peterson, K. Parts, K. Partel, E. Otsing, E. Nouhra, a. L. Njouonkou, R. H. Nilsson, L. N. Morgado, J. Mayor, T. W. May, L. Majuakim, D. J. Lodge, S. S. Lee, K.-H. Larsson, P. Kohout, K. Hosaka, I. Hiiesalu, T. W. Henkel, H. Harend, L. -d. Guo, A. Greslebin, G. Grelet, J. Geml, G. Gates, W. Dunstan, C. Dunk, R. Drenkhan, J. Dearnaley, A. De Kesel, T. Dang, X. Chen, F. Buegger, F. Q. Brearley, G. Bonito, S. Anslan, S. Abell, and K. Abarenkov. 2014a. Global diversity and geography of soil fungi. Science 346:1256688–1256688.

Tedersoo, L., M. Bahram, S. Polme, U. Koljalg, N. S. Yorou, R. Wijesundera, L. V. Ruiz, a. M. Vasco-Palacios, P. Q. Thu, a. Suija, M. E. Smith, C. Sharp, E. Saluveer, a. Saitta, M. Rosas, T. Riit, D. Ratkowsky, K. Pritsch, K. Poldmaa, M. Piepenbring, C. Phosri, M. Peterson, K. Parts, K. Partel, E. Otsing, E. Nouhra, a. L. Njouonkou, R. H. Nilsson, L. N. Morgado, J. Mayor, T. W. May, L. Majuakim, D. J. Lodge, S. S. Lee, K.-H. Larsson, P. Kohout, K. Hosaka, I. Hiiesalu, T. W. Henkel, H. Harend, L. -d. Guo, a. Greslebin, G. Grelet, J. Geml, G. Gates, W. Dunstan, C. Dunk, R. Drenkhan, J. Dearnaley, a. De Kesel, T. Dang, X. Chen, F. Buegger, F. Q. Brearley, G. Bonito, S. Anslan, S. Abell, and K. Abarenkov. 2014b. Global diversity and geography of soil fungi. Science.

Tesařová, M., K. Fiala, and V. Studený. 1982. Live and Dead Roots - Their Mass Ratio in Several Grassland Stands. Folia Geobotanica et Phytotaxonomica 17:427–430.

Tilman, D., and D. Wedin. 1991. Plant traits and resource reduction for five grasses growing on a nitrogen gradient. Ecology 72:685–700.

Treseder, K. 2004. A meta-analysis of mycorrhizal responses to nitrogen, phosphorus, and atmospheric CO2 in field studies. New Phytologist 164:347–355.

Treseder, K. K. 2008. Nitrogen additions and microbial biomass: a meta-analysis of ecosystem studies. Ecology letters 11:1111–20.

Truyens, S., N. Weyens, A. Cuypers, and J. Vangronsveld. 2015. Bacterial seed endophytes: genera, vertical transmission and interaction with plants. Environmental Microbiology Reports 7:40–50.

Turcotte, M. M., A. K. Lochab, N. E. Turley, and M. T. J. Johnson. 2015. Plant domestication slows pest evolution. Ecology Letters 18:907–915.

Turner, T. R., E. K. James, and P. S. Poole. 2013. The plant microbiome. Genome biology 14:209.

Turrini, A., T. Giordani, L. Avio, L. Natali, M. Giovannetti, and A. Cavallini. 2016. Large variation in mycorrhizal colonization among wild accessions, cultivars, and inbreds of sunflower (Helianthus annuus L.). Euphytica 207:331–342.

Uriarte, M., N. G. Swenson, R. L. Chazdon, L. S. Comita, W. John Kress, D. Erickson, J. Forero-Montaña, J. K. Zimmerman, and J. Thompson. 2010. Trait similarity, shared ancestry and the structure of neighbourhood interactions in a subtropical wet forest: implications for community assembly. Ecology Letters 13:1503–1514.

Vaidehi, B. K., C. Manoharachary, D. K. Purohit, S. Ram Reddy, S. Charya, and S. Girisham. 2002. Seed mycoflora of sunflower - a perspective. Pages 25–40Frontiers in microbial biotechnology and plant pathology. Scientific Publishers, Jodhpur.

Vandenkoornhuyse, P., A. Quaiser, M. Duhamel, A. Le Van, and A. Dufresne. 2015. The importance of the microbiome of the plant holobiont. New Phytologist 206:1196–1206.

Verbon, E. H., and L. M. Liberman. 2016. Beneficial Microbes Affect Endogenous Mechanisms Controlling Root Development. Trends in Plant Science 21:218–229.

Vieira-Silva, S., and E. P. C. Rocha. 2010. The systemic imprint of growth and its uses in ecological

(meta)genomics. PLoS genetics 6:e1000808.

Viketoft, M., C. Palmborg, B. Sohlenius, K. Huss-Danell, and J. Bengtsson. 2005. Plant species effects on soil nematode communities in experimental grasslands. Applied Soil Ecology 30:90–103.

Vorholt, J. A. 2012. Microbial life in the phyllosphere. Nature Reviews Microbiology 10:828–840.

de Vries, F. T., P. Manning, J. R. B. Tallowin, S. R. Mortimer, E. S. Pilgrim, K. A. Harrison, P. J. Hobbs, H. Quirk, B. Shipley, J. H. C. Cornelissen, J. Kattge, and R. D. Bardgett. 2012. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. Ecology Letters 15:1230–1239.

Wagg, C., S. F. Bender, F. Widmer, and M. G. A. van der Heijden. 2014. Soil biodiversity and soil community composition determine ecosystem multifunctionality. Proceedings of the National Academy of Sciences of the United States of America 111:5266–70.

Wagner, M. R., D. S. Lundberg, T. G. del Rio, S. G. Tringe, J. L. Dangl, and T. Mitchell-Olds. 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nature communications 7:1–15.

Waldrop, M. P., T. C. Balser, and M. K. Firestone. 2000. Linking microbial community composition to function in a tropical soil. Soil Biology & Biochemistry 32:1837–1846.

Wall, D. H., U. N. Nielsen, and J. Six. 2015. Soil biodiversity and human health. Nature 528:69–76.

Walters, W., E. R. Hyde, D. Berg-lyons, G. Ackermann, G. Humphrey, A. Parada, J. a Gilbert, and J. K. Jansson. 2015. Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems 1:e0009-15.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007a. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73:5261–5267.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007b. Naive Bayesian classifier for rapid assignment

of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology 73:5261–7.

Wang, R., Y. Balkanski, O. Boucher, P. Ciais, J. Peñuelas, and S. Tao. 2015. Significant contribution of combustion-related emissions to the atmospheric phosphorus budget. Nature Geoscience 8:48–54.

Ward, N. L., J. F. Challacombe, P. H. Janssen, B. Henrissat, P. M. Coutinho, M. Wu, G. Xie, D. H. Haft, M. Sait, J. Badger, R. D. Barabote, B. Bradley, T. S. Brettin, L. M. Brinkac, D. Bruce, T. Creasy, S. C. Daugherty, T. M. Davidsen, R. T. DeBoy, J. C. Detter, R. J. Dodson, a. S. Durkin, A. Ganapathy, M. Gwinn-Giglio, C. S. Han, H. Khouri, H. Kiss, S. P. Kothari, R. Madupu, K. E. Nelson, W. C. Nelson, I. Paulsen, K. Penn, Q. Ren, M. J. Rosovitz, J. D. Selengut, S. Shrivastava, S. a. Sullivan, R. Tapia, S. Thompson, K. L. Watkins, Q. Yang, C. Yu, N. Zafar, L. Zhou, and C. R. Kuske. 2009. Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. Applied and Environmental Microbiology 75:2046–2056.

Wardle, D. A. 2006. The influence of biotic interactions on soil biodiversity. Ecology Letters 9:870–886.

Wardle, D. A., R. D. Bardgett, J. N. Klironomos, H. Setälä, W. H. van der Putten, and D. H. Wall. 2004. Ecological linkages between aboveground and belowground biota. Science 304:1629–33.

Wei, C., Q. Yu, E. Bai, X. Lü, Q. Li, J. Xia, P. Kardol, W. Liang, Z. Wang, and X. Han. 2013. Nitrogen deposition weakens plant-microbe interactions in grassland ecosystems. Global change biology 19:3688–97.

Weinert, N., Y. Piceno, G. C. Ding, R. Meincke, H. Heuer, G. Berg, M. Schloter, G. Andersen, and K. Smalla. 2011. PhyloChip hybridization uncovered an enormous bacterial diversity in the rhizosphere of different potato cultivars: Many common and few cultivar-dependent taxa. FEMS Microbiology Ecology 75:497–506.

Wessén, E., K. Nyberg, J. K. Jansson, and S. Hallin. 2010. Responses of bacterial and archaeal ammonia oxidizers to soil organic and fertilizer amendments under long-term management. Applied Soil

Ecology 45:193–200.

Whipps, J. M., P. Hand, D. Pink, and G. D. Bending. 2008. Phyllosphere microbiology with special reference to diversity and plant genotype. Journal of applied microbiology 105:1744–55.

White, T. J., T. Bruns, S. Lee, and J. W. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315–322*in* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, editors.PCR Protocols: A Guide to Methods and Applications. Academic Press, San Diego.

Wieder, W. R., G. B. Bonan, and S. D. Allison. 2013. Global soil carbon projections are improved by modelling microbial processes. Nature Climate Change 3:909–912.

Wieland, G., R. Neumann, and H. Backhaus. 2001. Variation of microbial communities in soil, rhizosphere, and rhizoplane in response to crop species, soil type, and crop development. Applied and Environmental Microbiolgy 67:5849–5854.

Wissuwa, M., M. Mazzola, and C. Picard. 2009. Novel approaches in plant breeding for rhizosphere-related traits. Plant and Soil 321:409–430.

Wu, T., E. Ayres, R. D. Bardgett, D. H. Wall, and J. R. Garey. 2011. Molecular study of worldwide distribution and diversity of soil animals. Proceedings of the National Academy of Sciences 108:17720–17725.

Xing, X., A. M. Koch, A. M. P. Jones, D. Ragone, S. Murch, and M. M. Hart. 2012. Mutualism breakdown in breadfruit domestication. Proceedings of the Royal Society B: Biological Sciences 279:1122–1130.

Yang, C. H., D. E. Crowley, J. Borneman, and N. T. Keen. 2001. Microbial phyllosphere populations are more complex than previously realized. Proceedings of the National Academy of Sciences of the United States of America 98:3889–94.

Yashiro, E., R. N. Spear, and P. S. McManus. 2011. Culture-dependent and culture-independent assessment of bacteria in the apple phyllosphere. Journal of applied microbiology 110:1284–96.

Yoccoz, N. G., K. A. Bråthen, L. Gielly, J. Haile, M. E. Edwards, T. Goslar, H. Von Stedingk, A. K. Brysting, E. Coissac, F. Pompanon, J. H. Sonstebo, C. Miquel, A. Valentini, F. De Bello, J. Chave, W. Thuiller, P. Wincker, C. Cruaud, F. Gavory, M. Rasmussen, M. T. P. Gilbert, L. Orlando, C. Brochmann, E. Willerslev, and P. Taberlet. 2012. DNA from soil mirrors plant taxonomic and growth form diversity. Molecular Ecology 21:3647–3655.

Zhu, Y. G., S. E. Smith, A. R. Barritt, and F. A. Smith. 2001. Phosphorus (P) efficiencies and mycorrhizal responsiveness of old and modern wheat cultivars. Plant and Soil 237:249–255.

APPENDIX

CHAPTER II APPENDIX

Figure A2.1. Aerial photograph of the site at the Arnold Arboretum in Boston, MA, USA where ginkgo trees were sampled. Photograph from google.com/maps.

Figure A2.2. Photograph of Ginkgo_2.

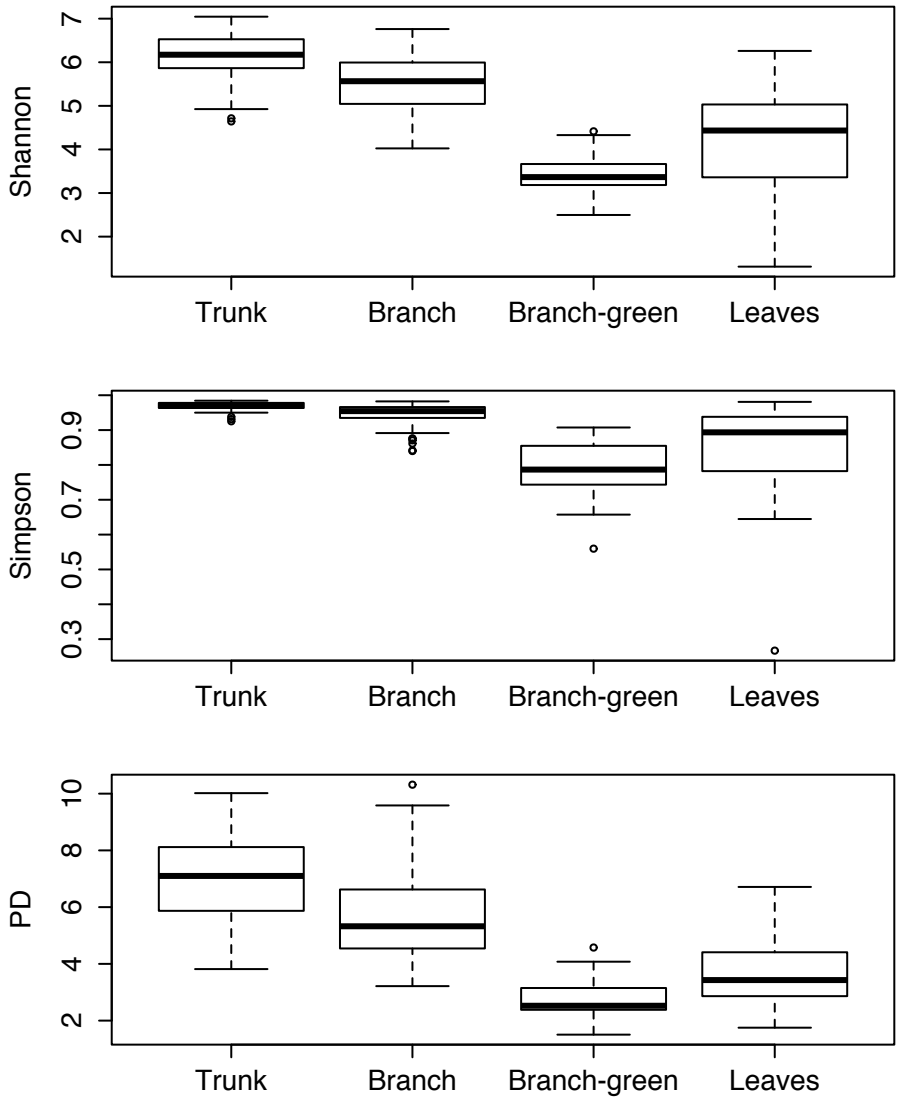

Figure A2.3. Boxplots showing values for three different bacterial diversity metrics across the various ginkgo locations sampled. PD = phylogenetic diversity.

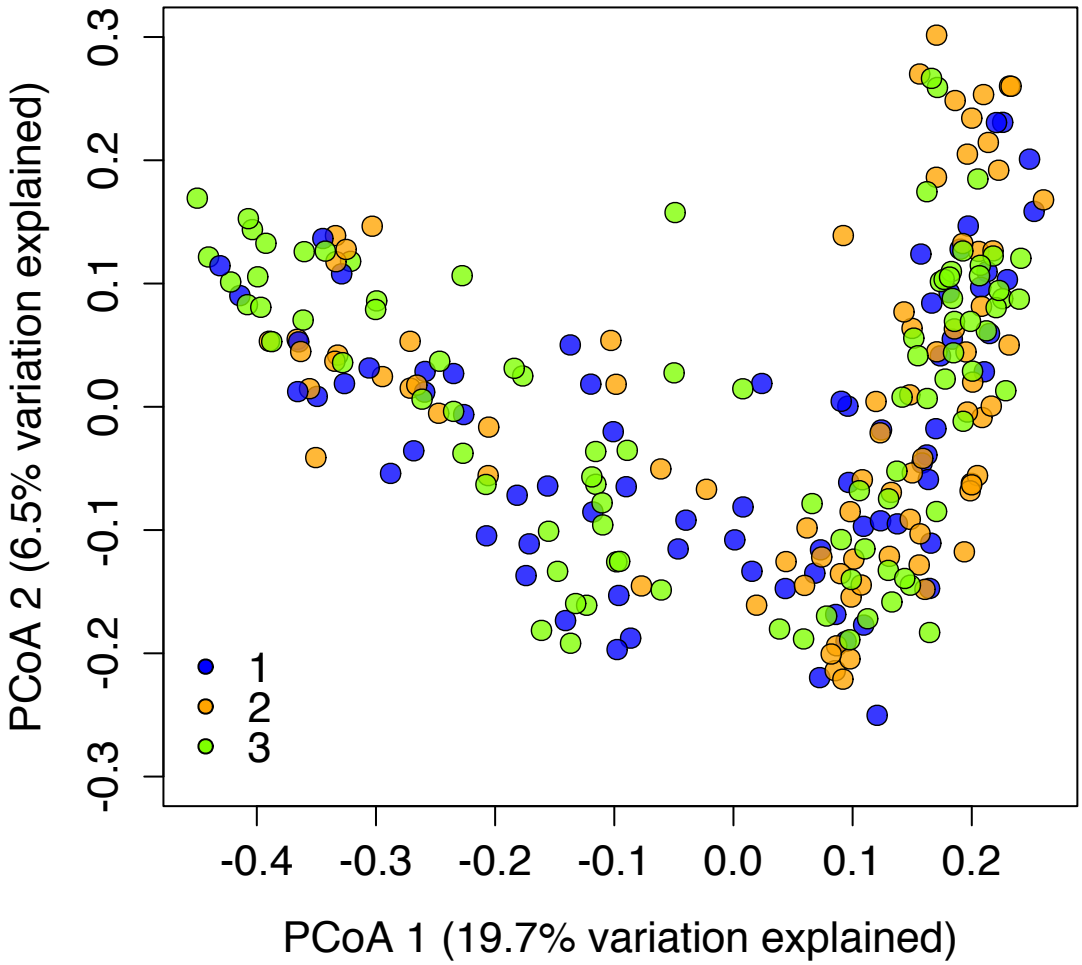

Figure A2.4. Principal coordinate analysis plot showing all samples as points colored by individual tree. This ordination was created using unweighted UniFrac.

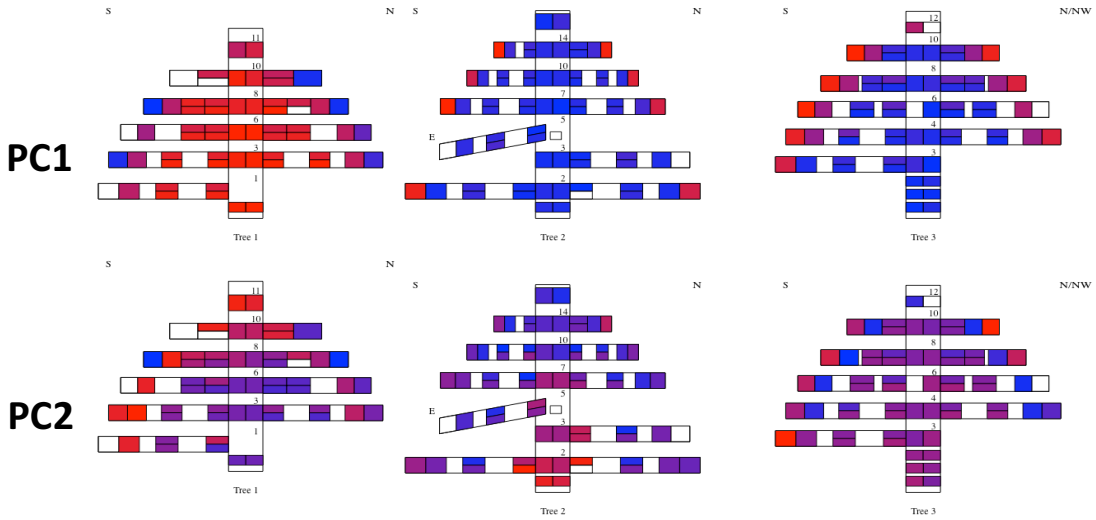

Figure A2.5. Spatial heat maps on schematic diagrams of the three replicate ginkgo trees sampled showing overall levels of bacterial community similarity calculated using principal coordinate axes scores (PC1 and PC2). Different colors represent different principal coordinate axis 1 and axis 2 scores. Trees 1 and 3 were male and tree 2 was female.

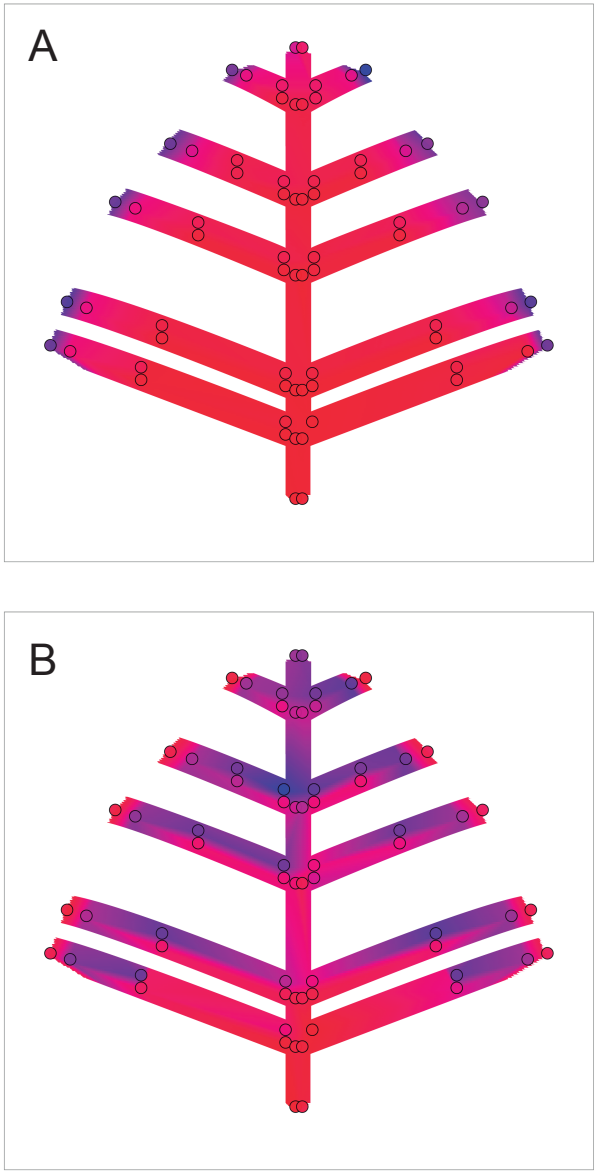

Figure A2.6. Spatial variation of bacterial community composition on bark. More dissimilar colors represent more dissimilar communities using principal coordinate analysis scores from (A) the first coordinate and (B) the second coordinate. Shading is based on linear interpolation between average sample values (indicated with circles). Representation of sample locations were adjusted to accommodate slight variations among the three replicate trees prior to computing average values for a given location, and the left represents south facing branches and the right represents north facing branches. Distances based on unweighted UniFrac distances.

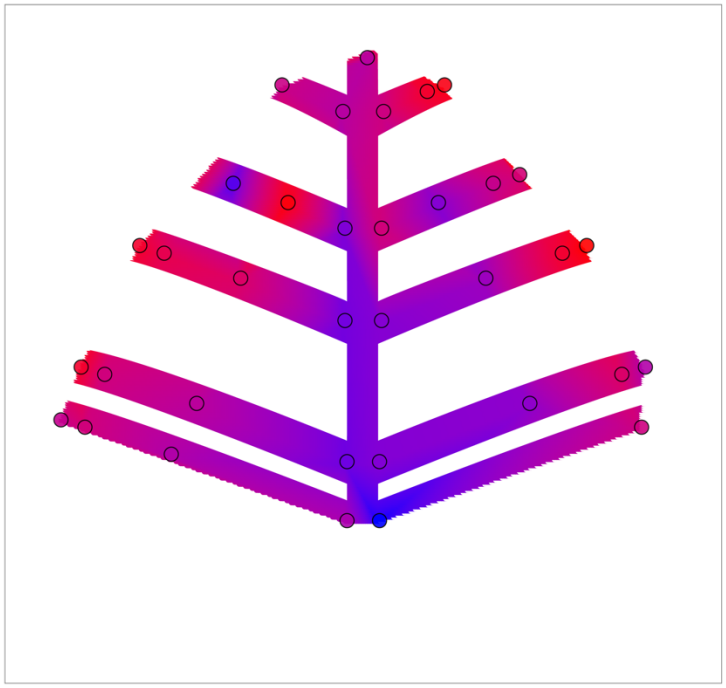

Figure A2.7. Spatial variation of bacterial community composition on leaves (as opposed to the bark results shown in Supplementary Figure 6). More dissimilar colors represent more dissimilar communities using principal coordinate analysis scores from the first coordinate.

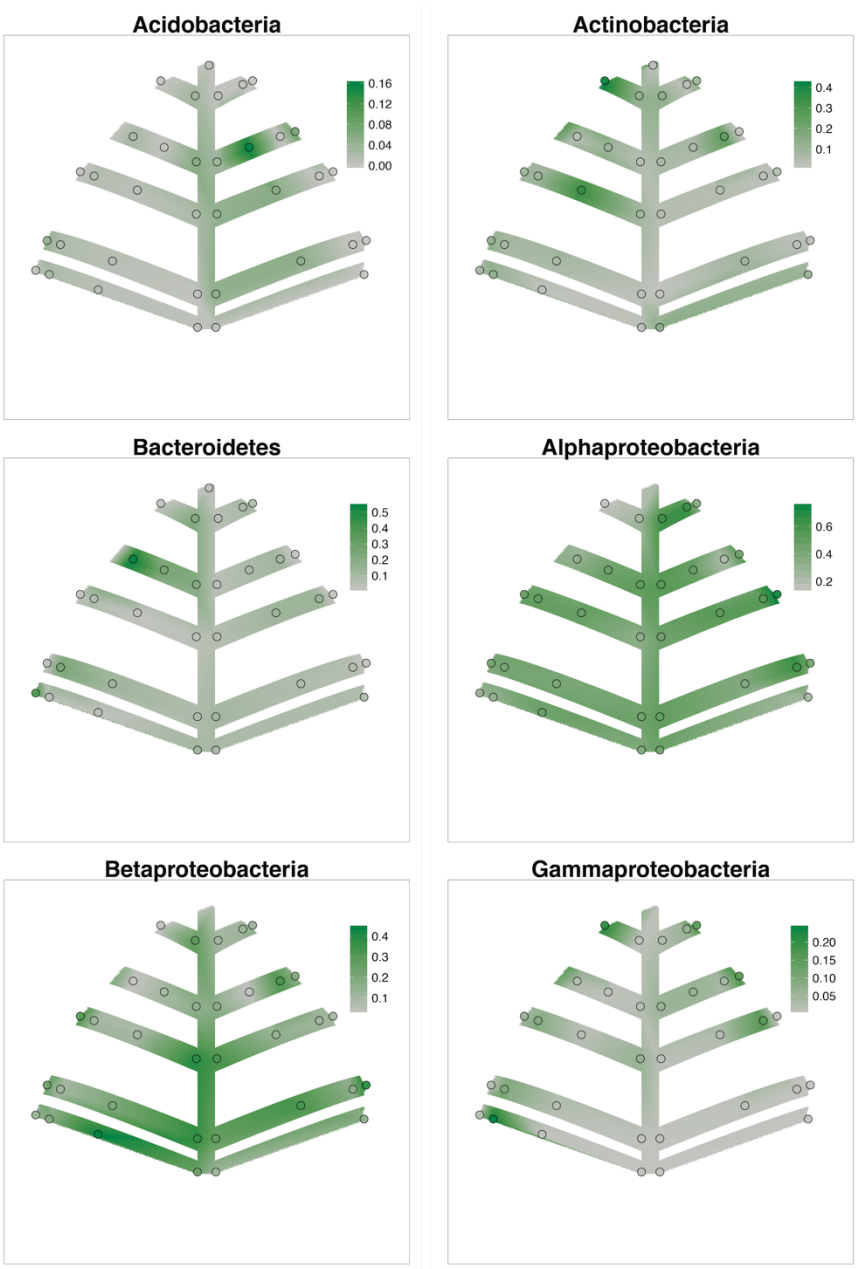

Figure A2.8. Schematic of leaf sampling sites colored by the relative abundance of the six dominant phyla and classes associated with leaves. Green sites indicate higher relative abundances, and gray sites indicate lower relative abundances. The left side of the subset panels represents south facing branches and the right side represents north-facing branches.

Table A2.1 Phylotypes that had a median relative abundance (%) of at least 1% in each of the sample types.

Sample type	Greengenes OTU ID	OTU taxonomy	Median relative abundance (%)
Branch			
	994849	*Hymenobacter* sp. 1	7.7
	357993	unclassified Rhizobiales 1	5.6
	1052108	*Hymenobacter* sp. 2	2.6
	1097610	unclassified Beijerinckiaceae	2.6
	959928	*Sphingomonas* sp. 1	2.1
	1063682	unclassified Acidobacteriaceae 1	1.7
	1021582	unclassified Acidobacteriaceae 2	1.6
	362511	*Hymenobacter* sp. 3	1.2
	163173	unclassified Sphingobacteriaceae	1.2
	104326	unclassified Acidobacteriaceae 3	1
	319666	*Sphingomonas* sp. 2	1
Branch-green			
	509212	unclassified Oxalobacteraceae 1	31.7
	220250	unclassified Oxalobacteraceae 2	6.6
	444896	*Hymenobacter* sp. 4	4.7
	568926	*Hymenobacter* sp. 5	2.8
	357993	unclassified Rhizobiales 1	1.7
	2910306	unclassified Oxalobacteraceae 3	1.7
	1043102	*Hymenobacter* sp. 6	1.2
Leaves			
	357993	unclassified Rhizobiales 1	3.7
	509212	unclassified Oxalobacteraceae 1	2.6
Trunk			
	994849	*Hymenobacter* sp. 1	7.3
	357993	unclassified Rhizobiales 1	3.5
	362511	*Hymenobacter* sp. 3	3.3
	959928	*Sphingomonas* sp. 1	3.1
	1052108	*Hymenobacter* sp. 2	2.2
	1021582	unclassified Acidobacteriaceae 2	1.4
	1063682	unclassified Acidobacteriaceae 1	1.4
	240940	*Actinomycetospora* sp.	1.4
	1097610	unclassified Beijerinckiaceae	1.3
	104326	unclassified Acidobacteriaceae 3	1.2
	542913	*Friedmanniella* sp.	1.2
	319666	*Sphingomonas* sp. 2	1
	568926	*Hymenobacter* sp. 5	1
	242089	unclassified Rhizobiales 3	1
	163173	unclassified Sphingobacteriaceae	1
	de novo OTU	unclassified Rhizobiales 2	1
	1043102	*Hymenobacter* sp. 6	1

Table A2.2. Summary statistics for the proportion of sequences that were identified as chloroplast, and the number of quality-filtered sequences obtained per sample after removal of the chloroplast sequences. Only samples with ≥50 sequences were included.

Sample type	N	Mean proportion chloroplast sequences	Mean sequence count	Standard deviation of sequence count	Minimum sequence count	Maximum sequence count
Trunk	47	0.0002	3042	1085	145	5644
Branch	125	0.0001	2956	1050	498	7132
Branch-green	28	0.0000	2898	2535	161	8954
Leaves	103	0.0233	1253	1370	51	6857

CHAPTER III APPENDIX

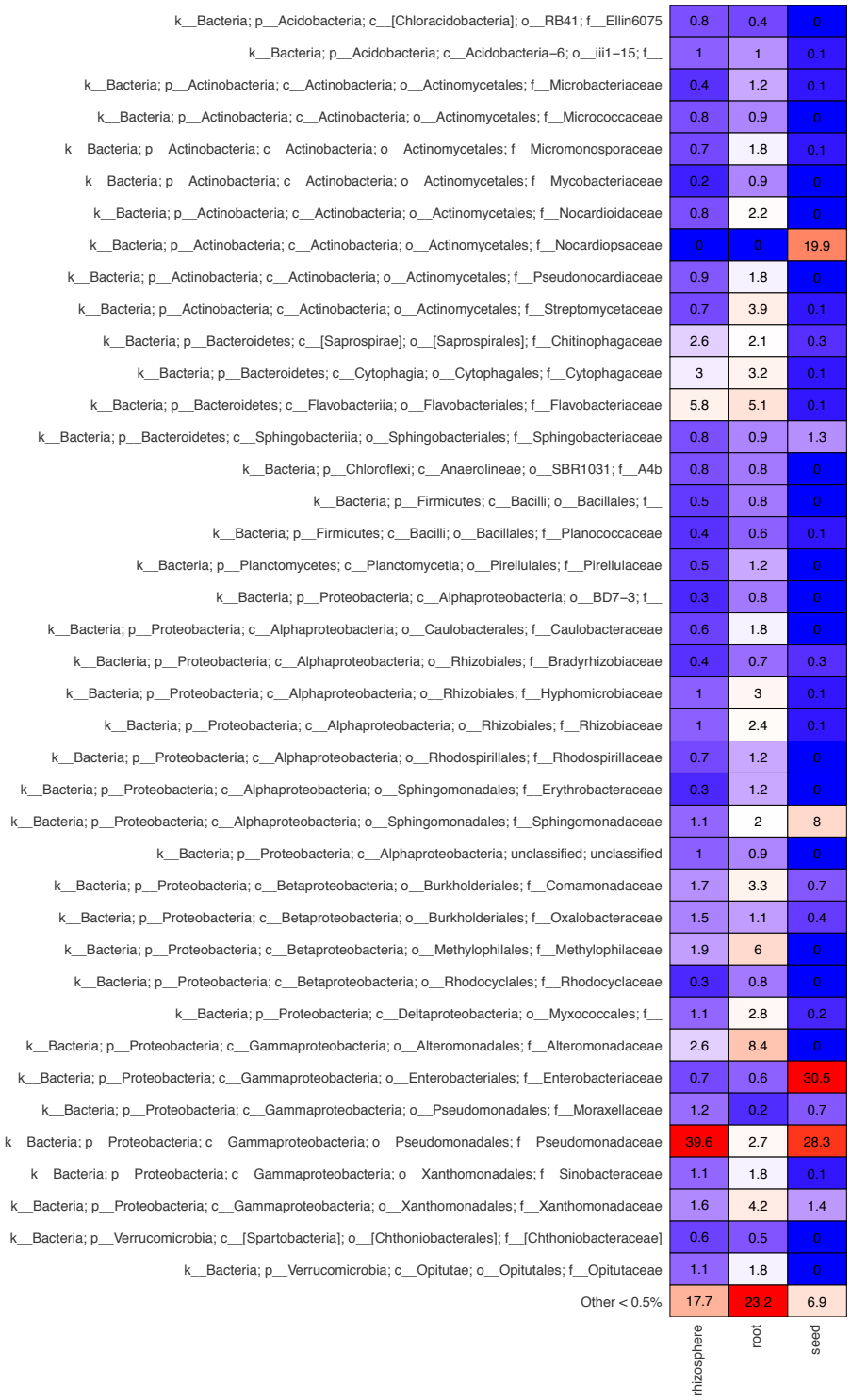

Figure A3.1 Heat map showing bacterial family relative abundances across sample types. Values represent mean relative abundances for each family and sample type. Red colors indicate higher relative abundances, and blue colors indicate lower relative abundances.

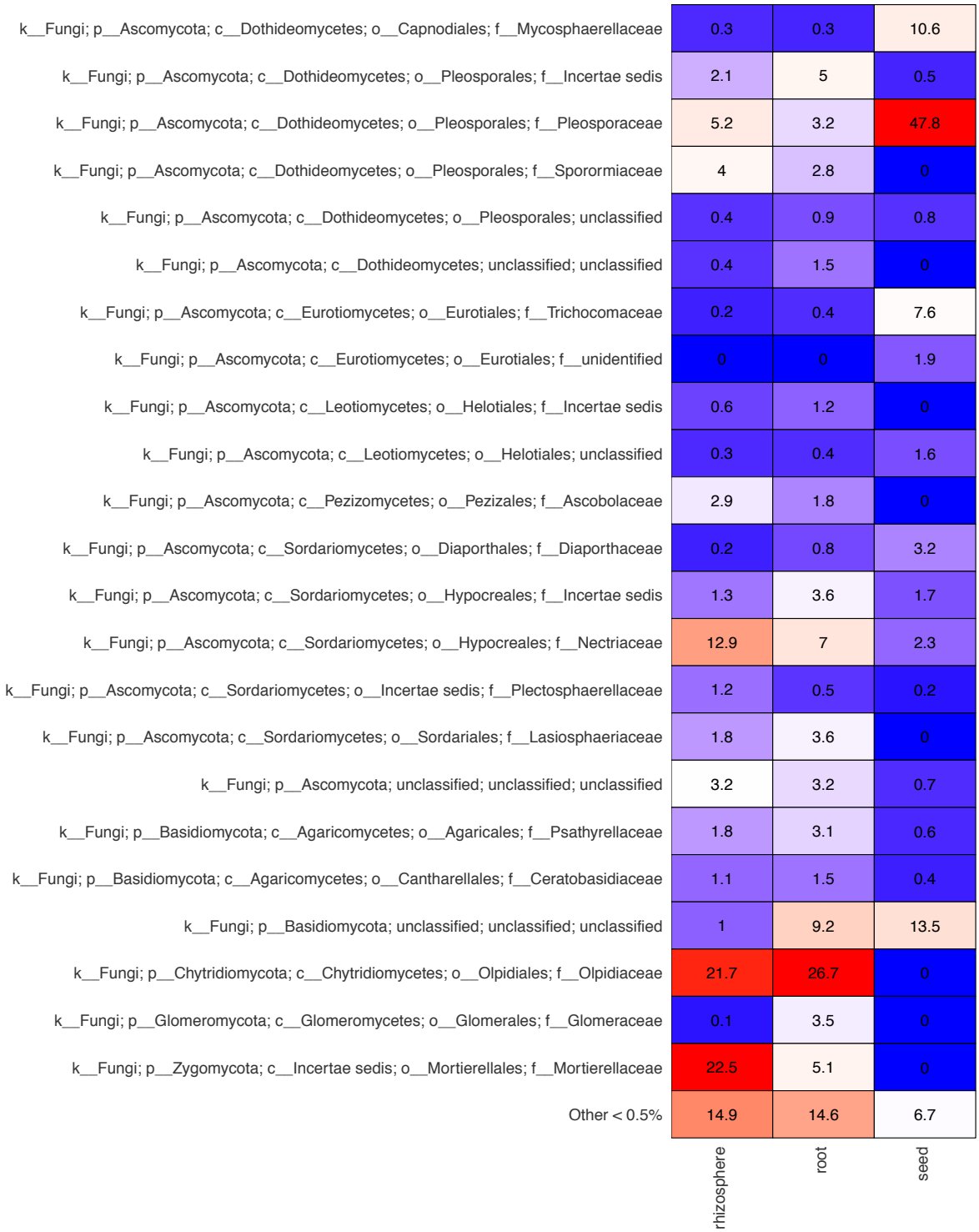

Figure A3.2 Heat map showing fungal family relative abundances across sample types. Values represent mean relative abundances for each family and sample type. Red colors indicate higher relative abundances, and blue colors indicate lower relative abundances.

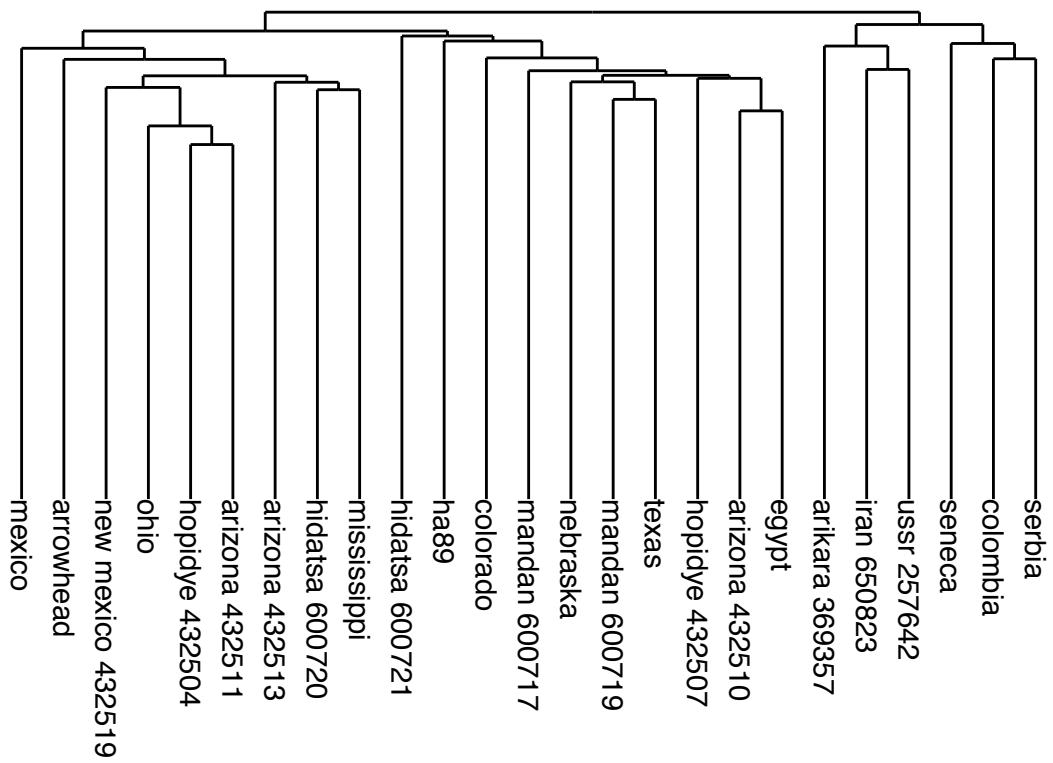

Figure A3.3 Cluster diagram showing differences in rhizosphere fungal community composition across sunflower (*Helianthus annuus* L.) strains. Lengths of branches between strains are proportional to the mean dissimilarity of their community compositions. The diagram was created using means of pairwise Bray-Curtis dissimilarities between samples from each pair of sunflower strains. Dissimilarities were calculated from rarefied square-root transformed phylotype relative abundances.

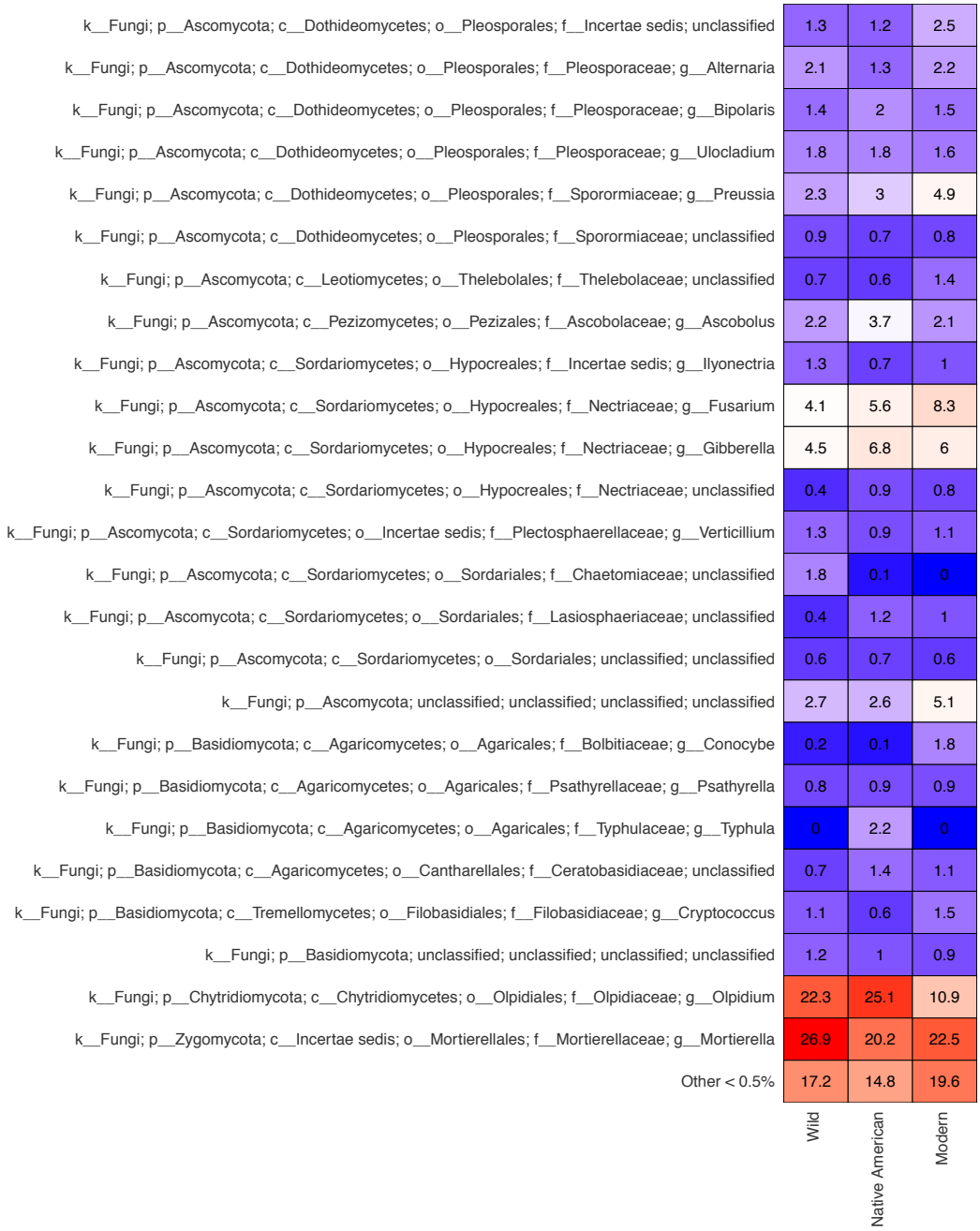

Figure A3.4 Heat map showing fungal genus relative abundances of rhizosphere communities across domestication levels. Values represent mean relative abundances for each family and sample type. Red colors indicate higher relative abundances, and blue colors indicate lower relative abundances.

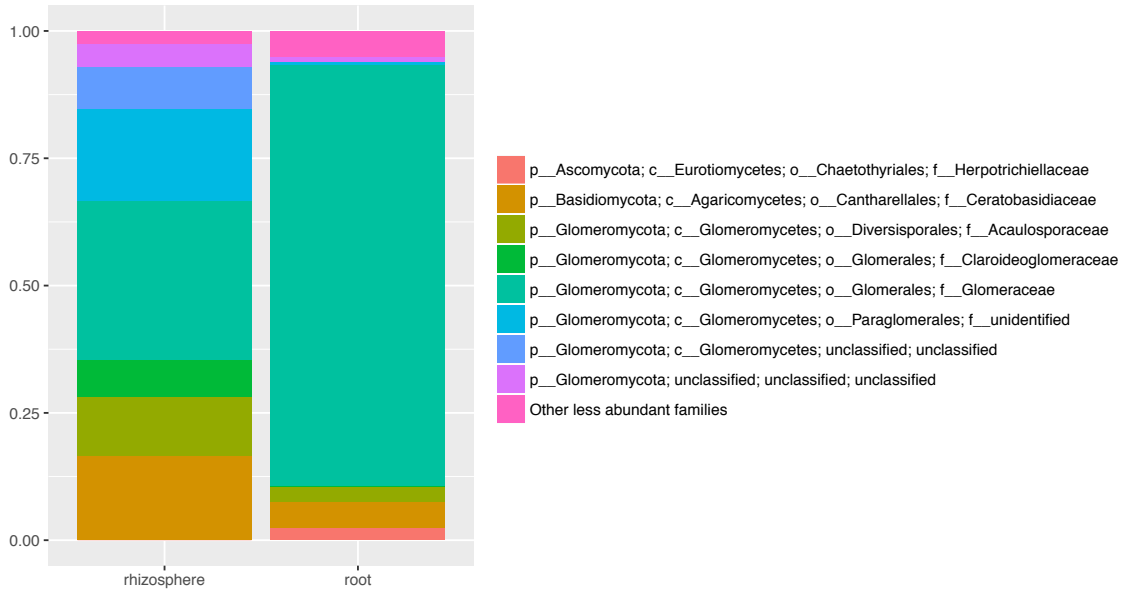

Figure A3.5 The composition of fungi identified as symbiotic in root and rhizosphere samples. Stacked bar plots represent mean relative abundances of phylotypes identified as symbiotic and summed within families. Only samples with ≥ 1% of their sequences identified as symbionts were included.

Table A3.1 The sunflower (*Helianthus annuus* L.) strains used in this study and characteristics (mean values) at the time of sample collection.

Plant strain	Accession	Domesti-cation level[1]	Height (cm)	Number nodes	Number branches	MRFELL[2] (cm)	MRFELW[2] (cm)	Stem diameter (cm)
arikara 369357	PI 369357	NAm	41.8	8.4	3	48.6	44.6	8.4
arizona 432510	PI 432510	NAm	47.4	7.6	3	24.4	43.4	5.2
arizona 432511	PI 432511	NAm	43.4	6.6	2	31.6	54.6	6.8
arizona 432513	PI 432513	NAm	42.2	6	4.4	32.4	49.6	7
arkansas	PI 613727	Wild	88.4	7.2	4.6	59.6	24.2	4
arrowhead	PI 650649	NAm	63.6	8.6	4	17.8	49.4	6.4
colombia	PI 265499	Modern	57.25	7.25	3.5	27.5	46.5	8.25
colorado	PI 435376	Wild	37.75	8.25	4	52.75	29.25	4.25
egypt	PI 250542	Modern	47.8	6.8	2.2	21.2	48.2	6.6
ha89	PI 543743	Modern	11.2	7	2	51.6	43.4	5.4
hidatsa 600720	PI 600720	NAm	43	8	2	54.6	44.2	5.2
hidatsa 600721	PI 600721	NAm	40.6	8	2	49.4	50.4	6
hopidye 432504	PI 432504	NAm	70.67	7.33	2	34	28	11
hopidye 432507	PI 432507	NAm	53.5	7.5	3	44.25	54.75	9
hopidye 432508	PI 432508	NAm	42.67	6.67	2	19.67	49.33	7.33
iowa	PI 613779	Wild	14.33	5.33	6	23.33	27.67	4.33
iran 650823	PI 650823	Modern	54.5	6	2	44.5	42	5.25
kansas	PI 586862	Wild	63.75	7.25	5.5	46.25	39	4.25
kentucky	PI 435613	Wild	27.2	7.8	7.2	55.2	32.2	4.4
mandan 600717	PI 600717	NAm	54.6	9	2	36.8	40.4	6.8
mandan 600719	PI 600719	NAm	15	6.2	2	40.8	43	5.2
mexico	PI 413121	Wild	82.6	8	6	61.4	34.8	3.8
mississippi	PI 664809	Wild	95.67	7	6	69.67	32.33	3.67
missouri	PI 413011	Wild	90.5	7.5	6.5	26.5	30.5	4
nebraska	PI 586867	Wild	45.2	7.8	7.6	36.2	45.4	5
new mexico 432519	PI 432519	NAm	45.8	7	2.2	48.6	51.6	7
north dakota 560147	PI 560147	Modern	56	5.67	2	52.67	23.33	4
Ohio	PI 649853	Wild	48.67	7.33	8.33	45.33	34	4.67
saskatchewan	PI 592315	Wild	96	8	9	60	31	6
seneca	PI 369360	NAm	25.75	3.75	2.75	42.75	16.5	6.75
serbia	PI 431554	Modern	48.4	6.2	2	51.4	20.4	3.6
texas	PI 599773	Modern	19.67	6	2	27.33	53.33	5
ussr 257642	PI 257642	Modern	39.6	7.2	2	22.2	40.4	6.4

[1]Nam = Native American

[2]MRFELL = most recent fully expanded leaf length; MRFELW = most recent fully expanded leaf width

Table A3.2. Spearman correlations between bacterial and fungal Shannon diversity in rhizosphere and root samples and plant characteristics measured at the time of sample collection. Relationships with P < 0.1 are indicated in bold.

	Rhizosphere		Root	
Plant characteristic[1]	Rho	*P*	Rho	*P*
Bacteria				
Height (cm)	**-0.21**	**0.019**	-0.01	0.874
Number of nodes	**-0.16**	**0.065**	0.09	0.328
Number of branches	-0.04	0.680	-0.02	0.787
MRFELL (cm)	**-0.23**	**0.010**	-0.03	0.776
MRFELW (cm)	**-0.16**	**0.076**	0.02	0.827
Stem diameter (cm)	-0.19	0.029	-0.04	0.652
Fungi				
Height (cm)	-0.08	0.400	-0.11	0.211
Number of nodes	**-0.22**	**0.016**	-0.09	0.318
Number of branches	0.01	0.915	-0.08	0.387
MRFELL (cm)	-0.09	0.306	**-0.23**	**0.011**
MRFELW (cm)	-0.08	0.381	**-0.20**	**0.024**
Stem diameter (cm)	-0.07	0.438	-0.09	0.320

[1]MRFELL = most recent fully expanded leaf length; MRFELW = most recent fully expanded leaf width

Table A3.3. Spearman correlations and Mantel tests assessing relationships between dissimilarity in bacterial and fungal community composition in rhizosphere and root samples and Euclidian distances between plant characteristics from different samples. Relationships with P < 0.1 are indicated in bold.

	Rhizosphere		Root	
Plant characteristic[1]	Rho	*P*	Rho	*P*
Bacteria				
Height (cm)	0.04	0.128	-0.034	0.760
Number nodes	-0.04	0.825	-0.021	0.639
Number branches	0.03	0.252	0.010	0.398
MRFELL (cm)	0.01	0.436	-0.029	0.741
MRFELW (cm)	0.05	0.107	-0.016	0.635
Stem diameter (cm)	0.04	0.179	-0.039	0.776
Fungi				
Height (cm)	0.03	0.174	0.008	0.380
Number nodes	0.01	0.418	**0.070**	**0.057**
Number branches	-0.01	0.561	0.040	0.183
MRFELL (cm)	0.03	0.215	0.005	0.433
MRFELW (cm)	0.03	0.214	**0.095**	**0.010**
Stem diameter (cm)	0.01	0.355	-0.003	0.554

[1]MRFELL = most recent fully expanded leaf length; MRFELW = most recent fully expanded leaf width

CHAPTER IV APPENDIX

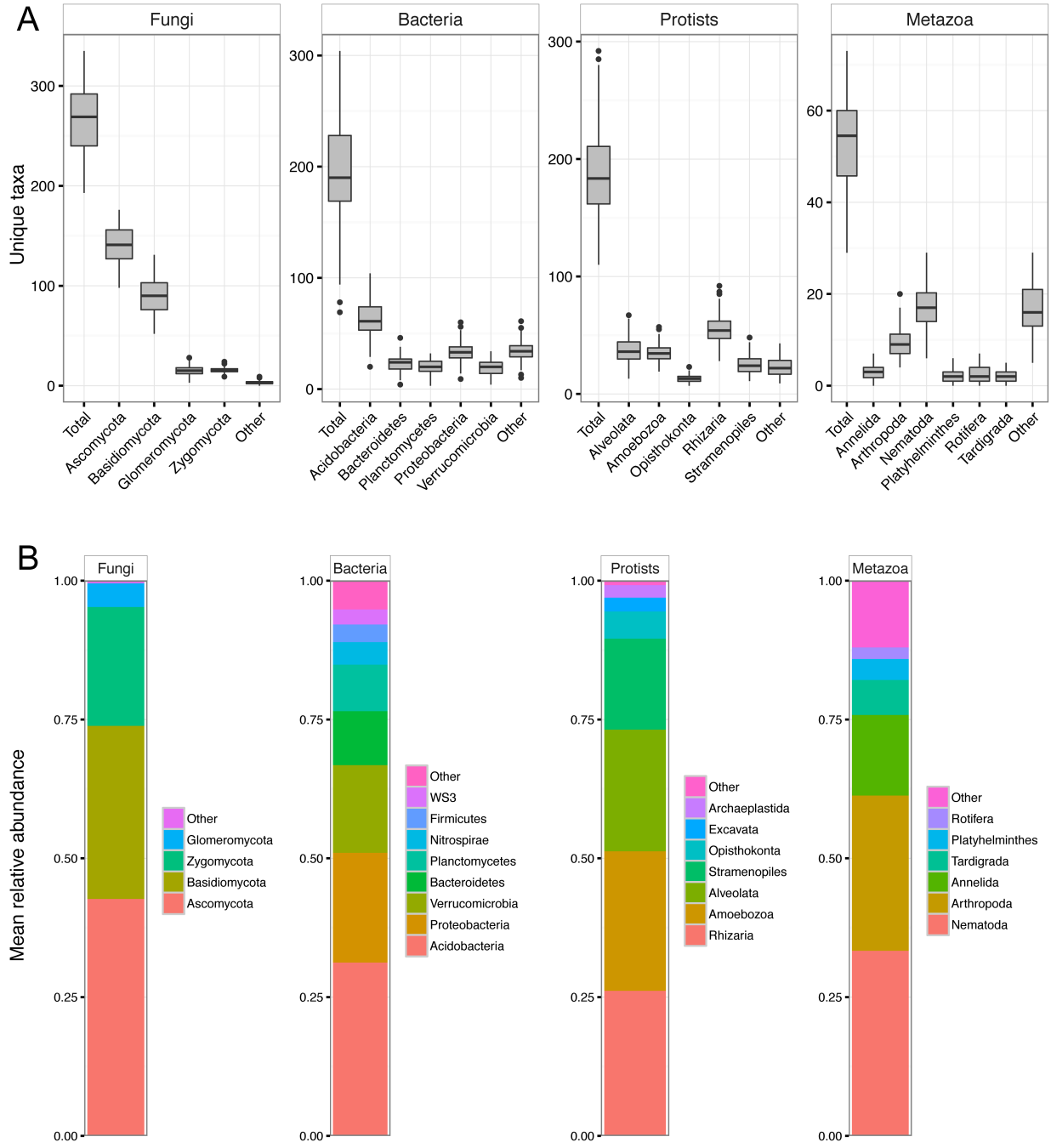

Figure A4.1. Overall community structure of mesocosm soil communities. The number of total unique taxa and unique taxa in the dominant phyla of each community (A). Mean relative abundances of dominant phyla across all samples (B).

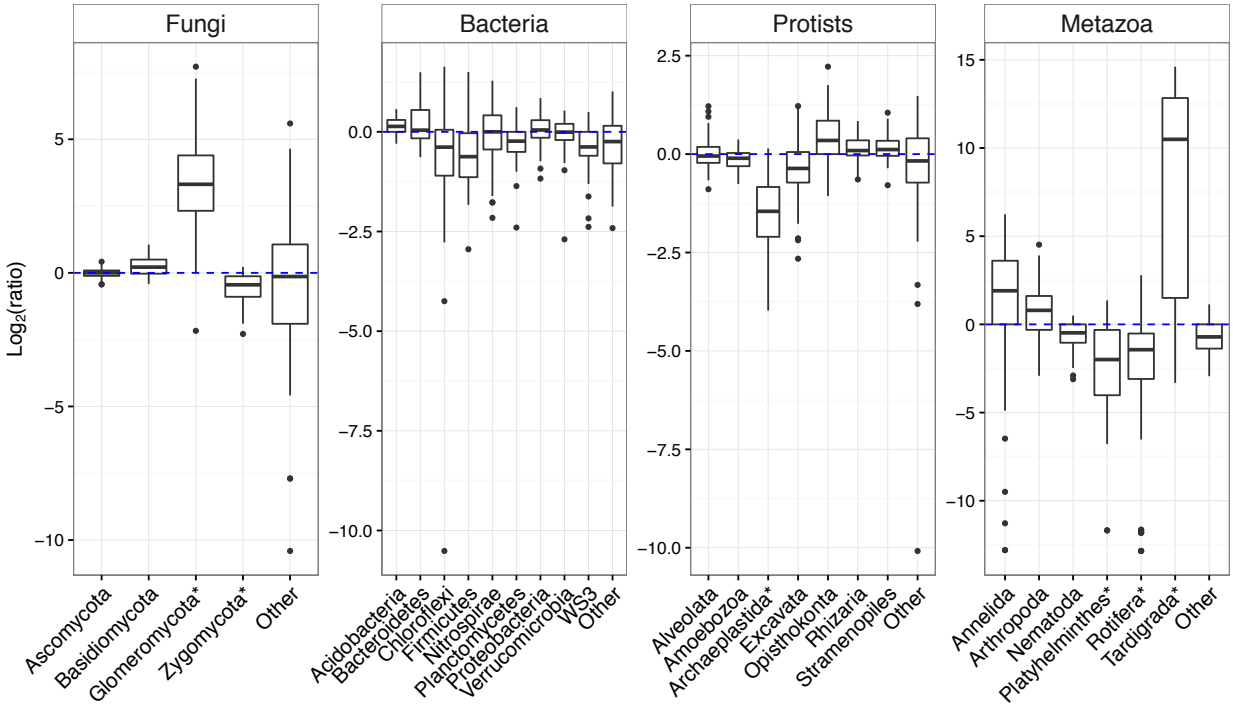

Figure A4.2. Ratios of phyla relative abundances in planted mesocosms relative to control mesocosms within blocks. Phyla with mean relative abundances <1% were grouped into "Other". Significant differences in phyla between planted and control mesocosms are indicated with an asterisk.

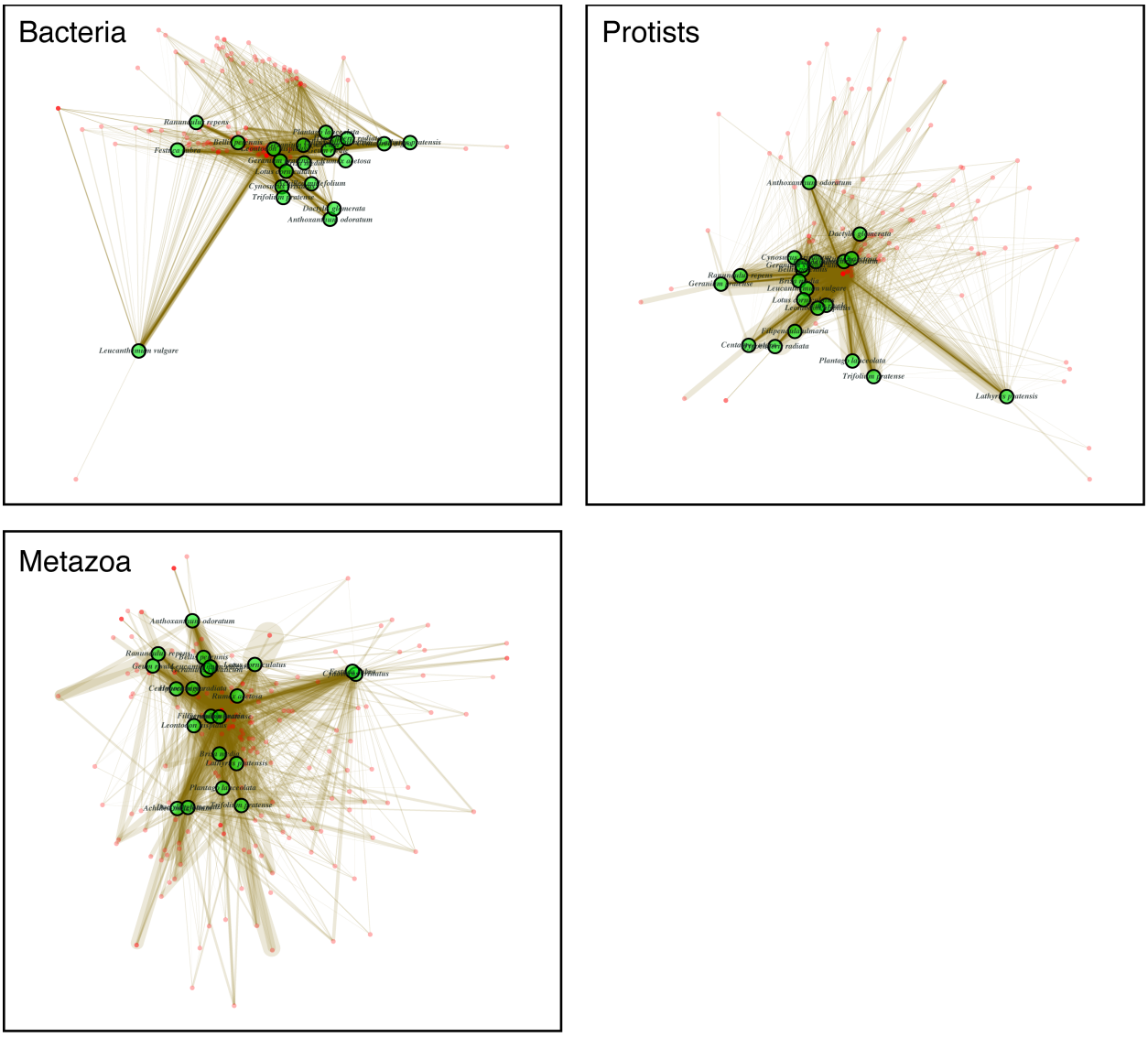

Figure A4.3. Bipartite network diagrams for bacterial, protistan, and metazoan communities as shown in Fig. 4.1C.

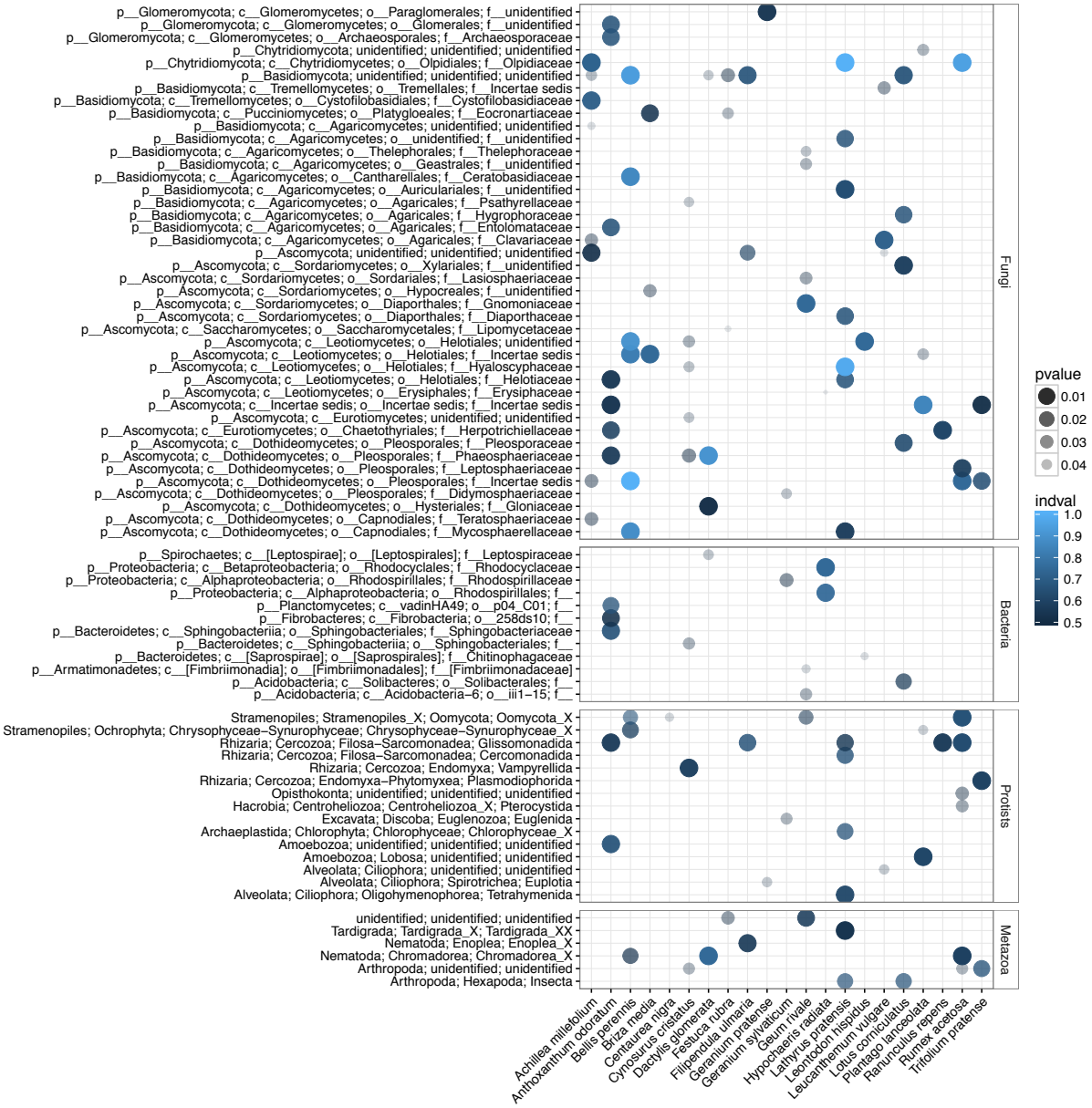

Figure A4.4. Indicator taxa for each of the plant species grown in the mesocosms. Point sizes show strength of indicator taxa (Indval). Points represent the maximum Indval and minimum *P* value for a plant species if multiple taxa represent the labeled taxonomy string. Only indicator taxa with Indval ≥0.5 and $P \leq 0.5$ are shown.

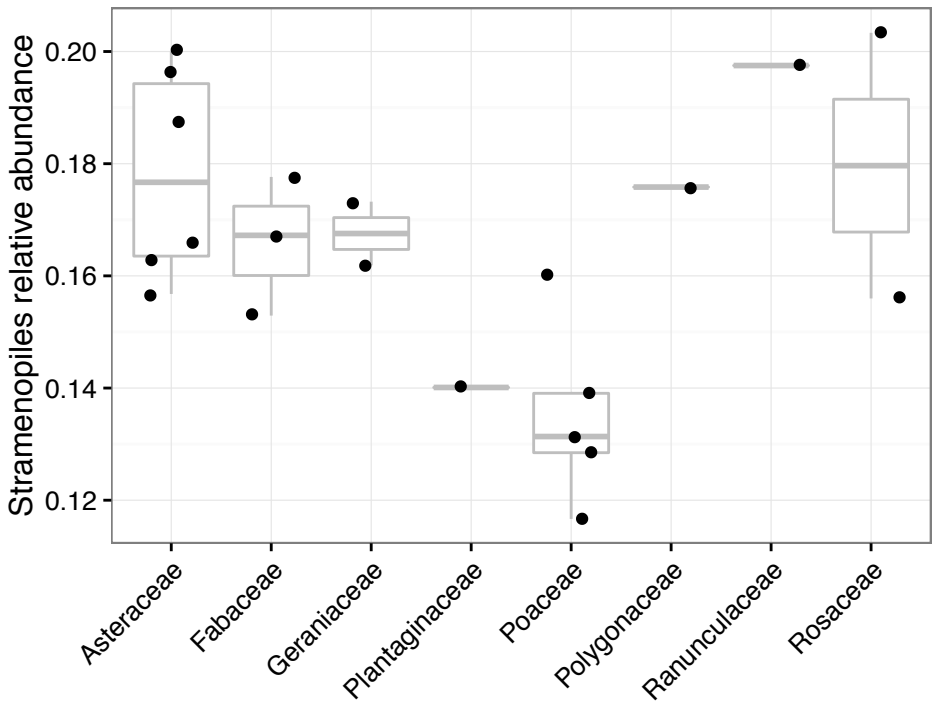

Figure A4.5. The mean relative abundance of protistan sequences representing the *Stramenopiles* for each plant species grown in the mesocosms. Plant species values are grouped by family, and boxplots for each family are shown in the background.

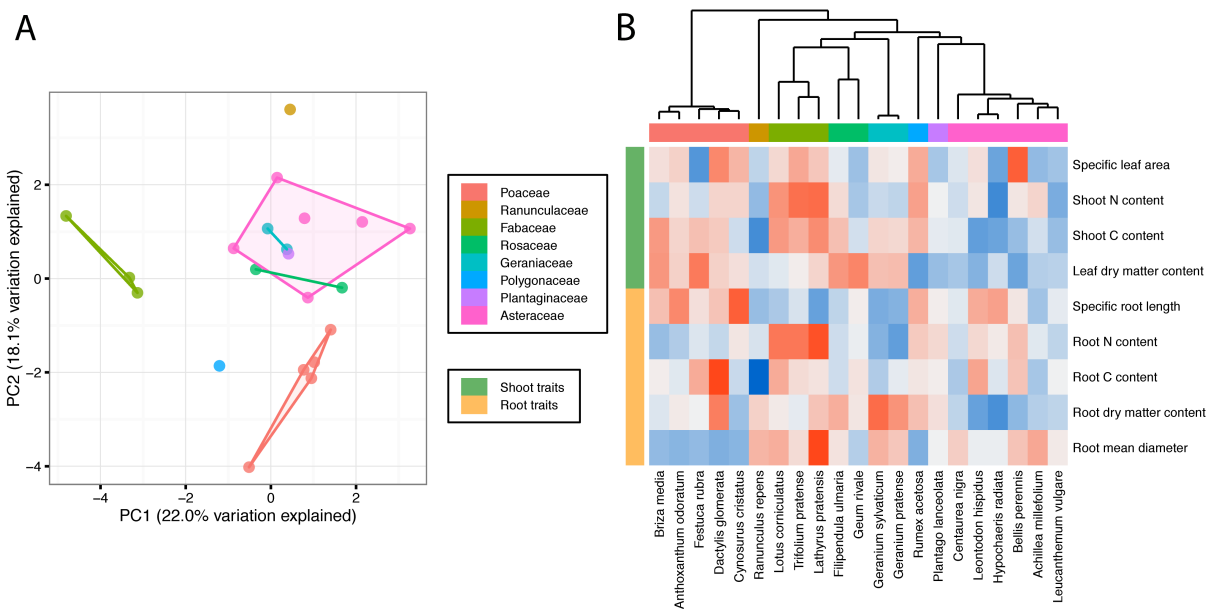

Figure A4.6. Traits of plant species grown in the mesocosms. Principal components analysis showing differences in species' collective mean trait values where points are colored by species family (A). Heatmap showing relative shoot and root trait values across plant species clustered by phylogeny (B). In both panels, trait values were transformed when necessary to approximate normal distributions and then standardized.

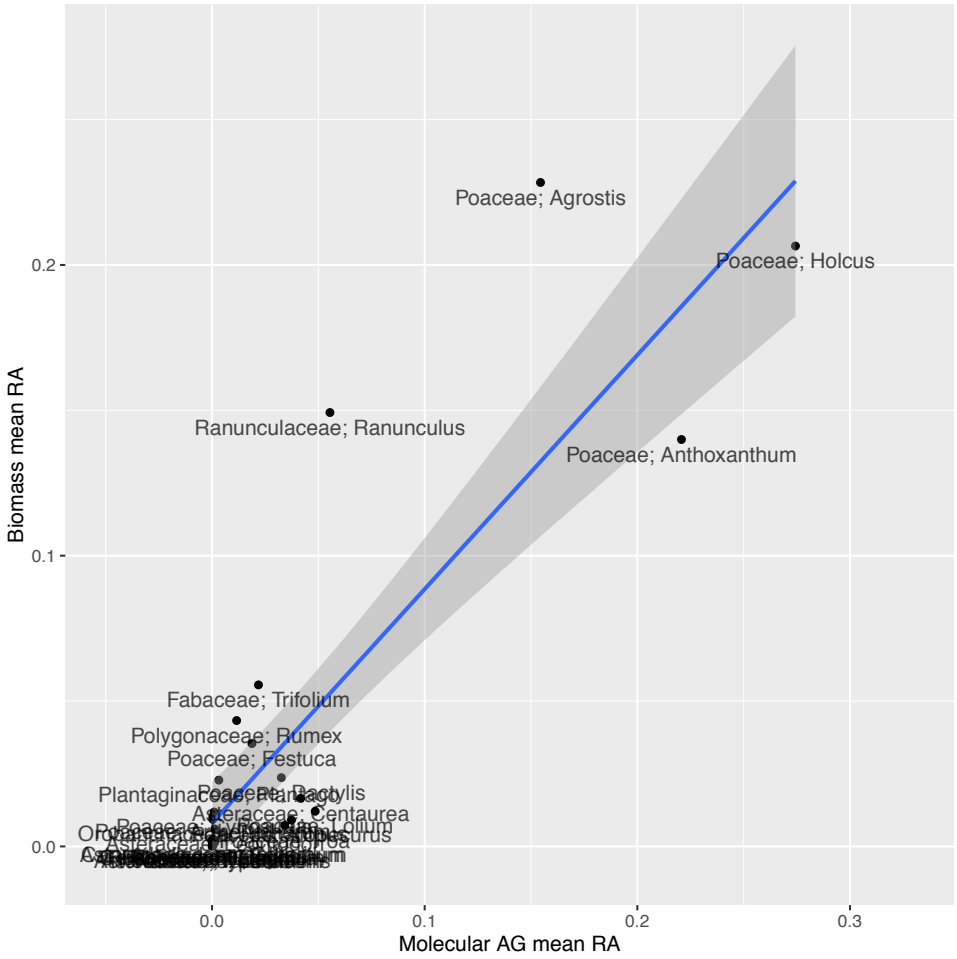

Figure A4.7. Molecular-identification-based relative abundance of aboveground plant biomass plant genera correlates with the relative abundance of the same plant genera as identified with visual identification. Points represent means from across the field samples.

A

Community	Soil characteristics[1] R^2	*P*	With plants (aboveground) R^2	*P* (plants)	With plants (soil) R^2	*P* (plants)
Fungi	0.23	0.001	0.25	0.01	0.26	0.001
Bacteria	0.29	0.001	0.29	0.57	0.36	0.001
Protists	0.22	0.001	0.22	0.30	0.24	0.015
Metazoa	0.13	0.001	0.14	0.06	0.13	0.190

[1] Terms included in models including only soil variables:
Fungi: Soil N, pH, moisture
Bacteria: Moisture, pH, soil N, soil C
Protists: Moisture, pH, soil N
Metazoa: pH, soil N

B

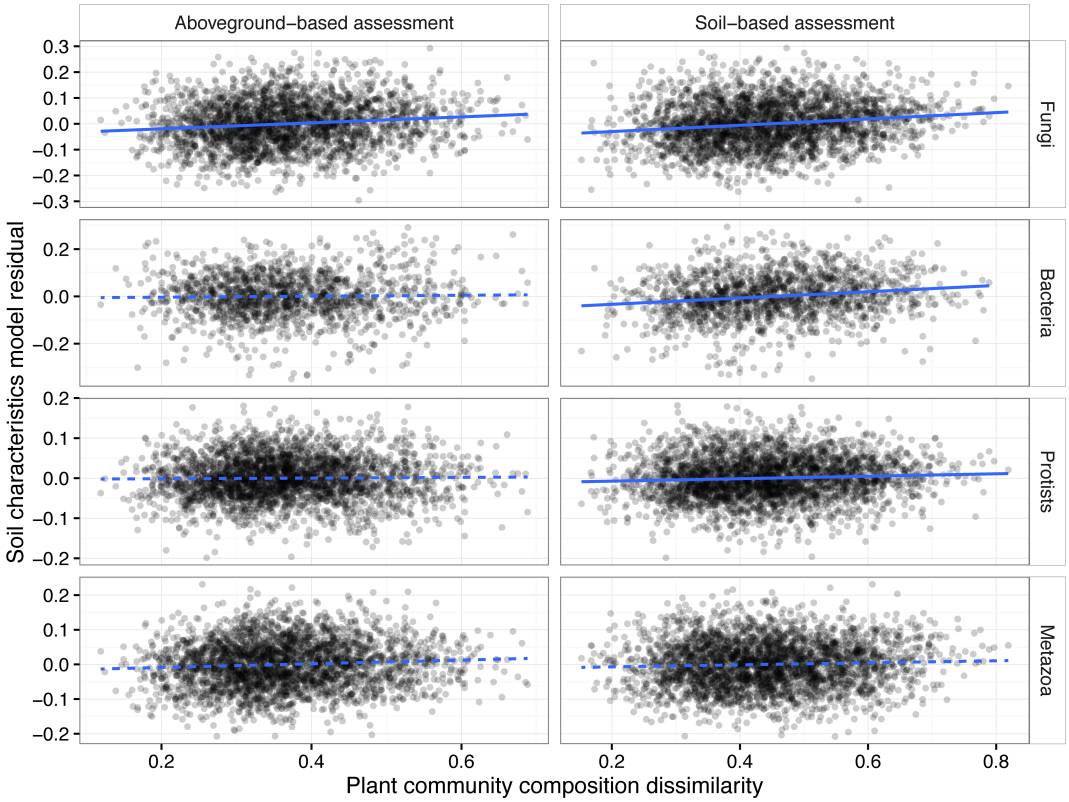

Figure A4.8. Differences in plant community composition improve models of overall differences in soil community composition in certain cases. Table comparing the best multiple regression on distance matrices models containing only soil variables and those where plant community dissimilarities were added as a predictor (A). The effects of including both aboveground-based and soil DNA-based plant community assessment are shown. P values after the addition of plant community dissimilarities refer to whether those dissimilarities added significant explanatory power. Relationships between residuals from soil-characteristics-only models and plant community composition dissimilarities (B). Solid lines indicate significant ($P < 0.05$) relationships.

Table A4.1. Mantel test statistics indicating the relationship between differences in soil community compositions and soil variables across the field samples.

	Fungi		Bacteria		Protists		Metazoa	
Soil variable	rho	*P*	rho	*P*	rho	*P*	rho	*P*
Moisture content	**0.29**	**0.001**	**0.34**	**0.002**	**0.35**	**0.001**	**0.20**	**0.005**
pH	**0.33**	**0.001**	**0.35**	**0.001**	**0.36**	**0.001**	**0.31**	**0.001**
Soil C	**0.33**	**0.001**	**0.42**	**0.001**	**0.26**	**0.001**	**0.23**	**0.001**
Soil N	**0.38**	**0.001**	**0.42**	**0.001**	**0.26**	**0.001**	**0.23**	**0.001**
DIN[1]	0.00	0.508	0.00	0.465	0.00	0.467	0.00	0.471
NH_4^+	-0.02	0.615	-0.01	0.558	0.00	0.431	0.01	0.437
NO_3^-	-0.01	0.586	-0.03	0.631	0.00	0.483	-0.04	0.701
DON[2]	-0.08	0.945	-0.05	0.774	-0.08	0.923	-0.09	0.946
Total soluble N	-0.07	0.906	-0.05	0.778	-0.07	0.905	-0.05	0.820
N mineralization	-0.01	0.580	0.03	0.319	0.00	0.499	0.03	0.338

[1]Dissolved inorganic N
[2]Dissolved organic N

CHAPTER V APPENDIX

Appendix A5 Methods

Site characteristics and experimental design

All 25 sites used in this study (Table A5.2) were from temperate-zone grasslands in Africa, Australia, Europe, and North America, and all were part of the Nutrient Network experiment (Borer et al. 2014) and have been described by (Prober et al. 2015). In brief, the sites ranged in many environmental characteristics including mean annual precipitation (262 - 1898 mm $year^{-1}$), mean annual temperature (0–18°C), elevation (50 - 2320 m), soil pH (4.5 - 8.4), total soil P (1 - 253 ppm), soil %N (0.03 - 1.5%), and aboveground plant productivity (15 - 1482 gm^{-2} $year^{-1}$).

Identical full factorial N and P addition experiments were established at each of the 25 sites (Borer et al. 2014). We used samples from 3 +N, 3 +P, and 3 +N/+P 1 m^2 plots, and 6 1 m^2 plots with no added nutrients. Nutrients were added annually as 10 g N m^{-2} yr^{-1} timed-release urea ($(NH_2)_2CO$) and 10 g P m^{-2} yr^{-1} triple-super phosphate ($Ca(H_2PO_4)_2$). Samples were treated between two to four years prior to collection (Table A5.2), but this timing did not significantly relate to the proportional changes in the major microbial taxon relative abundances ($P > 0.5$ in all cases). Fences to exclude herbivores surrounded 3 of the plots at each site (Borer et al. 2014), but since fencing had no effect on belowground microbial communities ($P > 0.1$), we included these plots as controls.

Plant community composition and biomass were assessed and soil samples were collected during the growing season of 2011 or 2012. Plant species were identified within the plots, and plant biomass and soil cores were sampled directly adjacent to the plots to determine soil characteristics and assess microbial community structure (Borer et al. 2014). Soil cores were shipped on ice to a central processing facility (Corvallis, Oregon, USA) immediately following collection, and samples for microbial community analysis were preserved at -20°C. Analyses to determine soil pH, C content, N content, and P content are described in (Prober et al. 2015).

Assessment of microbial community composition and diversity

Microbial community diversity and composition were assessed using targeted marker gene surveys focusing on the 16S rRNA gene for the Bacteria and Archaea domains and the internal transcribed spacer (ITS) region for Fungi as in Prober et al. (2015). DNA was extracted by inserting a sterile swab into each soil sample and cutting the swab tip off into a well in a bead plate of the PowerSoil-htp 96 well DNA extraction kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA). The downstream extraction procedure was followed according to the manufacturer's instructions, and two wells on each extraction plate were left empty to serve as blanks. Marker genes in isolated DNA were PCR amplified and barcoded in triplicate reactions for both the 16S rRNA gene (using the 515f/806r primer pair) and the ITS1 region (using the ITS1-F/ITS2 primer pair). PCR products from the triplicate reactions were combined and visualized on an agarose gel to ensure successful amplification and to verify no amplification from blanks. PCR product from samples that were successfully amplified was combined in equimolar ratios for each of the marker genes. This procedure was conducted separately for two sets of samples, one including samples from 19 of the sites and the other including the samples from the remaining 6 sites. For the first set, 16S rRNA and ITS amplicon pools were each sequenced on separate Illumina HiSeq 2000 lanes using 100 bp paired-end sequencing, and for the second set, each amplicon type was sequenced on a separate Illumina MiSeq run using 151 bp paired-end sequencing. All sequencing was conducted at the University of Colorado at Boulder (Boulder, CO, USA).

Raw sequences from 16S rRNA and ITS amplicons were processed using the UPARSE pipeline (Edgar 2013). Sequences were demultiplexed according to the raw sequenced barcodes, and since there were generally many more sequences for samples sequenced on the HiSeq run, a random subset of the sequences from each of these samples were included in further processing such that the number of sequences per sample was generally within two orders of magnitude across samples. Sequences from both sequencing runs were combined after trimming to a uniform length of 100 bp. A *de novo* database

of ≥ 97% similar sequence clusters was created in USEARCH v. 7 (Edgar 2010) by: (1) quality filtering sequences using a 'maxee' value of 0.5 (i.e. sequences with a predicted error rate of 0.5 bases per sequence were discarded), (2) dereplicating identical sequences, (3) removing singleton sequences, (4) clustering those sequences, and (5) filtering poor quality sequences by removing sequences that were not ≥ 75% similar to any sequence in Greengenes 13_5 (McDonald et al. 2012b) or UNITE 12_11 (Abarenkov et al. 2010) databases for 16S rRNA and ITS1 sequences, respectively. Raw demultiplexed sequences were then mapped against these *de novo* databases to generate counts of sequences matching clusters (i.e. phylotypes) for each sample. Taxonomy was assigned to each phylotype using the RDP classifier with a confidence threshold of 0.5 (Wang et al. 2007b) and trained on the databases indicated above. Among 16S rRNA sequences, phylotypes classified as bacteria were separated from those classified as archaea prior to further processing. Mycorrhizal fungi phylotypes were identified by comparing to known mycorrhizal ITS sequences as detailed in (Prober et al. 2015). To normalize the sequencing depth across samples, samples were rarefied to 18,000 bacterial, 100 archaeal, and 485 fungal sequences per sample. Due to insufficient sequence coverage, bacterial data from 4 samples, archaeal data from 90 samples, and fungal data from 29 samples were discarded. One sample ("NN10") was removed from downstream analyses due to a large disparity in diversity and community composition from all other sequences. The raw sequence data are available in the Sequence Read Archive at the National Center for Biotechnology Information (BioProject accession: PRJNA272747).

Shotgun metagenomic analyses

All replicate samples from a subset of six of the sites (90 samples; Fig. A5.1) used for the compositional analyses were prepared and submitted for shotgun metagenomic sequencing at the Joint Genome Institute. The sites were selected as they represented a broad range in environmental characteristics and exhibited the largest bacterial community composition responses to N and P. For this analysis, DNA was extracted using the same method as above except that approximately 0.25 g of each

sample was loaded into the first extraction plate. Library preparation and sequencing were conducted at the Department of Energy Joint Genome Institute (Walnut Creek, CA, USA) using their standard protocols. Briefly, 270 bp DNA fragments were prepared and sequenced on a HiSeq instrument using 150 bp paired end sequencing. Sequencing efforts yielded a total of 850 million read pairs across all samples (range = 428 thousand to 65 million). Raw sequence data can be accessed at IMG (http://img.jgi.doe.gov) and are referenced in the Genomes Online Database (GOLD Study ID: Gs0053063).

Sequence processing was conducted in a method similar to Fierer et al. (2013). Paired end sequences were first merged using FLASH v1.2.6 (Magoč and Salzberg 2011) with the parameters, '-r 150 –f 210 –s 21'. Merged sequences were quality filtered using PRINSEQ-lite v0.20.3 (Schmieder and Edwards 2011) to remove sequences containing > 5 undefined bases and exact duplicates. On average, 49% (range = 23 - 62%) of sequences were successfully merged and passed quality filtering from those samples that were successfully sequenced. Sequences from samples with high sequence coverage were then randomly subsampled to reduce computational processing time. The merged and quality filtered sequences were annotated by mapping them against the IMG v350 amino acid sequences (Markowitz et al. 2012) using BLAT (Kent 2002). Sequences that could be mapped within a minimum percentage identity of 0.55 and an e-value cutoff of 1e-03 were provided KEGG Orthology (KO) identifiers. Of the quality filtered sequences, 28.7 - 32.7% could be annotated with specific gene category information, a similar percentage as noted in previous studies exploring soil metagenomes (Fierer et al. 2011) but larger than studies using comparatively short sequences (Fierer et al. 2012b). Only sequences that were annotated were considered for further analysis, and samples were rarefied to an equal sequence depth of 500,000 annotated sequences per sample. This resulted in a total of 75 shotgun metagenomic samples spread across treatments and sites. When KO identifiers were categorized into coarser functional categories (Kanehisa et al. 2012), those that belonged to multiple categories were counted in

each of those categories. Functional diversity and composition calculations were performed identically to those used for taxonomic analyses except that KO identifiers were used instead of phylotypes.

To determine what proportion of the shotgun data represented bacterial genes, we compared merged and quality filtered reads to the SILVA 111 database (Quast et al. 2013), which contains representative SSU rRNA sequences from all three domains of life. Shotgun reads were mapped to this database at the 94% similarity level using UCLUST (Edgar 2010). We used a low similarity threshold to account for the diverse soil taxa that are not represented in the database.

Minimum generation time analysis

To estimate mean minimum generation times (i.e. community aggregated growth rates) from 16S rRNA sequence data, minimum generation times were estimated from published genomes and linked to bacterial community composition by matching 16S rRNA gene sequences. Fasta files of all open reading frames from 1136 genomes were obtained from the Joint Genome Institute's IMG database. These genomes were selected to only include non-host-associated organisms based upon information available in the Genomes OnLine Database (GOLD). Both IMG and GOLD were accessed in September of 2012. Minimum generation time (MGT) was estimated for each genome based upon the methods described by (Vieira-Silva and Rocha 2010) and python code graciously provided by the authors. Briefly, two measures of codon usage bias (ENC' and S) were combined via regression to generate a predictive model of MGT. MGTs were linked to phylotypes observed through 16S rRNA data by matching the 16S rRNA sequences in our dataset to the 16S rRNA sequences from the genomes at a 97% similarity threshold using the 'usearch_global' function in USEARCH (Edgar 2010) with 'maxaccepts' and 'maxrejects' set to 0 to disable approximate clustering. Mean MGT was calculated on a per sample basis by calculating proportional abundances of the phylotypes that matched represented genomes, excluding all other phylotypes, and multiplying it's estimated MGT by its proportional abundance. The mean MGT was the sum of these values.

Statistical analysis

Microbial diversity was calculated using Shannon diversity since this has been found to be a more reliable metric than species richness with microbial sequence data from complex communities (Haegeman et al. 2013). However, richness values are provided for interpretation and comparative purposes. Kruskal-Wallis tests were used to test whether there were significant differences in diversity among sites. Linear mixed effect models were used to test for differences in diversity with nutrient additions (+N, +P, +NP), using Genstat 14.0. We excluded the +NP term if not significant and treated site as a random variable.

We tested for significant shifts in overall community composition or functional gene composition across sites or with nutrient additions using PERMANOVA implemented in the 'adonis' function in the vegan package in R v. 3.0.2 (R Core Team 2013). Community or genetic composition was represented by Bray-Curtis dissimilarity matrices computed from square-root transformed abundance tables. N and P addition were used as predictor variables and the geographic site was included as 'strata', which restricts permutations to within sites. N × P interactions were found to be non-significant in all cases and therefore removed from final PERMANOVA models. When visualizing potential treatment differences, we used constrained ordination as implemented in the ‘capscale’ function in the vegan package in R. To test for significant differences in the relative abundance of specific taxa, gene categories, and MGTs between control and nutrient treatment plots, we used linear mixed effects models where each nutrient treatment (N and P additions) was a fixed factor, and site was a random effect. Relative abundances were rank transformed in order to meet the assumptions of the model. Tests were run for each taxon represented by a median of at least 1% of the sequences in any of the nutrient treatments, and false discovery rate (FDR) corrections were used to account for the multiple comparisons.

Relationships between pre-treatment environmental variables and changes in the relative abundance of taxa were assessed using Pearson correlations with mean difference from control plots (on a per site basis) as the response variable. We chose environmental variables to test based on their relationships with one another and avoided including multiple variables that were highly correlated with one another ($r^2 > 0.8$). Differences in post-treatment soil pH were investigated using a linear mixed effects model with N and P additions as fixed effects and site as a random effect. This analysis was conducted using the 'nlme' package in R.

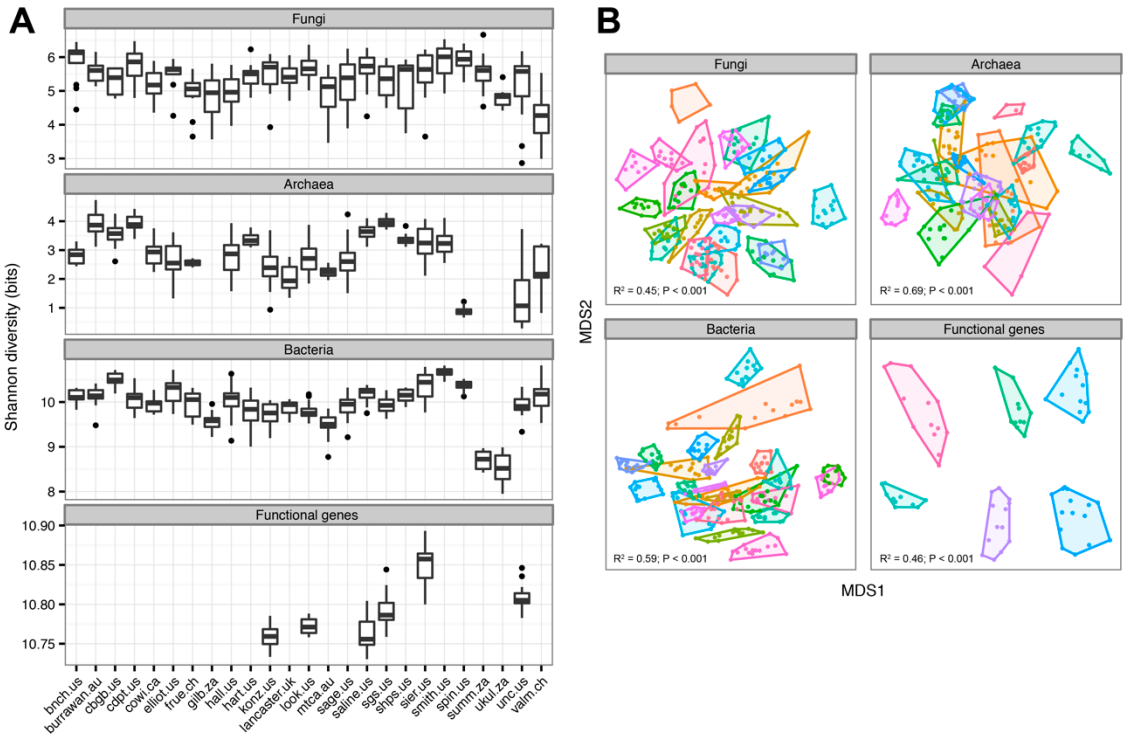

Figure A5.1. Microbial community structure and functional gene structure across sites. Boxplots showing Shannon diversity (A) and Nonmetric multidimensional scaling ordinations showing compositional differences (B) across sites for the three microbial taxonomic groups and the functional genes. Points in (B) represent individual samples and are colored by site. PERMANOVA statistics refer to site effects.

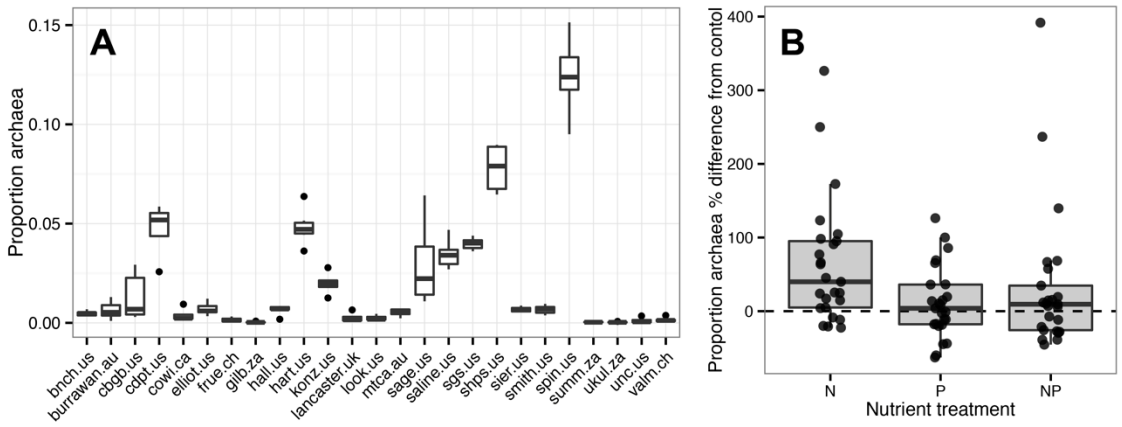

Figure A5.2. The proportion of 16S rRNA gene sequences identified as archaeal (versus bacterial) across sites (A) and changes in this proportion with nutrient additions within sites (B). Proportions were calculated after rarefying the full 16S rRNA dataset (including bacterial and archaeal sequences) to 18,000 sequences per sample.

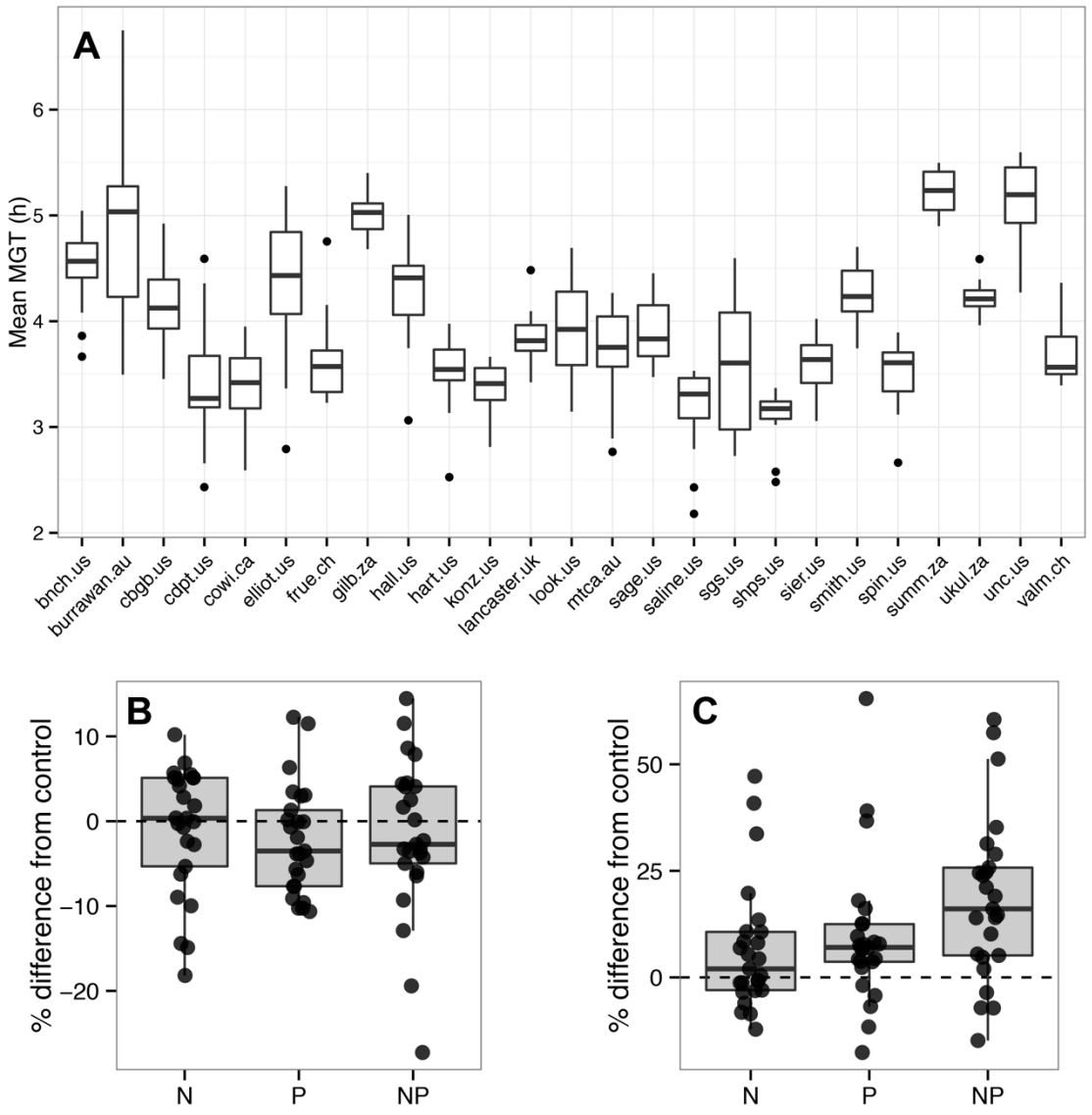

Figure A5.3. Minimum generation times (MGTs) estimated by matching 16S rRNA gene sequences to whole genomes. Variability across sites (A) and differences from control plots versus plots receiving nutrient additions within sites (B). Differences in the proportion of sequences matching whole genomes with nutrient additions within sites (C).

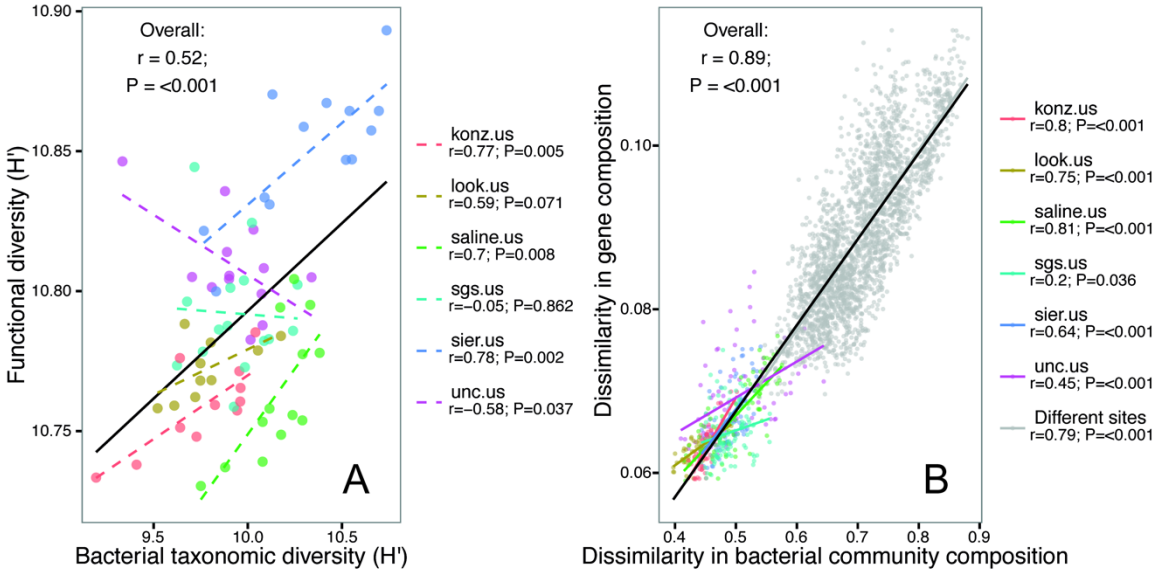

Figure A5.4. Relationships between functional gene diversity and bacterial taxonomic diversity (A) and between pairwise differences in functional gene composition and pairwise differences in bacterial community composition across all samples (B). Colored lines show within-site relationships, and black lines show cross-site relationships.

Table A5.1. Summary statistics for phylotype richness and the relative abundances of major phyla (mean ≥ 2%) from the three microbial groups. Values were calculated from all samples including those from nutrient treated plots. Bacterial, archaeal, and fungal samples were rarefied to 18,000, 100, and 485 sequences, respectively.

	Mean	**SD**	**Min**	**Max**
Fungal richness	105	14	78	134
Archaeal richness	17	5	7	26
Bacterial richness	3533	442	2560	4410
Relative abundances				
Fungal phyla				
Ascomycota	0.558	0.122	0.306	0.809
Basidiomycota	0.134	0.054	0.037	0.262
Glomeromycota	0.034	0.024	0.002	0.072
Zygomycota	0.169	0.124	0.010	0.465
Unclassified fungi	0.097	0.044	0.039	0.245
Archaeal phyla				
Crenarchaeota	0.922	0.066	0.769	0.994
Euryarchaeota	0.043	0.047	0.002	0.181
[Parvarchaeota]	0.033	0.035	0.000	0.121
Bacterial phyla				
Acidobacteria	0.203	0.059	0.130	0.379
Actinobacteria	0.112	0.044	0.058	0.227
Bacteroidetes	0.105	0.041	0.031	0.202
Chloroflexi	0.036	0.021	0.016	0.106
Firmicutes	0.033	0.019	0.003	0.070
Planctomycetes	0.043	0.012	0.023	0.074
Proteobacteria	0.205	0.041	0.114	0.274
Verrucomicrobia	0.182	0.092	0.036	0.433

Table A5.2. Location of study sites and selected environmental data. NA = Not available.

Site code	Site name	Continent	Country	Ecosystem	Latitude	Longitude	Elev (m)	Years treated	Used in meta-genomic analysis
bnch.us	Bunchgrass (Andrews LTER)	N. America	USA	Montane grassland	44.28	-121.97	1318	4	
burrawan.au	Burrawan	Australasia	Australia	Semiarid grassland	-27.73	151.14	425	4	
cbgb.us	Chichaqua Bottoms	N. America	USA	Tallgrass prairie	41.79	-93.39	275	2	
cdpt.us	Cedar Point Biological Station	N. America	USA	Shortgrass prairie	41.2	-101.63	965	4	
cowi.ca	Cowichan	N. America	Canada	Old field	48.46	-123.38	50	4	
elliot.us	Elliott Chaparral	N. America	USA	Annual grassland	32.88	-117.05	200	3	
frue.ch	Fruebuel	Europe	Switzerland	Pasture	47.11	8.54	995	3	
gilb.za	Mt Gilboa	Africa	South Africa	Montane grassland	-29.28	30.29	1748	2	
hall.us	Hall's Prairie	N. America	USA	Tallgrass prairie	36.87	-86.7	194	4	
hart.us	Hart Mountain	N. America	USA	Shrub steppe	42.72	-119.5	1508	4	
konz.us	Konza LTER	N. America	USA	Tallgrass prairie	39.07	-96.58	440	4	Y
lancaster.uk	Lancaster	Europe	UK	Mesic grassland	53.99	-2.63	180	3	
look.us	Lookout (Andrews LTER)	N. America	USA	Montane grassland	44.21	-122.13	1500	4	Y
mtca.au	Mt. Caroline	Australasia	Australia	Savanna	-31.78	117.61	285	3	
sage.us	Sagehen Creek UCNRS	N. America	USA	Montane grassland	39.43	-120.24	1920	4	
saline.us	Saline Experimental Range	N. America	USA	Mixedgrass prairie	39.05	-99.1	440	4	Y
sgs.us	Shortgrass Steppe LTER	N. America	USA	Shortgrass prairie	40.82	-104.77	1650	4	Y
shps.us	Sheep Experimental Station	N. America	USA	Shrub steppe	44.24	-112.2	910	4	
sier.us	Sierra Foothills REC	N. America	USA	Annual grassland	39.24	-121.28	197	4	Y
smith.us	Smith Prairie	N. America	USA	Mesic grassland	48.21	-122.62	62	4	
spin.us	Spindletop	N. America	USA	Pasture	38.14	-84.5	271	4	
summ.za	Summerveld	Africa	South Africa	Mesic grassland	-29.81	30.72	679	2	
ukul.za	Ukulinga	Africa	South Africa	Mesic grassland	-29.67	30.4	842	3	
unc.us	Duke Forest	N. America	USA	Old field	36.01	-79.02	141	4	Y
valm.ch	Val Mustair	Europe	Switzerland	Alpine grassland	46.63	10.37	2320	3	

Table A5.2, continued

Site code	Mean annual temp. (ºC)	Mean annual precip. (mm)	Soil C (%)	Soil N (%)	Soil P (ppm)	Soil pH	Sand (%)	Silt (%)	Clay (%)
bnch.us	5.5	1647	8.3	0.6	13.1	5.6	70.4	26.5	2.9
burrawan.au	18.4	683	1.2	0.1	18.0	5.6	82.2	9.1	8.6
cbgb.us	9	855	0.6	0.1	63.0	6.1	88.7	6.4	4.8
cdpt.us	9.5	445	1.5	0.1	31.9	6.7	68.0	21.7	10.2
cowi.ca	9.8	764	5.3	0.4	41.1	5.6	31.5	40.2	28.3
elliot.us	17.2	331	2.0	0.1	16.3	5.7	54.2	25.6	20.1
frue.ch	6.5	1355	3.8	0.4	69.3	5.5	38.2	41.0	20.6
gilb.za	13.1	926	20.4	1.2	17.9	5.1	NA	NA	NA
hall.us	13.6	1282	1.4	0.1	33.2	5.2	25.1	59.3	15.5
hart.us	7.4	272	1.1	0.1	64.6	7.2	47.7	22.7	29.5
konz.us	11.9	877	NA	NA	NA	NA	NA	NA	NA
lancaster.uk	8	1322	20.9	1.1	33.2	4.7	50.2	31.1	18.6
look.us	4.8	1898	16.7	1.2	54.9	5.1	70.0	29.1	0.8
mtca.au	17.3	330	1.3	0.1	8.7	5.2	82.0	11.3	6.6
sage.us	5.7	882	8.9	0.7	35.5	6.1	44.8	30.9	24.1
saline.us	11.8	607	NA	NA	NA	NA	NA	NA	NA
sgs.us	8.4	365	0.8	0.1	65.9	6.1	73.0	15.1	11.8
shps.us	5.5	262	2.2	0.2	35.0	8.0	50.7	37.8	11.5
sier.us	15.6	935	2.1	0.2	14.7	6.0	38.7	42.4	18.8
smith.us	9.8	597	7.5	0.6	76.3	6.1	78.0	15.1	6.8
spin.us	12.5	1140	2.7	0.3	233.4	6.4	29.3	49.8	20.8
summ.za	18.2	939	6.8	0.3	12.5	5.1	NA	NA	NA
ukul.za	18.1	880	5.0	0.3	9.3	5.8	18.3	37.0	44.6
unc.us	14.6	1163	2.2	0.2	21.3	5.3	56.0	22.6	21.3
valm.ch	0.3	1098	7.4	0.6	46.0	5.5	57.6	29.1	13.3

Table A5.3. Median relative abundances (%) of higher-level taxa among the control and nutrient treatment samples. Only taxa ≥ 0.5% of sequences within at least one treatment are shown. *P*-value corrections were made within taxonomic level for Fungi.

	P (FDR corrected)	Control	N	P	NP
Fungi					
Glomeromycota	0.000	2.5	2.3	2.9	1.2
Glomeraceae	0.000	1.9	1.3	1.4	0.8
Ascomycota	0.027	54.2	58.1	56.7	61.2
Nectriaceae	0.001	4.3	6.4	5.2	6.8
Mycosphaerellaceae	0.011	0.4	0.4	0.4	0.6
Plectosphaerellaceae	0.084	0.6	1.0	1.2	0.7
Pleosporaceae	0.229	1.6	1.4	1.6	1.8
Basidiomycota	0.058	12.0	10.6	11.1	10.4
Filobasidiaceae	0.005	0.4	0.8	0.6	0.9
Zygomycota	0.910	12.0	10.9	13.6	11.4
Mortierellaceae	0.867	11.3	10.1	11.5	11.2
Bacteria					
Alphaproteobacteria	0.002	7.8	8.0	8.0	8.5
Deltaproteobacteria	0.002	4.3	3.8	4.3	3.8
Gammaproteobacteria	0.002	2.9	3.2	3.2	3.7
Acidobacteria	0.007	19.8	18.4	20.0	18.0
Planctomycetes	0.006	4.1	4.0	4.3	3.7
Actinobacteria	0.014	10.2	10.8	10.1	10.3
Bacteroidetes	0.064	9.7	9.3	10.8	9.8
Verrucomicrobia	0.455	16.7	16.7	15.7	16.6
Chloroflexi	0.501	2.9	2.9	2.7	2.9
Firmicutes	0.516	2.2	2.7	2.7	2.2
Gemmatimonadetes	0.576	1.6	1.6	1.6	1.5
Betaproteobacteria	0.570	3.6	3.6	3.7	3.7
Archaea					
Crenarchaeota	0.000	94.0	98.0	94.0	98.0
Euryarchaeota	0.000	2.0	1.0	2.5	1.0
[Parvarchaeota]	0.001	2.0	0.5	1.0	1.0

Table A5.4. Correlations between change in environmental variables and changes in the relative abundances of taxonomic groups with N and P additions. Correlations were assessed using mean values for each site. Variables refer to percent differences from mean values in control plots. Numbers in bold type identify correlations that $P \leq 0.05$.

	Fungi											
	Ascomycota		Glomeromycota									
Variable	r	*P*	r	*P*								
(soil pH change)^2	-0.38	0.065	0.01	0.95								
ln(soil P change)	-0.33	0.119	0.17	0.428								
sqrt(soil N change)	0.37	0.075	-0.16	0.462								
sqrt(live plant biomass change)	-0.15	0.491	-0.17	0.431								
	Archaea											
	[Parvarchaeota]		Crenarchaeota		Euryarchaeota							
Variable	r	*P*	r	*P*	r	*P*						
(soil pH change)^2	-0.01	0.98	0.35	0.126	-0.01	0.972						
ln(soil P change)	0.44	0.071	0.06	0.793	-0.11	0.656						
sqrt(soil N change)	**0.56**	**0.015**	-0.06	0.796	0.3	0.202						
sqrt(live plant biomass change)	0.16	0.519	**-0.5**	**0.024**	0.36	0.122						
	Bacteria											
	Acidobacteria		Actinobacteria		Alphaproteobacteria		Deltaproteobacteria		Gammaproteobacteria		Planctomycetes	
Variable	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*
(soil pH change)^2	0.03	0.889	0.06	0.794	-0.21	0.32	-0.1	0.644	-0.23	0.283	0.17	0.422
ln(soil P change)	-0.04	0.847	0.38	0.065	**0.45**	**0.027**	-0.18	0.406	0.11	0.605	-0.12	0.592
sqrt(soil N change)	-0.34	0.101	0.07	0.729	-0.01	0.954	-0.14	0.513	0.35	0.098	0.02	0.929
sqrt(live plant biomass change)	0.11	0.602	0.32	0.131	0.25	0.233	-0.07	0.741	-0.11	0.602	0.1	0.635

Table A5.5. Median percent differences in relative abundances of gene categories of nutrient treated samples from within-site control plots.

			Difference from control (%)		
Gene category	***P***	***P* (FDR-corrected)**	**+N**	**+P**	**+N,+P**
Metabolism (unclassified)	0.002	0.048	-1.00	-0.31	-0.76
Translation	0.016	0.180	0.04	0.70	0.80
Carbohydrate Metabolism	0.017	0.129	0.72	0.02	0.37
Genetic Information Processing	0.031	0.178	0.83	-1.82	-1.39
Enzyme Families	0.032	0.145	-0.89	-1.64	-1.16
Replication and Repair	0.069	0.264	-0.07	0.54	0.96
Nucleotide Metabolism	0.111	0.364	0.22	0.44	0.63
Folding, Sorting and Degradation	0.120	0.344	-0.17	0.25	-0.75
Metabolism of Other Amino Acids	0.262	0.669	-0.12	-0.50	0.05
Cell Motility	0.331	0.762	-0.82	0.74	1.46
Energy Metabolism	0.334	0.699	0.44	0.62	0.53
Metabolism of Terpenoids and Polyketides	0.345	0.660	0.80	-0.38	0.46
Cellular Processes and Signaling	0.373	0.660	-1.19	0.46	0.05
Xenobiotics Biodegradation and Metabolism	0.463	0.760	0.18	-0.11	0.12
Amino Acid Metabolism	0.480	0.737	0.01	0.39	0.67
Lipid Metabolism	0.536	0.770	0.20	-0.72	-0.21
Transcription	0.542	0.733	0.73	-0.75	-0.64
Glycan Biosynthesis and Metabolism	0.711	0.908	-0.22	0.44	0.29
Metabolism of Cofactors and Vitamins	0.742	0.898	-0.13	0.10	0.01
Biosynthesis of Other Secondary Metabolites	0.806	0.927	0.42	-0.86	0.32
Membrane Transport	0.808	0.885	-0.62	0.16	0.25
Signal Transduction	0.863	0.902	-0.58	-0.16	-0.26
Poorly Characterized	0.918	0.918	-0.15	-0.17	-0.15