UNDERSTANDING DRIVERS OF PLANT MICROBIOME IN MICHIGAN AGRICULTURE: STUDIES OF THE APPLE ROOT ZONE AND COMMON BEAN SEEDS

By

Ari Fina Bintarti

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Crop and Soil Sciences – Doctor of Philosophy

2022

ABSTRACT

UNDERSTANDING DRIVERS OF PLANT MICROBIOME IN MICHIGAN AGRICULTURE: STUDIES OF THE APPLE ROOT ZONE AND COMMON BEAN SEEDS

By

Ari Fina Bintarti

Plant-associated microbial communities are crucial for plant health and fitness, and may enhance plant tolerance to various environmental stresses. As global climate change threatens crop production and increases demands on sustainable agriculture, harnessing the plant microbiome has become one potential strategy to address these issues. Thus, it is fundamental to understand the relative contributions of both the host plant as well as the environment in shaping the plant microbiome. Moreover, the response of plant microbiomes to stress and any consequences of microbiome stress responses for the host plants are poorly understood, though this information is critical to achieve a basis of knowledge for plant microbiome engineering. My research aimed to contribute to this knowledge by investigating the factors that structure root- and seed-associated microbial communities of two valuable crops for Michigan's agricultural economy: apple and common bean.

The first chapter of my dissertation aimed to assess the biogeography of bacterial, archaeal, fungal, and nematode communities in the root zone of apple trees, and to determine their relationships with each other and their changes over natural abiotic gradients across orchards. I also assessed the influence of plant cultivar on microbiome structure in the root zone. I found that root zone microbiome community structure was strongly affected by geographic location and edaphic properties of soil. The next chapter of my dissertation investigated the variability of seed endophyte community of common bean (*Phaseolus vulgaris* L.). My results showed that plant-to-plant variability under controlled growth conditions exceeded within-plant

variability among seeds from different pods. My study developed protocols and added insights to the growing toolkit of approaches to understand the plant-microbiome engagements that support the health of agricultural and environmental ecosystems. The last chapter assessed the responses of common bean seed endophytes to drought stress in the field across two growing locations and four genotypes of common bean. To summarize, this work advances foundational knowledge of the seed microbiome as a critical component of the plant microbiome, and in the context of two key crops for Michigan agriculture.

This dissertation is dedicated to my late mother (Ibu) and to my father (Bapak) who keeps me in his prayers in the middle of the night

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support and help I have received from many people. First of all, I would like to thank my PhD advisor, Dr. Ashley Shade. She interviewed me directly and accepted me as her graduate student as well as a research assistant in her lab. Therefore, I did not need to spend my first year doing lab rotation and that was a plus for my doctoral program. She mentored and taught me to conduct good research and developed a new set of skills that would be essential for my future career. She worked with me and supported me during the most challenging time of graduate school. Moreover, she also supported me to pursue my career in research. I admire her optimism and trust in me, even when I doubt myself.

I would also like to thank my guidance committee members, Dr. Gregory Bonito, Dr. Karen Cichy, and Dr. Edward Walker for their invaluable science feedbacks and thoughtful insights in building my research and supporting me to be a better scientist. Thank you to the past and present members of Shade Lab: Nejc Stopnisek, Pat Kearns, Alan Bowsher, Taylor Dunivin, Jackson Sorensen, John Chodkowski, Jacqueline Carroll, Joanna Colovas, Abby Sulesky, Keean Dolan, Bryce Davis, Lille Cunic, Sreejata Bandopadhyay, Marco Llontop, Samuel Barnett, and Oishi Bagchi for your help and support. Special thanks to the Shade Lab manager, Keara Grady, for your patience to help me during my research.

I would like to thank MSU OISS and IIE (The Fulbright Foreign Student Program) for the funding and the chance I awarded to pursue a doctoral degree at Michigan State University to chase my dream as a scientist. I want to thank the Department of Plant, Soil and Microbial

Sciences for accepting me as a graduate student and providing me with their professional development.

Finally, I am forever grateful for my family and friends for their relentless support. I especially want to thank my father, my mom, and my brothers, Bimantara Gilang Anggara and Zuhanida Satria Pamungkas for always checking upon me.

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES xi

KEY TO ABBREVIATIONS xiii

CHAPTER 1: Introduction 1
Plant – microbe interactions 2
Root-associated microbiome in perennial tree crops 3
Seed-associated microbiome and its implications for plant fitness 7
Response of plant microbiome under stress – an overview of biotic v. abiotic stress 9
Overview of the study 12
REFERENCES 15

CHAPTER 2: Biogeography and diversity of multi-trophic root zone microbiomes in Michigan apple orchards: analysis of rootstock, scion, and local growing region 26
Abstract 27
Introduction 28
Materials and Methods 29
Survey design, soil sample collection, and apple rootstocks and scions 29
Nematode, oligochaete, and mycorrhizal fungal quantification 30
Microbial DNA extraction and PCR amplification 31
Sequencing data analysis and OTU clustering 32
Microbial community analysis 34
Network analysis 36
Data and code availability 36
Results 37
Sequencing summary 37
Bacterial and fungal alpha diversity among sites, rootstocks, and scions 37
Nematodes and other groups: alpha diversity 42
Microbial beta diversity 45
Microbial community composition in apple root zone 47
Microbial network of apple root zone 51
Discussion 53
APPENDICES 57
APPENDIX A: Supplemental Tables 58
APPENDIX B: Supplemental Figures 98
APPENDIX C: Supplemental Information 108
REFERENCES 115

CHAPTER 3: Endophytic microbiome variation among single plant seeds 122
Abstract 123

Introduction 124
Materials and Methods 126
Growth conditions for parent plants 126
Study design 127
Seed harvest and endophyte microbial DNA extraction 127
PCR amplification and amplicon sequencing 130
Sequence analysis 131
Microbial community analysis 132
Data and code availability 135
Results 135
Sequencing summary and microbiome coverage 135
Microbiome diversity 138
Bean seed microbiome composition 144
Shared taxa among seeds and plants 148
Discussion 148
APPENDICES 153
APPENDIX A: Supplemental Table 154
APPENDIX B: Supplemental Figures 162
APPENDIX C: Supplemental Information, Results, and Protocols 166
REFERENCES 185

CHAPTER 4: Responses of seed endophytes under drought stress: Field study 192
Abstract 193
Introduction 193
Materials and methods 196
Plant cultivars 196
Field study design 198
Seed preparation and endophyte microbial DNA extraction 199
PCR amplification and amplicon sequencing 199
Sequencing analysis and OTU clustering 200
Seed-associated microbial community analysis 201
Data and code availability 203
Results and Discussion 203
Future Directions 208
APPENDIX 209
REFERENCES 212

CHAPTER 5: Conclusions and Future Directions 219
Summary 220
Future Directions 222
REFERENCES 226

LIST OF TABLES

Table 2.1. Statistical analysis of microbial richness and Shannon diversity index among sites (n = 20) and rootstocks (n = 8) using Kruskal-Wallis and one-way (ANOVA)[a] 41

Table 2.2. Environmental factors that have explanatory value for the bacterial/archaeal and fungal communities were fitted into principal coordinate analysis plot (beta diversity was calculated using Bray-Curtis dissimilarity indices) and tested using permutation test using "envfit" function in vegan package (v2.5-4)[a] 46

Table 2.3. Soil physical and chemical properties of 45 soil samples taken from 20 different sites of apple orchards in Michigan 59

Table 2.4. Comparison of bacterial and archaeal richness and Pielou's evenness among sites using post-hoc Dunn's test with Benjamini-Hochberg correction of p-values and post-hoc Tukey's HSD test, respectively 67

Table 2.5. Comparison of bacterial and archaeal richness and Pielou's evenness among rootstocks using post-hoc Dunn's test with Benjamini-Hochberg correction of p-values and post-hoc Tukey's HSD test, respectively 73

Table 2.6. Pearson's correlation analysis between alpha diversity metrics and soil parameters. r = correlation coefficient. Significant correlations (p-val < 0.05) are in bold 74

Table 2.7. Comparison using post-hoc Tukey's HSD test of fungal richness among sites..... 75

Table 2.8. Comparison using post-hoc Tukey's HSD test of fungal richness among rootstocks . 81

Table 2.9. The abundance of nematodes, mycorrhizal fungi, and Oligochaetes from 45 soil samples..... 82

Table 2.10. Prevalence percentage of nematodes and other soil microorganisms among 45 soil samples taken from 20 different sites 84

Table 2.11. Comparison using post-hoc Tukey's HSD test of total absolute abundance of nematodes among rootstocks 85

Table 2.12. Comparison using post-hoc Tukey's HSD test of oligochaetes abundance among sites 86

Table 2.13. Comparison using post-hoc Tukey's HSD test of mycorrhizal fungi abundance among sites..... 92

Table 3.1. Parent plant yield information and seed samples used in microbiome analyses 129

Table 3.2. List of microbial taxa identified in more than half of total seed samples (occupancy > 0.5, n = 47 for bacteria/archaea); and microbial taxa shared across plants (occupancy = 1, n = 3) and across seeds within plant (occupancy > 0.5) 155

Table 4.1. Description of the common bean cultivars used in this study 197

LIST OF FIGURES

Figure 2.1. Alpha diversity of apple root zone microbiome among rootstocks. 40

Figure 2.2. Absolute abundances of apple root zone multi-trophic levels among sites and rootstocks. 44

Figure 2.3. Relative abundances of apple root zone microbiome among sites and rootstocks..... 48

Figure 2.4. Relative abundances of apple root zone microbiome core taxa. 50

Figure 2.5. Co-occurrence network of apple root zone multi-trophic levels. 52

Figure 2.6. Map of the sampling location in Michigan apple orchards. 99

Figure 2.7. Rarefaction curves. 100

Figure 2.8. Alpha diversity of apple root zone microbiome among sites. 101

Figure 2.9. The linear regression relationship between bacterial/archaeal alpha diversity and soil parameters and nematodes. 102

Figure 2.10. The linear regression relationship between fungal alpha diversity and soil parameters. 103

Figure 2.11. The linear regression relationship between bacterial/archaeal and nematodes alpha diversity..... 104

Figure 2.12. PCoA plot of apple root zone microbiome among sites. 105

Figure 2.13. PCoA plot of apple root zone microbiome among rootstocks. 106

Figure 2.14. Occupancy vs. abundance plots of apple root zone microbiome 107

Figure 3.1. Rarefaction curves of common bean seed microbiome. 136

Figure 3.2. Alpha diversity of common bean seed microbiome among plants and pods. 139

Figure 3.3. Beta diversity visualizations of common bean seed microbiome based on Jaccard index. 142

Figure 3.4. Relative abundances of common bean seed microbiome. 145

Figure 3.5. Power analysis. 147

Figure 3.6. The proportion of plant reads. 163

Figure 3.7. Beta diversity visualization of the common bean seed microbiome based on Bray-Curtis dissimilarities. 164

Figure 4.1. Plant yield among cultivars in both locations..... 205

Figure 4.2. Plant yield between treatments within cultivar in both locations..... 206

KEY TO ABBREVIATIONS

ABA	Abscisic acid
ACC	1-aminocyclopropane-1-carboxylic acid
ANOVA	Analysis of variance
AOA	Ammonia-oxidizing archaea
AOB	Ammonia-oxidizing bacteria
BCMV	Bean common mosaic virus
BR	Broad-Range
Ca	Calcium
CBB	Common bacterial blight
CONSTAX	Consensus taxonomy
CSS	Cumulative sum scaling
DNA	Deoxyribonucleic acid
dsDNA	Double stranded deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
HSD	Honest significant difference
IAA	Indole-3-acetic acid
ITS	Internal transcribed spacer
K	Potassium
LMM	Linear mixed-effects model
M	Mean
MAFFT	Multiple alignment using fast fourier transform

MENA	Molecular ecological network analysis
Mg	Magnesium
N	Nitrogen
NCBI	National Center for Biotechnology Information
NH_4N	Ammonium
NO_3N	Nitrate
OM	Organic matter
OTU	Operational taxonomic unit
P	Phosphorus
PBS	Phosphate-buffered saline
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PDA	Potato dextrose agar
PERMANOVA	Permutational multivariate analysis of variance
PNA	Peptide nucleic acid
PROTEST	Procrustean randomization test
RDP	Ribosomal Database Project
RMT	Random matrix theory
rRNA	Ribosomal ribonucleic acid
RTSF	Research Technology Support Facility
SD	Standard deviation
SPNL	Soil and Plant Nutrient Laboratory
SRA	Sequence read archive

TE	Tris-EDTA
TSA	Trypticase soy agar
UPREC	Upper Peninsula Research and Extension Center
V4	16S fourth hypervariable

CHAPTER 1: Introduction

Plant – microbe interactions

Plant microbiota are defined as microbes including archaea, bacteria, fungi, and protists that associate with plants and inhabit different plant microhabitats including rhizosphere, phyllosphere, and endosphere (1). Aside from those three most common microhabitats, microbes are also present in other plant compartments such as flower (anthosphere), fruit (carposphere), stem (caulosphere), root surface (rhizoplane), germinating seed area (spermosphere), and the seed itself (2). It has been suggested that the plant and its microbial community co-evolve, and the evolutionary selection of the microbiome members and the host plant affects the system as a whole, which is referred to as the holobiont concept (3, 4). Plant microbiome also refers to the auxiliary genome of the plant, where the plant partially depends on their associated microbiota for specific functions and traits (5). Together, the interplay between the plant and its microbiota determines the structure and composition of the microbial communities and the characteristics of their local environment as well as the physiology of the hosts (6). It is widely known that plant microbiomes are essential for plant productivity and tolerance to various environmental stresses, for example, by providing active metabolites, such as enzymes and phytohormones (7). Meanwhile, the plant provides ecological niches for the associated microbes and contributes to the structure of the plant microbiome, for example, by producing root exudates or allelochemicals (8, 9).

Beneficial (mutualistic) plant microbiome members are fundamental for plant survival, although, plant microbiota possess a broad range of interactions with the host including those that are deleterious (pathogenic) as well as neutral (commensalistic). Previous studies have reported the beneficial function of plant microbiota for nutrient acquisition and inducing plant development by producing plant growth hormones (10), promoting plant tolerance to abiotic

stress (11), enhancing plant defense mechanisms to pathogen attack (12), and inducing flowering time (13). In addition, studies have investigated plant-microbe interactions, especially the rhizosphere and phyllosphere, to better understand how they can improve plant performance in a changing or stressful environment, such as during water limitation (14-17). Given these potential benefits of the plant microbiome, studies have focused on isolating and characterizing beneficial plant microbiome members and investigating their impact on plant productivity through culture-dependent and/or culture-independent methods (18-20).

Nowadays, studies on the assembly of plant-associated microbes as a community and the driving forces that structure the plant microbiome as well as their ecological function has grown tremendously in recent years with the rapid development of multi-omics technologies (21, 22). In addition, as global climate change threatens to limit crop production placing heightened demand on sustainable agriculture, harnessing the plant microbiome has become one potential crop management strategy to address these issues. Thus, this research is motivated to explore mechanisms of plant-microbiome interactions, and how these interactions affect plant performance, especially under environmental stresses.

Root-associated microbiome in perennial tree crops

The microbiome associated with plant roots and the rhizosphere, the intimate zone surrounding plant roots enriched in microbial activity, is the most well-studied plant microbiome because of its tremendous potential for plant fitness and health (5, 23, 24). The rhizosphere is sampled from the soil that remains closely adhered to the root system, which is heavily influenced by plant chemicals, making it a nutrient-rich hotspot that enables the growth of diverse microorganisms (25). As an estimation, per gram of plant roots are colonized by

approximately 10^9 - 10^{11} bacterial cells (26). Root-associated microbiome members are mainly acquired through horizontal transmission from the soil as the main microbial reservoir (1, 4), although, a fraction of root-associated microbiota may also be seed-borne or vertically transmitted from the parent plant and remain in the root and/or inhabit the rhizosphere during seed germination and plant development (27, 28).

It is known that the rhizosphere is one of the most complex ecological niches inhabited by multi-trophic microorganisms, where they form a complex network with each member of the community and significantly contribute to carbon and nitrogen cycling, and organic matter decomposition (29, 30). Previous studies reported that plants recruit their rhizosphere microbiome members by releasing a wide variety of compounds, such as alkaloids, sugars, flavonoids, amino acids, phenolics, enzymes, vitamins, and carbohydrates derived from plant metabolism and photosynthesis processes (31, 32). It is estimated that up to 40% of plant photosynthates are allocated to the rhizosphere (23, 33) and released into the soil through different mechanisms including secretion, diffusion, and cell lysis (32). These diverse groups of root metabolites help regulate the structure of the rhizosphere microbial community by acting as a chemical signal to mediate microbe-microbe and/or plant-microbe communication and interactions, suppressing the growth of competitor or plant pathogens by their antimicrobial activity, enhancing the growth of beneficial microbes, and altering the soil physicochemical properties (31, 33, 34).

The continuous influence of root-derived chemicals on the surrounding soil results in enrichment or loss of a subset of microbiota. Thus, some studies have revealed that the rhizosphere has lower microbial diversity compared to the surrounding bulk soil (35-37) and even much lower for the microbial diversity of root endophytes (36). In this aspect, root-derived

compounds may serve as a selective force that selects and shapes the root-associated microbiome members. The selective property of the host is also related to the plant life cycle as one of the important factors affecting the rhizosphere microbiome (e.g., perennial vs annual). In this case, we consider perennial as a woody tree, which can be distinguished from an herbaceous perennial.

The interactions between the rhizosphere microbiome and perennial trees are distinct from the interactions that occur in annual plants due to their longevity. It has been suggested that the host effect (host selective property) on the rhizosphere microbiome is much stronger in perennial trees than in annual plants (38, 39). Moreover, the rhizosphere microbial community of perennials is characterized by changes in both richness and composition over the plant lifetime (38, 40). It could be assumed that rhizosphere microbiota associated with long-lived perennial trees are consecutively affected by the selective nature of plants over a long period of the plant's age, which eventually alters the structure of the rhizosphere microbiome throughout the plant lifespan. The dynamic changes in the rhizosphere microbiome structure of perennial trees are the result of the constant adaptation of the host plant to the environment due to seasonal variations (39).

Unlike annuals, the cultivation of perennials is often undisturbed by anthropogenic activities, such as crop rotation and soil tillage, resulting in a steady and persistent flow of photosynthates to their associated microbiota that may favor a particular subset of either beneficial or pathogenic microbes (39). Because of these established and prolonged interactions between root-associated microbiota and perennial trees, it can be expected that the productivity of perennial crops is profoundly influenced by these beneficial or pathogenic microbes. For instance, perennial crop growth and productivity were reported to be significantly enhanced by mycorrhizal colonization compared to annual crops (41). In the particular case of negative

impact, in perennial tree crops with neither replanting nor crop rotation practices, the accumulation of pathogens in the root zone caused a negative soil feedback phenomenon which ultimately results in reduced yields (42, 43).

The diversity and composition of the rhizosphere microbiome are also driven by environmental factors (e.g. climate, weather, rainfall) and soil characteristics (44, 45). In addition, plant genotype also has notable explanatory value on rhizosphere microbiome of woody crops, as previously observed in apple trees, where rootstock genotype determines the structure and composition of rhizosphere microbiome (46), and the same case was also found in grapevine (47). Analysis of rhizosphere microbiome of *Populus* tree showed the differences between wild-type and transgenic line which indicates the effect of plant genotype (48). However, it has been reported that soil edaphic and environmental factors have a stronger impact on the rhizosphere microbiome than plant genotype or species, especially under field conditions (37, 49, 50).

Understanding the driving forces that structure the rhizosphere microbiome of perennial tree crops, as well as the dynamic changes of the microbial community over the lifetime of perennials, are an essential part of harnessing plant microbiome for enhancing crop production. However, the study focusing on the interaction between root-associated microbiome with perennial tress is still scarce relative to annual plants, mainly due to the natural longevity of perennials. Moreover, since one growing season of a perennial tree does not represent the entire plant lifetime, a long-term (temporal) study is desired to better understand the variations of perennial root-associated microbial communities.

Seed-associated microbiome and its implications for plant fitness

In contrast to the rhizosphere, the microbial community in other plant habitats, especially in the seed, is relatively less studied. This is because, unlike rhizosphere, seed bears relatively low microbial biomass or is even believed to be the lowest among all plant compartments (51). In addition, the seed microbiome is often neglected due to the assumption that soil is the only main source of plant microbiota through horizontal transmission (52). Nevertheless, the study on seed microbiome has been increasing in recent years, although the functional aspects of seed microbiome have been largely unexamined.

Seed is essential for plants, especially spermatophytes, because it initiates plant growth and development, carries plant genetic information, which is then expressed in new plant generation, and acts as an ecological tool for the plant and microbial dispersal. Study on seed-associated microbiota is primarily encouraged by the later proven assumption that plant microbiome members are vertically transmitted from parent plants to their offsprings through seed (53, 54). Moreover, seeds are attractive because they represent a starting point for plant microbiome assembly (2), where seed-associated microbiome members can be considered early colonizers, potentially influencing the plant microbiome's structure during plant development(55). When a seed germinates, the seed-borne endophytes, which are mostly believed to be dormant inside the seed (2, 56), will be active and colonize the seedling, and together with the seed surrounding (spermosphere) microbiota, play an essential role in driving the plant microbiome ecology and determining the host plant physiology and function (57, 58).

Other than vertically transmitted microbiota, microbes from the surrounding environment (e.g., soil, leaves and fruit surfaces, residues) also colonize seed through horizontal transmission (2, 59). It is believed that early colonization determines successful colonization, and seed

endophytes are considered as microbes with successful and established colonization within the seed tissues. Successful colonization means seed endophyte candidates are capable of establishing an association with plant tissues inside seed compartments (e.g., seed coat, endosperm, and embryo) without causing visible harm to the host plant (60). Moreover, endophytes inside the seed are assumed to be unique or possess distinct features from endophytes in other plant compartments (60, 61). The bottleneck of seed endophyte colonization is the unfavorable seed environment, meaning that, a successful colonizer must be able to survive and cope with high desiccation, long exposure to high osmotic pressure, antimicrobial compounds, and starch contents inside the seed (54, 60, 62). Seed endophytes, such as some Firmicutes and *Bacillus* can form endospore that protects them in changing environment during seed maturation (63). Cell motility is another unique feature possessed by seed endophytes, which allows them to enter seeds before they harden (60, 64). In addition, it has been suggested that seed endophytes can use starch through amylase activity (62), as well as phytate as the main form of phosphorus in the seed (65).

Seed microbiome supports plant fitness, especially during the early stages of plant development (66-68). The plant likely selects beneficial microbiota and stores them in the seed (69), which are then passed to its progeny over generations to ensure beneficial and successful colonization from the earliest stage of plant development. Seed microbiome members provide benefits for plant (70). They can stimulate germination and promote seedling growth (71) by producing phytohormones, such as auxin (IAA) (20, 72), cytokinin (66), and gibberellin (73), or through phosphate solubilization (67), and producing siderophores (20). A study revealed that seed-endophytes removal from rice seeds reduces seedlings development and re-inoculation of seed-endophyte isolates recovers the seedling development (67) and a similar scenario has also

been recently reported in pearl millet (74). Seed endophytes also offer protection against soil-borne pathogens (75, 76), for example, by producing cell wall-degrading enzymes (β-1,3-glucanase, cellulase, chitinase) (20). A previous study showed that volatile compounds produced by seed endophytes of wild cabbage not only promote seed germination but also inhibit the growth of pathogenic fungi and increase the mortality of cabbage moth larvae (77). Another recent study observed that pearl millet seed endophytes can produce lipopeptides that have antifungal activity against fungal phytopathogens (74). Moreover, several seed endophytes isolated from peanuts are reported to be able to produce ACC deaminase, which is important for lowering the ethylene level and alleviating salinity stress in the plant (78). On the other hand, it is known that seeds can also be inhabited by certain pathogens (seed-borne pathogens) (61). Indeed, previous studies on seed-associated microbiota have largely focused on the identification of seed-borne pathogens and transmission of certain pathogen species (70, 79, 80).

Similar to other beneficial plant microbiome, efforts are directed toward exploring and harnessing beneficial seed endophytes for crop production either by inoculation of beneficial endophytes into the plant or manipulating the native plant microbiota. To be able to engineer seed endophytes, it is important to dig deeper into the driving factors that structure seed microbiome assembly as well as the mechanisms behind the transmission and preservation of beneficial seed endophytes over plant generations.

Response of plant microbiome under stress – an overview of biotic v. abiotic stress

It is widely known that the plant microbiome is a major component of plant health including protection against a wide range of environmental stresses. Environmental stresses, which can be classified into biotic and abiotic stresses, are the main challenge in agriculture and

crop production worldwide as they cause significant yield reductions. It is estimated that abiotic stress leads to more than 50 % crop yield reduction (81). Moreover, this condition is exacerbated by the presence of pathogen and disease attacks that become more severe due to uncertain climate changes. Even though plants have adaptive mechanisms to cope with particular stress (82, 83), it has been reported that plant microbiota support plants to mitigate stresses, such as drought (84) and soil-borne pathogen attack (85). Due to the potential benefits of plant-associated microbiota, manipulating the plant microbiome become a promising and more sustainable approach to increase plant stress tolerance and overcome the negative impacts of biotic and abiotic stresses. To be able to harness plant microbiome for supporting plant growth and productivity, it is fundamental to address a complex interplay between plants and their microbiota as well as with the environment where they live.

Because of the extensive interactions with reciprocal impacts between the plant and its native microbiota, any perturbations that affect the plant may also then affect its microbial communities (14, 86). Environmental factors have been reported to be a major driver influencing plant-associated microbiota and drought. Previous studies found that drought stress induces a shift in root-associated microbial community composition and diversity of various angiosperm species (87), several grass species (88), sorghum (89), rice (90), and *Populus* (91). Studies revealed that the effects of drought stress are more pronounced on endophytes than rhizosphere microbiome which is related to the close interaction between the endophytes with the host plant (87, 90). Moreover, drought also leads to enrichment of certain taxa that belong to Actinobacteria and it is assumed that drought exposure results in changing plant root traits which favors certain taxa (88). This assumption was later supported by another study showing that enriched taxa (Actinobacteria and Firmicutes) had increased transporter gene activity for specific

root metabolites produced under drought exposure (89). It has been suggested that changes in the structure of the plant microbiome are caused either by the direct effect of environmental stresses or the indirect effect through plant stress responses (88, 91, 92). Drought stress reduces plant photosynthetic activity and changes plant metabolites production (91), which in turn affects root microbiota because they are highly influenced by photosynthates and root exudates. Furthermore, selective enrichment of a specific plant microbial taxa under stress potentially offers beneficial effects to the host plant. For instance, Actinobacteria may produce antimicrobial compounds to inhibit pathogens and phytohormones that are important for plant survival and resilience during stress exposure (93).

Microbe-microbe interaction is important in shaping plant microbiome assembly through different mechanisms, including mutualism, parasitism, and competition (44). Shifts in root-associated microbial community structure of different plant species have been reported after pathogen attacks (94-96). Soil-borne fungal pathogen invasion altered the composition of the rhizosphere microbial community by increasing the abundance of beneficial taxa, including Actinobacteria known for their biocontrol properties (94) and enrichment of fluorescent pseudomonads that have antifungal activity (95), suggesting that plants actively select and recruit beneficial groups of taxa under pathogen invasion. Other studies identified changes in root metabolite profile after pathogen infection affect the root-associated microbiome structure (96, 97), which further activates disease-suppressive soil activity (96). Moreover, certain phenolic compounds induced by pathogen attack modify the microbial community structure and directly suppress the pathogen growth, indicating indirect and direct effects of plant biotic stress responses (97). Altogether, these studies show a complex interaction between pathogen, plant,

and its microbial community, demonstrating that shift in root metabolite profile and microbial competition plays a vital role in shaping plant microbiome under biotic stress.

Overview of the study

There is a knowledge gap in our understanding of the relative contributions of the host plant versus the environment in shaping plant microbiome. Moreover, the consequences of an altered plant microbiome are poorly understood, though this information is critical to achieve a basis of knowledge for plant microbiome engineering and modification. My research aimed to contribute to this knowledge by investigating the factors that structure root- and seed-associated microbial communities of two valuable crops for Michigan's agricultural economy: apple and common bean.

The soil surrounding roots is a nutrient-rich hotspot that enables the growth of diverse microorganisms (25). Because the root zone is a complex ecosystem, multi-trophic interactions among bacteria, archaea, fungi, and nematodes are expected (5). However, these multiple players are rarely investigated together. In addition, changes in soil edaphic factors (e.g., pH, texture, and organic matter content) are known to drive microbiome assembly over space, including in managed systems such as fruit orchards (45). Hence, the objectives of the first chapter of my dissertation were to assess the biogeography of bacterial, archaeal, fungal, and nematode communities in the root zone of apple trees, and to determine their relationships with each other and their changes over natural abiotic gradients across 20 orchards that represent the main Michigan apple-producing region. I also assessed the influence of plant cultivar (different cultivars of the rootstock and scion) on microbiome structure. I hypothesized that the host plant, as well as abiotic soil characteristics and biotic, multi-player microbial interactions engaged in

feedbacks that ultimately shaped the microbial community and determined its interactions with the plant.

The next two research chapters of my dissertation investigate the endophytic microbiome associated with common bean (*Phaseolus vulgaris* L.) seed. Because seed endophytes are present at the very early developmental stages of the plant (seed to seedling), they are targets to understand their potential to provide beneficial traits to plants (60). Furthermore, plants may transfer these seed microbes to the next generation through vertical transmission (2). Therefore, it is important to understand how seeds may facilitate critical, early stages of plant microbiome assembly and also enable vertical transmission of microbiome members over plant generations. Studying seed endophytes is challenging for several reasons. First, there is low-microbial biomass inside the seed and it is difficult to extract. Second, seeds that begin to activate can release exudates that can select for or against particular microbial populations, therefore biasing observation. Also, host tissue disruption can lead to high plastid contamination in cultivation-independent approaches. Thus, the objectives of my second chapter were to 1) determine the appropriate observational unit of endophytic seed microbiome assessment for common bean by examining seed-to-seed, pod-to-pod, and plant-to-plant variability; and 2) to develop a robust protocol for individual seed microbiome extraction that could be generally applied to other plants that have similarly medium- to large-sized seeds.

My final research chapter is focused on the common bean seed microbiome and its response to drought. As one of the most damaging abiotic stressors in crop cultivation, drought can cause complete crop failure and yield loss (98). Climate change projections predict increasing drought severity and duration in several regions in the world that cultivate common bean as a staple (e.g., in parts of South and Central America and Africa) (99). One potential

mechanism by which plants may promote long-term drought stress tolerance is via a selection of beneficial members of the seed microbiome. However, the functions and persistence of seed microbiome members are still poorly understood. Moreover, the impact of drought on seed microbiome members is unknown and their contribution in determining plant stress tolerance to drought stress is unclear. Because managing the seed microbiome is one potential mechanism that could be used to improve plant tolerance to abiotic stress, I wanted to understand how drought impacted the seed microbiome. In this chapter, I am investigating this by collaborating to leverage a field experiment organized and executed by MSU bean breeders. The experiment was conducted in a field setting in two Michigan locations in Catham (Upper Peninsula) and East Lansing, MI (Lower Peninsula) that have organic and conventional farming managements. I hypothesized that drought alters the structure of seed endophytes and leads to enrichment and/or depletion of subsets of taxa.

REFERENCES

REFERENCES

1. Compant S, Samad A, Faist H, Sessitsch A. 2019. A review on the plant microbiome: ecology, functions and emerging trends in microbial application. Journal of advanced research 19:29-97.

2. Shade A, Jacques MA, Barret M. 2017. Ecological patterns of seed microbiome diversity, transmission, and assembly. Curr Opin Microbiol 37:15-22.

3. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. 2020. Plant–microbiome interactions: from community assembly to plant health. Nature Reviews Microbiology 18:607-621.

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

5. Mendes R, Garbeva P, Raaijmakers JM. 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. FEMS Microbiology Reviews 37:634-663.

6. Foo JL, Ling H, Lee YS, Chang MW. 2017. Microbiome engineering: Current applications and its future. Biotechnol J 12.

7. Eichmann R, Richards L, Schäfer P. 2021. Hormones as go-betweens in plant microbiome assembly. The Plant Journal 105:518-541.

8. Sasse J, Martinoia E, Northen T. 2018. Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? Trends Plant Sci 23:25-41.

9. Jones P, Garcia BJ, Furches A, Tuskan GA, Jacobson D. 2019. Plant Host-Associated Mechanisms for Microbial Selection. Frontiers in Plant Science 10.

10. Rascovan N, Carbonetto B, Perrig D, Díaz M, Canciani W, Abalo M, Alloati J, González-Anta G, Vazquez MP. 2016. Integrated analysis of root microbiomes of soybean and wheat from agricultural fields. Scientific Reports 6:28084.

11. Kumar Meena R, Kumar Singh R, Pal Singh N, Kumari Meena S, Singh Meena V. 2015. Isolation of low temperature surviving plant growth – promoting rhizobacteria (PGPR) from pea (Pisum sativum L.) and documentation of their plant growth promoting traits. Biocatalysis and Agricultural Biotechnology 4:806-811.

12. Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JH, Piceno YM, DeSantis TZ, Andersen GL, Bakker PA, Raaijmakers JM. 2011. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. Science 332:1097-100.

13. Panke-Buisse K, Poole AC, Goodrich JK, Ley RE, Kao-Kniffin J. 2015. Selection on soil microbiomes reveals reproducible impacts on plant function. The ISME Journal 9:980-989.

14. Naylor D, Coleman-Derr D. 2018. Drought Stress and Root-Associated Bacterial Communities. Frontiers in Plant Science 8.

15. Ortiz N, Armada E, Duque E, Roldán A, Azcón R. 2015. Contribution of arbuscular mycorrhizal fungi and/or bacteria to enhancing plant drought tolerance under natural soil conditions: Effectiveness of autochthonous or allochthonous strains. Journal of Plant Physiology 174:87-96.

16. Stone BWG, Weingarten EA, Jackson CR. 2018. The role of the phyllosphere microbiome in plant health and function, p 533-556, Annual Plant Reviews online doi:https://doi.org/10.1002/9781119312994.apr0614.

17. Ulrich DEM, Sevanto S, Ryan M, Albright MBN, Johansen RB, Dunbar JM. 2019. Plant-microbe interactions before drought influence plant physiological responses to subsequent severe drought. Sci Rep 9:249.

18. Anguita-Maeso M, Olivares-García C, Haro C, Imperial J, Navas-Cortés JA, Landa BB. 2020. Culture-Dependent and Culture-Independent Characterization of the Olive Xylem Microbiota: Effect of Sap Extraction Methods. Frontiers in Plant Science 10.

19. Kavamura VN, Robinson RJ, Hayat R, Clark IM, Hughes D, Rossmann M, Hirsch PR, Mendes R, Mauchline TH. 2019. Land Management and Microbial Seed Load Effect on Rhizosphere and Endosphere Bacterial Community Assembly in Wheat. Frontiers in Microbiology 10.

20. Bziuk N, Maccario L, Straube B, Wehner G, Sørensen SJ, Schikora A, Smalla K. 2021. The treasure inside barley seeds: microbial diversity and plant beneficial bacteria. Environmental Microbiome 16:20.

21. Zancarini A, Westerhuis JA, Smilde AK, Bouwmeester HJ. 2021. Integration of omics data to unravel root microbiome recruitment. Current Opinion in Biotechnology 70:255-261.

22. Yamazaki S, Mardani-korrani H, Kaida R, Ochiai K, Kobayashi M, Nagano AJ, Fujii Y, Sugiyama A, Aoki Y. 2021. Field multi-omics analysis reveals a close association

between bacterial communities and mineral properties in the soybean rhizosphere. Scientific Reports 11:8878.

23. Berendsen RL, Pieterse CMJ, Bakker PAHM. 2012. The rhizosphere microbiome and plant health. Trends in Plant Science 17:478-486.

24. Bakker PA, Berendsen RL, Doornbos RF, Wintermans PC, Pieterse CM. 2013. The rhizosphere revisited: root microbiomics. Front Plant Sci 4:165.

25. Hartmann A, Rothballer M, Schmid M. 2008. Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research. Plant and Soil 312:7-14.

26. Berg G, Rybakova D, Grube M, Köberl M. 2016. The plant microbiome explored: implications for experimental botany. J Exp Bot 67:995-1002.

27. Kong HG, Song GC, Ryu C-M. 2019. Inheritance of seed and rhizosphere microbial communities through plant–soil feedback and soil memory. Environmental Microbiology Reports 11:479-486.

28. Hardoim PR, Hardoim CCP, van Overbeek LS, van Elsas JD. 2012. Dynamics of seed-borne rice endophytes on early plant growth stages. PloS one 7:e30438-e30438.

29. Ahkami AH, Allen White R, Handakumbura PP, Jansson C. 2017. Rhizosphere engineering: Enhancing sustainable plant ecosystem productivity. Rhizosphere 3:233-243.

30. Van Der Heijden MGA, Bardgett RD, Van Straalen NM. 2008. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecology Letters 11:296-310.

31. Bais HP, Loyola-Vargas VM, Flores HE, Vivanco JM. 2001. Root-specific metabolism: The biology and biochemistry of underground organs. In Vitro Cellular & Developmental Biology - Plant 37:730-741.

32. Jones DL, Nguyen C, Finlay RD. 2009. Carbon flow in the rhizosphere: carbon trading at the soil–root interface. Plant and Soil 321:5-33.

33. Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. Annual Review of Plant Biology 57:233-266.

34. Nardi S, Concheri G, Pizzeghello D, Sturaro A, Rella R, Parvoli G. 2000. Soil organic matter mobilization by root exudates. Chemosphere 41:653-658.

35. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, Tremblay J, Engelbrektson A, Kunin V, Rio TGd, Edgar RC, Eickhorst T, Ley RE, Hugenholtz P, Tringe SG, Dangl JL. 2012. Defining the core Arabidopsis thaliana root microbiome. Nature 488:86-90.

36. Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, Rauf P, Huettel B, Reinhardt R, Schmelzer E, Peplies J, Gloeckner FO, Amann R, Eickhorst T, Schulze-Lefert P. 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 488:91-95.

37. Liu F, Hewezi T, Lebeis SL, Pantalone V, Grewal PS, Staton ME. 2019. Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. BMC Microbiology 19:201.

38. Habiyaremye JdD, Goldmann K, Reitz T, Herrmann S, Buscot F. 2020. Tree Root Zone Microbiome: Exploring the Magnitude of Environmental Conditions and Host Tree Impact. Frontiers in Microbiology 11.

39. Mercado-Blanco J, Abrantes I, Barra Caracciolo A, Bevivino A, Ciancio A, Grenni P, Hrynkiewicz K, Kredics L, Proença DN. 2018. Belowground Microbiota and the Health of Tree Crops. Frontiers in Microbiology 9.

40. Goldmann K, Ammerschubert S, Pena R, Polle A, Wu B-W, Wubet T, Buscot F. 2020. Early Stage Root-Associated Fungi Show a High Temporal Turnover, but Are Independent of Beech Progeny. Microorganisms 8:210.

41. Jayne B, Quigley M. 2014. Influence of arbuscular mycorrhiza on growth and reproductive response of plants under water deficit: a meta-analysis. Mycorrhiza 24:109-19.

42. Vukicevich E, Lowery T, Bowen P, Úrbez-Torres JR, Hart M. 2016. Cover crops to increase soil microbial diversity and mitigate decline in perennial agriculture. A review. Agronomy for Sustainable Development 36:48.

43. Mazzola M, Manici LM. 2012. Apple replant disease: role of microbial ecology in cause and control. Annu Rev Phytopathol 50:45-65.

44. Dastogeer KMG, Tumpa FH, Sultana A, Akter MA, Chakraborty A. 2020. Plant microbiome–an account of the factors that shape community composition and diversity. Current Plant Biology 23:100161.

45. Deakin G, Tilston EL, Bennett J, Passey T, Harrison N, Fernández-Fernández F, Xu X. 2018. Spatial structuring of soil microbial communities in commercial apple orchards. Applied Soil Ecology 130:1-12.

46. Van Horn C, Somera TS, Mazzola M. 2021. Comparative Analysis of the Rhizosphere and Endophytic Microbiomes across Apple Rootstock Genotypes in Replant Orchard Soils. Phytobiomes Journal 5:231-243.

47. Berlanas C, Berbegal M, Elena G, Laidani M, Cibriain JF, Sagües A, Gramaje D. 2019. The Fungal and Bacterial Rhizosphere Microbiome Associated With Grapevine Rootstock Genotypes in Mature and Young Vineyards. Frontiers in Microbiology 10.

48. Hacquard S, Schadt CW. 2015. Towards a holistic understanding of the beneficial interactions across the Populus microbiome. New Phytologist 205:1424-1430.

49. Pérez-Jaramillo JE, Mendes R, Raaijmakers JM. 2016. Impact of plant domestication on rhizosphere microbiome assembly and functions. Plant Mol Biol 90:635-44.

50. Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V. 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. Proceedings of the National Academy of Sciences 112:E911-E920.

51. Chartrel V, Dugat-Bony E, Sarthou A-S, Huchette S, Bonnarme P, Irlinger F. 2021. The microbial community associated with pea seeds (Pisum sativum) of different geographical origins. Plant and Soil 462:405-427.

52. Wassermann B, Cernava T, Müller H, Berg C, Berg G. 2019. Seeds of native alpine plants host unique microbial communities embedded in cross-kingdom networks. Microbiome 7:108.

53. Johnston-Monje D, Gutiérrez JP, Lopez-Lavalle LAB. 2021. Seed-Transmitted Bacteria and Fungi Dominate Juvenile Plant Microbiomes. Frontiers in Microbiology 12.

54. Mitter B, Pfaffenbichler N, Flavell R, Compant S, Antonielli L, Petric A, Berninger T, Naveed M, Sheibani-Tezerji R, von Maltzahn G, Sessitsch A. 2017. A New Approach to Modify Plant Microbiomes and Traits by Introducing Beneficial Bacteria at Flowering into Progeny Seeds. Frontiers in Microbiology 8.

55. Walsh CM, Becker-Uncapher I, Carlson M, Fierer N. 2021. Variable influences of soil and seed-associated bacterial communities on the assembly of seedling microbiomes. The ISME Journal 15:2748-2762.

56. Bacon CW, Hinton DM. 2019. Fungal and Bacterial Maize Kernel Interactions with the Vertically Transmitted Endophytic State of Fusarium verticillioides, p 191-209. *In* Verma SK, White JJF (ed), Seed Endophytes: Biology and Biotechnology doi:10.1007/978-3-030-10504-4_10. Springer International Publishing, Cham.

57. Nelson EB. 2004. MICROBIAL DYNAMICS AND INTERACTIONS IN THE SPERMOSPHERE. Annual Review of Phytopathology 42:271-309.

58. Berg G, Raaijmakers JM. 2018. Saving seed microbiomes. The ISME Journal 12:1167-1170.

59. Barret M, Briand M, Bonneau S, Préveaux A, Valière S, Bouchez O, Hunault G, Simoneau P, Jacques M-A. 2015. Emergence Shapes the Structure of the Seed Microbiota. Applied and Environmental Microbiology 81:1257-1266.

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

61. Rodríguez CE, Mitter B, Barret M, Sessitsch A, Compant S. 2018. Commentary: seed bacterial inhabitants and their routes of colonization. Plant and Soil 422:129-134.

62. Mano H, Tanaka F, Watanabe A, Kaga H, Okunishi S, Morisaki H. 2006. Culturable Surface and Endophytic Bacterial Flora of the Maturing Seeds of Rice Plants (Oryza sativa) Cultivated in a Paddy Field. Microbes and Environments 21:86-100.

63. Compant S, Mitter B, Colli-Mull JG, Gangl H, Sessitsch A. 2011. Endophytes of Grapevine Flowers, Berries, and Seeds: Identification of Cultivable Bacteria, Comparison with Other Plant Parts, and Visualization of Niches of Colonization. Microbial Ecology 62:188-197.

64. Shahzad R, Khan AL, Bilal S, Asaf S, Lee I-J. 2018. What Is There in Seeds? Vertically Transmitted Endophytic Resources for Sustainable Improvement in Plant Growth. Frontiers in Plant Science 9.

65. López-López A, Rogel MA, Ormeño-Orrillo E, Martínez-Romero J, Martínez-Romero E. 2010. Phaseolus vulgaris seed-borne endophytic community with novel bacterial species such as Rhizobium endophyticum sp. nov. Syst Appl Microbiol 33:322-7.

66. Goggin DE, Emery RJN, Kurepin LV, Powles SB. 2015. A potential role for endogenous microflora in dormancy release, cytokinin metabolism and the response to fluridone in Lolium rigidum seeds. Annals of botany 115:293-301.

67. Verma SK, Kingsley K, Irizarry I, Bergen M, Kharwar RN, White JF, Jr. 2017. Seed-vectored endophytic bacteria modulate development of rice seedlings. J Appl Microbiol 122:1680-1691.

68. White JF, Kingsley KL, Butterworth S, Brindisi L, Gatei JW, Elmore MT, Verma SK, Yao X, Kowalski KP. 2019. Seed-Vectored Microbes: Their Roles in Improving Seedling Fitness and Competitor Plant Suppression, p 3-20. *In* Verma SK, White JJF (ed), Seed Endophytes: Biology and Biotechnology doi:10.1007/978-3-030-10504-4_1. Springer International Publishing, Cham.

69. Bergna A, Cernava T, Rändler M, Grosch R, Zachow C, Berg G. 2018. Tomato Seeds Preferably Transmit Plant Beneficial Endophytes. Phytobiomes Journal 2:183-193.

70. Nelson EB. 2018. The seed microbiome: Origins, interactions, and impacts. Plant and Soil 422:7-34.

71. Shearin ZRC, Filipek M, Desai R, Bickford WA, Kowalski KP, Clay K. 2018. Fungal endophytes from seeds of invasive, non-native Phragmites australis and their potential role in germination and seedling growth. Plant and Soil 422:183-194.

72. Shahzad R, Waqas M, Khan AL, Al-Hosni K, Kang SM, Seo CW, Lee IJ. 2017. Indoleacetic acid production and plant growth promoting potential of bacterial endophytes isolated from rice (Oryza sativa L.) seeds. Acta Biol Hung 68:175-186.

73. Shahzad R, Waqas M, Khan AL, Asaf S, Khan MA, Kang S-M, Yun B-W, Lee I-J. 2016. Seed-borne endophytic Bacillus amyloliquefaciens RWL-1 produces gibberellins and regulates endogenous phytohormones of Oryza sativa. Plant Physiology and Biochemistry 106:236-243.

74. Kumar K, Verma A, Pal G, Anubha, White JF, Verma SK. 2021. Seed Endophytic Bacteria of Pearl Millet (Pennisetum glaucum L.) Promote Seedling Development and Defend Against a Fungal Phytopathogen. Frontiers in Microbiology 12.

75. White JF, Kingsley KI, Kowalski KP, Irizarry I, Micci A, Soares MA, Bergen MS. 2018. Disease protection and allelopathic interactions of seed-transmitted endophytic pseudomonads of invasive reed grass (Phragmites australis). Plant and Soil 422:195-208.

76. Khalaf EM, Raizada MN. 2018. Bacterial Seed Endophytes of Domesticated Cucurbits Antagonize Fungal and Oomycete Pathogens Including Powdery Mildew. Frontiers in Microbiology 9.

77. Tyc O, Putra R, Gols R, Harvey JA, Garbeva P. 2020. The ecological role of bacterial seed endophytes associated with wild cabbage in the United Kingdom. Microbiologyopen 9:e00954.

78. Pal KK, Dey R, Sherathia DN, Devidayal, Mangalassery S, Kumar A, Rupapara RB, Mandaliya M, Rawal P, Bhadania RA, Thomas M, Patel MB, Maida P, Nawade BD, Ahmad S, Dash P, Radhakrishnan T. 2021. Alleviation of Salinity Stress in Peanut by Application of Endophytic Bacteria. Frontiers in Microbiology 12.

79. Barret M, Guimbaud JF, Darrasse A, Jacques MA. 2016. Plant microbiota affects seed transmission of phytopathogenic microorganisms. Mol Plant Pathol 17:791-5.

80. Torres-Cortés G, Garcia BJ, Compant S, Rezki S, Jones P, Préveaux A, Briand M, Roulet A, Bouchez O, Jacobson D, Barret M. 2019. Differences in resource use lead to coexistence of seed-transmitted microbial populations. Scientific Reports 9:6648.

81. Mahajan S, Tuteja N. 2005. Cold, salinity and drought stresses: An overview. Archives of Biochemistry and Biophysics 444:139-158.

82. Ramegowda V, Senthil-Kumar M. 2015. The interactive effects of simultaneous biotic and abiotic stresses on plants: Mechanistic understanding from drought and pathogen combination. Journal of Plant Physiology 176:47-54.

83. Fang Y, Xiong L. 2015. General mechanisms of drought response and their application in drought resistance improvement in plants. Cellular and Molecular Life Sciences 72:673-689.

84. Naveed M, Mitter B, Reichenauer TG, Wieczorek K, Sessitsch A. 2014. Increased drought stress resilience of maize through endophytic colonization by Burkholderia phytofirmans PsJN and Enterobacter sp. FD17. Environmental and Experimental Botany 97:30-39.

85. Narendra Babu A, Jogaiah S, Ito S-i, Kestur Nagaraj A, Tran L-SP. 2015. Improvement of growth, fruit weight and early blight disease protection of tomato plants by rhizosphere bacteria is correlated with their beneficial traits and induced biosynthesis of antioxidant peroxidase and polyphenol oxidase. Plant Science 231:62-73.

86. Berg G, Grube M, Schloter M, Smalla K. 2014. Unraveling the plant microbiome: looking back and future perspectives. Frontiers in microbiology 5:148-148.

87. Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ. 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. Proceedings of the National Academy of Sciences 115:E1157-E1165.

88. Naylor D, DeGraaf S, Purdom E, Coleman-Derr D. 2017. Drought and host selection influence bacterial community dynamics in the grass root microbiome. The ISME Journal 11:2691-2704.

89. Xu L, Naylor D, Dong Z, Simmons T, Pierroz G, Hixson KK, Kim YM, Zink EM, Engbrecht KM, Wang Y, Gao C, DeGraaf S, Madera MA, Sievert JA, Hollingsworth J, Birdseye D, Scheller HV, Hutmacher R, Dahlberg J, Jansson C, Taylor JW, Lemaux PG, Coleman-Derr D. 2018. Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. Proc Natl Acad Sci U S A 115:E4284-e4293.

90. Santos-Medellín C, Liechty Z, Edwards J, Nguyen B, Huang B, Weimer BC, Sundaresan V. 2021. Prolonged drought imparts lasting compositional changes to the rice root microbiome. Nature Plants 7:1065-1077.

91. Timm CM, Carter KR, Carrell AA, Jun S-R, Jawdy SS, Vélez JM, Gunter LE, Yang Z, Nookaew I, Engle NL, Lu T-YS, Schadt CW, Tschaplinski TJ, Doktycz MJ, Tuskan GA, Pelletier DA, Weston DJ. 2018. Abiotic Stresses Shift Belowground Populus-Associated Bacteria Toward a Core Stress Microbiome. mSystems 3:e00070-17.

92. Haichar FeZ, Cernava T, Liu J, Timm CM. 2021. Editorial: Novel Insights Into the Response of the Plant Microbiome to Abiotic Factors. Frontiers in Plant Science 12.

93. Santos-Medellín C, Edwards J, Liechty Z, Nguyen B, Sundaresan V, Ausubel FM. 2017. Drought Stress Results in a Compartment-Specific Restructuring of the Rice Root-Associated Microbiomes. mBio 8:e00764-17.

94. Lazcano C, Boyd E, Holmes G, Hewavitharana S, Pasulka A, Ivors K. 2021. The rhizosphere microbiome plays a role in the resistance to soil-borne pathogens and nutrient uptake of strawberry cultivars under field conditions. Scientific Reports 11:3188.

95. Dudenhöffer J-H, Scheu S, Jousset A. 2016. Systemic enrichment of antifungal traits in the rhizosphere microbiome after pathogen attack. Journal of Ecology 104:1566-1575.

96. Yuan J, Zhao J, Wen T, Zhao M, Li R, Goossens P, Huang Q, Bai Y, Vivanco JM, Kowalchuk GA, Berendsen RL, Shen Q. 2018. Root exudates drive the soil-borne legacy of aboveground pathogen infection. Microbiome 6:156.

97. Gu Y, Wei Z, Wang X, Friman V-P, Huang J, Wang X, Mei X, Xu Y, Shen Q, Jousset A. 2016. Pathogen invasion indirectly changes the composition of soil microbiome via shifts in root exudation profile. Biology and Fertility of Soils 52:997-1005.

98. Lesk C, Rowhani P, Ramankutty N. 2016. Influence of extreme weather disasters on global crop production. Nature 529:84-87.

99. Beebe SE, Rao IM, Blair MW, Acosta-Gallegos JA. 2013. Phenotyping common beans for adaptation to drought. Frontiers in physiology 4:35-35.

CHAPTER 2: Biogeography and diversity of multi-trophic root zone microbiomes in Michigan apple orchards: analysis of rootstock, scion, and local growing region

Work presented in the chapter has been published as

Bintarti AF, Wilson JK, Quintanilla-Tornel MA, Shade A. Biogeography and Diversity of Multi-Trophic Root Zone Microbiomes in Michigan Apple Orchards: Analysis of Rootstock, Scion, and Local Growing Region. *Phytobiomes Journal* **4**, 122-132 (2020).

Abstract

Soil is a highly heterogeneous environment with many physical and chemical factors that are expected to vary within and across fruit orchards, and many of these factors also drive changes in the soil microbiome. To understand how biogeography influences apple root microbiomes, we characterized the bacterial and archaeal, fungal, nematode, oligochaete, and mycorrhizal communities of the root zone soil (soil adjacent to the tree trunk and expected to be influenced by the plant) across 20 sites that represent the main Michigan apple-producing region. Amplicon sequencing of the 16S rRNA and ITS genes were performed, as well as direct quantification of nematodes, oligochaetes, and mycorrhizal fungi with microscopy. The microbiome community structures were affected by site and rootstock, but not by scion. Microbiomes had taxa typical of soil, including an archaeal taxon affiliated with family Nitrososphaeraceae, bacterial phyla Proteobacteria and Acidobacteria, and fungal phyla Ascomycota and Basidiomycota. While many taxa were detected in all samples and collectively composed 41.55% of the relative abundances, they had average relative abundances each of less than 1%, with no notable dominance. We used network analysis to understand potential for inter-trophic interactions, but detected few cross-kingdom associations. Together, these results show the complexity of the apple root zone microbiome and did not identify obvious biotargets that may universally associate with tree health. This suggests that the key attributes of the apple root zone community may be in the community-level functional traits that are shared and distributed across the membership, rather than by its composition.

Introduction

Soil microbial communities have been known to play an important role for plant growth and fitness (1), enhancing plant nutrient acquisition (2), inducing flowering time (3), improving plant tolerance against abiotic stresses (4) and promoting pathogen resistance (5). The plant and soil-associated microbiome includes numerous players that are expected to interact with each other either directly or through trophic cascades and multi-trophic interactions (e.g., (6, 7)). Multi-trophic phytobiome interactions can involve bacteria, archaea, fungi, and nematodes that reside in plant-associated soils and on or in the plant itself. However, these multiple players are rarely investigated as a system within the same study. Additionally, changes in soil edaphic factors, such as pH, texture, and organic matter content, are known to drive microbiome assembly over space, including in managed systems such as fruit orchards (8, 9). Taken together, it is expected that the host plant, as well as abiotic soil characteristics and biotic, multi-player microbial interactions engage in feedbacks that ultimately shape the microbial community and determine its interactions with the plant.

The objective of our study is to assess the biogeography of bacterial, archaeal, fungal, and nematode communities in the root zone of apple trees, and to determine their relationships with each other and their changes over natural abiotic gradients across orchards. We collected root zone samples from 20 mature commercial apple orchards in Michigan. Although Michigan is ranked third in the United States in terms of apple production with 1.07 billion pounds of apples valued on average at $297 million per year (10), microbiome-apple relationships have not been investigated here. We used high-throughput amplicon sequencing to assess bacterial and archaeal, and fungal communities, and microscopy to identify nematodes, oligochaetes, and mycorrhizal fungi. The results uncover possible interactions between these important apple root

zone community members and provide foundational baseline information on microbiome diversity and putative phytobiome interactions prior to anticipated apple tree removal and replant on these farms.

Materials and Methods

Survey design, soil sample collection, and apple rootstocks and scions

Forty-five root zone soil samples were collected from 20 mature (i.e. at least 10 years old) apple orchards in Michigan in June 2017 (**Figure 2.6 Appendix B**). These orchards were selected first because they are representative of an area considered to be prime orchard ground where apples have been grown for as many as six generations on family farms, and second because they offered a key comparison across known local differences in soil edaphic characteristics. Each orchard was considered as an experimental unit for understanding biogeography, and statistical comparisons were made across orchards to assess spatial dynamics and both within and across orchards to assess influence of different rootstock and scion varieties. Within an orchard, each distinct combination of apple rootstock and scion was planted in a different tree row. Soil cores were collected from the bases of each of ten trees in a single tree row (e.g., trees with the same rootstock/scion combination). These soil cores were composited into a homogenized soil sample to represent the tree row and rootstock/scion combination. Soil cores (20 cm depth x 2.5 cm diameter) were used to collect root zone samples and were cored within 15-20 cm of the base of a tree trunk. By "root zone", we mean the local soil surrounding and adjacent to the plant and its root structures that is expected to be chemically and physically influenced by the plant via exudates, stemflow, etc. As a qualifier, the exact quantities and rates of root exudation were not measured. The soil corer was sprayed and wiped with 70% ethanol

before sample collection to prevent cross-site microbial contamination. Soil core composites were collected into Whirl-Pak® sample bags and immediately placed on ice for transport. In the laboratory, soils were sieved through 4 mm mesh to remove large pieces of rocks, and root tissue, and other plant residues. Sieved soil was stored at -80˚C until microbial DNA extraction. One hundred grams of each soil sample was sent to the Michigan State University Soil and Plant Nutrient Laboratory (SPNL) for soil physicochemical testing. The analysis of soil chemistry and characteristics were conducted to provide us more information and knowledge in assessing root zone microbial community in apple orchards. Soil chemistry and characteristics also allow us to decipher the influence of environmental factors to the microbial community structure and diversity. Soil parameters including pH, lime index, phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), nitrate (NO_3N), ammonium (NH_4N), organic matter (OM), sand, silt, clay, and soil type were measured from all samples. Full metadata, including growing locations, soil environmental characteristics, and scion and rootstock information can be found in **Table 2.3 Appendix A**.

Nematode, oligochaete, and mycorrhizal fungal quantification

The abundances of nematodes, oligochaetes, and mycorrhizal fungi were assessed using standard protocols of the Michigan State Plant and Pest Diagnostic Laboratory. Nematodes, mycorrhizal fungi, and oligochaetes were removed from the soil with a modified centrifugation and flotation method (11). One hundred grams of soil was suspended in water and then poured over sieves that were nested in mesh size, allowing soil particles to pass through but capturing the nematodes, oligochaetes, and mycorrhizal fungi spores. These samples then were centrifuged, and water was decanted and replaced with 61.5% sucrose. The sample was

centrifuged again to capture the microbial groups in the sucrose gradient to separate them from any remaining soil particles. Nematodes, mycorrhizal spores, and oligochaetes were removed from the sucrose, rinsed, and then finally identified using inverted microscopy at 200X magnification. Nematodes were identified to the lowest level of taxonomic classification possible, with most identifications possible at the genus level. Mycorrhizal spores and oligochaetes individuals were counted but not taxonomically identified.

Microbial DNA extraction and PCR amplification

Microbial DNA extraction was carried out for 0.25 g of each soil sample using the manufacturer's protocol of PowerSoil®DNA Isolation Kit (MoBio Laboratories, Solana Beach, CA, United States). The soil DNA was then quantified using Qubit™dsDNA BR Assay Kit (ThermoFisher Scientific, Waltham, MA, United States) to determine the concentration. Polymerase Chain Reaction (PCR) was conducted to verify the amplification of the V4 region of bacterial and archaeal 16S rRNA gene using 515f (5'-GTGCCAGCMGCCGCGGTAA-3') and 806r (5'-GGACTACHVGGGTWTCTAAT-3') universal primers (12). The 16S rRNA gene amplification was conducted under following condition: 94°C for 3 min, followed by 35 cycles of 94°C (45 s), 50°C (60 s), and 72°C (90 s), with a final extension at 72°C (10 min). The amplification was performed in 25 µl mixture containing 12.5 µl GoTaq®Green Master Mix (Promega, Madison, WI, United States), 0.625 µl of each primer (20 mM), 1 µl of DNA template (~ 15 nanogram per µl), and 10.25 µl nuclease free water. The amplicons were diluted to the concentration of 10-20 nanogram per µl then sequenced using Illumina MiSeq platform at the Research Technology Support Facility (RTSF) Genomics Core, Michigan State sequencing facility.

Fungal communities were detected by PCR amplification of ITS1 region using ITS1f (5'-CTTGGTCATTTAGAGGAAGTAA-3′) and ITS2 (5'- GCTGCGTTCTTCATCGATGC-3') primer pair (13) with addition of index adapters as required by the RTSF Genomics Core (https://rtsf.natsci.msu.edu/genomics/sample-requirements/illumina-sequencing-sample-requirements/). The PCR condition for ITS1 amplification as following: 94°C for 3 min, followed by 35 cycles of 94°C (30 s), 63.5°C for (30 s), and 72°C for (30 s), with a final extension at 72°C for 10 min. The amplification was performed in 50 µl mixture containing 20 µl GoTaq®Green Master Mix (Promega, Madison, WI, United States), 1 µl of each primer (10 mM), 1 µl of DNA template (~ 15 nanogram per µl), and 27 µl nuclease free water. PCR products were purified using the manufacturer's protocol of Wizard®SV Gel and PCR Clean-Up System (Promega, Madison, WI, United States), then quantified using Qubit™dsDNA BR Assay Kit (ThermoFisher Scientific, Waltham, MA, United States). Purified PCR products with the concentration range 2-10 nanogram per µl were sequenced at the RTSF Genomics Core using Ilumina MiSeq platform. The 16S and ITS libraries were prepared using the Illumina TruSeq® Nano DNA Library Prep Kit. Ilumina MiSeq was run using v2 Standard and paired-end reads sequencing format (2 x 250 bp).

Sequencing data analysis and OTU clustering

Bacterial and archaeal raw reads produced from Illumina MiSeq were processed using USEARCH (v10.0.240). Preparation of raw reads was performed using the protocol established in the USEARCH pipeline followed by Operational Taxonomic Units (OTUs) clustering using UPARSE method (14), then further analyses were conducted using QIIME1 (v1.9.1) (15). Read preparation and processing used in this study consisted of paired end reads merging, filtering the

low-quality sequences, dereplication to find unique sequences, singleton removal, denoising (pre-clustering) via *cluster_fast* command which implements UNOISE algorithm and chimera checking (16). Operational taxonomic unit picking was conducted using open reference strategy as described in the previous study (17). First, closed reference OTU picking was performed at 97 % identity threshold by clustering quality filtered reads against the SILVA database (v1.32) (18) using *usearch_global* command. Later, reads that failed to hit the SILVA reference were clustered de novo at 97% identity using *cluster_otus* command which also detected chimera. Thus, the OTUs generated by closed reference and de novo OTU picking were combined to make the full set of OTU representative sequences. Finally, all pre-dereplicated sequences were mapped back to the full set of OTU representative sequences to build an OTU table. The next analyses, performed in QIIME1, included taxonomic assignment to the reference data sets of SILVA (v.1.32) database using UCLUST algorithm at a minimum confidence of 0.9 (the default method; (19)), and eukaryotic (chloroplast and mitochondria) sequence removal from OTU table. Read quality control and filtering generated 1,786,268 bacterial/archaeal reads in total. Rarefaction to the lowest sequencing depth (27,716 bacterial/archaeal reads) was conducted to standardize the sampling efforts using *single_rarefaction.py* command in QIIME1 (20, 21).

The processing of fungal ITS raw reads was also conducted using USEARCH (v10.0.240) pipeline. Reads pre-processing including reads merging, primer removal using cutadapt (v1.17) (22), filtering and trimming using *fastq_filter* command, and reads dereplication to find unique sequences. Quality filtered reads then clustered into OTUs using the same approach as described above which was open reference OTU picking using UNITE fungal ITS database (v.7.2) (23) with 97% of identity threshold. The OTU table was built by mapping pre-dereplicated sequences back to the full set of OTU representative sequences obtained from open

reference OTU picking. Fungal taxonomic classification was performed using CONSTAX tool (24) which compares three taxonomic assignment tools for fungal sequence data: RDP Classifier (25, 26), UTAX (14, 19), and SINTAX (27). The CONSTAX tool generates an improved consensus taxonomic file which is a combination among those three programs and the reference used for taxonomic assignment in this tool is UNITE fungal ITS database (v.7.2). The ITS gene taxonomic classification was performed at a minimum confidence of 0.8. Read quality control and filtering generated 4,240,062 fungal reads in total. Subsamples of sequence was conducted by rarefying to the lowest number of sequence (56,240 fungal reads) (20, 21) using *single_rarefaction.py* command in QIIME1 (v1.9.1).

Microbial community analysis

Microbial community analyses were conducted in the R environment for statistical computing (v3.5.1) (R Core Development Team). Microbial composition and relative abundance of each sample and block was analyzed using Phyloseq package (v1.26.1) (28). Alpha diversity indices (Pielou's evenness, total species number or richness) and beta diversity of microbial community were calculated on the rarefied OTU table using the vegan package (v2.5-4) (29). We chose these to metrics because they complement one another in the information they provide: richness reveals the total number of taxa without accounting for their differences in abundances, while Pielou's evenness considers the equitability of taxon abundance. However, we found that the overarching patterns of these two metrics largely agree.

The difference of bacterial and archaeal, and fungal alpha diversity among sites, rootstocks, and scions were evaluated using one-way analysis of variance (ANOVA) and Tukey's HSD post-hoc test. The normality and homoscedasticity of the data were verified using

Saphiro-Wilk and Levene's test, respectively. Non-parametric Kruskal-Wallis test and post-hoc Dunn's test with Benjamini-Hochberg correction for *p*-values were performed when the normality assumptions of one-way ANOVA were not met. Welch's ANOVA and Games-Howell post-hoc tests were conducted for the data that did not meet the homoscedasticity assumption. The difference of nematode, oligochaete, and mycorrhizal fungi abundances among sites, rootstocks, and scions were assessed using one-way ANOVA and Tukey's HSD post-hoc test. Nematodes and oligochaetes count data were square-root-transformed. The mycorrhizal fungi count data were log10-transformed to meet test assumptions. Alpha diversity metrics of nematodes including richness and Pielou's evenness were calculated using untransformed count-data. Pearson correlation and regression analysis were conducted to see the relationship between microbial alpha diversity and all parameters (soil characteristics, nematodes, oligochaetes, and mycorrhizal fungi abundances).

Beta diversity was calculated using Bray-Curtis dissimilarity indices and visualized with principal coordinate analysis (PCoA). The environmental variables were fitted into PCoA plot and tested for their significance using permutation tests using 'envfit' function in vegan package (v2.5-4) (29). Permutational multivariate analysis of variance (PERMANOVA) was performed to assess the effects of different factors on the microbial community structure. We performed Mantel test to assess the correlation between geographic distance with bacterial/archaeal, fungal, and nematode distance matrices (30). The PCoA ordinates of bacterial/archaeal, fungal, and nematode communities were also compared and tested using Procrustes rotation with PROTEST (31) to analyze the congruence between two community ordinations. Core microbiota of apple root-zone soil was also assessed by assessing the microbial OTUs' abundances versus their occupancies (32).

Network analysis

The network was constructed based on Random Matrix Theory (RMT) methods (33). We combined the bacterial and archaeal, and fungal OTUs and nematodes, mycorrhizal fungi, and oligochaetes count number and ran the data through the Molecular Ecological Network Analysis (MENA) Pipeline (33) on the website (http://ieg4.rccc.ou.edu/mena) of the University of Oklahoma's Institute for Environmental Genomics. The network construction was conducted as the following setting: OTUs detected in 23 out of 45 total samples were selected (~ 50% occupancy), blanks with paired valid values were filled with 0.01, logarithm values were calculated, Pearson Correlation Coefficient was used for similarity matrix method, calculation order was conducted by decreasing the cutoff from top and only Poisson regression was used. We used greedy modularity optimization for separation method and calculate Zi (within-module connectivity) and Pi (among-module connectivity) values to identify the modularity. Module and network hubs, peripherals, and connectors of the network were determined by Zi and Pi value of 2.5 and 0.62, respectively. The visualization of network was conducted using Cytoscape software (v.3.7.1) (34).

Data and code availability

The computational workflows for sequence processing and ecological statistics are available on GitHub (https://github.com/ShadeLab/PAPER_Bintarti_2020_Phytobiomes/). Raw sequence data of bacteria/archaea and fungi have been deposited in the Sequence Read Archive (SRA) NCBI database under accession number PRJNA507629.

Results

Sequencing summary

A total of 1,786,268 and 4,240,062 quality controlled bacterial/archaeal and fungal reads were obtained from 45 root zone soil samples across 20 orchards in Michigan. Each community was subsampled to the minimum number of quality sequences observed to construct the taxon table (27,716 and 56,240 of bacterial/archaeal and fungal reads, respectively). Operational taxonomic units (OTUs) were defined at 97% sequence identity resulted in 22,510 and 3,553 of bacterial and archaeal, and fungal total OTUs respectively. Rarefaction curves indicated that the sequencing depth was sufficient to observe all taxa and microbial community in the sample (**Figure 2.7 Appendix B**).

Bacterial and fungal alpha diversity among sites, rootstocks, and scions

We assessed the alpha (within-sample) diversity among sites, rootstocks, and scions for the dominant trophic groups within the microbial ecosystem, including bacteria and archaea, and fungi assessed using sequencing of phylogenetic marker genes, and nematodes, oligochaetes, and mycorrhizal fungi assessed using traditional soil microscopy and counting. This allowed us to relate players from different trophic levels to one another in their occurrences and use them as explanatory variables for the biogeographic patterns observed.

There were overall differences in bacterial and archaeal richness and Pielou's evenness among sites and rootstocks (**Figure 2.8A Appendix B**, **Figure 2.1**, Kruskal-Wallis and ANOVA results **Table 2.1**). Site-by-site comparisons revealed that there were alpha diversity differences among five orchards, sites 1, 3, 13, 15, and 18 (**Table 2.4 Appendix A**). Rootstock-by-rootstock comparisons for richness were not significant, suggesting that any differences detected were

marginal (**Table 2.5 Appendix A**); however, there were pairwise rootstock differences in Pielou's evenness (**Table 2.5 Appendix A**). Specifically, root zone soil of Bud 9 and M7 rootstocks had higher bacterial and archaeal Pielou's evenness compared to M26 and M9.

There were correlations between bacterial and archaeal alpha diversity with soil texture, and soil chemical properties as indicated by linear regression model (**Table 2.6 Appendix A, Figure 2.9 Appendix B).** Sand content positively correlated with the alpha diversity, meanwhile silt and clay content negatively correlated with the alpha diversity. Moreover, soil type also had an impact on the bacterial and archaeal richness (Welch's ANOVA, F-stat = 13.568, df = 2, n = 3, *p*-val = 0.01). Sandy loam soil had higher bacterial and archaeal richness compared to loam (Games-Howell post-hoc test, *p*-val = 0.002) and sandy clay loam soil (Games-Howell post-hoc test, *p*-val = 0.008). These results suggested that soil with coarser and sandy texture are likely to harbor more diverse microbes than soil with finer texture. Among soil chemical properties tested, bacterial and archaeal alpha diversity positively correlated with P and negatively correlated with K and Ca, indicated that these communities likely play a role in macro and micro-nutrient cycles in soil, including in P solubilization.

Similar to the bacterial and archaeal alpha diversity patterns, there were differences of fungal richness among sites and rootstocks, but there were no differences of Pielou's evenness (**Table 2.1, Figure 2.8B Appendix B, Figure 2.1B**). Specifically, Site 17 had higher fungal richness than almost half of other sites (**Table 2.7 Appendix A**). Soil taken from M26 root zones had higher fungal richness than Bud9. On balance, M126 rootstock had higher fungal richness compared to most of the rootstocks (**Table 2.8 Appendix A**).

Soil chemistry and texture only correlated with fungal Pielou's evenness and not richness (**Table 2.6 Appendix A, Figure 2.10 Appendix B**). Similar to what was observed for bacterial

and archaeal alpha diversity, fungal evenness was positively correlated with sand content and negatively correlated with silt content. Soil type also affected fungal evenness (one-way ANOVA, F-stat = 4.027, df = 2, n = 3, *p*-val = 0.02), and, again, similar to bacteria and archaea, sandy loam soil had higher fungal evenness than loam soil (Tukey's HSD post-hoc test, *p*-val = 0.02). Among soil chemical properties tested, fungal Pielou's evenness negatively correlated with K and Mg content.

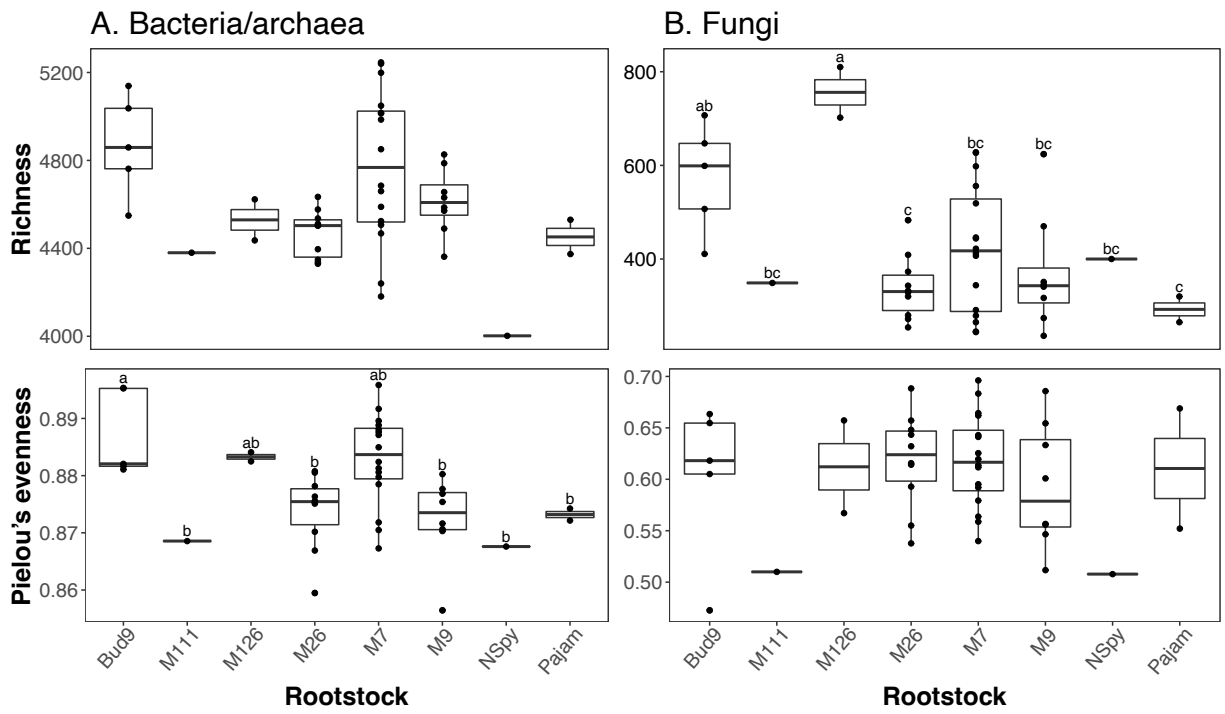

Figure 2.1. Alpha diversity of apple root zone microbiome among rootstocks.

Alpha diversity metrics of **A,** bacteria/archaea and **B,** fungi: richness (operational taxonomic number, clustered at 97 % identity threshold) and Pielou's evenness among rootstocks. For each box plot, circles represent measurement for each sample. The central horizontal lines represent the mean of measurements. Boxes labelled with different letters are identified as significantly different based on Tukey's honestly significant difference post-hoc test. Boxes without label are not significantly different.

Table 2.1. Statistical analysis of microbial richness and Shannon diversity index among sites (n = 20) and rootstocks (n = 8) using Kruskal-Wallis and one-way (ANOVA)[a]

Kruskal-Wallis test	Bacterial/ archaeal Richness		Fungal Richness	One-way ANOVA test	Bacterial/ archaeal Richness	Bacterial/ archaeal Pielou's evenness			Fungal Richness		Fungal Pielou's evenness		
	Site	Root-stock	Scion		Scion	Site	Root-stock	Scion	Site	Root-stock	Site	Root-stock	Scion
chi-squared	32.16	16.37	12.82	**F-stat**	1.19	4.17	4.07	0.82	3.82	5.08	1.24	1.16	0.79
df	19	7	13	**df**	13	19	7	13	19	7	19	7	13
				R^2_{adj}	0.05	0.58	0.33	-0.057	0.55	0.39	0.09	0.02	-0.06
p-value	**0.03**	**0.02**	0.46 (ns)	**p-value**	0.33 (ns)	**0.0005**	**0.002**	0.64(ns)	**0.001**	**0.0004**	0.3 (ns)	0.35 (ns)	0.6 (ns)

[a] Significant results ($P < 0.05$) appear in bold. ns = not significant.

Nematodes and other groups: alpha diversity

Other soil trophic levels counted from the apple root zone included eleven trophic groups of nematodes that are classified based on their feeding habits (including bacterivores, herbivores, omnivores, carnivores, and fungivores), mycorrhizal fungi, and oligochaetes. A total count of 31,820 nematodes, 3,420 mycorrhizal fungi, and 544 oligochaetes were observed. Mean of total count showed that the nematode group Rhabditidae had the highest absolute abundance compared to others (560.9±480.86, SD as well as highest prevalence (100%)) (**Table 2.9 Appendix A, Table 2.10 Appendix A**). In addition, the distribution of nematodes and other soil trophic levels revealed that mycorrhizal fungi were found in all soil samples (100%), followed by *Tylenchus* (97.77 %), Dorylaimidae (88.88 %), *Aphelenchus* (86.66 %), oligochaetes (82.22 %), *Xiphinema* spp. of nematodes (aka: dagger nematodes; 73.33 %), and *Pratylenchus* spp. of nematodes (aka: lesion nematodes; 53.33 %) (**Table 2.10 Appendix A**). Among these groups, *Xiphinema* spp. and *Pratylenchus* spp. belong to plant-parasitic nematodes.

Total absolute nematode abundances among sites showed *p*-val of 0.057 (ANOVA). Significant differences in the total absolute abundance of nematodes were also detected among rootstocks (ANOVA, F-stat = 2.69, df = 7, n = 8, *p*-val = 0.02) but not scions (ANOVA, *p*-val > 0.05). The comparison analysis showed that NSpy and Bud9 root zone had higher nematode abundances than Pajam (Tukey's HSD post-hoc test, *p*-val = 0.02 and 0.03, respectively) (**Figure 2.2, Table 2.11 Appendix A**). Moreover, significant differences in oligochaetes and mycorrhizal fungi abundances were detected among sites (ANOVA, F-stat = 3.45 and 3.46, respectively, df = 19, n = 20, *p*-val = 0.002) and Site 17 had higher oligochaetes abundance than almost half of other sites (Tukey's HSD post-hoc test, *p*-val < 0.05) (**Figure 2.2, Table 2.12 Appendix A**). Site to site comparison showed that Site 8 had lower mycorrhizal fungi abundances than several other

sites (Tukey's HSD post-hoc test, p-val < 0.05) (**Figure 2.2, Table 2.13 Appendix A**). The differences of absolute oligochaetes abundance among rootstocks had a p-val of 0.054 (ANOVA).

We further tested the correlation of nematode, mycorrhizal fungi, and oligochaetes abundances to the microbial alpha diversity. The nematode group *Tylenchus* negatively correlated with bacterial and archaeal alpha diversity, meanwhile Rhabditidae positively correlated with bacterial and archaeal evenness (**Figure 2.9 Appendix B**). In contrast, there were no correlations identified between fungal alpha diversity and absolute abundances of other trophic levels. Nematode alpha diversity showed no differences among sites, rootstocks, or scions (ANOVA, p-val > 0.05). Furthermore, we tested the correlation between nematodes and bacterial and archaeal and fungal alpha diversity metrics. Nematode Pielou's evenness negatively correlated with bacterial and archaeal Pielou's evenness (**Figure 2.11 Appendix B**).

From all the analyses, there were no differences between bacterial and archaeal, fungal, and nematode alpha diversity among scions (ANOVA, p-val > 0.05)**,** suggesting that the above-ground scion has marginal influence on the root zone microbiome. Those results indicated that root zone microbial communities were varied among different apple orchard locations and were also influenced by the variety of the rootstock planted by the growers.

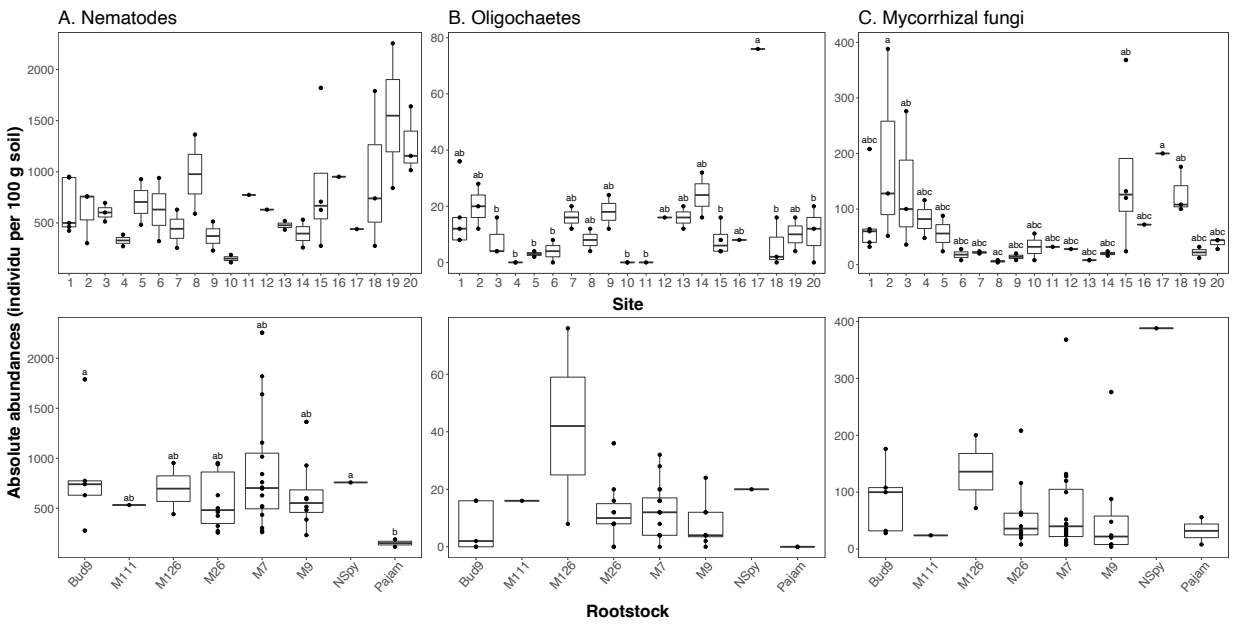

Figure 2.2. Absolute abundances of apple root zone multi-trophic levels among sites and rootstocks.

Absolute **A,** nematode, **B,** oligochaetes, and **C,** mycorrhizal fungi abundances across sites and rootstocks. For each box plot, circles represent measurement for each sample. The central horizontal lines represent the mean of measurements. Samples labelled with different letters are identified as significantly different based on Tukey's HSD post-hoc test. Samples with no letters are not significantly different.

Microbial beta diversity

Soil pH, lime index, Ca, Mg, sand, and silt content as well as a nematode group, *Tylenchus*, had explanatory value for describing the patterns in beta diversity among bacterial and archaeal communities (**Table 2.2**, **Figure 2.12A Appendix B**). For fungi, soil pH, lime index, Mg content, and nematode groups of *Pratylenchus* spp. and *Tylenchus* had explanatory value (**Table 2.2**, **Figure 2.12B Appendix B**). There were no distinct separations of microbial communities by orchard site or apple rootstock. However, on balance, communities from similar sites or rootstocks were proximal to each other, especially for bacterial and archaeal PCoA plot (**Figure 2.12 Appendix B, Figure 2.13 Appendix B**). Permutated multivariate analysis of variance (PERMANOVA) revealed that bacterial/archaeal and fungal community structure and composition were influenced by site (F-stat = 1.88, R^2 = 0.58, *p*-val = 0.001 and F-stat = 1.68, R^2 = 0.56, *p*-val = 0.001, respectively) and rootstock (F-stat = 1.48, R^2 = 0.21, *p*-val = 0.004 and F-stat = 1.35, R^2 = 0.20, *p*-val = 0.006, respectively), but not by scion (*p*-val > 0.05). We also calculated the effect of interaction between variables (site and rootstock) to the microbial beta diversity. We detected differences of microbial beta diversity among different sites with the same rootstock (PERMANOVA, *p*-val < 0.05), in contrast, there were no differences of microbial beta diversity among different rootstocks in the same site (PERMANOVA, *p*-val > 0.05). These results indicated that site or orchard location had a stronger effect on microbial diversity than rootstock. Among the three variables tested (site, rootstock, scion), only site had explanatory value for nematode community structure (PERMANOVA, F-stat = 1.50, R^2 = 0.53, *p*-val = 0.017). In addition, nitrate-nitrogen (NO_3N) was the only environmental factor that had explanatory value for the nematode community (**Figure 2.12C Appendix B, Figure 2.13C Appendix B).**

Table 2.2. Environmental factors that have explanatory value for the bacterial/archaeal and fungal communities were fitted into principal coordinate analysis plot (beta diversity was calculated using Bray-Curtis dissimilarity indices) and tested using permutation test using "envfit" function in vegan package (v2.5-4)[a]

Variables	Bacteria/archaea		Fungi	
	R^2	*p*-val	R^2	*p*-val
Soil physicochemistry				
pH	0.64	0.001***	0.49	0.002**
lime index	0.66	0.001***	0.52	0.001***
Ca	0.42	0.005**	0.20	0.133
Mg	0.34	0.025*	0.42	0.008**
Sand	0.30	0.035*	0.17	0.178
Silt	0.36	0.016*	0.20	0.135
Nematode				
Pratylenchus spp	0.12	0.302	0.42	0.006**
Tylenchus	0.39	0.007**	0.32	0.018*
Scion	0.64	0.023*	0.66	0.038*
Rootstock	0.62	0.001***	0.58	0.014*

[a] *** indicates significance at 0.001; ** indicates significance at 0.01; and * indicates significance at 0.05.

To test for biogeographic signal (e.g., distance decay), we performed Mantel tests on the bacterial/archaeal, fungal, and nematode Bray-Curtis dissimilarity matrices against geographic distance. There were no significant correlations between geographic distance and microbial or nematode beta diversity (all p-val > 0.05). We also tested for patterns of synchrony in beta diversity among bacterial/archaeal, fungal, and nematode communities, and found that the bacterial/archaeal community correlated with both fungal community (PROTEST, p-val = 0.001) and nematode community (PROTEST, p-val = 0.009). However, there was no correlation detected between fungal and nematode communities (PROTEST, p-val > 0.05).

Microbial community composition in apple root zone

The bacterial/archaeal 16S rRNA and fungal ITS gene sequences were classified into 43 phyla (146 classes) and 16 phyla (48 classes), respectively. The overall composition of bacterial/archaeal, and fungal communities across sites and rootstocks were comparable with relatively minor variation in the relative abundances of each phylum (**Figure 2.3**). Based on the mean relative abundance, the bacterial/archaeal communities in all samples were dominated by Proteobacteria (31.45%), Acidobacteria (18.5%), Bacteroidetes (11.18%), Verrucomicrobia (10.11%), Planctomycetes (6.89%), and Actinobacteria (6.25%). Meanwhile, Ascomycota (43.47%), Basidiomycota (31.49%), Mortierellomycota (14.7%), and an unidentified phylum (9.04%) were the most dominant fungal phyla in the apple root zone. These are typical soil taxa and these bacterial and fungal phyla have been identified in apple root zone soil previously (35-38).

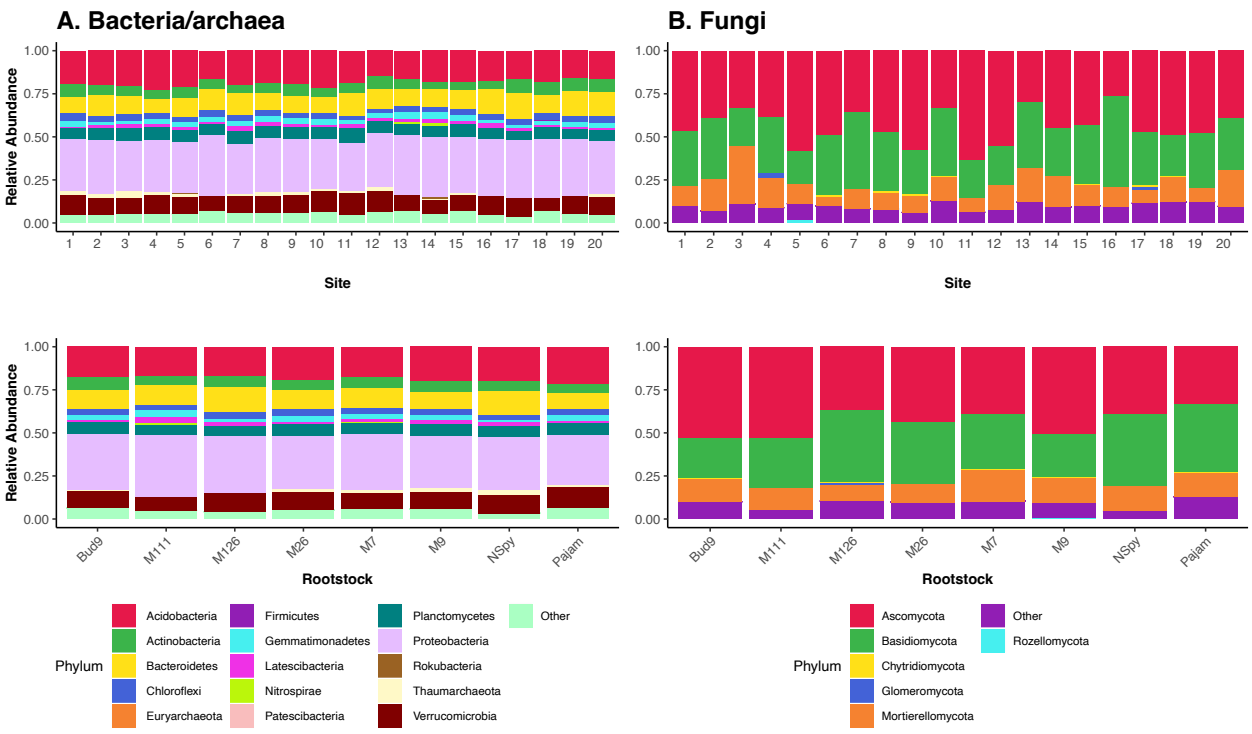

Figure 2.3. Relative abundances of apple root zone microbiome among sites and rootstocks.

The relative abundance of **A,** bacterial/archaeal and **B,** fungal taxa grouped at phylum level across sites and rootstocks. Taxa with relative abundance less than 0.01 and unassigned taxa or those for which taxonomic assignment could not be made past Domain were defined as "other". Bars are color-coded by phylum.

We prioritized members of the core microbiome of the apple root zone by exploring the relationship between taxon occupancy (e.g., the proportion of samples in which the taxa were detected) and abundance (32). Here, core microbiomes were defined as OTUs detected in all samples (occupancy = 1; **Figure 2.14 Appendix B**). There were 383 bacterial or archaeal core taxa found in all samples that belonged to 15 phyla and represented 41.55% of the total relative abundance. Some bacterial and archaeal core taxa existed in high abundance in this study including uncultured archaeon (Nitrososphaeria) (0.91%), bacterial taxa from phylum Acidobacteria (Subgroup 6) (0.84%), class Deltaproteobacteria (0.71%), family Chitinophagaceae (Bacteroidia) (0.69%), genus *Candidatus* Udaeobacter (Verrucomicrobiae) (0.66%), *Pseudomonas* (Gammaproteobacteria) (0.62%) and *Bradyrhizobium* (Alphaproteobacteria) (0.59%) (**Figure 2.4A**). Among all core bacterial and archaeal phyla, Proteobacteria (138 taxa), Acidobacteria (74 taxa), Bacteroidetes (53 taxa), and Verrucomicrobia (39 taxa) were the most abundant. Fungal core microbiome members consisted of 27 OTUs representing 60.26% of the total relative abundance that belong to Ascomycota (11 taxa), Basidiomycota (8 taxa), Mortierellomycota (3 taxa), and unidentified phyla (5 taxa) (**Figure 2.4B**). Fungal core taxa with high abundance were *Tetracladium* (15.6%), *Solicoccozyma* (12.1%), *Cystofilobasidiales* (5.9%), *Mortierella* (4.9%), *Exophiala* (2.3%), and *Alternaria* (1.2%) (**Figure 2.4B**). Finally, members of genus *Fusarium* registered at 100% occupancy. The high distribution of *Fusarium* in this study, members of which are common fungal pathogens in a variety of perennial crops, suggests that they are regionally cosmopolitan among Michigan apple orchards.

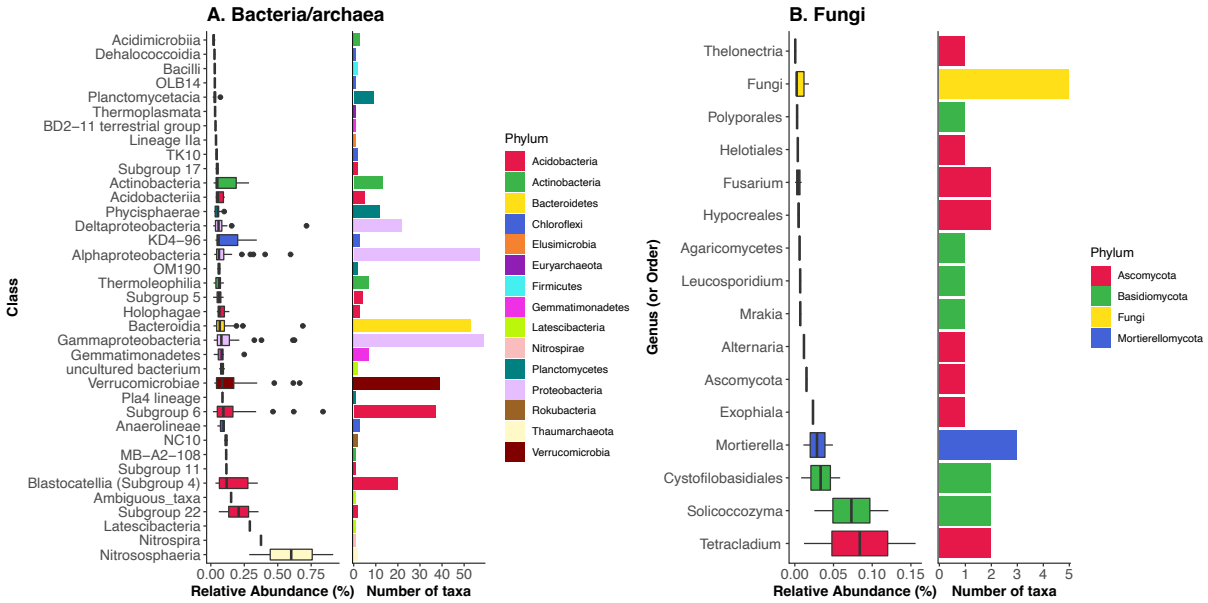

Figure 2.4. Relative abundances of apple root zone microbiome core taxa.

Bacterial/archaeal and fungal taxa with occupancy of 1 that were defined as core microbiome members. Bacterial/archaeal and fungal core taxa were classified into 15 and 3 phyla, respectively. Boxplots represent the percentage relative abundance of **A,** bacterial/archaeal core taxa that are grouped by class and **B,** fungal taxa that are grouped by genus or order. Bars represent the number of bacterial/archaeal and fungal core taxa are color coded by phylum.

Microbial network of apple root zone

We constructed correlation networks to better understand the complex associations within and between bacteria and archaea, fungi, and nematode, mycorrhizal fungi, and oligochaete communities in the apple root zone. We used 3,321 total OTUs that were detected in more than half of the samples (23 out of 45 total samples). The network was scale-free (R-square of power-law = 0.89), and it had 426 nodes and 615 edges. The network showed no significant correlation of nematodes, mycorrhizal fungi, and oligochaetes with 16S rRNA gene bacterial and archaeal taxa, or ITS fungal taxa. Among all nodes, 376 nodes were bacteria, 4 nodes were archaea, and 46 nodes were fungi. Bacteria-bacteria interactions dominated the networks and there were few bacteria-fungi, bacteria-archaea, and fungi-fungi associations (**Figure 2.5A**). Positive associations were generally separated in the network from negative associations. Among the few bacteria-fungi correlations observed, the associations tended to be negative. There were 12 taxa defined as module hubs ($Zi>2.5$ and $Pi<0.62$), and 5 taxa were connectors across modules ($Zi<2.5$ and $Pi>0.62$ **Figure 2.5B**). Module hubs included Gammaproteobacteria (4 OTUs), Verrucomicrobia (2 OTUs), Acidobacteria (2 OTUs), Chloroflexi (2 OTUs), Bacteroidetes (1 OTU), and Alphaproteobacteria (1 OTU). Meanwhile, there were three OTUs belonging to Acidobacteria, one Deltaproteobacteria, and one unclassified OTU that identified as connectors. The majority of taxa (409 OTUs) were peripheral ($Zi<2.5$ and $Pi<0.62$). Notably, there were no network hubs detected, indicating that there were no taxa with many interactions within and among modules ($Zi>2.5$ and $Pi>0.62$). This agrees with our prior abundance-occupancy and beta diversity analyses that suggest no strongly dominant taxa and substantial orchard-to-orchard variability.

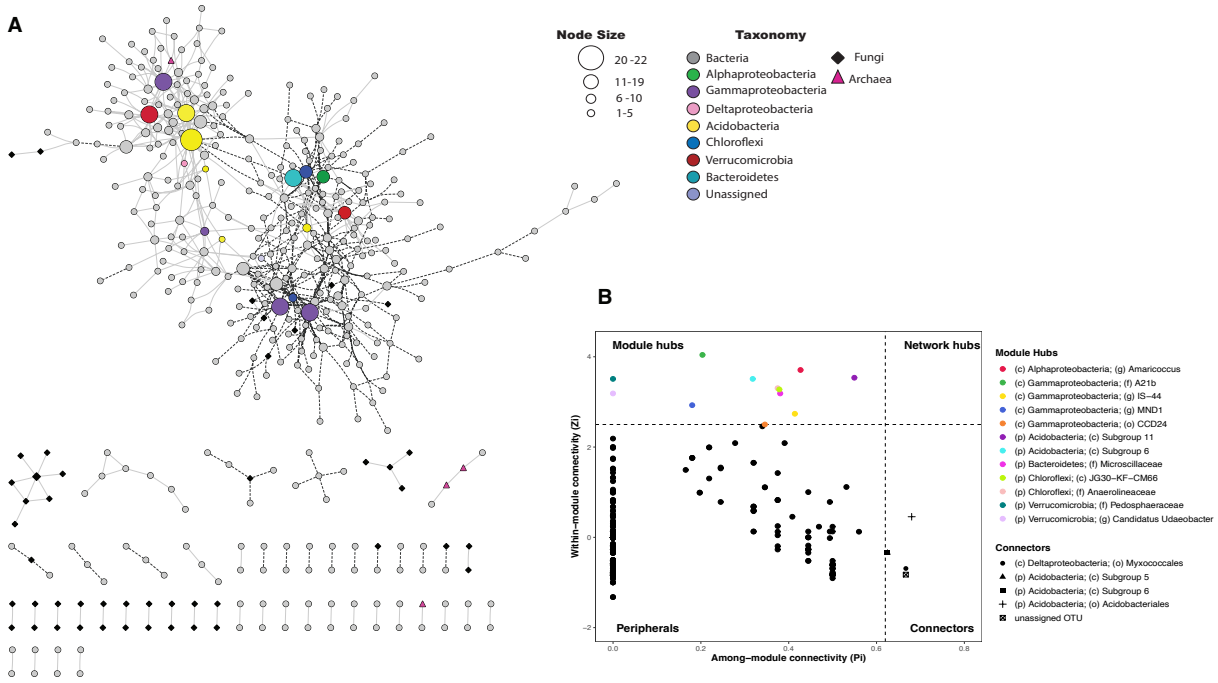

Figure 2.5. Co-occurrence network of apple root zone multi-trophic levels.

Co-occurrence network of trophic levels in the apple root zone was dominated by bacteria-bacteria interactions and there were few bacteria-fungi and bacteria-archaea interactions. There were no interactions of nematode, mycorrhizal fungi, and oligochaetes detected in the network. Solid and dash lines indicate positive and negative interactions, respectively. Node size is determined by the number of connecting edges. **A,** Colored nodes are taxa belonging to the module hub and connector which may play an important role for microbial network structure. **B,** The within (Zi) and among- (Pi) module connectivity plot revealed 12 and 5 module hub and connector taxa, respectively (B). All module hub and connector taxa were bacteria and the majority of taxa were peripherals.

Discussion

This study assessed multi-trophic microbial communities in apple root zones in an important U.S. apple growing region and revealed their association with each other and with environmental factors. Our results show that differences in microbial community structure in the apple root zone were mainly explained by the differences in orchard location, while the edaphic properties of particular soils were associated with bacterial and fungal alpha diversity. This result is consistent with the study of (8) which reported that different orchard locations determined soil microbial community composition and structure in the United Kingdom, and with (9), which reported that soil properties and orchard location influenced microbial composition in orchards in Bohai Gulf, China. Therefore, our study confirms that different geographical sites reflect the differences of soil properties and characteristics and lead to variation in microbial diversity in apple root zones.

We found that microbial community structure in the apple root zone was also affected by the rootstock but not the scion. However, this effect was minor compared with the effect of orchard location. These results agree with previous findings reported by (39), (40), and (41) that bacterial and fungal composition in apple root zones were influenced by rootstock cultivar, and this has also been reported for other plants (e.g. grapevine (42), tomato (43)). The rootstock is by definition in direct contact with the soil and associated soil microbes, and likely through root exudation (44) shapes the microbial diversity and composition in the root zone. Transcriptomic analysis of apple rootstock planted into disease-conducive soil revealed upregulated genes involved in secondary metabolism and plant defense, such as flavonoid, phenylpropanoid, and phenolic compounds that indicate a response to biotic stress (45). Moreover, phenolic compounds and rhizodeposits of apple rootstocks have been proposed as a contributing factor to

microbial community composition in the root zone (46, 47). In contrast, the scion is the particular cultivar grafted to the rootstock above ground and likely contributes little to the microbial communities in the root zone, though a recent study reported that different genotype combinations of scion and rootstock influenced fungal endophytic community composition of apple trees (48). Here, our study agrees again with prior work because we also detected a weak scion-rootstock interaction. Together, these results suggest that biogeography is more important in determining apple root zone communities, but that rootstock can also explain some of the variation after accounting for location.

We identified core microbiome members as bacterial, archaeal, or fungal taxa that were present in all samples. We discuss in more detail the composition of this core and what is known about their roles in soil or associations with apple in the Supplementary Information. This was a diverse group composed of over 400 taxa, which is a relatively large cohort for a core microbiome, as compared to our other studies that applied the abundance-occupancy approach (32, 49). Also striking was that there were no strongly dominating taxa in the bacterial core, with high mean relative abundances greater than 1%, and most would be considered members of the rare biosphere. Therefore, while we expected to identify a handful of tens of taxa that may be core to the apple root zone, we could not prioritize a few taxa from this large cohort. This suggests either that there are no universal bacterial and archaeal members of root zone microbiome, or that functional aspects of the collection of microbes is instead more important than the membership.

Perennial tree crops, such as apple, are assumed to develop more stable interactions with microbial communities in the rhizosphere due to the relative longevity of perennial plants and lack of soil disturbances like annual rotation (50). Moreover, it has been suggested that microbial

communities in the root zone of perennial trees are persistently affected by root exudates, which can eventually shape the microbial communities in distinctive manner from annual plants (50). Our study revealed that microbial community in the apple root zone had high bacterial diversity and also evenness. In addition, taxonomic identity of the core microbiome members revealed no dominance of particular taxa, which agrees with our observation of high evenness. An analysis of the root zone microbial community of apple and other fruit tree crops also reported high bacterial diversity in apple root zone (51). Greater diversity and evenness was also detected in a study comparing a perennial grass (switchgrass) as compared to an annual one (corn) (52). Together, our study and these others hint that high diversity and evenness, and also lack of dominant core microbiome members, may be characteristic of perennial crop microbiomes.

An advance that our study offers for characterizing the healthy apple root zone microbiome is that we have quantified several trophic levels and related their dynamics to one another, including bacteria and archaea, fungi, and nematodes. On balance, we did not find evidence of strong correlative associations between trophic levels. We observed concordance in the beta diversity and overarching biogeographic patterns between bacteria and fungi and between bacteria and nematodes (but not between fungi and nematodes), but the network analyses suggested few associations between particular taxa from these communities. For example, there were no interactions detected between nematode and microbial taxa, even though the PCoA showed that nematodes *Tylenchus* and *Pratylenchus* spp. had explanatory value for microbial community structure. The few network associations between bacteria and fungi were negative. However, network analyses should be interpreted carefully because they are hypothesis generating tools (53), and second, may not capture known biological interactions (54). However,

our study does not provide evidence that there are many or strong inter-trophic relationships that define the apple root zone microbiome.

In conclusion, our assessment of the microbial community structure and network of apple root zones revealed the complex associations among microbial members. Our study showed that the microbial community in apple root zones was strongly influenced by orchard location. Rootstock was also a minor but significant factor that contributed to the microbial community structure. In this study, we identified key belowground players and their possible interactions in Michigan apple orchards. The apple root zone microbial community showed diversity and structure typical of perennial crops, with high diversity and high evenness and many rare core microbiome members. However, we did not detect particular taxa and/or specific patterns of inter-trophic interactions that were characteristic of apple root zone soil. This is the first study to evaluate multiple trophic levels of apple orchard microbiome community through network analysis. This work can be used to inform “baseline” microbiome community structure and biogeography in the root zone, and in the future could be compared with unhealthy trees to determine any site-specific taxonomic shifts that are associated with tree health.

APPENDICES

APPENDIX A: Supplemental Tables

Table 2.3. Soil physical and chemical properties of 45 soil samples taken from 20 different sites of apple orchards in Michigan

Sample	Site	Lat.	Long.	Scion	Roots tock	pH	Lime index	P	K	Ca	Mg	NO3N	NH4N	OM	Sand	Silt	Clay	Soil type
								(ppm)							(%)			
F01	1	43.095927	-85.67709	Red	M26	6.7	71	41	178	1377	148	0.2	4.1	3.5	40.2	42	18	Loam
F02	1	43.095826	-85.677088	Empire	M26	6.3	70	84	189	901	89	0.2	3.7	2.8	51.2	34	15	Loam
F03	1	43.095557	-85.677091	Jona Gold	M26	5.7	68	78	137	615	57	0.2	4.2	2.5	63.1	25	12	Sandy Loam
F04	1	43.095475	-85.677066	Ida Red	M26	6.4	70	108	207	765	81	0.2	4.4	2.4	61.2	27	12	Sandy Loam
F05	1	43.09539	-85.677057	Ginger	M26	6	70	100	164	660	58	0.2	4.4	2.2	67.3	22	11	Sandy Loam

Table 2.3 (cont'd)

F06	2	43.102094	-85.72892	NSpy	M7	6.8	NA	34	135	1201	122	0.2	3.4	2.5	54.3	29	17	Sandy Loam
F07	2	43.102924	-85.728835	Rome	M7	6.6	72	67	115	862	57	0.2	2.5	1.6	74.3	16	10	Sandy Loam
F08	2	43.103401	-85.728604	NSpy	NSpy	7.2	NA	52	252	1908	148	0.4	5	4.4	41	36	23	Loam
F09	3	43.103714	-85.710496	Jonathan	M7	7.2	NA	31	162	1699	132	0.3	3.7	3.4	46.9	29	24	Loam
F10	3	43.103712	-85.710331	Red	M7	7.2	NA	19	204	1651	135	0.3	3.2	3.2	49	28	23	Sandy Clay Loam
F11	3	43.103724	-85.711774	Empire	M9	6.8	NA	8	75	1453	75	0.2	2.5	2.7	55	24	21	Sandy Clay Loam

Table 2.3 (cont'd)

F12	4	43.074743	-85.722812	Golden	M9	6.5	71	108	75	1203	208	0.2	3.5	3.1	41	38	21	Loam
F13	4	43.074724	-85.722356	Cameo	M26	6.3	70	29	103	1208	200	0.3	3.2	3.1	37	42	21	Loam
F14	5	43.059277	-85.745972	Jona Gold	M9	7	NA	78	153	1350	193	0.2	2.7	3.4	51	33	16	Loam
F15	5	43.059248	-85.74565	Fuji	M9	6.9	NA	54	113	1270	190	0.1	2.8	2.8	55	31	14	Sandy Loam
F16	6	43.11753	-85.766846	MacIntosh	M26	7.8	NA	37	117	3904	136	0.3	3.8	3.4	33	45	22	Loam
F17	6	43.117345	-85.766833	Red	M26	7.3	NA	67	101	2835	127	0.4	3.8	3	33	46	21	Loam

Table 2.3 (cont'd)

F18	7	43.117647	-85.760891	MacIntosh	M26	7.5	NA	47	112	1975	257	0.4	3.4	3.3	39	41	20	Loam
F19	7	43.117662	-85.761353	Red	M26	7.5	NA	41	132	1896	237	0.4	3.2	2.5	45	35	20	Loam
F20	8	43.121553	-85.826545	Jonathan	M9	6.7	71	96	160	1163	158	0.2	2.9	2.7	49	36	15	Loam
F21	8	43.121532	-85.826289	Golden	M9	6.3	71	131	222	1104	192	1	6.7	2.6	54	32	14	Sandy Loam
F22	9	43.121538	-85.825868	Jonathan	M9	6.7	71	114	181	1259	184	0.3	3.8	3	49	36	15	Loam
F23	9	43.12154	-85.825685	Golden	M9	6.4	70	117	208	1067	195	0.4	10.4	2.9	49	36	15	Loam

Table 2.3 (cont'd)

F24	10	42.932279	-85.786714	NSpy	Pajam	6.2	69	76	197	908	112	4.8	4.6	3.3	50	35	15	Loam
F25	10	42.932285	-85.786915	Paula Red	Pajam	7.1	NA	74	202	1233	141	0.3	3	2.4	51	34	15	Loam
F26	11	43.555967	-85.941927	Ida Red	Bud9	7	NA	86	226	2010	265	0.5	3.6	4.5	32	47	21	Loam
F27	12	43.554082	-85.941843	Ida Red	Bud9	6.6	71	40	120	1681	217	0.2	4.5	4.1	37.3	43	19	Loam
F28	13	43.548203	-85.971142	Rome	M7	6.5	71	157	119	1108	148	0.1	3	3.5	57.3	28	14	Sandy Loam
F29	13	43.54824	-85.97205	Red	M7	6.7	71	143	147	1174	181	0.2	3.7	3.8	59.3	25	15	Sandy Loam

Table 2.3 (cont'd)

F30	14	43.444984	-85.012165	MacIntosh	M7	6.9	NA	90	124	988	103	0.3	2.5	1.9	66.3	21	12	Sandy Loam
F31	14	43.444996	-85.011829	Red	M111	6.7	71	117	148	929	149	0.2	2.5	2.1	59.3	25	15	Sandy Loam
F32	15	43.238994	-85.741697	Empire	M7	6.1	70	135	87	622	106	0.2	2.4	1.6	71.3	18	10	Sandy Loam
F33	15	43.239027	-85.742739	Ida Red	M7	6.5	71	157	115	657	106	0.2	3	1.8	65.2	24	11	Sandy Loam
F34	15	43.238955	-85.742712	Jona Gold	M7	6.9	NA	124	111	814	145	0.2	2.9	2	65.2	23	12	Sandy Loam
F35	15	43.238943	-85.741731	NSpy	M7	6.7	72	127	105	809	116	0.2	2.6	1.9	71.2	18	11	Sandy Loam

Table 2.3 (cont'd)

F36	16	43.013931	-85.363285	Red	M126	6.5	71	38	218	1218	152	0.2	4.5	3.6	51.2	32	17	Loam
F37	17	43.012489	-85.362903	Red	M126	6.3	71	12	112	1231	193	0.3	4.3	3.6	53.2	28	19	Sandy Loam
F38	18	43.125647	-85.363552	Jona Gold	Bud9	7.1	NA	127	133	1093	135	0.4	3.6	2.6	61.2	27	12	Sandy Loam
F39	18	43.125644	-85.363781	Gala	Bud9	7.3	NA	96	145	1199	100	0.2	2.6	2.4	60.2	27	13	Sandy Loam
F40	18	43.125642	-85.364023	Golden	Bud9	6.8	NA	84	159	949	92	0.2	2.9	2.2	61.2	27	12	Sandy Loam
F41	19	43.233103	-85.746267	Red	M7	6.9	NA	157	179	1485	166	0.4	3.9	3.9	53.2	33	14	Sandy Loam

Table 2.3 (cont'd)

F42	19	43.233111	-85.746738	Ida Red	M7	6.9	NA	192	185	1497	175	0.8	4.7	3.6	61.3	24	14	Sandy Loam
F43	20	43.231152	-85.7498	Golden	M7	6.9	NA	152	205	1578	193	0.2	5	3.2	39.3	42	18	Loam
F44	20	43.231148	-85.749929	Jonathan	M7	6.9	NA	186	214	1483	184	0.4	4.4	3.9	43.3	37	19	Loam
F45	20	43.231109	-85.750031	Red	M7	7	NA	179	264	1872	235	0.4	4.3	3.7	35.3	43	21	Loam

Table 2.4. Comparison of bacterial and archaeal richness and Pielou's evenness among sites using post-hoc Dunn's test with Benjamini-Hochberg correction of p-values and post-hoc Tukey's HSD test, respectively

Site Comparison	Bacterial and archaeal Richness (post-hoc Dunn's test)			Bacterial and archaeal Pielou's evenness (post-hoc Tukey's HSD test)
	Z-test	P.unadj	P.adj	*p*-val
1 - 10	0.0819	0.9347	0.9812	1.000
1 - 11	-0.4587	0.6464	0.8773	0.959
10 - 11	-0.4663	0.6410	0.8826	0.999
1 - 12	-1.2233	0.2212	0.7374	0.919
10 - 12	-1.1501	0.2501	0.7665	0.998
11 - 12	-0.5922	0.5537	0.8623	1.000
1 - 13	-2.4207	0.0155	0.2943	**0.017**
10 - 13	-2.0938	0.0363	0.3829	0.227
11 - 13	-1.2433	0.2137	0.7252	0.990
12 - 13	-0.5595	0.5758	0.8615	0.996
1 - 14	-0.6916	0.4892	0.8768	0.999
10 - 14	-0.6472	0.5175	0.8625	1.000
11 - 14	-0.0622	0.9504	0.9814	1.000
12 - 14	0.6217	0.5342	0.8529	1.000
13 - 14	1.4466	0.1480	0.6695	0.422
1 - 15	-3.0191	0.0025	0.1204	**0.005**
10 - 15	-2.4177	0.0156	0.2698	0.210
11 - 15	-1.3620	0.1732	0.7001	0.998
12 - 15	-0.6129	0.5399	0.8549	0.999
13 - 15	0.0000	1.0000	1.0000	1.000
14 - 15	-1.6704	0.0948	0.6006	0.433
1 - 16	-0.8063	0.4201	0.8968	0.779
10 - 16	-0.7771	0.4371	0.8930	0.985
11 - 16	-0.2692	0.7878	0.9473	1.000
12 - 16	0.3230	0.7467	0.9458	1.000
13 - 16	0.9325	0.3511	0.8552	0.999
14 - 16	-0.2487	0.8036	0.9484	0.998
15 - 16	1.0215	0.3070	0.7991	1.000
1 - 17	0.2363	0.8132	0.9479	0.896
10 - 17	0.1554	0.8765	0.9854	0.997

Table 2.4 (cont'd)

11 - 17	0.5384	0.5903	0.8695	1.000
12 - 17	1.1306	0.2582	0.7666	1.000
13 - 17	1.8650	0.0622	0.5137	0.998
14 - 17	0.6838	0.4941	0.8612	0.999
15 - 17	2.0430	0.0411	0.4105	0.999
16 - 17	0.8076	0.4193	0.9054	1.000
1 - 18	-2.5995	0.0093	0.2957	**0.005**
10 - 18	-2.1546	0.0312	0.3704	0.169
11 - 18	-1.2089	0.2267	0.7301	0.992
12 - 18	-0.4835	0.6287	0.8783	0.997
13 - 18	0.1390	0.8894	0.9768	1.000
14 - 18	-1.4457	0.1483	0.6551	0.353
15 - 18	0.1661	0.8680	0.9935	1.000
16 - 18	-0.8792	0.3793	0.8683	0.999
17 - 18	-1.8682	0.0617	0.5331	0.998
1 - 19	-1.4652	0.1429	0.6787	0.344
10 - 19	-1.2944	0.1955	0.7145	0.904
11 - 19	-0.5906	0.5548	0.8570	1.000
12 - 19	0.0933	0.9257	0.9826	1.000
13 - 19	0.7995	0.4240	0.8952	0.998
14 - 19	-0.6472	0.5175	0.8702	0.986
15 - 19	0.9231	0.3559	0.8561	0.999
16 - 19	-0.2798	0.7797	0.9435	1.000
17 - 19	-1.2123	0.2254	0.7384	1.000
18 - 19	0.7368	0.4613	0.8853	0.998
1 - 2	-0.8271	0.4082	0.8914	1.000
10 - 2	-0.7368	0.4613	0.8764	1.000
11 - 2	-0.0879	0.9299	0.9816	0.998
12 - 2	0.6374	0.5239	0.8655	0.995
13 - 2	1.5569	0.1195	0.6678	0.115
14 - 2	-0.0278	0.9778	0.9882	1.000
15 - 2	1.8609	0.0628	0.4969	0.080
16 - 2	0.2418	0.8090	0.9488	0.967
17 - 2	-0.7473	0.4549	0.8910	0.992
18 - 2	1.5853	0.1129	0.6501	0.068
19 - 2	0.6811	0.4958	0.8563	0.795
1 - 20	-0.0973	0.9225	0.9847	0.467
10 - 20	-0.1529	0.8785	0.9818	0.979

Table 2.4 (cont'd)

11 - 20	0.3736	0.7087	0.9160	1.000
12 - 20	1.0990	0.2718	0.7484	1.000
13 - 20	2.1407	0.0323	0.3609	0.917
14 - 20	0.5560	0.5782	0.8582	0.999
15 - 20	2.5587	0.0105	0.2852	0.955
16 - 20	0.7033	0.4818	0.8803	1.000
17 - 20	-0.2857	0.7751	0.9501	1.000
18 - 20	2.2380	0.0252	0.3686	0.895
19 - 20	1.2650	0.2059	0.7380	1.000
2 - 20	0.6528	0.5139	0.8718	0.928
1 - 3	1.0843	0.2782	0.7552	0.999
10 - 3	0.7924	0.4282	0.8939	0.986
11 - 3	1.1209	0.2623	0.7668	0.663
12 - 3	1.8463	0.0649	0.4929	0.570
13 - 3	3.0860	0.0020	0.1285	**0.004**
14 - 3	1.5013	0.1333	0.6493	0.891
15 - 3	**3.6885**	**0.0002**	**0.0429**	**0.001**
16 - 3	1.4506	0.1469	0.6807	0.386
17 - 3	0.4616	0.6444	0.8808	0.529
18 - 3	3.2949	0.0010	0.0936	**0.001**
19 - 3	2.2103	0.0271	0.3676	0.095
2 - 3	1.7096	0.0873	0.5927	0.973
20 - 3	1.0568	0.2906	0.7668	0.127
1 - 4	-0.6006	0.5481	0.8606	0.999
10 - 4	-0.5710	0.5680	0.8565	1.000
11 - 4	0.0000	1.0000	1.0000	1.000
12 - 4	0.6838	0.4941	0.8692	1.000
13 - 4	1.5228	0.1278	0.6563	0.411
14 - 4	0.0761	0.9393	0.9806	1.000
15 - 4	1.7583	0.0787	0.5537	0.421
16 - 4	0.3108	0.7559	0.9326	0.998
17 - 4	-0.6217	0.5342	0.8601	0.999
18 - 4	1.5291	0.1262	0.6663	0.342
19 - 4	0.7233	0.4695	0.8745	0.985
2 - 4	0.1112	0.9115	0.9840	1.000
20 - 4	-0.4726	0.6365	0.8827	0.999
3 - 4	-1.4179	0.1562	0.6596	0.898
1 - 5	-0.8281	0.4076	0.9005	0.937

Table 2.4 (cont'd)

10 - 5	-0.7614	0.4464	0.8836	0.999
11 - 5	-0.1554	0.8765	0.9913	1.000
12 - 5	0.5284	0.5972	0.8596	1.000
13 - 5	1.3324	0.1827	0.7085	0.791
14 - 5	-0.1142	0.9091	0.9870	1.000
15 - 5	1.5386	0.1239	0.6727	0.848
16 - 5	0.1554	0.8765	0.9972	1.000
17 - 5	-0.7771	0.4371	0.9027	1.000
18 - 5	1.3206	0.1866	0.7092	0.754
19 - 5	0.5330	0.5941	0.8616	1.000
2 - 5	-0.0973	0.9225	0.9902	0.999
20 - 5	-0.6811	0.4958	0.8486	1.000
3 - 5	-1.6264	0.1039	0.6167	0.517
4 - 5	-0.1903	0.8490	0.9777	1.000
1 - 6	-0.1456	0.8842	0.9768	0.978
10 - 6	-0.1903	0.8490	0.9836	1.000
11 - 6	0.3108	0.7559	0.9387	1.000
12 - 6	0.9947	0.3199	0.7997	1.000
13 - 6	1.9035	0.0570	0.5155	0.687
14 - 6	0.4568	0.6478	0.8729	1.000
15 - 6	2.1979	0.0280	0.3541	0.739
16 - 6	0.6217	0.5342	0.8674	1.000
17 - 6	-0.3108	0.7559	0.9449	1.000
18 - 6	1.9461	0.0516	0.4906	0.635
19 - 6	1.1040	0.2696	0.7533	0.999
2 - 6	0.5282	0.5973	0.8470	1.000
20 - 6	-0.0556	0.9557	0.9815	1.000
3 - 6	-1.0009	0.3169	0.8028	0.641
4 - 6	0.3807	0.7034	0.9154	1.000
5 - 6	0.5710	0.5680	0.8633	1.000
1 - 7	0.4914	0.6231	0.8770	0.975
10 - 7	0.3426	0.7319	0.9333	1.000
11 - 7	0.7460	0.4557	0.8834	1.000
12 - 7	1.4298	0.1528	0.6597	1.000
13 - 7	2.4364	0.0148	0.3131	0.699
14 - 7	0.9898	0.3223	0.7952	1.000
15 - 7	2.8134	0.0049	0.1863	0.751
16 - 7	1.0568	0.2906	0.7776	1.000

Table 2.4 (cont'd)

17 - 7	0.1243	0.9011	0.9839	1.000
18 - 7	2.5300	0.0114	0.2709	0.648
19 - 7	1.6370	0.1016	0.6229	0.999
2 - 7	1.1121	0.2661	0.7661	0.999
20 - 7	0.5282	0.5973	0.8533	1.000
3 - 7	-0.4170	0.6767	0.8991	0.628
4 - 7	0.9137	0.3609	0.8465	1.000
5 - 7	1.1040	0.2696	0.7645	1.000
6 - 7	0.5330	0.5941	0.8682	1.000
1 - 8	-1.0101	0.3124	0.8022	1.000
10 - 8	-0.9137	0.3609	0.8571	1.000
11 - 8	-0.2798	0.7797	0.9496	0.994
12 - 8	0.4041	0.6861	0.9053	0.984
13 - 8	1.1801	0.2379	0.7535	0.116
14 - 8	-0.2665	0.7899	0.9439	1.000
15 - 8	1.3627	0.1730	0.7144	0.096
16 - 8	0.0311	0.9752	0.9908	0.932
17 - 8	-0.9014	0.3674	0.8512	0.978
18 - 8	1.1538	0.2486	0.7743	0.078
19 - 8	0.3807	0.7034	0.9217	0.730
2 - 8	-0.2641	0.7917	0.9401	1.000
20 - 8	-0.8480	0.3965	0.8862	0.880
3 - 8	-1.7932	0.0729	0.5330	0.999
4 - 8	-0.3426	0.7319	0.9396	1.000
5 - 8	-0.1523	0.8790	0.9766	0.996
6 - 8	-0.7233	0.4695	0.8832	0.999
7 - 8	-1.2563	0.2090	0.7354	0.999
1 - 9	-1.5107	0.1309	0.6544	1.000
10 - 9	-1.3324	0.1827	0.7233	1.000
11 - 9	-0.6217	0.5342	0.8749	0.999
12 - 9	0.0622	0.9504	0.9868	0.999
13 - 9	0.7614	0.4464	0.8929	0.253
14 - 9	-0.6852	0.4932	0.8758	1.000
15 - 9	0.8792	0.3793	0.8580	0.239
16 - 9	-0.3108	0.7559	0.9512	0.989
17 - 9	-1.2433	0.2137	0.7384	0.998
18 - 9	0.6950	0.4870	0.8813	0.193
19 - 9	-0.0381	0.9696	0.9905	0.925

Table 2.4 (cont'd)

2 - 9	-0.7228	0.4698	0.8666	1.000
20 - 9	-1.3067	0.1913	0.7128	0.986
3 - 9	-2.2520	0.0243	0.3851	0.979
4 - 9	-0.7614	0.4464	0.9024	1.000
5 - 9	-0.5710	0.5680	0.8703	1.000
6 - 9	-1.1421	0.2534	0.7643	1.000
7 - 9	-1.6751	0.0939	0.6154	1.000
8 - 9	-0.4188	0.6754	0.9037	1.000

Table 2.5. Comparison of bacterial and archaeal richness and Pielou's evenness among rootstocks using post-hoc Dunn's test with Benjamini-Hochberg correction of p-values and post-hoc Tukey's HSD test, respectively

Rootstock Comparison	Bacterial and archaeal Richness (post-hoc Dunn's test)			Bacterial and archaeal Pielou's evenness (Tukey's HSD test)
	Z-test	P.unadj	P.adj	*p*-val
Bud9 - M111	1.752	0.080	0.373	0.308
Bud9 - M126	1.429	0.153	0.390	0.998
M111 - M126	-0.591	0.555	0.740	0.715
Bud9 - M26	2.780	0.005	0.152	**0.037**
M111 - M26	-0.377	0.706	0.760	0.996
M126 - M26	0.423	0.673	0.785	0.701
Bud9 - M7	0.819	0.413	0.608	0.945
M111 - M7	-1.454	0.146	0.408	0.554
M126 - M7	-1.034	0.301	0.562	1.000
M26 - M7	-2.736	0.006	0.087	0.068
Bud9 - M9	1.412	0.158	0.368	**0.022**
M111 - M9	-1.050	0.294	0.588	0.999
M126 - M9	-0.494	0.622	0.791	0.563
M26 - M9	-1.513	0.130	0.405	0.999
M7 - M9	0.890	0.373	0.581	**0.038**
Bud9 - NSpy	2.308	0.021	0.196	0.250
M111 - NSpy	0.431	0.667	0.812	1.000
M126 - NSpy	1.088	0.277	0.596	0.648
M26 - NSpy	0.958	0.338	0.557	0.991
M7 - NSpy	2.045	0.041	0.286	0.472
M9 - NSpy	1.624	0.104	0.417	0.998
Bud9 - Pajam	1.884	0.060	0.334	0.330
M111 - Pajam	-0.280	0.780	0.809	0.999
M126 - Pajam	0.381	0.703	0.788	0.857
M26 - Pajam	0.069	0.945	0.945	1.000
M7 - Pajam	1.542	0.123	0.431	0.641
M9 - Pajam	0.975	0.329	0.577	1.000
NSpy - Pajam	-0.777	0.437	0.612	0.998

Table 2.6. Pearson's correlation analysis between alpha diversity metrics and soil parameters. r = correlation coefficient. Significant correlations (p-val < 0.05) are in bold

Soil physiochemical properties	Bacteria/archaea				Fungi			
	Richness		Pielou's evenness		Richness		Pielou's evenness	
	r	*p*-val	r	*p*-val	r	*p*-val	r	*p*-val
pH	-0.17	0.25	0.13	0.38	-0.05	0.71	-0.1	0.5
P	0.47	**0.001**	0.43	**0.002**	0.16	0.26	0.04	0.79
K	-0.39	**0.006**	-0.19	0.21	0.02	0.87	-0.3	**0.04**
Ca	-0.38	**0.008**	-0.05	0.71	-0.1	0.49	-0.14	0.33
Mg	-0.16	0.26	0.15	0.31	0.09	0.52	-0.32	**0.02**
NO3N	-0.17	0.24	-0.07	0.63	-0.11	0.46	-0.19	0.19
NH4N	-0.24	0.09	-0.14	0.34	-0.02	0.87	-0.17	0.25
OM	-0.42	**0.003**	-0.01	0.91	0.07	0.6	-0.22	0.13
Sand	0.56	**4.38E-05**	0.22	0.13	0.18	0.23	0.32	**0.02**
Silt	-0.45	**0.001**	-0.15	0.31	-0.17	0.23	-0.31	**0.03**
Clay	-0.67	**3.38E-07**	-0.32	**0.03**	-0.14	0.33	-0.27	0.06

Table 2.7. Comparison using post-hoc Tukey's HSD test of fungal richness among sites

Site Comparison	Fungal Richness
	p-val
1 - 10	1.00
1 - 11	0.73
1 - 12	0.99
1 - 13	1.00
1 - 14	1.00
1 - 15	0.16
1 - 16	0.21
1 - 17	**0.03**
1 - 18	0.19
1 - 19	1.00
1 - 2	1.00
1 - 20	1.00
1 - 3	1.00
1 - 4	1.00
1 - 5	1.00
1 - 6	1.00
1 - 7	1.00
1 - 8	0.98
1 - 9	1.00
10 - 11	0.57
10 - 12	0.95
10 - 13	1.00
10 - 14	1.00
10 - 15	0.18
10 - 16	0.16
10 - 17	**0.02**
10 - 18	0.19
10 - 19	0.99
10 - 2	1.00
10 - 20	0.97
10 - 3	1.00
10 - 4	1.00
10 - 5	1.00
10 - 6	1.00

Table 2.7 (cont'd)

10 - 7	1.00
10 - 8	0.90
10 - 9	1.00
11 - 12	1.00
11 - 13	0.85
11 - 14	0.65
11 - 15	1.00
11 - 16	1.00
11 - 17	0.99
11 - 18	1.00
11 - 19	1.00
11 - 2	0.50
11 - 20	1.00
11 - 3	0.75
11 - 4	0.68
11 - 5	0.98
11 - 6	0.85
11 - 7	0.75
11 - 8	1.00
11 - 9	0.49
12 - 13	1.00
12 - 14	0.97
12 - 15	1.00
12 - 16	0.99
12 - 17	0.79
12 - 18	1.00
12 - 19	1.00
12 - 2	0.93
12 - 20	1.00
12 - 3	0.99
12 - 4	0.98
12 - 5	1.00
12 - 6	1.00
12 - 7	0.99
12 - 8	1.00
12 - 9	0.91
13 - 14	1.00
13 - 15	0.52

Table 2.7 (cont'd)

13 - 16	0.35
13 - 17	0.07
13 - 18	0.51
13 - 19	1.00
13 - 2	1.00
13 - 20	1.00
13 - 3	1.00
13 - 4	1.00
13 - 5	1.00
13 - 6	1.00
13 - 7	1.00
13 - 8	1.00
13 - 9	1.00
14 - 15	0.24
14 - 16	0.19
14 - 17	**0.03**
14 - 18	0.25
14 - 19	0.99
14 - 2	1.00
14 - 20	0.99
14 - 3	1.00
14 - 4	1.00
14 - 5	1.00
14 - 6	1.00
14 - 7	1.00
14 - 8	0.95
14 - 9	1.00
15 - 16	1.00
15 - 17	0.82
15 - 18	1.00
15 - 19	0.99
15 - 2	0.08
15 - 20	0.96
15 - 3	0.26
15 - 4	0.27
15 - 5	0.90
15 - 6	0.52
15 - 7	0.35

Table 2.7 (cont'd)

15 - 8	1.00
15 - 9	0.13
16 - 17	1.00
16 - 18	1.00
16 - 19	0.82
16 - 2	0.11
16 - 20	0.74
16 - 3	0.23
16 - 4	0.21
16 - 5	0.65
16 - 6	0.35
16 - 7	0.25
16 - 8	0.94
16 - 9	0.12
17 - 18	0.90
17 - 19	0.30
17 - 2	**0.01**
17 - 20	0.21
17 - 3	**0.03**
17 - 4	**0.03**
17 - 5	0.18
17 - 6	0.07
17 - 7	**0.04**
17 - 8	0.46
17 - 9	**0.02**
18 - 19	0.98
18 - 2	0.10
18 - 20	0.94
18 - 3	0.28
18 - 4	0.28
18 - 5	0.89
18 - 6	0.51
18 - 7	0.35
18 - 8	1.00
18 - 9	0.13
19 - 2	0.97
19 - 20	1.00
19 - 3	1.00

Table 2.7 (cont'd)

19 - 4	1.00
19 - 5	1.00
19 - 6	1.00
19 - 7	1.00
19 - 8	1.00
19 - 9	0.96
2 - 20	0.93
2 - 3	1.00
2 - 4	1.00
2 - 5	1.00
2 - 6	1.00
2 - 7	1.00
2 - 8	0.85
2 - 9	1.00
20 - 3	1.00
20 - 4	0.99
20 - 5	1.00
20 - 6	1.00
20 - 7	1.00
20 - 8	1.00
20 - 9	0.93
3 - 4	1.00
3 - 5	1.00
3 - 6	1.00
3 - 7	1.00
3 - 8	0.98
3 - 9	1.00
4 - 5	1.00
4 - 6	1.00
4 - 7	1.00
4 - 8	0.96
4 - 9	1.00
5 - 6	1.00
5 - 7	1.00
5 - 8	1.00
5 - 9	1.00
6 - 7	1.00
6 - 8	1.00

Table 2.7 (cont'd)

6 - 9	1.00
7 - 8	0.98
7 - 9	1.00
8 - 9	0.84

Table 2.8. Comparison using post-hoc Tukey's HSD test of fungal richness among rootstocks

Rootstock Comparison	Fungal Richness
	***p*-val**
Bud9 - M111	0.63
Bud9 - M126	0.56
Bud9 - M26	**0.01**
Bud9 - M7	0.18
Bud9 - M9	0.06
Bud9 - NSpy	0.86
Bud9 - Pajam	0.094
M111 - M126	0.102
M111 - M26	1.000
M111 - M7	0.999
M111 - M9	1.000
M111 - NSpy	1.000
M111 - Pajam	1.000
M126 - M26	**0.001**
M126 - M7	**0.01**
M126 - M9	**0.003**
M126 - NSpy	0.21
M126 - Pajam	**0.01**
M26 - M7	0.65
M26 - M9	1.00
M26 - NSpy	1.00
M26 - Pajam	1.00
M7 - M9	0.97
M7 - NSpy	1.00
M7 - Pajam	0.81
M9 - NSpy	1.00
M9 - Pajam	0.99
NSpy - Pajam	0.99

Table 2.9. The abundance of nematodes, mycorrhizal fungi, and Oligochaetes from 45 soil samples

Sample code	Site	Root stock	Absolute abundance (individuals per 100-gram dry soil)												
			Pratylenchus spp. (lesion)	*Xiphinema* spp. (dagger)	Criconematidae (ring)	*Paratylenchus* spp. (pin)	*Tylenchorhynchus* (stunt)	*Helicotylenchus* spp. (spiral)	*Tylenchus*	*Aphelenchus*	Dorylaimidae	Monochiodae	Bacterial Feeders	Mycorrhizal Fungi	Oligochaetes
F01	1	M26	4	8	0	0	0	0	60	8	4	20	840	208	36
F02	1	M26	4	4	0	0	0	0	84	12	8	0	312	64	12
F03	1	M26	14	0	0	0	0	0	244	16	4	4	180	32	8
F04	1	M26	0	4	0	0	0	0	92	36	4	0	816	60	8
F05	1	M26	2	4	0	66	0	0	256	8	0	0	164	40	16
F06	2	M7	0	14	0	0	0	0	40	8	4	4	232	128	28
F07	2	M7	2	16	0	0	50	8	28	8	28	8	612	52	12
F08	2	NSpy	8	48	6	0	0	200	176	8	12	4	296	388	20
F09	3	M7	12	16	0	2	0	0	148	28	8	0	300	100	4
F10	3	M7	0	6	0	2	0	0	228	32	8	0	420	36	16
F11	3	M9	8	74	0	0	0	0	112	20	24	0	364	276	4
F12	4	M9	0	22	0	0	0	4	92	40	8	4	216	48	0
F13	4	M26	0	0	0	0	0	0	56	16	20	0	180	116	0
F14	5	M9	0	20	6	60	0	0	60	0	4	8	324	88	2
F15	5	M9	4	32	0	0	0	0	140	36	12	4	700	24	4
F16	6	M26	10	18	0	0	0	0	20	4	36	4	848	28	0
F17	6	M26	4	10	0	36	0	0	40	4	8	8	212	8	8
F18	7	M26	0	0	0	4	4	0	32	20	20	0	176	24	12
F19	7	M26	4	0	0	2	0	0	28	8	32	16	540	20	20
F20	8	M9	0	2	0	10	0	0	32	8	8	8	1296	4	12
F21	8	M9	0	12	0	46	0	0	24	24	0	4	480	8	4

Table 2.9 (cont'd)

F22	9	M9	0	4	0	26	0	0	24	4	0	0	172	20	12
F23	9	M9	0	22	0	28	0	0	92	0	4	0	368	8	24
F24	10	Pajam	10	0	0	0	0	0	0	12	4	4	84	8	0
F25	10	Pajam	0	18	0	2	0	0	112	0	8	0	48	56	0
F26	11	Bud9	2	48	0	360	0	0	28	12	4	0	320	32	0
F27	12	Bud9	10	24	4	96	0	0	140	8	8	8	332	28	16
F28	13	M7	4	38	0	0	0	0	4	8	16	0	364	8	12
F29	13	M7	4	12	0	0	0	0	8	8	8	8	472	8	20
F30	14	M7	0	4	0	0	0	0	44	8	0	4	200	16	32
F31	14	M111	0	0	0	0	0	0	20	8	12	0	492	24	16
F32	15	M7	0	4	2	2	0	0	60	44	16	0	500	24	4
F33	15	M7	0	0	0	0	0	0	20	0	0	0	1800	368	8
F34	15	M7	12	4	0	4	0	0	84	16	44	12	532	132	16
F35	15	M7	0	0	0	0	0	0	64	0	8	0	204	120	4
F36	16	M126	18	18	0	0	0	0	32	28	8	4	844	72	8
F37	17	M126	4	140	0	0	0	0	96	12	8	0	180	200	76
F38	18	Bud9	0	0	10	10	0	0	80	4	8	20	608	100	0
F39	18	Bud9	0	0	0	0	0	0	100	40	10	0	1640	108	16
F40	18	Bud9	4	4	0	0	0	0	68	8	8	0	184	176	2
F41	19	M7	4	26	0	0	0	0	120	36	8	16	632	12	4
F42	19	M7	2	4	0	0	0	0	20	60	10	0	2160	32	16
F43	20	M7	0	4	0	0	0	0	16	16	4	0	976	44	12
F44	20	M7	4	0	0	0	0	0	64	40	24	0	1024	44	20
F45	20	M7	0	0	0	0	0	0	60	0	20	0	1560	28	0
TOTAL			**154**	**684**	**28**	**756**	**54**	**212**	**3348**	**716**	**492**	**172**	**25204**	**3420**	**544**
MEAN			**3.42**	**15.20**	**0.62**	**16.80**	**1.20**	**4.71**	**74.40**	**15.91**	**10.93**	**3.82**	**560.09**	**76.00**	**12.09**
STDEV			**4.53**	**24.64**	**1.99**	**56.25**	**7.46**	**29.80**	**61.91**	**14.53**	**9.84**	**5.52**	**480.86**	**89.78**	**13.25**

Table 2.10. Prevalence percentage of nematodes and other soil microorganisms among 45 soil samples taken from 20 different sites

Group	Characteristic	Prevalence (%)
Non-nematodes		
Mycorrhizal Fungi	Beneficial fungi - non plant-parasitic	100
Oligochaetes	Detritivore - non plant-parasitic	82.22
Nematodes		
Rhabditidae	Bacterivores - non plant-parasitic	100
Tylenchus	Fungivores - non plant-parasitic	97.77
Aphelenchus	Fungivores - non plant-parasitic	86.66
Dorylaimidae	Omnivores - non plant-parasitic	88.88
Monochidae	Carnivore - non plant-parasitic	46.66
Xiphinema spp. (dagger)	Herbivore - plant-parasitic	73.33
Pratylenchus spp. (lesion)	Herbivore - plant-parasitic	53.33
Paratylenchus spp. (pin)	Herbivore - plant-parasitic	37.77
Criconematidae (ring)	Herbivore - plant-parasitic	11.11
Helicotylenchus spp. (spiral)	Herbivore - plant-parasitic	6.66
Tylenchorhynchus (stunt)	Herbivore - plant-parasitic	4.44

Table 2.11. Comparison using post-hoc Tukey's HSD test of total absolute abundance of nematodes among rootstocks

Rootstock Comparison	Nematodes Abundance
	***p*-val**
Bud9 - M111	0.70
Bud9 - M126	1.00
Bud9 - M26	0.66
Bud9 - M7	0.95
Bud9 - M9	0.97
Bud9 - NSpy	0.89
Bud9 - Pajam	**0.03**
M111 - M126	0.90
M111 - M26	0.99
M111 - M7	0.92
M111 - M9	0.93
M111 - NSpy	0.32
M111 - Pajam	0.99
M126 - M26	0.99
M126 - M7	1.00
M126 - M9	1.00
M126 - NSpy	0.86
M126 - Pajam	0.19
M26 - M7	0.98
M26 - M9	0.99
M26 - NSpy	0.34
M26 - Pajam	0.24
M7 - M9	1.00
M7 - NSpy	0.53
M7 - Pajam	0.07
M9 - NSpy	0.56
M9 - Pajam	0.11
NSpy - Pajam	**0.02**

Table 2.12. Comparison using post-hoc Tukey's HSD test of oligochaetes abundance among sites

Site Comparison	Oligochaetes Abundance
	***p*-val**
1 - 10	0.136
1 - 11	0.507
1 - 12	1.000
1 - 13	1.000
1 - 14	1.000
1 - 15	0.998
1 - 16	1.000
1 - 17	0.158
1 - 18	0.834
1 - 19	1.000
1 - 2	1.000
1 - 20	0.999
1 - 3	0.999
1 - 4	0.136
1 - 5	0.908
1 - 6	0.790
1 - 7	1.000
1 - 8	1.000
1 - 9	1.000
10 - 11	1.000
10 - 12	0.615
10 - 13	0.304
10 - 14	0.092
10 - 15	0.681
10 - 16	0.955
10 - 17	**0.002**
10 - 18	0.989
10 - 19	0.739
10 - 2	0.092
10 - 20	0.782
10 - 3	0.773
10 - 4	1.000
10 - 5	0.998

Table 2.12 (cont'd)

10 - 6	1.000
10 - 7	0.304
10 - 8	0.846
10 - 9	0.233
11 - 12	0.810
11 - 13	0.627
11 - 14	0.313
11 - 15	0.934
11 - 16	0.988
11 - 17	**0.010**
11 - 18	0.999
11 - 19	0.928
11 - 2	0.362
11 - 20	0.958
11 - 3	0.955
11 - 4	1.000
11 - 5	1.000
11 - 6	1.000
11 - 7	0.627
11 - 8	0.966
11 - 9	0.543
12 - 13	1.000
12 - 14	1.000
12 - 15	1.000
12 - 16	1.000
12 - 17	0.581
12 - 18	0.993
12 - 19	1.000
12 - 2	1.000
12 - 20	1.000
12 - 3	1.000
12 - 4	0.615
12 - 5	0.994
12 - 6	0.980
12 - 7	1.000
12 - 8	1.000
12 - 9	1.000
13 - 14	1.000

Table 2.12 (cont'd)

13 - 15	1.000
13 - 16	1.000
13 - 17	0.338
13 - 18	0.944
13 - 19	1.000
13 - 2	1.000
13 - 20	1.000
13 - 3	1.000
13 - 4	0.304
13 - 5	0.962
13 - 6	0.902
13 - 7	1.000
13 - 8	1.000
13 - 9	1.000
14 - 15	0.928
14 - 16	0.999
14 - 17	0.658
14 - 18	0.592
14 - 19	0.995
14 - 2	1.000
14 - 20	0.940
14 - 3	0.945
14 - 4	0.092
14 - 5	0.683
14 - 6	0.543
14 - 7	1.000
14 - 8	0.981
14 - 9	1.000
15 - 16	1.000
15 - 17	**0.038**
15 - 18	1.000
15 - 19	1.000
15 - 2	0.967
15 - 20	1.000
15 - 3	1.000
15 - 4	0.681
15 - 5	1.000
15 - 6	1.000

Table 2.12 (cont'd)

15 - 7	1.000
15 - 8	1.000
15 - 9	0.998
16 - 17	0.238
16 - 18	1.000
16 - 19	1.000
16 - 2	1.000
16 - 20	1.000
16 - 3	1.000
16 - 4	0.955
16 - 5	1.000
16 - 6	1.000
16 - 7	1.000
16 - 8	1.000
16 - 9	1.000
17 - 18	**0.013**
17 - 19	0.118
17 - 2	0.398
17 - 20	**0.046**
17 - 3	**0.048**
17 - 4	**0.002**
17 - 5	**0.021**
17 - 6	**0.014**
17 - 7	0.338
17 - 8	0.084
17 - 9	0.410
18 - 19	1.000
18 - 2	0.651
18 - 20	1.000
18 - 3	1.000
18 - 4	0.989
18 - 5	1.000
18 - 6	1.000
18 - 7	0.944
18 - 8	1.000
18 - 9	0.888
19 - 2	0.999
19 - 20	1.000

Table 2.12 (cont'd)

19 - 3	1.000
19 - 4	0.739
19 - 5	1.000
19 - 6	0.999
19 - 7	1.000
19 - 8	1.000
19 - 9	1.000
2 - 20	0.975
2 - 3	0.977
2 - 4	0.092
2 - 5	0.756
2 - 6	0.607
2 - 7	1.000
2 - 8	0.995
2 - 9	1.000
20 - 3	1.000
20 - 4	0.782
20 - 5	1.000
20 - 6	1.000
20 - 7	1.000
20 - 8	1.000
20 - 9	0.998
3 - 4	0.773
3 - 5	1.000
3 - 6	1.000
3 - 7	1.000
3 - 8	1.000
3 - 9	0.999
4 - 5	0.998
4 - 6	1.000
4 - 7	0.304
4 - 8	0.846
4 - 9	0.233
5 - 6	1.000
5 - 7	0.962
5 - 8	1.000
5 - 9	0.922
6 - 7	0.902

Table 2.12 (cont'd)

6 - 8	1.000
6 - 9	0.833
7 - 8	1.000
7 - 9	1.000
8 - 9	1.000

Table 2.13. Comparison using post-hoc Tukey's HSD test of mycorrhizal fungi abundance among sites

Site Comparison	Mycorrhizal fungi Abundance
	***p*-val**
1 - 10	0.97
1 - 11	1.00
1 - 12	1.00
1 - 13	0.24
1 - 14	0.95
1 - 15	1.00
1 - 16	1.00
1 - 17	1.00
1 - 18	1.00
1 - 19	0.95
1 - 2	1.00
1 - 20	1.00
1 - 3	1.00
1 - 4	1.00
1 - 5	1.00
1 - 6	0.78
1 - 7	0.98
1 - 8	0.09
1 - 9	0.63
10 - 11	1.00
10 - 12	1.00
10 - 13	1.00
10 - 14	1.00
10 - 15	0.66
10 - 16	1.00
10 - 17	0.70
10 - 18	0.62
10 - 19	1.00
10 - 2	0.53
10 - 20	1.00
10 - 3	0.80
10 - 4	0.98
10 - 5	1.00

Table 2.13 (cont'd)

10 - 6	1.00
10 - 7	1.00
10 - 8	0.97
10 - 9	1.00
11 - 12	1.00
11 - 13	0.99
11 - 14	1.00
11 - 15	1.00
11 - 16	1.00
11 - 17	0.97
11 - 18	0.99
11 - 19	1.00
11 - 2	0.98
11 - 20	1.00
11 - 3	1.00
11 - 4	1.00
11 - 5	1.00
11 - 6	1.00
11 - 7	1.00
11 - 8	0.94
11 - 9	1.00
12 - 13	1.00
12 - 14	1.00
12 - 15	0.99
12 - 16	1.00
12 - 17	0.95
12 - 18	0.97
12 - 19	1.00
12 - 2	0.95
12 - 20	1.00
12 - 3	0.99
12 - 4	1.00
12 - 5	1.00
12 - 6	1.00
12 - 7	1.00
12 - 8	0.97
12 - 9	1.00
13 - 14	1.00

Table 2.13 (cont'd)

13 - 15	0.06
13 - 16	0.73
13 - 17	0.17
13 - 18	0.07
13 - 19	1.00
13 - 2	**0.05**
13 - 20	0.80
13 - 3	0.12
13 - 4	0.39
13 - 5	0.76
13 - 6	1.00
13 - 7	1.00
13 - 8	1.00
13 - 9	1.00
14 - 15	0.58
14 - 16	1.00
14 - 17	0.65
14 - 18	0.55
14 - 19	1.00
14 - 2	0.46
14 - 20	1.00
14 - 3	0.74
14 - 4	0.96
14 - 5	1.00
14 - 6	1.00
14 - 7	1.00
14 - 8	0.98
14 - 9	1.00
15 - 16	1.00
15 - 17	1.00
15 - 18	1.00
15 - 19	0.58
15 - 2	1.00
15 - 20	0.95
15 - 3	1.00
15 - 4	1.00
15 - 5	1.00
15 - 6	0.35

Table 2.13 (cont'd)

15 - 7	0.69
15 - 8	**0.02**
15 - 9	0.23
16 - 17	1.00
16 - 18	1.00
16 - 19	1.00
16 - 2	1.00
16 - 20	1.00
16 - 3	1.00
16 - 4	1.00
16 - 5	1.00
16 - 6	0.98
16 - 7	1.00
16 - 8	0.51
16 - 9	0.94
17 - 18	1.00
17 - 19	0.65
17 - 2	1.00
17 - 20	0.93
17 - 3	1.00
17 - 4	1.00
17 - 5	0.99
17 - 6	0.47
17 - 7	0.72
17 - 8	0.08
17 - 9	0.37
18 - 19	0.55
18 - 2	1.00
18 - 20	0.93
18 - 3	1.00
18 - 4	1.00
18 - 5	1.00
18 - 6	0.33
18 - 7	0.65
18 - 8	**0.02**
18 - 9	0.23
19 - 2	0.46
19 - 20	1.00

Table 2.13 (cont'd)

19 - 3	0.74
19 - 4	0.96
19 - 5	1.00
19 - 6	1.00
19 - 7	1.00
19 - 8	0.98
19 - 9	1.00
2 - 20	0.87
2 - 3	1.00
2 - 4	1.00
2 - 5	0.99
2 - 6	0.26
2 - 7	0.56
2 - 8	**0.02**
2 - 9	0.18
20 - 3	0.99
20 - 4	1.00
20 - 5	1.00
20 - 6	1.00
20 - 7	1.00
20 - 8	0.50
20 - 9	0.99
3 - 4	1.00
3 - 5	1.00
3 - 6	0.50
3 - 7	0.82
3 - 8	**0.04**
3 - 9	0.37
4 - 5	1.00
4 - 6	0.86
4 - 7	0.98
4 - 8	0.19
4 - 9	0.75
5 - 6	0.99
5 - 7	1.00
5 - 8	0.49
5 - 9	0.97
6 - 7	1.00

Table 2.13 (cont'd)

6 - 8	1.00
6 - 9	1.00
7 - 8	0.96
7 - 9	1.00
8 - 9	1.00

APPENDIX B: Supplemental Figures

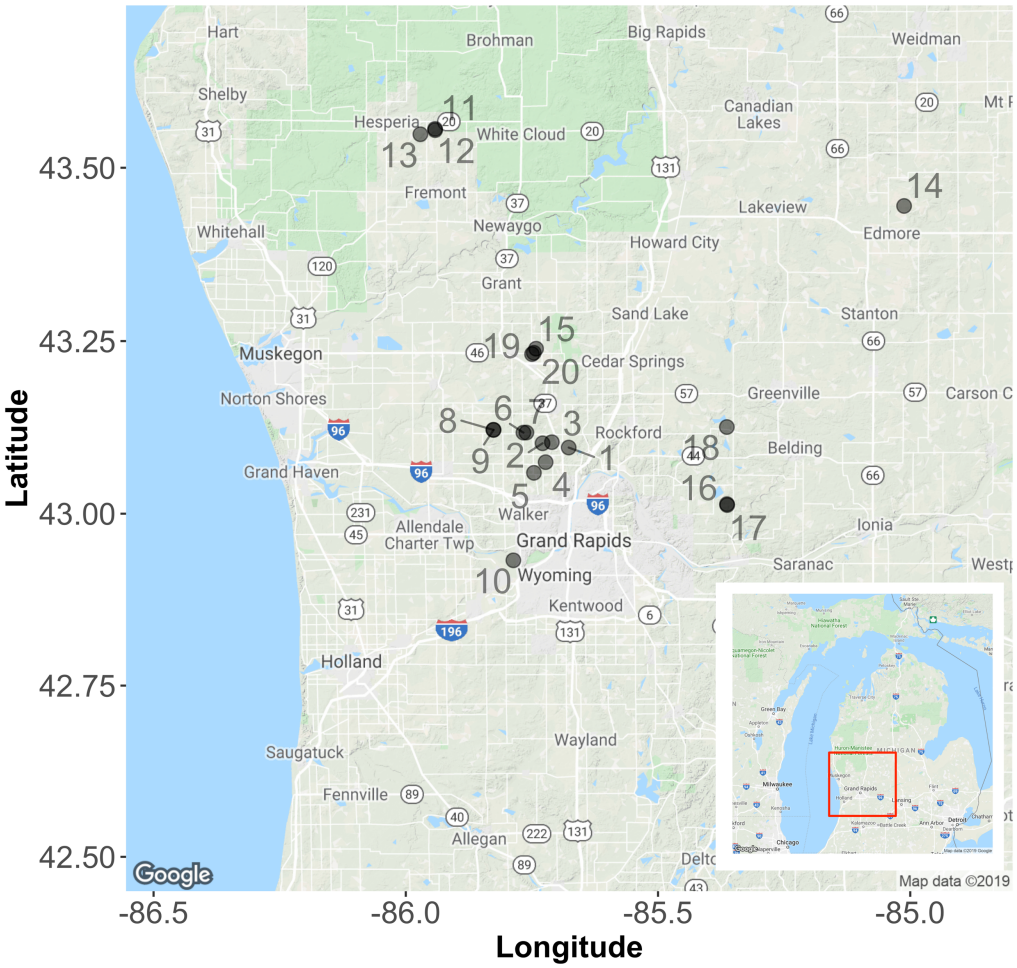

Figure 2.6. Map of the sampling location in Michigan apple orchards.

Each number represents the sampling sites. There are 20 different sites and some of them are overlap because their locations are close to each other. The map was constructed using ggmap package (v.3.0.0).

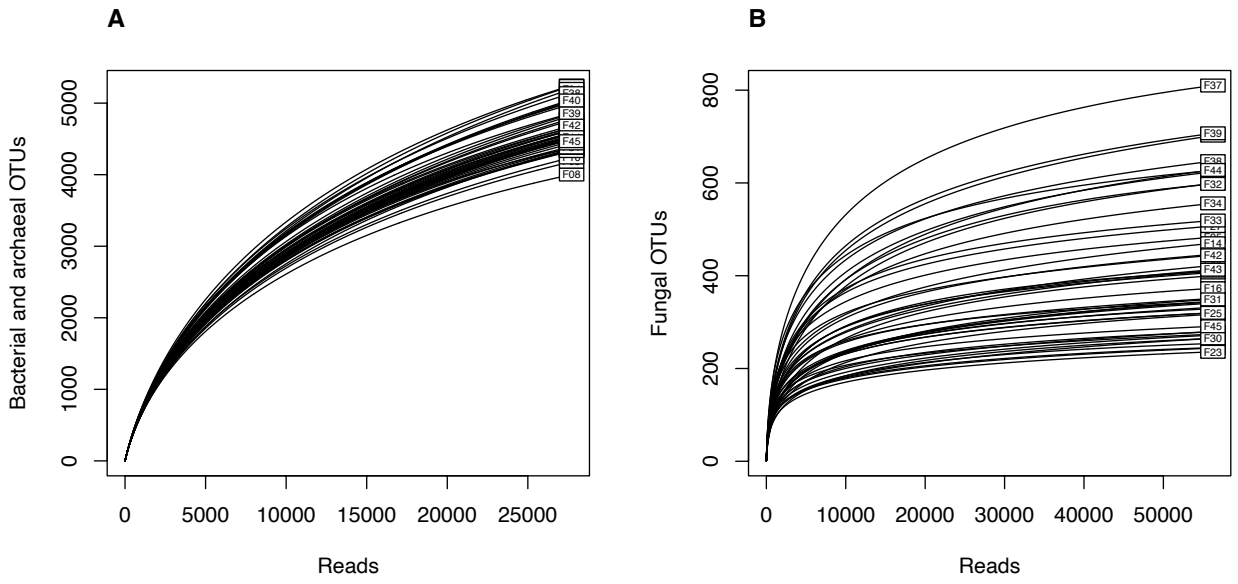

Figure 2.7. Rarefaction curves.

Rarefaction curves of bacteria/archaea (A) and fungi (B) from 45 soil samples (marked) at 97 % of clustering threshold were constructed by plotting the OTU number to the sequence (read) number. The rarefaction curves were constructed using vegan package (v2.5-4).

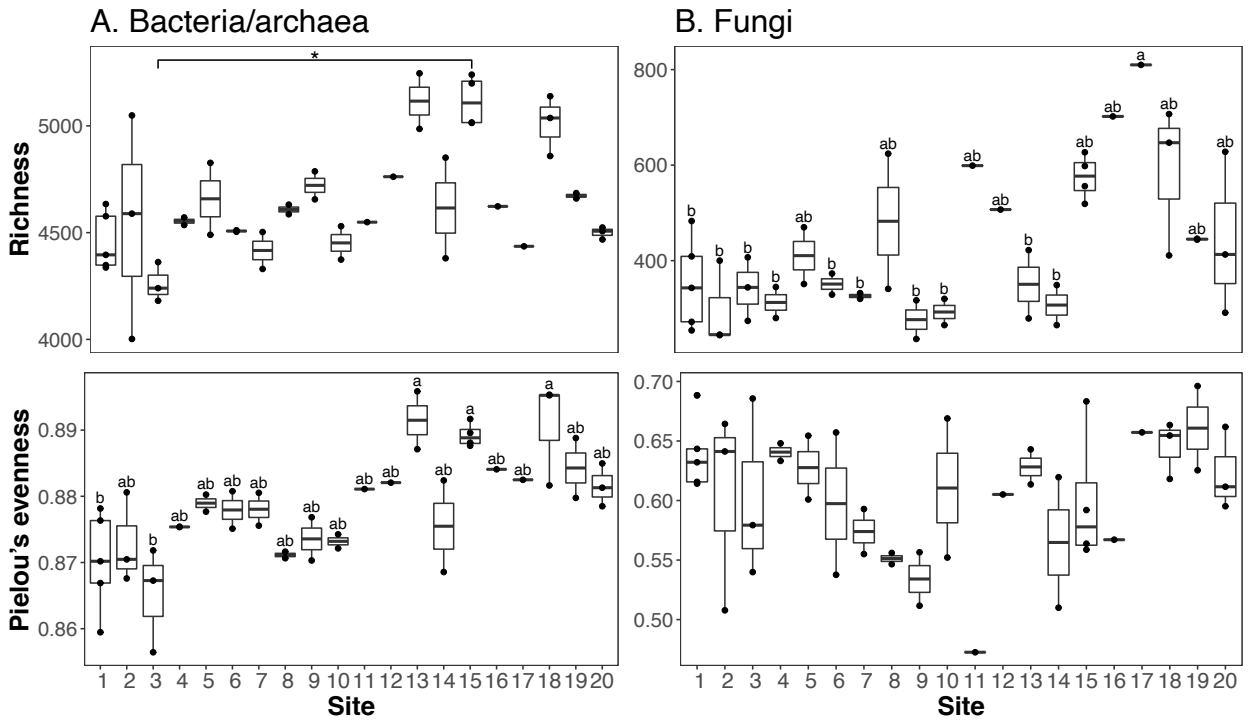

Figure 2.8. Alpha diversity of apple root zone microbiome among sites.

Alpha diversity metrics of bacteria/archaea and fungi: richness (OTU number, clustered at 97 % identity threshold) and Pielou's evenness among sites. For each box plot, circles represent measurement for each sample. The central horizontal lines represent the mean of measurements. Asterisks indicated significant differences between two sites (site 3 and 15) based on post-hoc Dunn's test multiple comparison with Benjamini Hochberg correction. Boxes labelled with different letters were identified as significantly different based on Tukey's HSD post-hoc test. Boxes without label were not significantly different (ANOVA, *p-val* > 0.05).

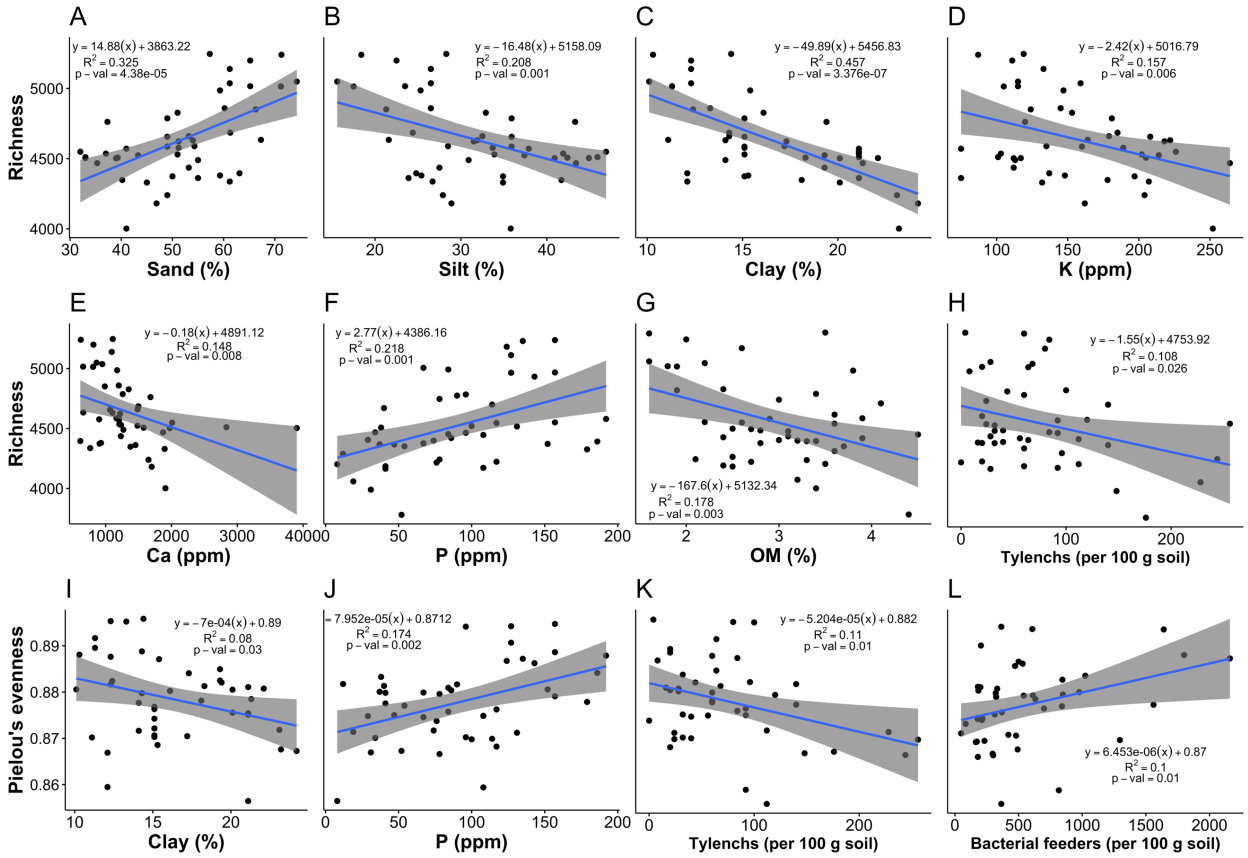

Figure 2.9. The linear regression relationship between bacterial/archaeal alpha diversity and soil parameters and nematodes.

Bacterial/archaeal alpha diversity (richness and Pielou's evenness index) as dependent variables and soil parameters and nematodes as independent variables.

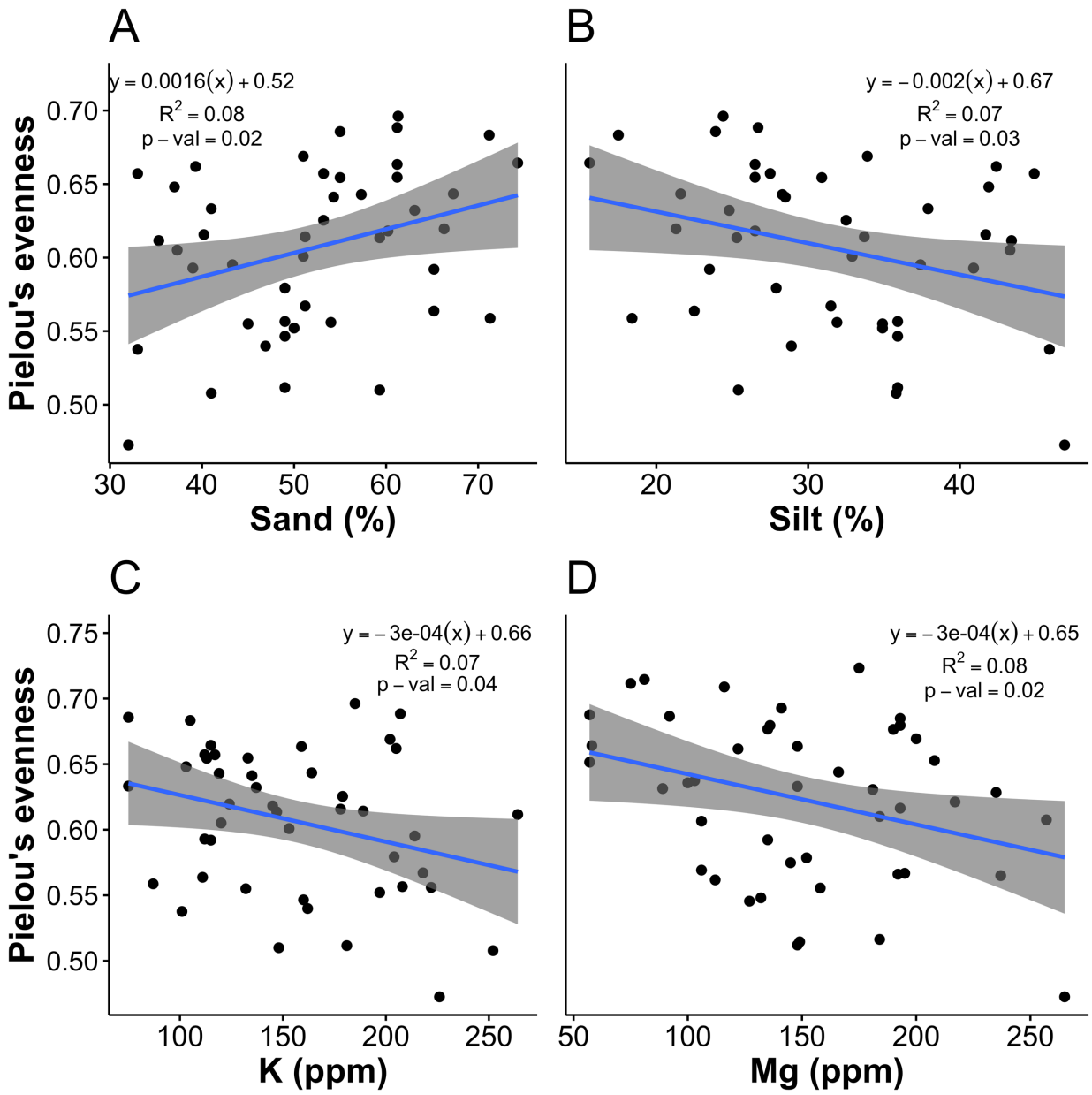

Figure 2.10. The linear regression relationship between fungal alpha diversity and soil parameters.

Fungal alpha diversity (Pielou's evenness index) as dependent variables and soil parameters as independent variables.

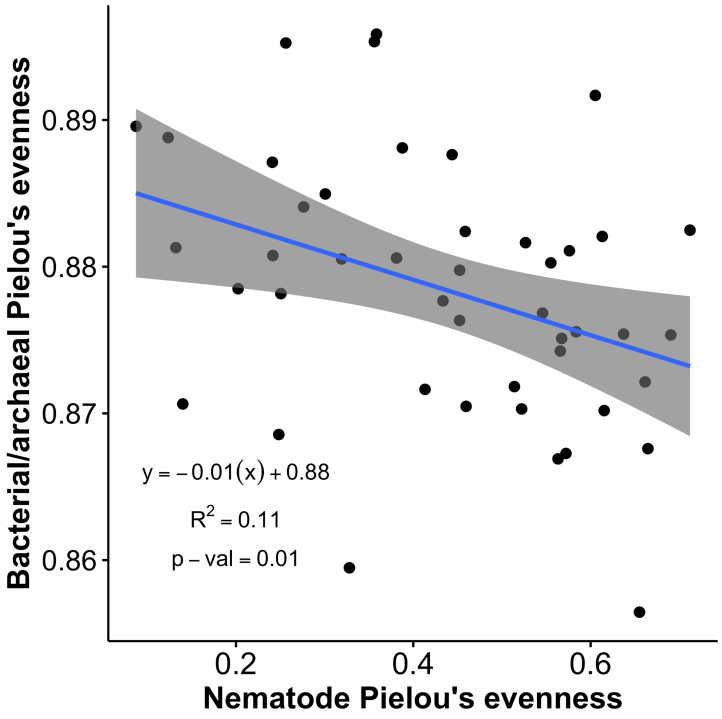

Figure 2.11. The linear regression relationship between bacterial/archaeal and nematodes alpha diversity.

Bacterial and archaeal Pielou's evenness and nematodes Pielou's evenness as dependent and explanatory variables, respectively.

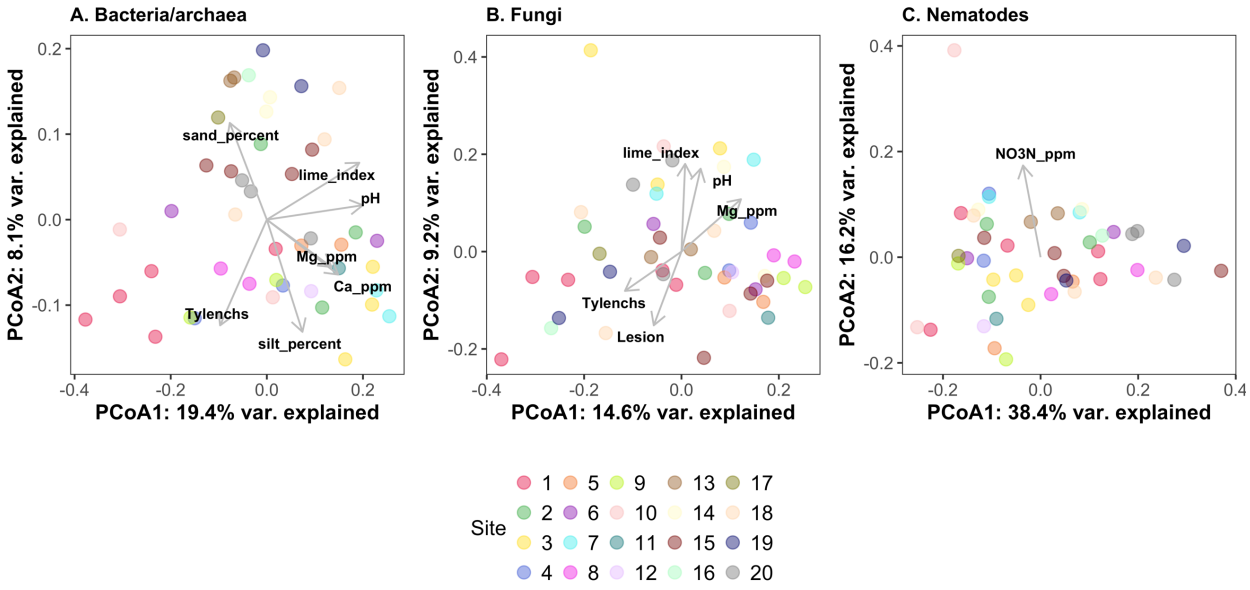

Figure 2.12. PCoA plot of apple root zone microbiome among sites.

Principal coordinate analysis (PCoA) plot based on Bray-Curtis dissimilarities of bacterial/archaeal and fungal OTUs and nematodes square root-transformed count data of 45 soil samples. The color represents 20 different sites. The environmental variables and nematodes that significantly correlated with the microbial community structure are indicated by the arrows (*p*-val < 0.05).

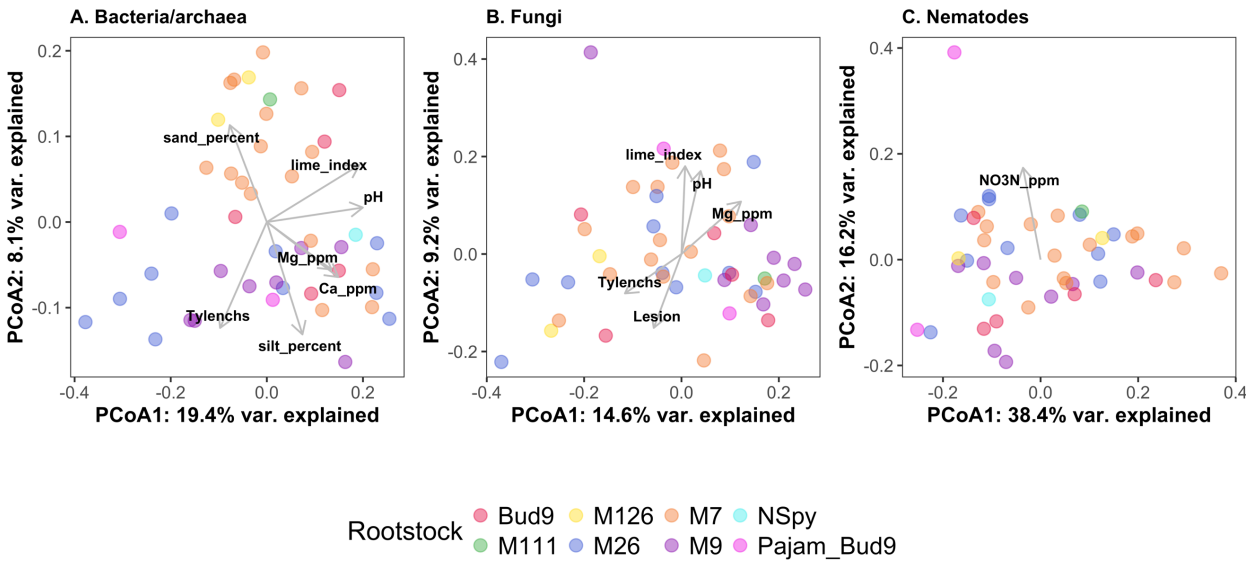

Figure 2.13. PCoA plot of apple root zone microbiome among rootstocks.

Principal coordinate analysis (PCoA) plot based on Bray-Curtis dissimilarities of bacterial/archaeal and fungal OTUs and nematodes square root-transformed count data. The samples were plotted and grouped based on rootstocks as illustrated by the colored circles on the plot. The environmental variables and nematodes that significantly correlated with the microbial community structure are indicated by the arrows (p-val < 0.05).

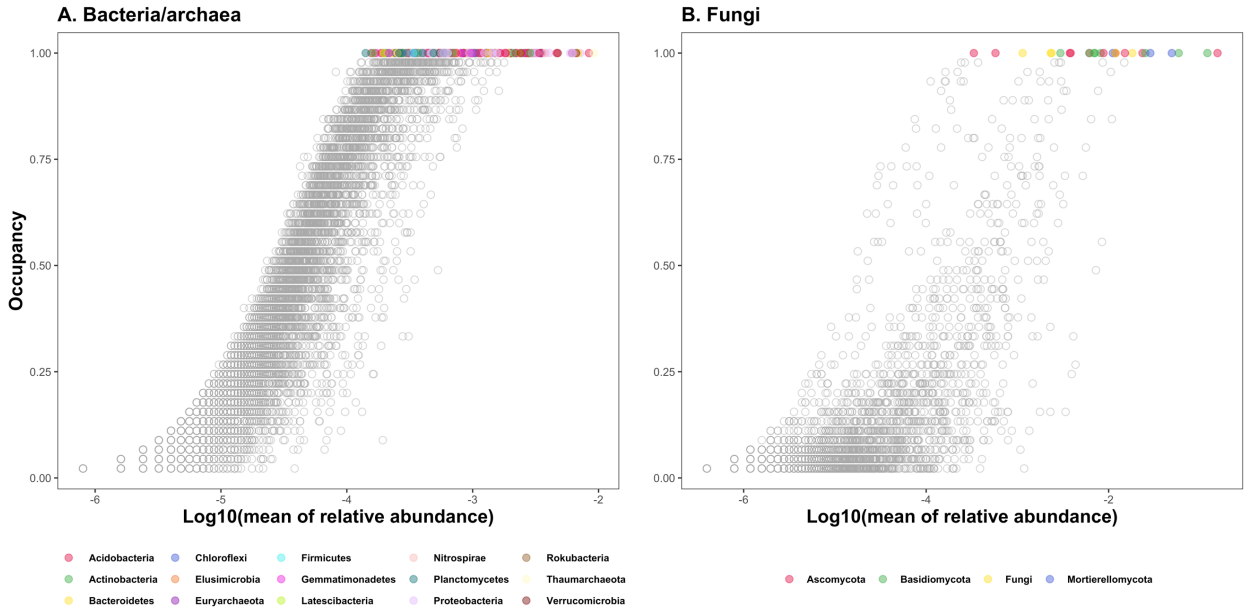

Figure 2.14. Occupancy vs. abundance plots of apple root zone microbiome

Occupancy vs. abundance plots of bacteria and archaea (A) and fungi (B). Each point represents OTU. Core microbiomes, the OTUs with occupancy of 1, are marked based on phylum.

APPENDIX C: Supplemental Information

Supplemental Information

Discussion of the core members of the apple root zone

Identification of core microbiomes provides us a valuable information of key player of microbial community in ecological niches. Even though each orchard site had different environmental condition and soil properties, there were microbial taxa that were prevalent in all soil samples. Core taxa shared among different orchard sites or rootstock genotypes are hypothesized to have important roles in plant-microbe and/or microbe-microbe interactions (1). Here, we highlight some of the major groups we detected among the core in the apple root zones in Michigan, and what is known about their relationships with apple specifically or with soil more generally.

I. Archaea and bacteria

The most abundant core taxon (though, still less than 1% relative abundance) identified in the study was uncultured archaeon of family Nitrososphaeraceae, phylum Thaumarchaeota. This archaeal taxa belongs to ammonia-oxidizing archaea (AOA) that commonly found in soil in high abundance and has important role in Nitrogen (N) cycling (2). Our study found that this archaeal core taxa dominated over bacterial core taxa. Even though AOA typically identified in high abundance in aquatic ecosystem (3, 4), previous study reported that AOA were more abundant than ammonia-oxidizing bacteria (AOB) in all pristine and agricultural soil samples and it suggested that AOA may represent the most abundant ammonia-oxidizing microbes in soil (5). Bacterial core taxa that play important roles for N cycling identified in this study such as N-fixers, *Bradyrhizobium*; nitrite-oxidizers, *Nitrospira*; ammonia-oxidizers, Nitrosomonadaceae; and nitrate-reducers, *Opitutus*.

We found Acidobacteria subgroup 6 included in core taxa and Navarrete et al. (6) described this group as one of the most abundant Acidobacteria in soil. As mentioned above that soil characteristics strongly impact microbial communities, the abundance of Acidobacteria in soil may regulate by soil pH (7). In this study, we found that soil pH had explanatory value to the microbial communities in apple root zone and the differences of Acidobacteria abundance among sites may reflect the differences of soil pH. Family Chitinophagaceae had been found in high abundance among bacterial core taxa. Comparison of bacterial abundance between rhizosphere soil of healthy and putative replant disease apple trees revealed that Acidobacteria and Chitinophagaceae were present in significantly greater abundance of rhizosphere soil from healthy trees (8). Thus, these bacterial groups may have positive roles in maintaining healthy soil.

Member of genus *Pseudomonas* are commonly associated with plants and have various relationships from antagonistic to beneficial (9-11). Specific to apple and multi-trophic interactions, a recent study showed that *Pseudomonas* had the ability to reduce plant parasitic nematode abundance, *Pratylenchus penetrans*, in apple seedlings (12). *Pseudomonas* and Burkholderiaceae belong to Gammaproteobacteria which is a class of core bacteria with the highest number of taxa in our results.

There were several bacterial lineages in the core about which we have less knowledge. For example, Candidatus *Udaeobacter* belongs to phylum Verrucomicrobia and it is one of the most abundant taxa found in soil (13) as well as in this study. We also found a potential antagonist, *Arthrobacter*, that had been reported present in greater abundance in rhizosphere of apple orchard under a replanting system rather than a perennial system (14). Finally, we also detected Candidatus *Xiphinematobacter* in bacterial core taxa. Candidatus *Xiphinematobacter* is

bacterial endosymbiont that closely associated with plant-parasitic dagger nematode, *Xiphinema americanum* (15).This suggests the presence of the nematode in the apple root zone. Several studies found the *X. americanum* nematode to be abundant in apple orchards and can cause root necrosis and suppress the growth of young trees (16-18).

II. Fungi

There were two species of *Fusarium*, *F. oxysporum* (mean relative abundance of 0.88%) and *Fusarium* sp. (mean relative abundance of 0.03%), that were present in all soil samples and belong to fungal core taxa. The *Fusarium* genus is diverse and contains species that range from highly pathogen to beneficial for plant growth (19), and some members of *Fusarium* have been associated with apple replant disease or found in abundance in soils replanted with young apple trees (19-21). While the functional role of *Fusarium* for apple trees remains unclear, the consistent detection of *Fusarium* in apple orchards (21, 22) and all root zone soils included in our study indicates that several *Fusarium* are likely core member.

We also detected *Tetracladium*, *Solicoccozyma*, Cystofilobasidiales, and *Mortierella* in high abundance among all fungal core taxa. Similar with *Fusarium*, *Mortierella* is composed of numerous species with wide range of type of association with plant, thus there is more to learn about their relationship with the apple tree.

III. Nematodes

We assessed nematodes communities in apple rhizosphere. We identified groups of nematodes that represented various dietary preferences, including herbivores, fungal feeders, bacterial feeders, omnivores, and carnivores. A group of nematodes, Rhabditidae, had the highest abundance and occupancy across soil samples, followed by fungal feeders. As bacteria are the most abundant soil microorganisms, nematodes that feed on bacteria are also commonly found in

soil. The study on the functional diversity of nematodes also showed that bacterial feeders were the most abundant group in the Fynbos, South Africa (23). *Tylenchus* and *Aphelenchus* are fungal feeders and they have been found in high abundance in cropping system soils (24). Several species of *Tylenchus* had been isolated and identified from different apple orchards in Europe (25). Dagger and Lesion are plant-parasitic nematodes that can cause plant diseases and we reported that they also had high distribution across apple rhizosphere soil samples. We found that nitrate nitrogen (NO_3N) had explanatory value for the nematode communities. This finding agrees with a previous study that reported that free-living nematodes such as bacterial and fungal feeders abundances greatly correlated with soil N and P, respectively, in the field (26).

Supplemental References

1. Eyre AW, Wang M, Oh Y, Dean RA. 2019. Identification and Characterization of the Core Rice Seed Microbiome. Phytobiomes Journal 3:148-157.

2. Zhalnina K, Dörr de Quadros P, AO Camargo F, Triplett EW. 2012. Drivers of archaeal ammonia-oxidizing communities in soil. Frontiers in microbiology 3:210.

3. Herrmann M, Saunders AM, Schramm A. 2008. Archaea dominate the ammonia-oxidizing community in the rhizosphere of the freshwater macrophyte Littorella uniflora. Applied and Environmental Microbiology 74:3279-3283.

4. Wuchter C, Abbas B, Coolen MJ, Herfort L, van Bleijswijk J, Timmers P, Strous M, Teira E, Herndl GJ, Middelburg JJ. 2006. Archaeal nitrification in the ocean. Proceedings of the National Academy of Sciences 103:12317-12322.

5. Leininger S, Urich T, Schloter M, Schwark L, Qi J, Nicol GW, Prosser JI, Schuster S, Schleper C. 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442:806-809.

6. Navarrete AA, Venturini AM, Meyer KM, Klein AM, Tiedje JM, Bohannan BJ, Nüsslein K, Tsai SM, Rodrigues JL. 2015. Differential response of Acidobacteria subgroups to forest-to-pasture conversion and their biogeographic patterns in the western Brazilian Amazon. Frontiers in microbiology 6:1443.

7. Kuramae EE, Yergeau E, Wong LC, Pijl AS, van Veen JA, Kowalchuk GA. 2012. Soil characteristics more strongly influence soil bacterial communities than land-use type. FEMS Microbiology Ecology 79:12-24.

8. Tilston EL, Deakin G, Bennett J, Passey T, Harrison N, O'Brien F, Fernández-Fernández F, Xu X. 2018. Candidate causal organisms for apple replant disease in the United Kingdom. Phytobiomes Journal 2:261-274.

9. Ganeshan G, Manoj Kumar A. 2005. Pseudomonas fluorescens, a potential bacterial antagonist to control plant diseases. Journal of Plant Interactions 1:123-134.

10. Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JH, Piceno YM, DeSantis TZ, Andersen GL, Bakker PA, Raaijmakers JM. 2011. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. Science 332:1097-100.

11. Yang J-i, Ruegger PM, McKenry MV, Becker JO, Borneman J. 2012. Correlations between Root-Associated Microorganisms and Peach Replant Disease Symptoms in a California Soil. PLOS ONE 7:e46420.

12. Watson TT, Forge TA, Nelson LM. 2018. Pseudomonads contribute to regulation of Pratylenchus penetrans (Nematoda) populations on apple. Canadian journal of microbiology 64:775-785.

13. Brewer TE, Handley KM, Carini P, Gilbert JA, Fierer N. 2016. Genome reduction in an abundant and ubiquitous soil bacterium 'Candidatus Udaeobacter copiosus'. Nature microbiology 2:1-7.

14. Jiang J, Song Z, Yang X, Mao Z, Nie X, Guo H, Peng X. 2017. Microbial community analysis of apple rhizosphere around Bohai Gulf. Scientific reports 7:1-9.

15. Brown AM, Howe DK, Wasala SK, Peetz AB, Zasada IA, Denver DR. 2015. Comparative genomics of a plant-parasitic nematode endosymbiont suggest a role in nutritional symbiosis. Genome biology and evolution 7:2727-2746.

16. Georgi LL. 1988. Effect of three plant species on population densities of Xiphinema americanum and X. rivesi. Journal of Nematology 20:474.

17. Jaffee B, Harrison M, Shaffer R, Strang M. 1987. Seasonal population fluctuation of Xiphinema americanum and X. rivesi in New York and Pennsylvania orchards. Journal of Nematology 19:369.

18. Mai W, Abawi G. 1981. Controlling replant diseases of pome and stone fruits in Northeastern United States by preplant fumigation. Plant Diseases 65:859-864.

19. Tewoldemedhin YT, Mazzola M, Botha WJ, Spies CF, McLeod A. 2011. Characterization of fungi (Fusarium and Rhizoctonia) and oomycetes (Phytophthora and Pythium) associated with apple orchards in South Africa. European Journal of Plant Pathology 130:215-229.

20. Franke-Whittle IH, Manici LM, Insam H, Stres B. 2015. Rhizosphere bacteria and fungi associated with plant growth in soils of three replanted apple orchards. Plant and Soil 395:317-333.

21. Wang G, Yin C, Pan F, Wang X, Xiang L, Wang Y, Wang J, Tian C, Chen J, Mao Z. 2018. Analysis of the Fungal Community in Apple Replanted Soil Around Bohai Gulf. Horticultural Plant Journal 4:175-181.

22. Kelderer M, Manici LM, Caputo F, Thalheimer M. 2012. Planting in the 'inter-row'to overcome replant disease in apple orchards: a study on the effectiveness of the practice based on microbial indicators. Plant and Soil 357:381-393.

23. Moroenyane I, Dong K, Singh D, Chimphango S, Adams J. 2016. Deterministic processes dominate nematode community structure in the Fynbos Mediterranean heathland of South Africa. Evolutionary Ecology 30:685-701.

24. Treonis AM, Unangst SK, Kepler RM, Buyer JS, Cavigelli MA, Mirsky SB, Maul JE. 2018. Characterization of soil nematode communities in three cropping systems through morphological and DNA metabarcoding approaches. Scientific Reports 8:2004.

25. Bert W, Claeys M, Borgonie G. 2006. The comparative cellular architecture of the female gonoduct among tylenchoidea (nematoda: tylenchina). Journal of nematology 38:362-375.

26. Ferris H, Griffiths BS, Porazinska DL, Powers TO, Wang K-H, Tenuta M. 2012. Reflections on plant and soil nematode ecology: past, present and future. Journal of nematology 44:115-126.

REFERENCES

REFERENCES

1. Sugiyama A, Bakker MG, Badri DV, Manter DK, Vivanco JM. 2013. Relationships between Arabidopsis genotype-specific biomass accumulation and associated soil microbial communities. Botany 91:123-126.

2. Weidner S, Koller R, Latz E, Kowalchuk G, Bonkowski M, Scheu S, Jousset A. 2015. Bacterial diversity amplifies nutrient-based plant–soil feedbacks. Functional Ecology 29:1341-1349.

3. Panke-Buisse K, Poole AC, Goodrich JK, Ley RE, Kao-Kniffin J. 2015. Selection on soil microbiomes reveals reproducible impacts on plant function. The ISME Journal 9:980-989.

4. Meena KK, Sorty AM, Bitla UM, Choudhary K, Gupta P, Pareek A, Singh DP, Prabha R, Sahu PK, Gupta VK, Singh HB, Krishanani KK, Minhas PS. 2017. Abiotic Stress Responses and Microbe-Mediated Mitigation in Plants: The Omics Strategies. Frontiers in Plant Science 8.

5. Chapelle E, Mendes R, Bakker PAH, Raaijmakers JM. 2016. Fungal invasion of the rhizosphere microbiome. The ISME journal 10:265-268.

6. Thakur MP, Geisen S. 2019. Trophic regulations of the soil microbiome. Trends in microbiology 27:771-780.

7. Wei Z, Yang T, Friman V-P, Xu Y, Shen Q, Jousset A. 2015. Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. Nature communications 6:1-9.

8. Deakin G, Tilston EL, Bennett J, Passey T, Harrison N, Fernández-Fernández F, Xu X. 2018. Spatial structuring of soil microbial communities in commercial apple orchards. Applied Soil Ecology 130:1-12.

9. Jiang J, Song Z, Yang X, Mao Z, Nie X, Guo H, Peng X. 2017. Microbial community analysis of apple rhizosphere around Bohai Gulf. Scientific reports 7:1-9.

10. USDA-NAAS. 2019. Noncitrus fruits and nuts June 2018 summary. https://www.nass.usda.gov/Publications/Todays_Reports/reports/ncit0619.pdf. Accessed September 30, 2019.

11. Jenkins WR. 1964. A rapid centrifugal-flotation technique for separating nematodes from soil. Plant disease reporter 48.

12. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences 108:4516-4522.

13. Smith DP, Peay KG. 2014. Sequence Depth, Not PCR Replication, Improves Ecological Inference from Next Generation DNA Sequencing. PLOS ONE 9:e90234.

14. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10:996-998.

15. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335-336.

16. Edgar RC, Flyvbjerg H. 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31:3476-82.

17. Lee S-H, Sorensen JW, Grady KL, Tobin TC, Shade A. 2017. Divergent extremes but convergent recovery of bacterial and archaeal soil communities to an ongoing subterranean coal mine fire. The ISME Journal 11:1447-1459.

18. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590-6.

19. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460-1.

20. Gihring TM, Green SJ, Schadt CW. 2012. Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. Environmental Microbiology 14:285-290.

21. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R. 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5:27.

22. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17:3.

23. Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Põldmaa K, Saag L, Saar I, Schüßler A, Scott JA, Senés C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiss M, Larsson K-H. 2013. Towards a unified paradigm for sequence-based identification of fungi. Molecular Ecology 22:5271-5277.

24. Gdanetz K, Benucci GMN, Vande Pol N, Bonito G. 2017. CONSTAX: a tool for improved taxonomic resolution of environmental fungal ITS sequences. BMC Bioinformatics 18:538.

25. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42:D633-42.

26. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology 73:5261-5267.

27. Edgar RC. 2016. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. bioRxiv doi:10.1101/074161:074161.

28. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.

29. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2019. vegan: Community Ecology Package. R package version 2.5-6.

30. Guillot G, Rousset F. 2013. Dismantling the Mantel tests. Methods in Ecology and Evolution 4:336-344.

31. Oksanen J. 2015. Multivariate analysis of ecological communities in R: vegan tutorial R Package Version 2.5-4. https://wwwmooreecologycom/uploads/2/4/2/1/24213970/vegantutorpdf.

32. Grady KL, Sorensen JW, Stopnisek N, Guittar J, Shade A. 2019. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. Nature communications 10:1-10.

33. Deng Y, Jiang Y-H, Yang Y, He Z, Luo F, Zhou J. 2012. Molecular ecological network analyses. BMC bioinformatics 13:1-20.

34. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 13:2498-2504.

35. Franke-Whittle IH, Manici LM, Insam H, Stres B. 2015. Rhizosphere bacteria and fungi associated with plant growth in soils of three replanted apple orchards. Plant and Soil 395:317-333.

36. Singh J, Silva KJP, Fuchs M, Khan A. 2019. Potential role of weather, soil and plant microbial communities in rapid decline of apple trees. PLoS One 14:e0213293.

37. Wang G, Yin C, Pan F, Wang X, Xiang L, Wang Y, Wang J, Tian C, Chen J, Mao Z. 2018. Analysis of the Fungal Community in Apple Replanted Soil Around Bohai Gulf. Horticultural Plant Journal 4:175-181.

38. Zhang Q, Sun J, Liu S, Wei Q. 2013. Manure refinement affects apple rhizosphere bacterial community structure: a study in sandy soil. PloS one 8:e76937.

39. Rumberger A, Yao S, Merwin IA, Nelson EB, Thies JE. 2004. Rootstock genotype and orchard replant position rather than soil fumigation or compost amendment determine tree growth and rhizosphere bacterial community composition in an apple replant soil. Plant and Soil 264:247-260.

40. Yao S, Merwin IA, Abawi GS, Thies JE. 2006. Soil fumigation and compost amendment alter soil microbial community composition but do not improve tree growth or yield in an apple replant site. Soil Biology and Biochemistry 38:587-599.

41. St. Laurent A, Merwin IA, Fazio G, Thies JE, Brown MG. 2010. Rootstock genotype succession influences apple replant disease and root-zone microbial community composition in an orchard soil. Plant and Soil 337:259-272.

42. Marasco R, Rolli E, Fusi M, Michoud G, Daffonchio D. 2018. Grapevine rootstocks shape underground bacterial microbiome and networking but not potential functionality. Microbiome 6:1-17.

43. Poudel R, Jumpponen A, Kennelly MM, Rivard CL, Gomez-Montano L, Garrett KA. 2019. Rootstocks shape the rhizobiome: rhizosphere and endosphere bacterial communities in the grafted tomato system. Applied and Environmental Microbiology 85:e01765-18.

44. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, Cho H, Karaoz U, Loqué D, Bowen BP. 2018. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nature microbiology 3:470-480.

45. Weiß S, Bartsch M, Winkelmann T. 2017. Transcriptomic analysis of molecular responses in Malus domestica 'M26'roots affected by apple replant disease. Plant molecular biology 94:303-318.

46. Leisso R, Rudell D, Mazzola M. 2017. Metabolic composition of apple rootstock rhizodeposits differs in a genotype-specific manner and affects growth of subsequent plantings. Soil Biology and Biochemistry 113:201-214.

47. Yim B, Winkelmann T, Ding G-C, Smalla K. 2015. Different bacterial communities in heat and gamma irradiation treated replant disease soils revealed by 16S rRNA gene analysis–contribution to improved aboveground apple plant growth? Frontiers in microbiology 6:1224.

48. Liu J, Abdelfattah A, Norelli J, Burchard E, Schena L, Droby S, Wisniewski M. 2018. Apple endophytic microbiota of different rootstock/scion combinations suggests a genotype-specific influence. Microbiome 6:1-11.

49. Shade A, Stopnisek N. 2019. Abundance-occupancy distributions to prioritize plant core microbiome membership. Current Opinion in Microbiology 49:50-58.

50. Mercado-Blanco J, Abrantes I, Barra Caracciolo A, Bevivino A, Ciancio A, Grenni P, Hrynkiewicz K, Kredics L, Proença DN. 2018. Belowground microbiota and the health of tree crops. Frontiers in microbiology 9:1006.

51. Si P, Shao W, Yu H, Yang X, Gao D, Qiao X, Wang Z, Wu G. 2018. Rhizosphere microenvironments of eight common deciduous fruit trees were shaped by microbes in Northern China. Frontiers in Microbiology 9:3147.

52. Hargreaves SK, Williams RJ, Hofmockel KS. 2015. Environmental filtering of microbial communities in agricultural soil shifts with crop growth. PLoS One 10:e0134345.

53. Faust K, Raes J. 2012. Microbial interactions: from networks to models. Nature Reviews Microbiology 10:538-550.

54. Hirano H, Takemoto K. 2019. Difficulty in inferring microbial community structure based on co-occurrence network approaches. BMC bioinformatics 20:1-14.

CHAPTER 3: Endophytic microbiome variation among single plant seeds

Work presented in the chapter has been published as

Bintarti AF, Sulesky-Grieb A, Stopnisek N, Shade A. Endophytic Microbiome Variation Among Single Plant Seeds. *Phytobiomes Journal* **6**, 45-55 (2022).

Abstract

Like other plant compartments, the seed harbors a microbiome. The members of the seed microbiome are the first to colonize a germinating seedling, and they may initiate the trajectory of microbiome assembly for the next plant generation. Therefore, the members of the seed microbiome are important for the dynamics of plant microbiome assembly and the vertical transmission of potentially beneficial symbionts. However, it remains challenging to assess the microbiome at the individual seed level (and, therefore, for the future individual plants) due to low endophytic microbial biomass, seed exudates that can select for particular members, and high plant and plastid contamination of resulting reads. Here, we report a protocol for extracting microbial DNA from an individual seed (common bean, *Phaseolus vulgaris*) with minimal disruption of host tissue, which we expect to be generalizable to other medium- and large-seed plant species. We applied this protocol to determine the 16S ribosomal RNA (rRNA) V4 and rRNA internal transcribed spacer (ITS)2 amplicon composition and examine the variability of individual seeds harvested from replicate common bean plants grown under standard, controlled conditions to maintain health. Using DNA extractions from individual seeds, we compared seed-to-seed, pod-to-pod, and plant-to-plant microbiomes, and found highest microbiome variability at the plant level. This suggests that several seeds from the same plant could be pooled for microbiome assessment, given experimental designs that apply treatments at the parent plant level. This study adds protocols and insights to the growing toolkit of approaches to understand the plant-microbiome engagements that support the health of agricultural and environmental ecosystems.

Introduction

Seed microbiomes offer a reservoir of microbiota that can be vertically passed from parent plants to offspring (1-3) and some of these members have plant-beneficial phenotypes (4-7). Therefore, the seed microbiome is expected to play a key role in plant health and fitness (8), and especially in the assembly and establishment of the developing plant's microbiome (9). This expected importance of the seed microbiome has fueled recent research activity to use high-throughput sequencing to characterize the seed microbiomes of various plants (e.g., (10-15)).

Seed microbiomes include microbial members that live on the seed surface as epiphytes and members that colonize inside the internal tissues of the seed as endophytes (16). Among these microbiome members, endophytes that closely associate with endosperm and embryo tissues are more likely to be transmitted to the next plant generations than are seed-associated epiphytes (16, 17). By itself, an endophytic association does not confirm that there is a functional benefit or co-evolutionary relationship between the plant and the microbiome member (16). However, endophytic microbes offer the first source of inoculum for the germinating seedling (as reviewed in (16); (18)), and given the potential for priority effects or pathogen exclusion, these members can have implications for the mature plant's microbial community composition or structure. Therefore, understanding the endophytic seed microbiome is expected to provide insights into how seeds can facilitate microbiome assembly and the vertical transmission of microbiome members over plant generations.

As is true for other plant compartments, different plant species or divergent crop lines, varieties, or cultivars often have different seed microbiome composition (taxonomic identities of members) or structure (relative contributions of taxa to the community) (7, 19-21). However, many seed microbiome studies have reported generally high variability across seed samples from

the same plant type and treatment (6, 7, 22), with strong explanatory value of either seed origin or seed lot, geographic region or soil edaphic conditions (10, 20, 21). While these insights may call into question the proportion of "inherited" versus acquired seed microbiome members, the high microbiome variability may be in part due to methods applied to extract the microbial DNA from the seed compartment, and different methods applied across studies. For instance, some studies surface sterilize the seeds while others do not, some germinate the seed prior to microbiome analysis while others do not, and so on. One source of microbiome variability could be the common practice of the pooling of many seeds from the same or different plants to produce a composite seed microbiome sample for DNA extraction. Because multiple seeds are investigated at once, it is unclear at what level the most microbiome variability is highest: the seed, the pod or fruit, the plant, or the field or treatment. This information is required to determine the necessary sample size in well-powered experimental designs. More importantly, the question of vertical transmission cannot directly be addressed without seed microbiome assessment of an individual.

Our study objectives were to (i) determine the appropriate observational unit of endophytic seed microbiome assessment for common bean (*Phaseolus vulgaris* L.) by examining seed-to-seed, pod-to-pod, and plant-to-plant variability in 16S ribosomal RNA (rRNA) V4 and rRNA internal transcribed spacer (ITS)2 amplicon analyses; and (ii) develop a robust protocol for individual seed microbiome extraction that could be generally applied to other plants that have similarly medium- to large- sized seeds. Here, we use a working definition of seed endophyte as the microbes internal to the ungerminated seed, including under the seed coat and within the internal compartments (cotyledon, radical, hypocotyl, plumule), but excluding those on the surface of the seed coat. Our rationale for applying this working definition is to

distinguish microbes that are more likely acquired via the parent plant from those that may have been acquired via the seed surface contact with the environment. We found that plant-to-plant variability under controlled growth conditions exceeded within-plant variability among different pods and conclude that seeds can be pooled by parent plant (but not across different plants) in study designs that aim to compare seed microbiomes resulting from treatments applied at the level of the individual plant (e.g., the experimental unit is one plant).

Materials and Methods

Growth conditions for parent plants

We used common bean *P. vulgaris* L. 'Red Hawk', a dark red kidney bean developed at Michigan State University (23) which belongs to the Andean lineage (24). The seeds used to grow the parental plants originated from Michigan State University's Agronomy Farm located in East Lansing, MI, U.S.A., and were harvested following standard agricultural practices.

Because we targeted the endophytic seed microbiome, surface sterilization of the bean seeds was conducted before germination and planting. To sterilize, seeds were soaked in a solution of 10% bleach with 0.1% Tween 20 for 15 minutes, then rinsed four times with sterile water. The final rinse water was plated on tryptic soy agar and potato dextrose agar plates to test for sterilization efficacy. Sterilized seeds were placed in Petri dishes on sterile tissue paper moistened with sterile water, and allowed to germinate in the dark for 4 days. After 4 days, the radicle had emerged and the germinated seeds were transferred to the growth chamber. The germinated seeds were planted in three 4.54-liter (1-gal.) pots filled with a 50:50 (vol/vol) mixture of agricultural bean field soil and vermiculite. The pots were placed in a BioChambers model SPC-37 growth chamber with a cycle of 14 h/day and 10 h/night cycle at 26°C and 22°C,

respectively, 260 mE light intensity, and 50% relative humidity. All plants received 300 ml of water every other day and 200 ml of half-strength Hoagland solution (25) once a week.

Study design

We planted three germinated seeds per pot and culled to one seedling per pot at the early vegetative growth stage. There were three plant replicates designated as A, B, and C, grown under the above-described conditions for normal, healthy growth. The three plants yielded different numbers of pods and seeds, and we aimed to balance and maximize the number of seeds used for analysis across plants (**Table 3.1**).

Seed harvest and endophyte microbial DNA extraction

Once the plants reached maturity at the R9 growth stage (yellowing leaves and dry pods), the seeds were harvested for endophytic microbiome analysis. Seeds were distinguished by plant and pod. The endophytic microbiome of each seed was extracted and sequenced individually. To extract the endophytic microbial DNA, a protocol was adapted from Barret et al. (8) and Rezki et al. (26). First, the seeds were surface-sterilized as above and the seed coat was carefully removed using sterilized forceps. Each seed was then soaked in 3 ml of phosphate-buffered saline solution with 0.05% Tween 20 (hereafter, "soaking solution") overnight at 4°C with constant agitation of 170 rpm. Because low levels of microbial biomass are expected in single seed extractions, positive and negative controls were included in the extraction protocol. This ensures that, if no extractable microbial DNA is present in a sample that it is representative of the sample, rather than the extraction methods. A mock community was used as a DNA extraction positive control by adding one, 75-µl aliquot of the ZymoBIOMICS Microbial Community

Standard (Zymo Research, Irvine, CA, U.S.A.) to 3 ml of the soaking solution immediately prior to conducting the extraction protocol. Sterile soaking solution (3 ml) was used as a negative DNA extraction control.

After soaking overnight, the samples were centrifuged at 4,500 x *g* for 60 min at 4°C to pellet any material that had been released from the seed tissues. After centrifugation, the seed was removed, and the pelleted material was resuspended in 1 to 2 ml of supernatant (soaking solution) and transferred to a microcentrifuge tube for DNA extraction using the E.Z.N.A Bacterial DNA Kit (Omega Bio-tek, Inc. Norcross, GA, U.S.A.). The manufacturer's centrifugation protocol was used with minor modifications. Specifically, the pelleted seed material was suspended in Tris-EDTA buffer (step 4), the incubation for the lysozyme step was extended to 20 min, 30 µl of elution buffer was used, and the elution step was extended to a 15-min incubation. These modifications were performed to maximally recover the limited amount of microbial DNA expected from a single seed. We detail the standard operating protocol and provide notes on the alternatives that we tested in optimizing this protocol in the Supplementary Material.

Table 3.1. Parent plant yield information and seed samples used in microbiome analyses

Plant	Number produced[a]		Sequencing samples[b]	
	Pods	Seed	Pods	Seeds
A	5	22	A1	4[c]
			A2	4
			A3	4
B	6	29	B1	4[c]
			B2	4
			B3	4
			B4	4
			B5	4
			B6	4
C	7	26	C5	3
			C6	4
			C7	4

[a] Number of seeds per pod varied from 2 to 6.

[b] Sequencing samples for 16S ribosomal RNA (rRNA) V4 and rRNA internal transcribed spacer (ITS)2 (n = 47 for bacteria or archaea and n = 45 for fungi) were grouped by plant.

[c] Unable to amplify rRNA ITS2 target DNA in one (of the four total) seed samples.

PCR amplification and amplicon sequencing

To confirm successful DNA extraction from the seed pellet, DNA quantification and target gene PCR assays were performed. First, the DNA extracted from the seed samples and the positive and negative controls were quantified using the Qubit dsDNA BR Assay Kit (ThermoFisher Scientific, Waltham, MA, U.S.A.). Then, PCR amplification and sequencing of the V4 region of 16S rRNA bacterial or archaeal gene and the ITS2 region of the fungal rRNA gene were performed. The V4 region of 16S rRNA gene amplification was conducted using 515f (5'-GTGCCAGCMGCCGCGGTAA-3') and 806r (5'-GGACTACHVGGGTWTCTAAT-3') universal primers (27) under the following conditions: 94°C for 3 min, followed by 35 cycles of 94°C for 45 s, 50°C for 60 s, and 72°C for 90 s; with a final extension at 72°C for 10 min. The amplification was performed in 25-µl mixtures containing 12.5 µl GoTaq Green Master Mix (Promega Corp., Madison, WI, U.S.A.), 0.625 µl of each primer (20 µM), 2 µl of DNA template (approximately 1 ng/µl), and 9.25 µl nuclease-free water. The amplicon DNA (concentration of approximately 1 ng/µl) was sequenced at the Research Technology Support Facility (RTSF) Genomics Core, Michigan State sequencing facility using the Illumina MiSeq platform v2 Standard flow cell. The sequencing was performed in a 2-by-250-bp paired-end format.

The PCR amplification of the rRNA ITS2 region was performed using ITS86f (5'-GTGAATCATCGAATCTTTGAA-3′) and ITS4 (5'- TCCTCCGCTTATTGATATGC-3') primers (28) with addition of index adapters by the RTSF Genomics Core. The PCR amplification of the rRNA ITS2 was conducted under the following conditions: 95°C for 2 min, followed by 40 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min; with a final extension at 72°C for 10 min. The amplification was performed in 50-µl mixture containing 20 µl GoTaq Green Master Mix (Promega Corp.), 1 µl of each primer (10 µM), 1 µl of DNA template

(approximately 1 ng/µl), and 27 µl nuclease-free water. The PCR products were purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Purified PCR products with a concentration range 6 to 10 ng/µl were sequenced at the RTSF Genomics Core using Illumina MiSeq platform v2 Standard flow cell and 2-by-250-bp paired-end format.

Sequence analysis

The USEARCH pipeline (v.10.0.240) was used to merge paired-end bacterial/archaeal raw reads, filter for low-quality sequences, dereplicate, remove singletons, denoise, and check for chimeras (29). An in-house open reference strategy was performed for operational taxonomic unit (OTU) clustering (30). First, closed-reference OTU picking was performed by clustering the quality filtered reads against the SILVA database (v.132) (31) at 97% identity using USEARCH algorithm (usearch_global command) (32). Then, a de novo OTU picking process was performed on the reads that failed to match the reference using UPARSE-OTU algorithm (cluster_otus command) (33) at 97% identity. Finally, closed-reference and de novo OTUs were combined into a full set of representative sequences. The merged sequences were then mapped back to the representative sequences using the usearch_global command.

Sequence alignment, taxonomy assignment, non-bacteria/archaea filtering, and phylogenetic diversity calculation were performed using QIIME 1.9.1 (34). The representative sequences were aligned against the SILVA database (v.132) (31) using PyNAST (35). The unaligned OTUs and sequences were excluded from the OTU table and the representative sequences file, respectively. Taxonomy assignment was performed using the default classifier method (UCLUST algorithm) at a minimum confidence of 0.9 (32) using SILVA database (v.132) as the reference. Plant contaminants (chloroplast and mitochondria) and unassigned taxa

were removed from the OTU table and the representative sequences using filter_taxa_from_otu_table.py and filter_fasta.py command (**Appendix C Figure 10**). Filtering the microbial contaminants from the OTU table was conducted in R (v.3.4.2; R Core Development Team) using the microDecon package (36). Reads were normalized using cumulative sum scaling (CSS) method in metagenomeSeq Bioconductor package on R (37).

The fungal ITS raw reads were processed using the USEARCH (v.10.0.240) pipeline. Read processing included merging paired-end reads, removing primers using cutadapt (v.2.1) (38), dereplication, and singleton removal. OTUs were picked and chimeras removed using de novo clustering at 97% identity threshold with the UPARSE-OTU algorithm (cluster_otus command) (33). Then, all merged sequences were mapped to the clustered reads using usearch_global command to generate an OTU table. Fungal taxonomic classification was performed in CONSTAX (39) using RDP Classifier (v.11.5) (40, 41) at a minimum confidence of 0.8 and with the UNITE reference database (release 12 January 2017). Plant and microbial contaminants removal and read normalization were performed in R (v.3.4.2). Plant contaminants were removed from the OTU table by filtering out OTUs that were assigned into Kingdom Plantae (**Figure 3.6 Appendix B**). Microbial contaminants were removed using the microDecon package (36). The CSS method from the metagenomeSeq Bioconductor package was performed to normalize the fungal reads (37).

Microbial community analysis

Microbiome statistical analyses were conducted in R (v.3.4.2) (R Core Development Team). Microbial alpha and beta diversity were calculated on the CSS-normalized OTU table using the vegan package (v.2.5-7) (42). Richness (count of observed OTUs) and Faith's

phylogenetic diversity were used to analyze the bacterial or archaeal alpha diversity. For fungal alpha diversity, we used richness. The evenness of the seed microbiomes was visualized using rank-abundance curves (Phyloseq package v.1.28.0) in R (43). Differences in alpha diversity among plants and pods were determined by fitting the linear mixed-effects model (LMM) using the lme function of the nlme package (v.3.1-152) (44). We performed LMM because the study has an unbalanced nested design with pod as the random factor, nested within plant as the fixed factor. Microbial composition and relative abundance were analyzed using the Phyloseq package (v.1.28.0) in R (43).

Beta diversity was calculated using Jaccard distances and visualized using principal coordinate analysis (PCoA) plot. We used the Jaccard index, which is based on presence-absence (unweighted), rather than a metric based on relativized abundance (weighted) because we reasoned that the seed microbiome members are likely to be dormant inside the seed prior to germination (45), and that any differences in relative abundances are not directly attributable to competitive fitness outcomes inside the seed. Furthermore, exponential growth would allow that any viable cell successfully packaged and passaged via the seed could, in theory, successfully colonize the new plant. Finally, consistent host-selection or enrichment (that may favor some taxa over others) cannot be assessed directly with our experimental design because we do not have data from multiple plant generations. For comparison, we also provide an assessment of beta diversity using the weighted Bray-Curtis dissimilarity (**Figure 3.7 Appendix B**), but caution against over-interpreting abundance-weighted analyses for the reasons listed above.

Nested permutational multivariate analysis of variance (PERMANOVA) using the function nested.npmanova from the BiodiversityR package (46) was performed to assess the microbial community composition and structure among plants and pods. We performed

multivariate analysis to check the homogeneity of dispersion (variance) among groups using the function betadisper (42). We performed PERMDISP to test the significant differences in dispersions between groups and Tukey's honestly significant difference (HSD) test to determine which groups differ in relation to the dispersions (variances).

Power analysis and sample size were calculated using the pwr.t.test function from the pwr package (v.1.3-0). We performed power analysis of two-category *t* test. Because the most microbiome variability was observed across plants, we pooled individual seed sequence profiles in silico at the parent plant level for this analysis. We calculated Cohen's *d* effect size given the information of mean (M) and standard deviation (SD) of bacterial/archaeal alpha diversity (richness and phylogenetic diversity) from three plant samples from this study: plant A (n = 12; richness: M = 30.58, SD = 6.42, phylogenetic diversity: M = 4.17, SD = 0.89), plant B (n = 24; richness: M = 18.21, SD = 7.35, phylogenetic diversity: M = 2.92, SD = 0.82) and plant C (n = 11; richness: M = 19.09, SD = 10.95, phylogenetic diversity: M = 3.09, SD = 1.39). We calculated the common SD (σ pool of all groups) using the above information; then, we calculated Cohen's *d* effect size for both richness and phylogenetic diversity. Cohen's *d* effect size was defined by calculating the difference between the largest and smallest means divided by the square root of the mean square error (or the common SD). Power analysis was run with Hedges's *g* effect size (corrected with Cohen's *d* effect size) and significance level of 0.05.

Here we defined shared microbiome members (sometimes referred to as a "core") as microbial taxa that were shared and detected in all considered samples. Taxon occupancy is the proportion of samples in which the taxa are detected, with an occupancy of 1, meaning that the taxon was detected in all samples (47). We report the taxa that were shared across seeds originating from different plants, and across seeds that originated from the same plant.

Data and code availability

The computational workflows for sequence processing and ecological statistics are available on GitHub (https://github.com/ShadeLab/Bean_seed_variability_Bintarti_2021). Raw sequence data of bacteria/archaea and fungi have been deposited in the Sequence Read Archive (SRA) NCBI database under Bioproject accession number PRJNA714251.

Results

Sequencing summary and microbiome coverage

In total, 5,056,769 16S rRNA V4 and 8,756,009 rRNA ITS2 quality reads were generated from 47 DNA samples isolated from individual seeds for bacteria or archaea, and from 45 samples for fungi. We removed >90% of reads that were plant contaminants (**Figure 3.6 Appendix B**), resulting in 17,128 and 67,878 16S rRNA bacterial or archaeal and rRNA ITS2 fungal reads, respectively. After removing plant and microbial contaminants, we determined 211 bacterial or archaeal and 43 fungal OTUs defined at 97% sequence identity. Although the majority of individual seeds from plants A and B had exhaustive to sufficient sequencing effort, some seeds from plant C did not **(Figure 3.1A)**. However, the fungal rarefaction curves reached asymptotes and had sufficient sequencing depth **(Figure 3.1B)**. Both bacterial or archaeal and fungal seed microbiomes were highly uneven, with few dominant and many rare taxa, as is typical for microbiomes **(Figure 3.1C and 3.1D)**.

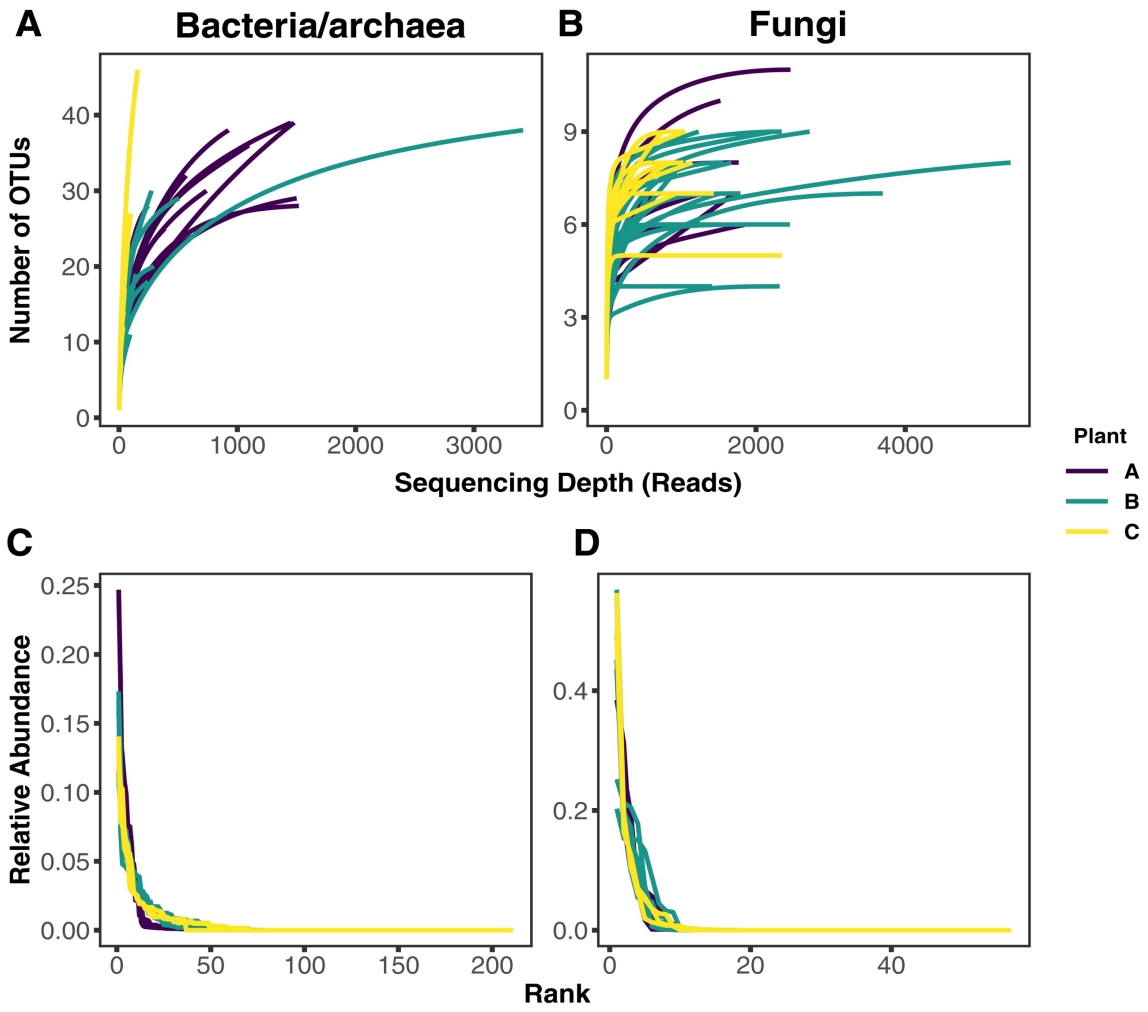

Figure 3.1. Rarefaction curves of common bean seed microbiome.

Rarefaction curves of **A,** bacteria or archaea and **B,** fungi from individual seeds (marked) at 97% of clustering threshold were constructed by plotting the operational taxonomic unit (OTU) number after decontamination (microbial and plant contaminants removed) to the sequence (read) number. Each curve represents microbiome sequence data from microbial DNA extraction from an individual seed. Rarefaction curves were constructed using the vegan package (v2.5-4). Rank abundance curve of decontaminated and normalized **C,** bacterial or archaeal and **D,** fungal

Figure 3.1 (cont'd)

OTU tables. Samples (n = 47 and n = 45 for bacteria or archaea and fungi, respectively) were grouped by plant.

Microbiome diversity

There were differences in bacterial or archaeal community richness among seeds from different plants (LMM; df = 2, *F* value = 6.91, *P* value = 0.015) (**Figure 3.2A**), where plant B and C had lower seed richness than plant A (Tukey's HSD post hoc test; *P* value = 0.001 and 0.006, respectively). However, bacterial or archaeal community richness among seeds from pods collected from the same plant were not different (LMM, *P* value > 0.05) (**Figure 3.2B**). Similarly, bacterial or archaeal phylogenetic diversity were different among seeds collected from different plants (LMMs; df = 2, *F* value = 6.56, *P* value = 0.003) (**Figure 3.2C**) but not among seeds from pods within the same plant (LMM, *P* value > 0.05) (**Figure 3.2D**). Plants B and C had lower seed microbiome bacterial or archaeal phylogenetic diversity compared with plant A (Tukey's HSD post hoc test, *P* value = 0.001 and 0.013, respectively). We observed no differences in fungal richness among seeds from different plants (LMM; df = 2, *F* value = 1.11, *P* value = 0.37) (**Figure 3.2E**) and among seeds from pods within the same plant (LMM, *P* value > 0.05) (**Figure 3.2F**). To summarize, these results suggest that seed bacterial or archaeal alpha diversity but not fungal diversity varied plant to plant.

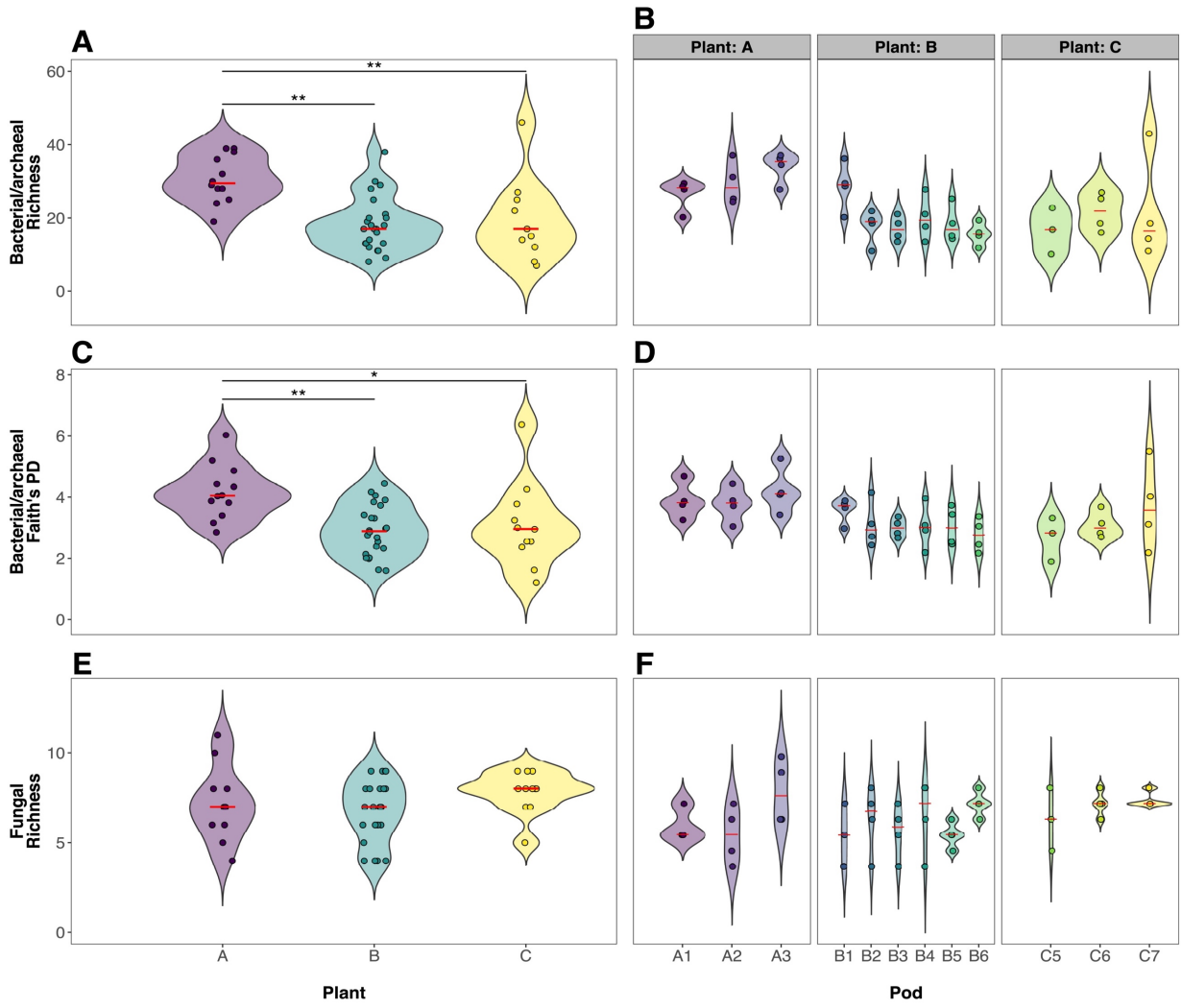

Figure 3.2. Alpha diversity of common bean seed microbiome among plants and pods.

A, Bacterial or archaeal richness in seeds among plants were different (linear mixed-effects model [LMM] *P* value = 0.015) **B,** but not among pods within plant (*P* value > 0.05). **C,** Bacterial or archaeal phylogenetic diversity in seeds among plants were different (LMM *P* value = 0.003) **D,** but not among pods within a plant (*P* value > 0.05). **E,** Fungal richness in seeds was not different among plants (LMM *P* value = 0.37) and **F,** among pods within a plant (*P* value >

Figure 3.2 (cont'd)

0.05). Here, each point represents the microbiome richness from a microbial DNA extraction from an individual seed.

We detected a difference in seed bacterial or archaeal community composition among plants (nested PERMANOVA, df = 2, *F* value = 2.94, R^2 = 0.12, *P* value = 0.002) (**Figure 3.3A**) but, again, not among pods from the same plant (nested PERMANOVA, df = 9, *F* value = 0.99, R^2= 0.18, *P* value = 0.63). Though separation among pods and plants are not obvious on the PCoA for the fungal seed microbiomes, we detected modest differences in fungal community composition among seeds from different plants (nested PERMANOVA, df = 2, *F* value = 1.69, R^2 = 0.09, *P* value = 0.004) (**Figure 3.3B**), but not among seeds from pods from the same plant (nested PERMANOVA, df = 9, *F* value = 1.17, R^2 = 0.25, *P* value = 0.11). An analysis of beta dispersion revealed that there were differences in seed microbiome dispersion across different plants for bacterial or archaeal communities (PERMDISP, df = 2, *F* value = 63.94, R^2 = 0.74, *P* value = 0.001) (**Figure 3.3C**) but not for fungal communities (PERMDISP, df = 2, *F* value = 0.89, R^2 = 0.05, *P* value = 0.4) (**Figure 3.3D**). These results are qualitatively the same for analyses based on community structure as assessed by Bray-Curtis dissimilarity (**Figure 3.7 Appendix B**). Therefore, statistical differences in the seed microbiome across plants for the bacteria or archaea may be attributed to either centroid or dispersion, whereas fungal seed communities were different by centroid.

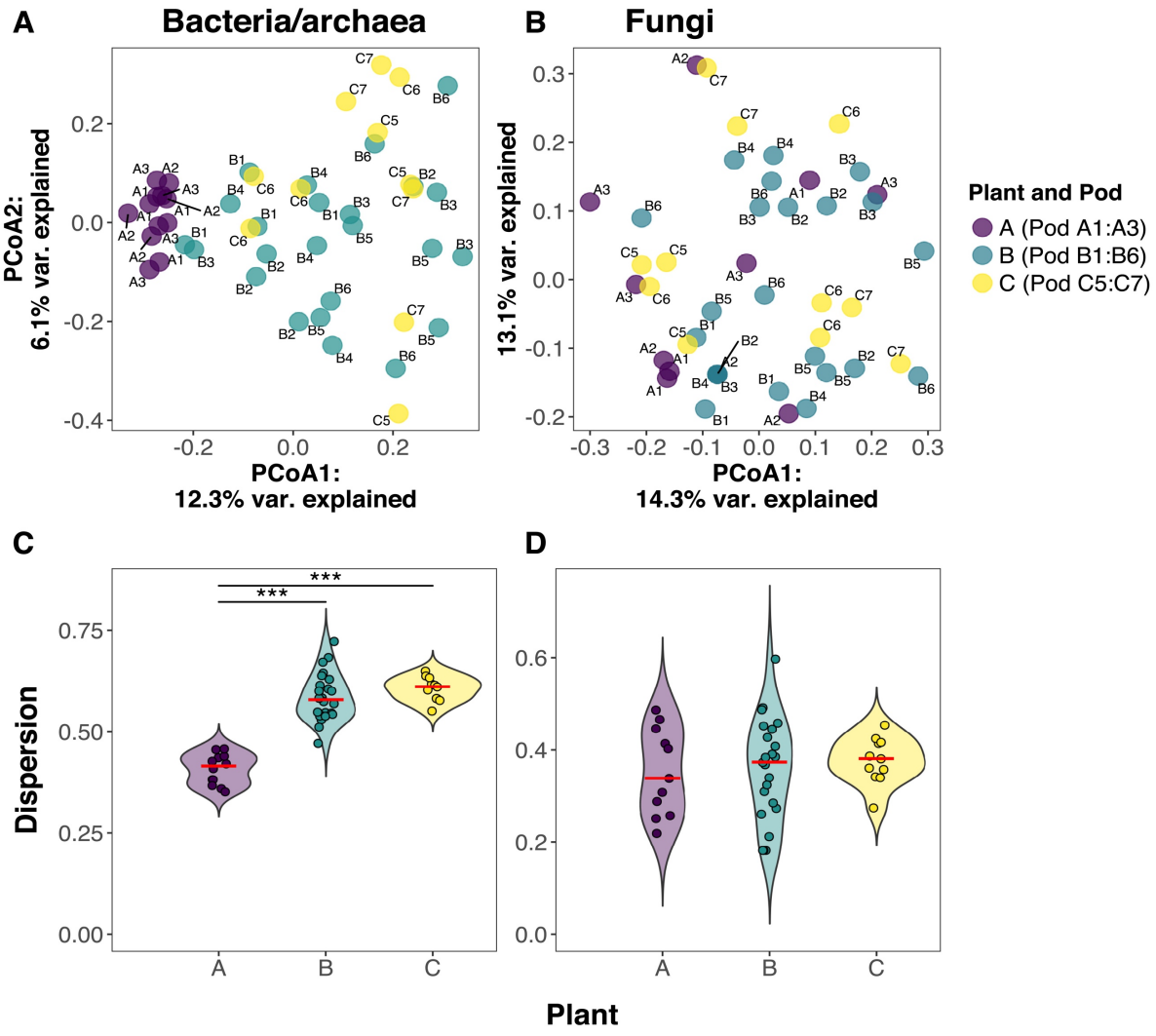

Figure 3.3. Beta diversity visualizations of common bean seed microbiome based on Jaccard index.

Principal coordinate analysis (PCoA) plot based on unweighted Jaccard dissimilarities of **A,** bacterial or archaeal and **B,** fungal microbiomes and dispersion to centroid for each **C,** bacterial or archaeal and **D,** fungal microbiome. Each point represents microbiome data from microbial DNA extraction from an individual seed. Samples were plotted and grouped by plant as illustrated by different colors. Each point represents a seed microbiome that is labeled by a plant

Figure 3.3 (cont'd)

letter and a pod number. Significant differences in distance to centroid among seeds from different plants (C and D) are indicated with asterisks (*** = *P* value < 0.001).

Bean seed microbiome composition

We identified 135 bacterial or archaeal and 49 fungal taxa at the genus level. The bacterial or archaeal individual seed communities were dominated by taxa from classes *Gammaproteobacteria* (50.47%), *Bacilli* (24.48%), *Alphaproteobacteria* (8.68%), and *Bacteroidia* (6.59%) (**Figure 3.4A**), and include *Pseudomonas* (13.58 %), *Bacillus* (10.2 %), *Acinetobacter* (9.5 %), *Raoultella* (7.09%), and *Escherichia-Shigella* (5.19%) as the major genera. Among members of the class Alphaproteobacteria, we also found genera *Bradyrhizobium* and *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* with relative abundance of 2.57 and 0.85 %, respectively. Although seed fungal community composition varied among plants and also pods within plant, the fungal community was dominated by taxa belonging to classes Dothideomycetes (22.77%), Agaricomycetes (16.61%), and Eurotiomycetes (14.44%) (**Figure 3.4B**), and the genera *Aspergillus* (14.44%), *Capnodiales* unidentified sp. 23791 (9.27%), and *Aureobasidium* (8.28%).

A key objective of this research was to understand the sources of variability in the individual bean seed microbiome to inform future study design. Because we found that the plant-to-plant seed microbiome variability was highest when grown in control conditions, we performed a power analysis to determine how many plants would be required to observe a treatment effect from seed samples pooled per plant. To detect the effect of treatment on bacterial or archaeal richness and phylogenetic diversity, pooled seeds from 9 and 12 plants are needed for 16S rRNA richness and phylogenetic diversity, respectively, to achieve power of 0.8; and 13 and 19 plants to achieve power of 0.95 (**Figure 3.5.**).

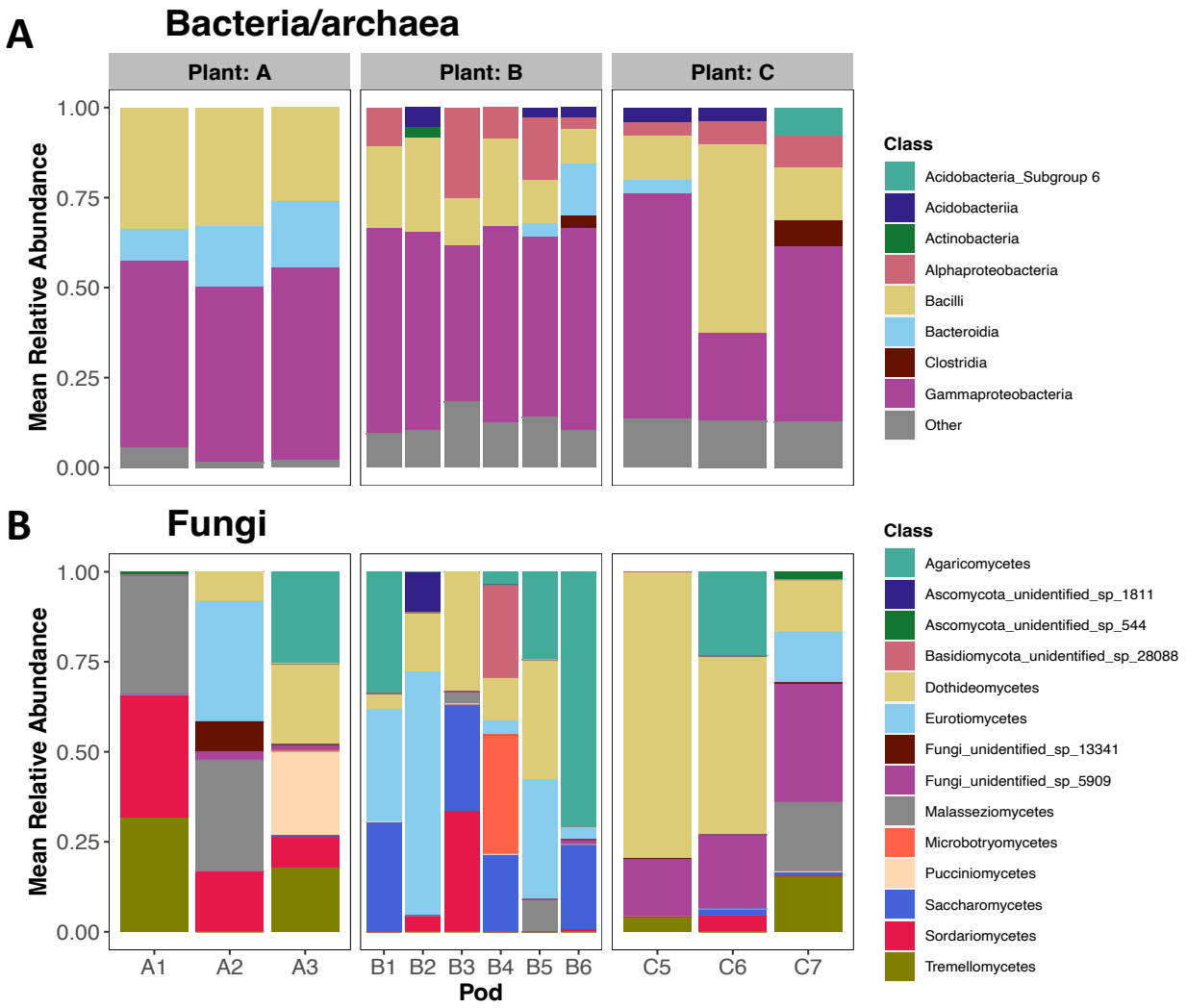

Figure 3.4. Relative abundances of common bean seed microbiome.

Bar plots represent mean relative abundances of **A,** bacterial or archaeal and **B,** fungal classes in seeds detected across plants. Each bar shows the average composition of individual seeds that were each extracted and analyzed from the same pod. For bacteria or archaea, each pod consisted of four seeds (except for C5; three seeds); and, for fungi, each pod consisted of three seeds (except for A3, B6, C6, and C7: four seeds). The endophyte microbiome was assessed from the DNA extracted from a single seed collected from each pod. Bacterial or archaeal and fungal

Figure 3.4 (cont'd)

classes with mean relative abundances of less than 10% were grouped into the "Other" classification, which includes many lineages (not monophyletic).

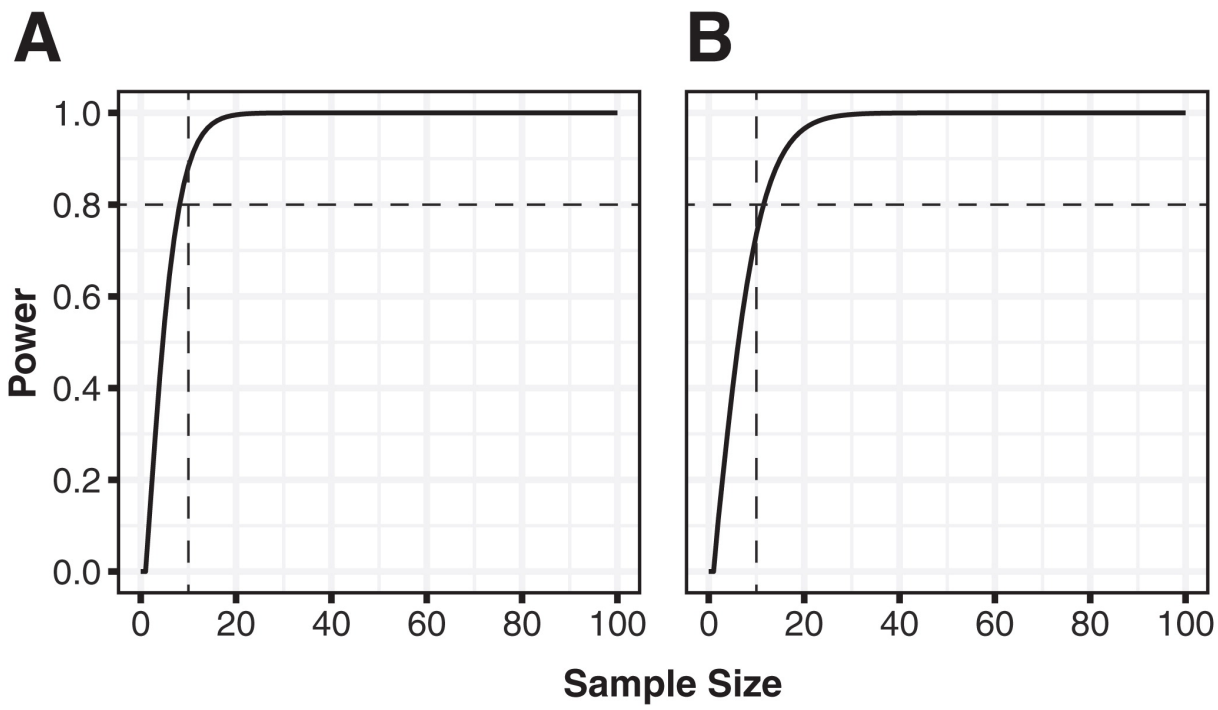

Figure 3.5. Power analysis.

Analysis of power revealed that an effect of treatment on the 16S ribosomal RNA bacterial or archaeal **A,** α diversity (richness) and **B,** phylogenetic diversity would be detectable in 12 plants at a power of 0.8. Because the highest seed microbiome variability was at the parent plant level, individual seed microbiome sequence profiles were pooled in silico by plant to perform this power analysis at the individual plant level.

Shared taxa among seeds and plants

We explored the microbial taxa shared across all seed samples, detected across all three plants, and also shared among all seeds within one plant (**Table 3.2 Appendix A**). Although there were no bacterial or archaeal taxa detected and shared among all seeds, there were 11 taxa detected in more than half of seed samples (occupancy > 0.5, $n = 47$), and taxa from genus *Bacillus* were most common. Other bacterial or archaeal taxa found in more than half of seeds were assigned to the genera *Stenotrophomonas*, *Raoultella*, *Pseudomonas*, *Lactobacillus*, *Acinetobacter*, *Listeria*, *Bradyrhizobium*, and *Entereococcus*. There were no fungal taxa shared among seeds. One fungal taxon from the genus *Aspergillus* was detected in ~30% of the seeds.

In all, 54 bacterial or archaeal taxa were detected and shared across all plants, and these belonged to belonged to phyla Proteobacteria (Gammaproteobacteria = 21 OTUs, and Alphaproteobacteria = 6 OTUs), Firmicutes (13 OTUs), Actinobacteria (7 OTUs), Acidobacteria (4 OTUs), Chloroflexi (1 OTU), Bacteroidetes (1 OTU), and Verrucomicrobia (1 OTU). There were seven fungal taxa detected and shared across all plants, and these belonged to classes Eurotiomycetes (1 OTU), Dothideomycetes (2 OTUs), Sordariomycetes (1 OTU), Malasseziomycetes (1 OTU), Agaricomycetes (1 OTU), and one unindentified fungal taxon (**Table 3.2 Appendix A**). Together, these results suggest the taxa that should be explored further to understand any importance to the host and their consistency and rates of transmission from plant parent to offspring.

Discussion

There remain gaps in our understanding of the persistence and assembly of seed microbiome members, especially across plant generations, and which microbiome members are

beneficial and actively selected by, or even coevolved with, the host. Here, we investigated the variability of the common bean microbiome at the resolution of the individual seed, which is the unit that delivers any vertically transmitted microbiome to the offspring. Because multiple legume seeds within a pod develop as a result of a single flower pollination, one simple hypothesis is that the individual seeds within a pod may harbor a highly similar microbiome if the floral pathway of assembly is prominent. However, recent work has suggested that the endophytic seed microbiome of green bean varieties of common bean likely colonize predominantly via the internal vascular pathway, and not the floral pathway (9), which may result in more homogeneity among seed microbiomes of the same plant. Our data support this finding, because seeds from different pods in the same plant (and, therefore, a common vascular pathway across pods) had relatively low microbiome variability, especially compared across plants. It is expected that the vascular pathway of seed microbiome assembly is more likely to colonize the internal seed compartments (e.g., embryo) and, therefore, more likely to be vertically transmitted (17). It is as-yet unclear whether plant species that have a stronger relative importance of the floral pathway in seed microbiome assembly may exhibit higher microbiome variability at the pod or fruit level. Such an outcome may indicate that the experimental unit should, instead, be the pod level rather than the plant level for plant species dominated by floral assembly pathways.

There are many challenges in analyzing the microbiome of seeds generally and of a single seed in particular, which may be why cultivation-independent studies of single seeds are few (48). Previous studies showed that seeds have low microbial biomass and diversity (4, 9, 49), especially relative to other plant compartments or soil. Therefore, many studies pool seeds to analyze the aggregated microbiome of many seeds and to get enough microbial biomass for

microbial DNA extraction (4, 6, 19-21, 50). Generally, microbiome samples that have low biomass have numerous challenges in sequence-based analysis, as discussed elsewhere (51, 52). First, unknown contaminants, either from nucleic acid kits or from mishandling of the samples, can have relatively high impact on the observed community composition and, thus, extraction and PCR controls are needed for assessment of contaminants and subtraction of suspected contaminants from the resulting community (53). Second, the sparse datasets (e.g., many zero observations for many taxa in many samples) generated from low-biomass samples often require special statistical consideration and data normalization (54, 55).

Plant host contamination of the microbiome sequence data are another consideration expected with analysis of the seed, and this challenge also applies to other plant compartments (56, 57). For 16S rRNA amplicon sequencing, the contaminant reads typically derive from host mitochondria and chloroplasts but rRNA ITS2 or 18S amplicon analysis may also have reads annotated as Plantae. Therefore, nucleic acid extractions may attempt minimal disturbance of the plant tissue that is the target of microbiome investigation; for example, grinding tissues to include in the extraction will result in higher plant DNA contamination than separating microbial biomass from intact tissue. The cost of this is that any microbes lodged tightly into the host tissue or persisting within host cells may be missed.

For our study, we wished to understand the microbiome with which a dormant seed begins. This is a key aspect of our approach, because it is known that seeds can exude both antimicrobials and attractants to select for particular microbial members early in microbiome assembly of the germinated seed and emerging seedling (9, 58), and there is an active zone of plant and microbiome activity at the seed-soil interface of a germinating seed (the spermosphere) (59).Therefore, to target the native endophytic seed microbiome without also allowing the seeds

or seedling to select or filter particular members, we used dormant seeds and took care to minimally disrupt their tissues. Notably, many protocols have opted to first germinate seeds and, therefore, include the outcome of early plant selection on the observed seed microbiome (6, 7, 19, 60). Though there are advantages and disadvantages to both germinated and dormant seed microbiome assessment, we reason that focus on the endophyte of the dormant seed is more likely to characterize taxa that have been transmitted from parent to seed.

Taking all of these methodological aspects into consideration, this study presents a protocol and analysis pipeline for endophyte microbiome DNA extraction from a single dormant seed that experiences minimal tissue disruption in the extraction process, includes both positive and negative sequencing controls, and includes bioinformatic steps to identify contamination and remove host signal from the marker gene amplification. Notably, we chose to perform microbiome analysis based on a presence or absence (unweighted) taxon table rather than a table with relativized (weighted) taxon abundances. This was done in consideration of the ecology of the seed endophyte microbiome members to likely be dormant until germination (45) and, therefore, the differences in relativized abundances do not reflect differences in fitness outcomes inside the dormant seed. We acknowledge that relative abundances could reflect differential microbiome member recruitment by the host plant during seed formation; however, this is not the objective of the study and would be best addressed with a different design to determine the multigeneration consistency and transmission rates of any observed enrichments, which would be supported by assessment of the seed microbiome within individual seeds, and across plant generations. Finally, we acknowledge that the relatively prominent size of the edible common bean seed was to our study's advantage, and that some other seeds (e.g., from some dicots) may

not be as accessible for sampling via this protocol as individual units because of their small size, structure, and challenges in removing the seed coat.

In conclusion, individual seed microbiome assessment provides improved precision in our understanding of plant microbiome assembly and sets the stage for studies of vertical transmission. We found that seeds produced by an individual bean plant can be considered as a unit (for comparative treatment study designs), and that seeds produced by different plants are expected to have slightly different microbiomes, even if grown under the same, controlled conditions and in the same soil source. Future work may consider whether functional redundancy in plant-beneficial phenotypes across seed microbiome members may provide one mechanism for consistent outcomes in beneficial plant microbiome establishment.

APPENDICES

APPENDIX A: Supplemental Table

Table 3.2. List of microbial taxa identified in more than half of total seed samples (occupancy > 0.5, n = 47 for bacteria/archaea); and microbial taxa shared across plants (occupancy = 1, n = 3) and across seeds within plant (occupancy > 0.5)

Across Total Seeds								
Bacterial/ archaeal OTU	**Occupancy (n = 47)**	**Domain**	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Species**
KY646001.1.1573	0.70	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
AB745637.1.1513	0.68	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Stenotrophomonas*	uncultured bacterium
EF528273.1.1512	0.66	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
KF625186.1.1741	0.64	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Raoultella*	*Raoultella ornithinolytica*
DQ234192.1.1572	0.60	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
AB362767.1.1576	0.57	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	*Lactobacillus*	*Lactobacillus fermentum*
EF517956.1.1666	0.57	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	*Acinetobacter* sp. IGCAR-9/07
LKHO01000001.198.1803	0.57	Bacteria	Firmicutes	Bacilli	Bacillales	Listeriaceae	*Listeria*	*Listeria monocytogenes*
FPLS01006697.30.1498	0.53	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	*Bradyrhizobium*	NA
FR746074.1.1400	0.53	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
JUOP01000215.81.1657	0.53	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	*Enterococcus*	*Enterococcus faecalis*

Table 3.2 (cont'd)

Across Plants								
Bacterial/ archaeal OTU	**Occupancy (n= 3)**	**Domain**	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Species**
AB252903.1.1522	1	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	*Delftia*	NA
AB362767.1.1576	1	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	*Lactobacillus*	*Lactobacillus fermentum*
AB491963.1.1519	1	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	*Alicycliphilus*	NA
AB672179.1.1470	1	Bacteria	Actinobacteria	Thermoleophilia	Gaiellales	Gaiellales	*Gaiellales*	uncultured bacterium
AB745637.1.1513	1	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Stenotrophomonas*	uncultured bacterium
AJ439344.1.1502	1	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	*Corynebacterium 1*	NA
AM749763.1.1392	1	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiales Family XVII	*Clostridiales Family XVII*	uncultured Clostridia bacterium
AOKA01000137.3650.5241	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
ARMF01000004.652094.653575	1	Bacteria	Acidobacteria	Acidobacteriia	Solibacterales	Solibacteraceae (Subgroup 3)	*Bryobacter*	uncultured Acidobacteria bacterium
BCWL01000265.646.2166	1	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Neisseriaceae	*Neisseria*	NA
CCPS01000022.154.1916	1	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Escherichia-Shigella*	*Escherichia coli*
CP001965.357388.358915	1	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Gallionellaceae	*Sideroxydans*	uncultured bacterium

Table 3.2 (cont'd)

CP002739.60209.61784	1	Bacteria	Firmicutes	Clostridia	Thermoanaerobacterales	Thermoanaerobacterales Family III	*Thermoanaerobacterium*	NA
CP009312.1832900.1834429	1	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	*Lawsonella*	uncultured bacterium
CR931997.108684.110210	1	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	*Corynebacterium 1*	NA
CTEN01000001.315310.316856	1	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	*Streptococcus*	NA
CZKG01000048.33782.35294	1	Bacteria	Actinobacteria	Thermoleophilia	Solirubrobacterales	Solirubrobacteraceae	*Solirubrobacteraceae*	NA
DQ234192.1.1572	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
EF517956.1.1666	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	*Acinetobacter* sp. IGCAR-9/07
EF528273.1.1512	1	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
EU797470.1.1498	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
FJ538159.1.1489	1	Bacteria	Actinobacteria	Actinobacteria	Frankiales	Acidothermaceae	*Acidothermus*	uncultured soil bacterium
FJ538164.1.1471	1	Bacteria	Acidobacteria	Acidobacteriia	Acidobacteriales	Acidobacteriaceae (Subgroup 1)	*Occallatibacter*	NA
FJ624896.1.1468	1	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	*Roseiarcus*	uncultured bacterium
FPID01000096.6.1486	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	metagenome
FPLP01001110.16.1473	1	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Paracoccus*	metagenome
FPLS01006697.30.1498	1	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	*Bradyrhizobium*	NA

Table 3.2 (cont'd)

FPLS01016296.23.1536	1	Bacteria	Proteobacteria	Gammaproteobacteria	Gammaproteobacteria Incertae Sedis	Gammaproteobacteria Incertae Sedis	*Acidibacter*	NA
FPLS01034054.18.1510	1	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	*Acetobacteraceae*	metagenome
FPLS01043838.10.1533	1	Bacteria	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	*Paenibacillus*	metagenome
FR746074.1.1400	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
GQ483458.1.1495	1	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	*Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*	NA
JF833468.1.1560	1	Bacteria	Proteobacteria	Gammaproteobacteria	Gammaproteobacteria Incertae Sedis	Gammaproteobacteria Incertae Sedis	*Acidibacter*	uncultured gamma proteobacterium
JN082536.1.1536	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	NA
JN868932.1.1483	1	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	*Xanthobacteraceae*	NA
JUOP01000215.81.1657	1	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	*Enterococcus*	*Enterococcus faecalis*
JX025749.1.1465	1	Bacteria	Acidobacteria	Acidobacteriia	Acidobacteriales	Acidobacteriales	*Acidobacteriales*	uncultured bacterium
KC502951.1.1538	1	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Rhodanobacteraceae	*Rhodanobacter*	NA
KF625186.1.1741	1	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Raoultella*	*Raoultella ornithinolytica*
KJ410541.1.1362	1	Bacteria	Acidobacteria	Acidobacteriia	Acidobacteriales	Acidobacteriales	*Acidobacteriales*	NA
KJ878597.1.1448	1	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Rhodanobacteraceae	*Dyella*	NA
KJ955641.1.1496	1	Bacteria	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	*Paenibacillus*	uncultured *Paenibacillus* sp.
KM200448.1.1512	1	Bacteria	Proteobacteria	Gammaproteobacteria	Gammaproteobacteria KF-JG30-C25	Gammaproteobacteria KF-JG30-C26	Gammaproteobacteria KF-JG30-C27	NA

Table 3.2 (cont'd)

KM210514.1.1481	1	Bacteria	Bacteroidetes	Bacteroidia	Sphingobacteriales	Sphingobacteriaceae	*Nubsella*	NA
KP735610.1.1442	1	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	*Bacillus subtilis*
KR026982.1.1447	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
KR029299.1.1517	1	Bacteria	Firmicutes	Bacilli	Bacillales	Planococcaceae	*Sporosarcina*	NA
KR088380.1.1569	1	Bacteria	Proteobacteria	Gammaproteobacteria	Aeromonadales	Aeromonadaceae	*Aeromonas*	*Aeromonas salmonicida*
KX753099.1.1468	1	Bacteria	Chloroflexi	Chloroflexi AD3	Chloroflexi AD3	Chloroflexi AD3	*Chloroflexi AD3*	uncultured bacterium
KY646001.1.1573	1	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
KY777463.1.1542	1	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	*Micrococcus*	*Micrococcus luteus*
LKHO01000001.198.1803	1	Bacteria	Firmicutes	Bacilli	Bacillales	Listeriaceae	*Listeria*	*Listeria monocytogenes*
MTIS01000005.1947007.1948557	1	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	*Bacillus alkalitelluris*
Y07576.1.1528	1	Bacteria	Verrucomicrobia	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	*Candidatus Udaeobacter*	NA
Fungal OTU	**Occupancy (n= 3)**	**Domain**	**Phylum**	**Class**	**Order**	**Family**	***Genus***	**Species**
OTU_14	1	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	*Aspergillus*	*Aspergillus sydowii*
OTU_22	1	Fungi	Ascomycota	Dothideomycetes	Dothideales	Aureobasidiaceae	*Aureobasidium*	*Aureobasidium pullulans*
OTU_26	1	Fungi	Fungi unidentified sp 5909	Fungi unidentified sp 5909	Fungi unidentified sp 5909	Fungi unidentified sp 5909	Fungi unidentified sp 5909	Fungi unidentified sp 5909
OTU_31	1	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariales fam Incertae sedis	*Phialemoniopsis*	*Phialemoniopsis curvata*
OTU_32	1	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	*Malassezia*	*Malassezia globosa*

Table 3.2 (cont'd)

OTU_7	1	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	*Phlebiopsis*	*Phlebiopsis* sp 16232
OTU_9	1	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Capnodiales unidentified sp 23791	Capnodiales unidentified sp 23791	Capnodiales unidentified sp 23791

Across Seeds Within Plant A

Bacterial/ archaeal OTU	Occupancy (n= 12)	Domain	Phylum	Class	Order	Family	Genus	Species
AB362767.1.1576	1	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	*Lactobacillus*	*Lactobacillus fermentum*
AB745637.1.1513	1	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Stenotrophomonas*	uncultured bacterium
CCPS01000022.154.1916	1	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Escherichia-Shigella*	*Escherichia coli*
EF528273.1.1512	1	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
EU797470.1.1498	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
FOVD01000013.327.1846	1	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	*Chryseobacterium*	NA
FR746074.1.1400	1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
JUOP01000215.81.1657	1	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	*Enterococcus*	*Enterococcus faecalis*
KF625186.1.1741	1	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Raoultella*	*Raoultella ornithinolytica*
KY646001.1.1573	1	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
LKHO01000001.198.1803	1	Bacteria	Firmicutes	Bacilli	Bacillales	Listeriaceae	*Listeria*	*Listeria monocytogenes*

Table 3.2 (cont'd)

Across Seeds Within Plant B								
Bacterial/ archaeal OTU	**Occupancy (n= 24)**	**Domain**	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Species**
DQ234192.1.1572	0.75	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA
AB745637.1.1513	0.71	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Stenotrophomonas*	uncultured bacterium
FPLS01006697.30.1498	0.67	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	*Bradyrhizobium*	NA
KY646001.1.1573	0.67	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
EF517956.1.1666	0.63	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	*Acinetobacter* sp. IGCAR-9/07
KF625186.1.1741	0.54	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Raoultella*	*Raoultella ornithinolytica*
LKHO01000001.198.1803	0.50	Bacteria	Firmicutes	Bacilli	Bacillales	Listeriaceae	*Listeria*	*Listeria monocytogenes*
Across Seeds Within Plant C								
Bacterial/ archaeal OTU	**Occupancy (n= 11)**	**Domain**	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Species**
EF528273.1.1512	0.73	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus*	NA
CP002739.60209.61784	0.55	Bacteria	Firmicutes	Clostridia	Thermoanaerobacterales	Thermoanaerobacterales_Family III	*Thermoanaerobacterium*	NA
FR746074.1.1400	0.55	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Pseudomonas*	NA

APPENDIX B: Supplemental Figures

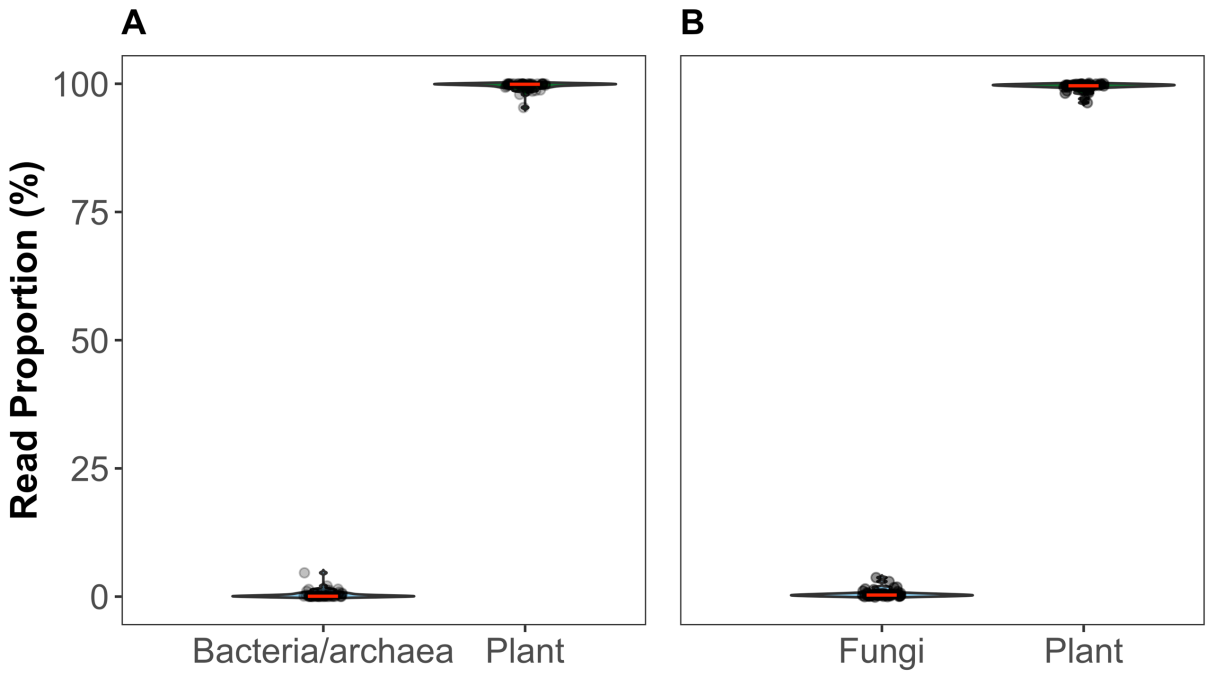

Figure 3.6. The proportion of plant reads.

The proportion of plant reads of the total bacterial/archaeal (a) and fungal (b) reads showed that more than 90 % reads obtained were plant contaminants

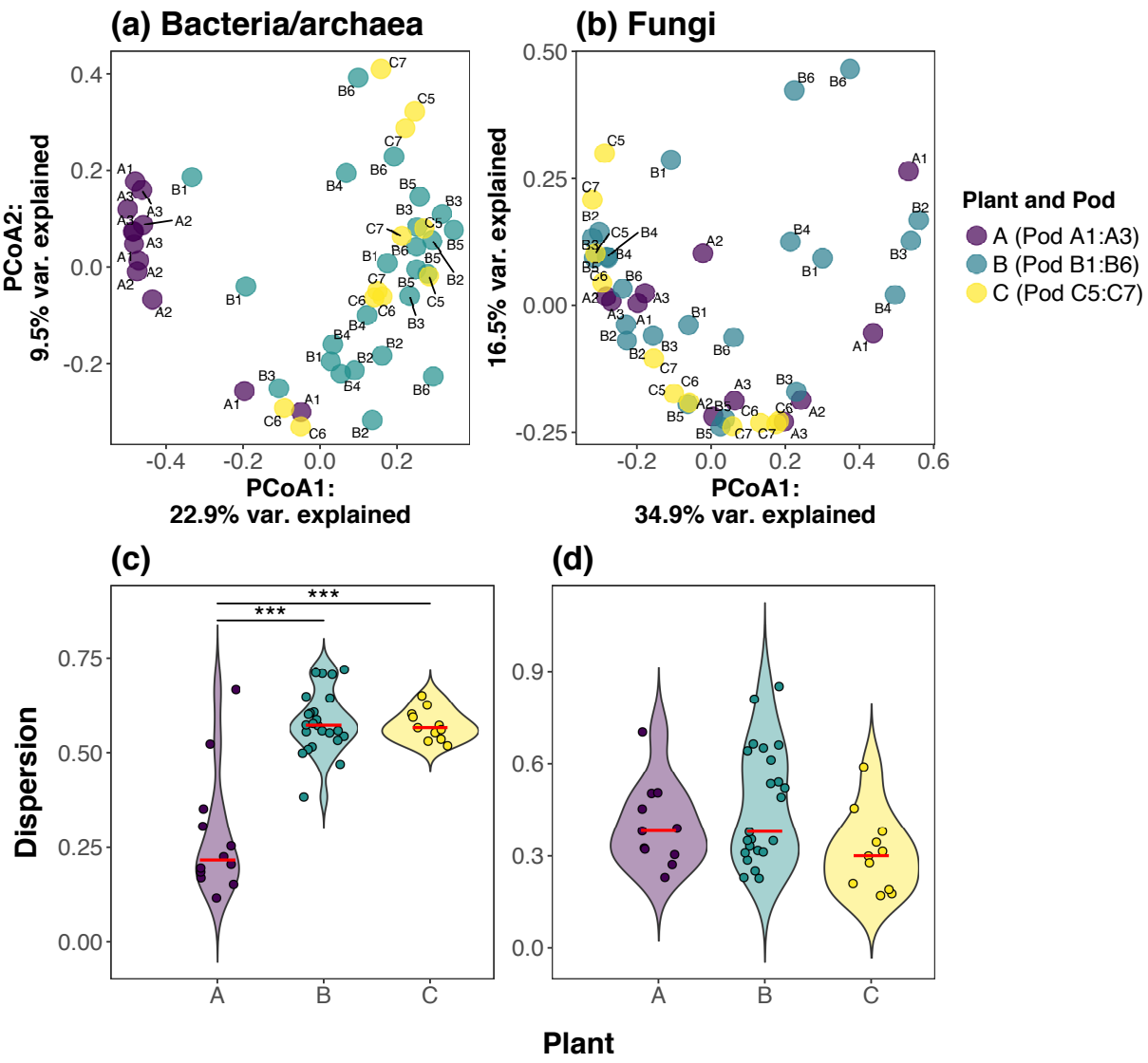

Figure 3.7. Beta diversity visualization of the common bean seed microbiome based on Bray-Curtis dissimilarities.

Principal coordinate analysis (PCoA) plot based on Bray-Curtis dissimilarities (a weighted resemblance metric) of bacterial/archaeal (a) and fungal (b) microbiomes, and dispersion to centroid for each bacterial/archaeal (c) and fungal (d) microbiomes. Each point represents microbiome data from microbial DNA extraction from an individual seed. The samples were plotted and grouped based on plant as illustrated by different colors. Each point represents a seed

Figure 3.6 (cont'd)

microbiome that is labelled by a plant letter and a pod number. For (c) and (d), significant differences in distance to centroid among seeds from different plants are indicated with asterisks (*** is p-value < 0.001).

APPENDIX C: Supplemental Information, Results, and Protocols

Supplemental Information

Background information about seed microbiome assessment

Seed microbiome assessment has been conducted either by culture-dependent or culture-independent methods or a combination of both. Culture-dependent approaches are limited by technical difficulties in isolating microbes from seeds because not all members can be cultivated on agar plates. Seed microbiome members are assumed to be in a dormant stage until the plant germinates (1), and this may contribute to its difficulties to cultivate because they may need specific nutrients and growth conditions to be able to grow. Thus, culture-based methods often fail to detect all the microbial community members and can lead to biased results. Community profiling approaches using next-generation high-throughput sequencing of marker genes, such as 16S rRNA and ITS genes for bacteria/archaea and fungi may provide a better assessment of microbial community in the seed and a more comprehensive picture of microbial community structure. A number of studies have shown variability of the microbial community in the seed from different plant species and genotypes (2, 3), different geographical sites (4), or even between seed developmental stages (5) and seed compartments (6).

To our knowledge, this is the first study to use marker gene high-throughput sequencing methods to assess the microbial community of individual seeds of common bean (*Phaseolus vulgaris*) to investigate its variability among plants and pods. Understanding on seed-to-seed, pod-to-pod, and plant-to-plant microbiome variability provides essential information on pooling biological samples and allows well-powered experimental design on the seed microbiome assessment under plant treatments. Extracting metagenomic microbial DNA from one seed is extremely difficult because the individual seed is considered as a low-microbial-biomass sample, and microbial DNA extraction from low-microbial-biomass samples can be a major challenge in

studying the microbial community and ecology. Hence, a robust and efficient DNA extraction from low microbial samples is a crucial step because reproducibility and accuracy of microbiome study with amplicon-based sequencing approaches will depend on the efficient DNA extraction from the sample (7).

Moreover, microbial DNA extraction from low-microbial-biomass and low diversity samples is prone to DNA contaminations from other microbes and/or plant contaminants, such as mitochondria and chloroplasts. Thus, it is necessary to set up strategies to minimize DNA contamination during extraction as well as in the downstream analysis. We developed and optimized protocols for microbial DNA extraction from individual seed samples of common bean (*P. vulgaris*) var. Red Hawk. The protocols described in this study were aimed to generate robust methods that can be generally implemented to study seed microbiomes.

Challenges to seed microbiome DNA extraction protocols

Total DNA extraction from one seed sample can be problematic and challenging. There are some key limitations that we need to consider and carefully assess before conducting DNA extraction from seed samples.

1.) Low diversity. Previous studies show that seeds have low microbial diversity (8-10) relative to other plant compartments, or rhizosphere and soil. Since seeds have low microbial diversity, it is important to include a mock microbial community as a positive control for assessing PCR amplification and sequencing efficiency. Since the expected composition of the mock microbial community is known, we can estimate any sequencing error (e.g., chimera), identify diversity biases, and determine microbial contaminations by including the mock community in a sequencing run (11).

2.) Low microbial biomass. Seeds, particularly individual seeds, contain low microbial biomass compared to soil or rhizosphere samples. One of the challenges in working with low microbial biomass samples (low DNA target) is the feasibility and efficiency of the DNA extraction method and exogenous microbial contamination. Our strategy to overcome this issue is to include whole-cell mock microbial community as the DNA extraction positive control to establish the DNA extraction procedure. We also include a negative control (buffer only) that is important for assessing microbial contamination in the samples.

3.) Plant anti-microbial chemicals (plant defense compounds). One hypothesis on why seeds have low microbial diversity is the occurrence of population bottlenecks in the seed (12), especially individual seeds, that is caused by the accumulation of anti- microbial compounds in the seed (9, 13). These compounds are activated when seeds are crushed or germinated. Thus, we performed buffer soaking methods instead of grinding the samples or germinating the seeds for the DNA extraction procedure to avoid anti-microbial compounds affecting our results.

4.) Plant contamination. Plant contamination in plant microbiome study is common because plant compartments including the seed contain plastids, chloroplasts, and mitochondria that share common ancestry and have sequence similarity with bacteria. There are three main approaches to minimize plant host contamination including modification of microbial DNA extraction to prevent the co-extraction of plant organelles, the application of PCR amplification blocking primers to block the amplification of plant sequences, and the use of specific mismatch primers (14). Lundberg et al. (15) constructed PNA clamps for plastids (pPNA) and mitochondria (mPNA) that can bind tightly to the contaminant sequences and block its amplification. Another approach is the use of anti-chloroplast primers 799F that can amplify 16S rRNA gene sequences but also avoid the amplification of plant sequences. Beckers et al. (14)

reported that primer pair 799F/1391R was the most efficient in eliminating plant DNA (very low amplification of plant DNA) and resulted in the highest amount of bacterial OTUs.

However, there is a scientific motivation to be able to directly compare microbiome data across studies (for instance, to compare with other studies that include soils, plants, potential sources of dispersal/immigration). Thus, the use of the popular Earth Microbiome Project 16S rRNA V45 primers is often desirable (https://earthmicrobiome.org/), despite that these primers co-amplify plant contaminants. Therefore, other steps to reduce host signal can be taken in the DNA extraction protocol. Specific to this seed microbiome study, we performed an adaptation of microbial DNA extraction to prevent the co-extraction of plant organelles. Instead of grinding the seed that can release plant organelles, we used a phosphate buffered saline (PBS) soaking procedure. This procedure has been used by previous studies in assessing seed microbiomes (5, 10, 16). By using a seed soaking procedure, microbial cells in the seed coat and funiculus will be released to the suspension (10, 17).

5.) Non-host DNA contamination. As we described above, DNA contamination introduced during the DNA extraction method is a major challenge in assessing microbial communities from low microbial biomass samples. There are different strategies in removing DNA contaminants before and after sequencing. In this study, we included blank or negative controls for the DNA extraction method as well as for PCR amplification. Strategies can also be performed after sequencing, for example, by removing any microbial taxa that have been published and identified as a common contaminant. However, this method cannot be implemented for all studies because the observed microbial community is different for each study. Another method is removing taxa that are also present in the negative control. However, this strategy may also remove the actual members of the microbial community because of

multiplexing artifacts that occur in the negative control (18). In our study we performed decontamination using an open-source R package called microDecon for identifying and removing contamination (19).

Supplemental Results: Microbiome beta-diversity analyses based on Bray-Curtis dissimilarity

We also calculated beta diversity using Bray-Curtis dissimilarity. We found that there were differences among plants in the bacterial/archaeal community structure (nested PERMANOVA, df = 2, F-value = 4.93, R^2 = 0.21, p-value = 0.003). There were no differences in bacterial/archaeal communities among pods from the same plant (nested PERMANOVA, df = 9, F-value = 1.23, R^2 = 0.19, p-value = 0.056) (**Figure 3.7a Appendix B**). Beta diversity analysis of seed fungal community structure using Bray-Curtis dissimilarity showed that there were no differences among plants (nested PERMANOVA, df = 2, F-value = 0.98, R^2 = 0.04, p-value = 0.39) nor pods (nested PERMANOVA, df = 9, F-value = 0.94, R^2 = 0.19, p-value = 0.60) (**Figure 3.7b Appendix B**). Permutated multivariate analysis of dispersion showed that there were differences of bacterial/archaeal community structure dispersion among plants (PERMDISP, df = 2, F-value = 38.04, R^2 = 0.63, p- value = 0.001) (**Figure 3.7c Appendix B**), but there were no significant differences of fungal community structure dispersion (PERMDISP, df = 2, F-value = 3.35, R^2 = 0.14, p-value = 0.056) (**Figure 3.7d Appendix B**).

Supplemental Protocols: Cultivation-independent native seed endophyte analysis

We performed surface sterilization of the seed samples before extracting the DNA because our study focused on the seed endophytic communities. Surface sterilization of sample is

a required procedure to study plant endophytes (20) because we need to completely remove the epiphytic microbes from the seed surfaces. The seed epiphytes are mostly derived from plant surfaces (e.g., leaves, stems, fruits) and/or environment (e.g., soil) (21). Surface sterilization also removes microbial contamination from human contact during harvesting, handling, and processing.

Part 1. Seed surface sterilization and overnight soaking procedures

Expected time: 20 minutes, overnight

Materials

1. Common bean seed (*P. vulgaris L.*, var. Redhawk) (approximately 0.6 gram per seed)
2. Trypticase Soy Agar (TSA) and Potato Dextrose Agar (PDA) plates
3. Sterilization solution: 10 % (v/v) bleach with 0.1% (v/v) Tween20
4. Sterile Phosphate Buffer Saline (PBS) 1 X with 0.05 % (v/v) Tween20

Equipment

1. 50 ml centrifuge tube (USA Scientific, VWR)
2. Beaker
3. Analytical balance
4. Sterile dissecting scalpel (size 20)
5. Sterile dissecting forceps
6. Sterile disposable Petri dishes
7. Orbital shaker
8. Plate spreader or plating beads

Procedure

1) Select healthy seeds with no disease symptoms from the stock and weigh the seeds to obtain seed mass data.

2) Place seed(s) into a sterile 50 ml centrifuge tube and immerse the seed in ~ 20- 25 ml sterilization solution (10 % (v/v) bleach with 0.1% (v/v) Tween20) for 10 minutes.

    - *A different volume of sterilization solution can be used, based on the number/size of seeds.*
    - *Shake the tube several time during incubation.*

3) Discard the sterilization solution and rinse/wash the seed with sterile water 5 times to remove bleach residue.

    - *To check the effectiveness of surface sterilization, spread 50-100 µl of the final rinse water on to TSA and PDA plates. Incubate the TSA and PDA plates at 30℃ for 2-3 days and 25-26℃ for 5 days, respectively. Discard associated sample if there is any microbial growth on the plates.*

4) Place sterile seed onto sterile plate and carefully dissect or open the seed in half long-ways on the natural division of cotyledon using sterile surgical blade and forceps.

    - *In this study, we removed the seed coat instead of dissecting the seed in half. The purpose of seed coat removal is because our study focused on seed endophytes, we assumed that removal the seed coat could increase the release of the endophytes located in*

the endosperm and embryo into the buffer solution. However, we observed high plant contamination in when the seed coat was removed (more than 90% of total reads). We also found that removing the seed coat is time consuming and produces plant debris that can interfere with the DNA extraction process and can be the source of chloroplast and mitochondria contamination. Thus, we propose to dissect/open the seed in half long-ways on the natural division of cotyledon instead of removing the seed coat. In our experience, this allows for the release of endophytes into the buffer and minimizes host contamination from seed coat removal.

5) Immerse and soak surface sterilized seed in sterile Phosphate Buffered Saline (PBS) 1X supplemented with 0.05% (v/v) Tween 20 (3 ml) under constant agitation (160 rpm) overnight at 4℃.

- *A different volume of buffer can be used based on the number/size of seed sample.*
- *We recommend to always include a DNA extraction positive control for low microbial biomass samples like seeds (e.g., a mock microbial community). We used the commercial ZymoBIOMICS Microbial Community Standard (catalog number: D6300) for this study by diluting 75 µl (1 prep) of the mock community into 3-5 ml PBS 1X with 0.05% (v/v) Tween20. Also, we created our own mock community in-house to include*

particular bacteria and fungi that reflect the expected composition of common seed microbial community members. The mock community included populations of type strains or isolates grown in the lab, and then combined at an equal ratio at a concentration of 10^8 cells/ml for bacteria (10^6 cells/ml for Streptomyces) and 10^7 cells/ml for fungi and stored in glycerol stock in the -80 ℃. Therefore, the positive control DNA extraction of our in-house mock-community would be performed directly on these cells and can be sequenced and checked for contamination from the expected composition.

- *We recommend to always include a DNA extraction negative control of extraction buffer only (3-5 ml PBS 1X with 0.05% (v/v) Tween20). This sample should be sequenced to check for contamination and to calculate a sequencing error rate (22).*

Part 2: Seed processing and pellet collection

Expected time: 90 minutes

Stopping points: It is recommended to either stop after the pellet collection step, or to go through the DNA extraction protocol in the same day

Materials

1. Overnight-soaked seed in sterile PBS 1X with 0.05%. (v/v) Tween20

Equipment

1. Swinging-bucket rotor centrifuge
2. Vortex

3. Sterile forceps
4. Beaker
5. Micropipette
6. Sterile barrier micropipette tips
7. Microcentrifuge tubes

Procedure

6) Centrifuge all samples and controls at 4500 x *g* for 60 minutes at 4℃.
    - *We used a centrifuge with swinging-bucket rotor rather than fixed-angle rotor so that the pellets will form at the bottom of the conical tube, thus, it is easier to resuspend and collect the pellets. The original protocol from previous study (Barret et al., 2015) stated that centrifugation was performed at 6000 x g for 10 minutes at 4℃. Since the maximum speed for swinging-bucket rotor centrifuge is 4500 x g, we extended the centrifugation time.*
7) Carefully remove seeds aseptically with sterile forceps, spin tubes again with bucket centrifugation at 4500 x *g* for 10 min at 4℃ to re-pellet any disturbed material. Carefully remove supernatant with sterile disposable pipette or micropipette until approximately 1-2 ml remain.
    - *Alternatively: After first hour of centrifugation, gently pour most of the supernatant out of the tubes and discard, then aseptically remove seeds with sterile forceps, leaving approx. 1-2 ml of supernatant in the tube.*

8) Resuspend pellet in remaining supernatant by vortexing for approximately 1 minute.

9) Transfer the suspension into 1.5- or 2-ml microcentrifuge tube and centrifuge at 20,000 x g for 10 minutes.

10) Discard the supernatant and keep the pellet for DNA extraction using E.Z.N.A.®Bacterial DNA Kit with centrifugation protocol.

- *Pellets can be stored at -20℃ until they are ready to be extracted.*

Part 3: Microbial DNA extraction from seed pellet with bead beating procedure using E.Z.N.A.® Bacterial DNA Kit with modification

Expected time: 4 hours active time, 3 hours of incubation time

Materials

1. Seed pellet collected from the previous step
2. E.Z.N.A.® Bacterial DNA Kit (D3350-02) (OMEGA Bio-Tek Inc., Norcross, GA, USA)
3. 100 % Ethanol
4. Tris-EDTA (TE) Buffer, Molecular Biology Grade (pH 8.0)

Equipment

1. Micropipette
2. Sterile barrier micropipette tips
3. Microcentrifuge tubes
4. Vortex
5. Beaker
6. Heat block or water bath

Before starting:

- Prepare HBC Buffer, DNA Wash Buffer, and Lysozyme kit components as instructed in the manufacturer's protocol
- Set a heatblock or water bath at 37℃
- Set a shaking heatblock or water bath at 55℃
- Set an incubator or a heatblock at 65℃ (can change the 37℃ to 65℃ later in the protocol, if necessary)
- Heat Elution Buffer to 65℃

Procedure

11) Add 100 µl TE buffer to the pellet and completely resuspend the pellet.
12) Add 10 µl Lysozyme resuspended with Elution Buffer (see bottle for instructions). Vortex to mix thoroughly. Incubate in 35℃ heat block for 1 hour.
    - *We used 1 hour incubation instead of 10 minutes as stated on the manufacturer's protocol to achieve complete digestion of the cell wall.*
13) While incubating, aseptically add 25 mg Glass Beads S (included with the kit) to new, labeled, 1.5 ml- tubes.
14) After incubation transfer entire sample, including any material that has precipitated out, into the corresponding tube with glass beads.
15) Vortex the bead-beating tubes at maximum speed for 10 minutes. After vortexing, allow tubes to rest a few minutes for glass beads to settle out. Transfer supernatant to clean 1.5 ml- tube.

- *We implemented a bead-beating step into the protocol for hard-to-lyse bacteria/archaea and fungi. This procedure yielded better results (higher DNA concentration) than extraction without a bead-beating step.*
- *We extended the vortexing time at maximum speed from 5 minutes to 10 minutes for optimal cell wall disruption.*

16) Add 100 µl TL Buffer and 20 µl Proteinase K Solution to all tubes. Pipette up and down to break up pellet, if present, and then vortex to mix thoroughly. Incubate at 55℃ in a shaking heat block for 2 hours (500 rpm). Alternatively, incubate in a stationary heat block and vortex every 20 minutes.

- *We used longer incubation time for optimal DNA yield.*

17) Add 5 µl RNase A. Invert tube several times to mix. Let sit at room temperature for 5 minutes.

18) Centrifuge at 10,000 x *g* for 2 minutes to pellet any undigested material.

19) Transfer the supernatant to a new 1.5 ml- microcentrifuge tube. Do not disturb the pellet. Discard pellet.

20) Add 220 µl BL Buffer. Vortex to mix. Incubate at 65℃ for 10 minutes. (Note: after this step, aliquot the needed amount of elution buffer into a tube and place in the 65℃ block to warm for later use).

21) Add 220 µl 100% ethanol. Vortex for 20 seconds at maximum speed to mix thoroughly. Break any precipitates by pipetting up and down 10 times.

22) Insert a HiBind® DNA Mini Column into a 2- ml Collection Tube. Transfer the entire sample to the HiBind® DNA Mini Column, including any precipitate that may have formed.

23) Centrifuge at 10,000 x *g* for 1 minute. Discard the filtrate and the collection tube.

24) Insert the HiBind® DNA Mini Column into a new 2- ml Collection Tube.

25) Add 500 µl HBC Buffer diluted with 100 % isopropanol (see the bottle for instructions). Centrifuge at 10,000 x *g* for 1 minute. Discard the filtrate and reuse the collection tube.

26) Add 700 µl DNA Wash Buffer diluted with 100 % ethanol (see the bottle for instructions). Centrifuge at 10,000 x *g* for 1 minute. Discard the filtrate and reuse the collection tube.

27) Repeat Step #26 for a second DNA Wash Buffer wash step.

28) Centrifuge the empty HiBind® DNA Mini Column at maximum speed (> 10,000 x *g*) for 2 minutes to dry the column.

    - *We used a centrifuge with maximum speed of 20,000 x g for optimal removal of trace ethanol.*

29) Insert the HiBind® DNA Mini Column into a new, nuclease-free 1.5- ml microcentrifuge tube.

30) Add 30 µl Elution Buffer heated to 65℃ to the center of the HiBind® matrix. Let sit for 10 minutes at room temperature.

- *We decreased Elution Buffer volume from 50-100 µl as stated on the manufacturer's protocol to 30 µl to increase DNA concentration.*
- *To obtain more yield, second elution can be conducted with the same Elution Buffer volume.*

31) Centrifuge at 10,000 x *g* for 1 minute to elute the DNA. Store the DNA at -20℃ for temporary storing or -80℃ for long-term storing.

- *We measured the DNA concentration using Qubit™ dsDNA HS (High Sensitivity) Assay Kit with the Qubit Fluorometer. This protocol yielded DNA with the concentration of approximately 0.7-1 ng/µl per gram of seed. Moreover, the PCR amplification of bacterial 16S V4 and fungal ITS2 also resulted in clear and specific bands.*
- *We tried the Qiagen DNeasy PowerSoil DNA Isolation Kit for the DNA extraction after collecting the seed pellets. However, the protocol using this kit was irreproducible. The DNA concentration was too low to be detected on the Qubit Fluorometer and the PCR amplification of bacterial 16S V4 and fungal ITS2 resulted in very weak or no specific bands. We assumed that the Qiagen DNeasy PowerSoil DNA Isolation Kit was not reliable enough to extract DNA from low microbial biomass samples, such as seeds or individual seeds, in particular. Thus, we chose to use the DNA extraction kit with optimum lysis*

that implement both enzymatic digestion of cell wall and physical disruption using bead-beating step.

Supplemental References

1. Cope-Selby N, Cookson A, Squance M, Donnison I, Flavell R, Farrar K. 2017. Endophytic bacteria in Miscanthus seed: implications for germination, vertical inheritance of endophytes, plant evolution and breeding. GCB Bioenergy 9:57-77.

2. Özkurt E, Hassani MA, Sesiz U, Künzel S, Dagan T, Özkan H, Stukenbrock EH. 2020. Seed-Derived Microbial Colonization of Wild Emmer and Domesticated Bread Wheat (Triticum dicoccoides and T. aestivum) Seedlings Shows Pronounced Differences in Overall Diversity and Composition. mBio 11.

3. Wassermann B, Cernava T, Müller H, Berg C, Berg G. 2019. Seeds of native alpine plants host unique microbial communities embedded in cross-kingdom networks. Microbiome 7:108.

4. Chartrel V, Dugat-Bony E, Sarthou A-S, Huchette S, Bonnarme P, Irlinger F. 2021. The microbial community associated with pea seeds (Pisum sativum) of different geographical origins. Plant and Soil 462:405-427.

5. Barret M, Briand M, Bonneau S, Préveaux A, Valière S, Bouchez O, Hunault G, Simoneau P, Jacques M-A. 2015. Emergence Shapes the Structure of the Seed Microbiota. Applied and Environmental Microbiology 81:1257-1266.

6. Eyre AW, Wang M, Oh Y, Dean RA. 2019. Identification and Characterization of the Core Rice Seed Microbiome. Phytobiomes Journal 3:148-157.

7. Davis A, Kohler C, Alsallaq R, Hayden R, Maron G, Margolis E. 2019. Improved yield and accuracy for DNA extraction in microbiome studies with variation in microbial biomass. Biotechniques 66:285-289.

8. Adam E, Bernhart M, Müller H, Winkler J, Berg G. 2018. The Cucurbita pepo seed microbiome: genotype-specific composition and implications for breeding. Plant and Soil 422:35-49.

9. Chesneau G, Torres-Cortes G, Briand M, Darrasse A, Preveaux A, Marais C, Jacques M-A, Shade A, Barret M. 2020. Temporal dynamics of bacterial communities during seed development and maturation. FEMS Microbiology Ecology 96.

10. Rezki S, Campion C, Iacomi-Vasilescu B, Preveaux A, Toualbia Y, Bonneau S, Briand M, Laurent E, Hunault G, Simoneau P, Jacques M-A, Barret M. 2016. Differences in stability of seed-associated microbial assemblages in response to invasion by phytopathogenic microorganisms. PeerJ 4:e1923-e1923.

11. Pollock J, Glendinning L, Wisedchanwet T, Watson M. 2018. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. Applied and Environmental Microbiology 84:e02627-17.

12. Newcombe G, Harding A, Ridout M, Busby PE. 2018. A Hypothetical Bottleneck in the Plant Microbiome. Frontiers in Microbiology 9.

13. Meldau S, Erb M, Baldwin IT. 2012. Defence on demand: mechanisms behind optimal defence patterns. Annals of Botany 110:1503-1514.

14. Beckers B, Op De Beeck M, Thijs S, Truyens S, Weyens N, Boerjan W, Vangronsveld J. 2016. Performance of 16s rDNA Primer Pairs in the Study of Rhizosphere and Endosphere Bacterial Microbiomes in Metabarcoding Studies. Frontiers in Microbiology 7.

15. Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL. 2013. Practical innovations for high-throughput amplicon sequencing. Nature Methods 10:999-1002.

16. Torres-Cortés G, Bonneau S, Bouchez O, Genthon C, Briand M, Jacques M-A, Barret M. 2018. Functional Microbial Features Driving Community Assembly During Seed Germination and Emergence. Frontiers in Plant Science 9.

17. Gitaitis R, Walcott R. 2007. The epidemiology and management of seedborne bacterial diseases. Annu Rev Phytopathol 45:371-97.

18. Karstens L, Asquith M, Davin S, Fair D, Gregory WT, Wolfe AJ, Braun J, McWeeney S. 2019. Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments. mSystems 4.

19. McKnight DT, Huerlimann R, Bower DS, Schwarzkopf L, Alford RA, Zenger KR. 2019. microDecon: A highly accurate read-subtraction tool for the post-sequencing removal of contamination in metabarcoding studies. Environmental DNA 1:14-25.

20. Saldierna Guzmán JP, Nguyen K, Hart SC. 2020. Simple methods to remove microbes from leaf surfaces. Journal of Basic Microbiology 60:730-734.

21. Shade A, Jacques MA, Barret M. 2017. Ecological patterns of seed microbiome diversity, transmission, and assembly. Curr Opin Microbiol 37:15-22.

22. Schloss PD, Gevers D, Westcott SL. 2011. Reducing the Effects of PCR Amplification and Sequencing Artifacts on 16S rRNA-Based Studies. PLOS ONE 6:e27310.

REFERENCES

REFERENCES

1. Mitter B, Pfaffenbichler N, Flavell R, Compant S, Antonielli L, Petric A, Berninger T, Naveed M, Sheibani-Tezerji R, von Maltzahn G, Sessitsch A. 2017. A New Approach to Modify Plant Microbiomes and Traits by Introducing Beneficial Bacteria at Flowering into Progeny Seeds. Frontiers in Microbiology 8.

2. Shade A, Jacques MA, Barret M. 2017. Ecological patterns of seed microbiome diversity, transmission, and assembly. Curr Opin Microbiol 37:15-22.

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

4. Adam E, Bernhart M, Müller H, Winkler J, Berg G. 2018. The Cucurbita pepo seed microbiome: genotype-specific composition and implications for breeding. Plant and Soil 422:35-49.

5. Berg G, Raaijmakers JM. 2018. Saving seed microbiomes. The ISME Journal 12:1167-1170.

6. Bergna A, Cernava T, Rändler M, Grosch R, Zachow C, Berg G. 2018. Tomato Seeds Preferably Transmit Plant Beneficial Endophytes. Phytobiomes Journal 2:183-193.

7. López-López A, Rogel MA, Ormeño-Orrillo E, Martínez-Romero J, Martínez-Romero E. 2010. Phaseolus vulgaris seed-borne endophytic community with novel bacterial species such as Rhizobium endophyticum sp. nov. Syst Appl Microbiol 33:322-7.

8. Barret M, Briand M, Bonneau S, Préveaux A, Valière S, Bouchez O, Hunault G, Simoneau P, Jacques M-A. 2015. Emergence Shapes the Structure of the Seed Microbiota. Applied and Environmental Microbiology 81:1257-1266.

9. Chesneau G, Torres-Cortes G, Briand M, Darrasse A, Preveaux A, Marais C, Jacques M-A, Shade A, Barret M. 2020. Temporal dynamics of bacterial communities during seed development and maturation. FEMS Microbiology Ecology 96.

10. Chartrel V, Dugat-Bony E, Sarthou A-S, Huchette S, Bonnarme P, Irlinger F. 2021. The microbial community associated with pea seeds (Pisum sativum) of different geographical origins. Plant and Soil 462:405-427.

11. Dai Y, Li X-Y, Wang Y, Li C-X, He Y, Lin H-H, Wang T, Ma X-R. 2020. The differences and overlaps in the seed-resident microbiome of four Leguminous and three Gramineous forages. Microbial Biotechnology 13:1461-1476.

12. Eyre AW, Wang M, Oh Y, Dean RA. 2019. Identification and Characterization of the Core Rice Seed Microbiome. Phytobiomes Journal 3:148-157.

13. Raj G, Shadab M, Deka S, Das M, Baruah J, Bharali R, Talukdar NC. 2019. Seed interior microbiome of rice genotypes indigenous to three agroecosystems of Indo-Burma biodiversity hotspot. BMC Genomics 20:924.

14. Rodríguez CE, Antonielli L, Mitter B, Trognitz F, Sessitsch A. 2020. Heritability and Functional Importance of the Setaria viridis Bacterial Seed Microbiome. Phytobiomes Journal 4:40-52.

15. Xing H-Q, Ma J-C, Xu B-L, Zhang S-W, Wang J, Cao L, Yang X-M. 2018. Mycobiota of maize seeds revealed by rDNA-ITS sequence analysis of samples with varying storage times. MicrobiologyOpen 7:e00609.

16. Nelson EB. 2018. The seed microbiome: Origins, interactions, and impacts. Plant and Soil 422:7-34.

17. Barret M, Guimbaud JF, Darrasse A, Jacques MA. 2016. Plant microbiota affects seed transmission of phytopathogenic microorganisms. Mol Plant Pathol 17:791-5.

18. Vujanovic V, Germida J. 2017. Seed endosymbiosis: a vital relationship in providing prenatal care to plants. Canadian Journal of Plant Science 97:972-981.

19. Wassermann B, Cernava T, Müller H, Berg C, Berg G. 2019. Seeds of native alpine plants host unique microbial communities embedded in cross-kingdom networks. Microbiome 7:108.

20. Klaedtke S, Jacques M-A, Raggi L, Préveaux A, Bonneau S, Negri V, Chable V, Barret M. 2016. Terroir is a key driver of seed-associated microbial assemblages. Environmental Microbiology 18:1792-1804.

21. Johnston-Monje D, Raizada MN. 2011. Conservation and Diversity of Seed Associated Endophytes in Zea across Boundaries of Evolution, Ethnography and Ecology. PLOS ONE 6:e20396.

22. Bintarti AF, Kearns PJ, Sulesky-Grieb A, Shade A. 2020. Abiotic treatment to common bean plants results in an altered endophytic seed microbiome. bioRxiv doi:10.1101/2020.06.05.134445:2020.06.05.134445.

23. Kelly JD, Hosfield GL, Varner GV, Uebersax MA, Long RA, Taylor J. 1998. Registration of 'Red Hawk' Dark Red Kidney Bean. Crop Science 38:cropsci1998.0011183X003800010056x.

24. Castro-Guerrero NA, Isidra-Arellano MC, Mendoza-Cozatl DG, Valdés-López O. 2016. Common Bean: A Legume Model on the Rise for Unraveling Responses and Adaptations to Iron, Zinc, and Phosphate Deficiencies. Frontiers in Plant Science 7.

25. Hoagland DR, Arnon DI. 1950. The water-culture method for growing plants without soil. Circular California agricultural experiment station 347.

26. Rezki S, Campion C, Simoneau P, Jacques M-A, Shade A, Barret M. 2018. Assembly of seed-associated microbial communities within and across successive plant generations. Plant and Soil 422:67-79.

27. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences 108:4516-4522.

28. Op De Beeck M, Lievens B, Busschaert P, Declerck S, Vangronsveld J, Colpaert JV. 2014. Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. PloS one 9:e97629.

29. Edgar RC, Flyvbjerg H. 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31:3476-82.

30. Rideout JR, He Y, Navas-Molina JA, Walters WA, Ursell LK, Gibbons SM, Chase J, McDonald D, Gonzalez A, Robbins-Pianka A, Clemente JC, Gilbert JA, Huse SM, Zhou HW, Knight R, Caporaso JG. 2014. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. PeerJ 2:e545.

31. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590-6.

32. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460-1.

33. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10:996-998.

34. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335-336.

35. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics 26:266-7.

36. McKnight DT, Huerlimann R, Bower DS, Schwarzkopf L, Alford RA, Zenger KR. 2019. microDecon: A highly accurate read-subtraction tool for the post-sequencing removal of contamination in metabarcoding studies. Environmental DNA 1:14-25.

37. Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for microbial marker-gene surveys. Nature methods 10:1200-1202.

38. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17:3.

39. Gdanetz K, Benucci GMN, Vande Pol N, Bonito G. 2017. CONSTAX: a tool for improved taxonomic resolution of environmental fungal ITS sequences. BMC Bioinformatics 18:538.

40. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42:D633-42.

41. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology 73:5261-5267.

42. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2019. vegan: Community Ecology Package. R package version 2.5-7.

43. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.

44. Pinheiro J, Bates D, DebRoy SS, Sarkar D, Team RC. 2021. nlme: Linear and Nonlinear Mixed Effects Models. R package version 31-152 3:1-113.

45. Cope-Selby N, Cookson A, Squance M, Donnison I, Flavell R, Farrar K. 2017. Endophytic bacteria in Miscanthus seed: implications for germination, vertical inheritance of endophytes, plant evolution and breeding. GCB Bioenergy 9:57-77.

46. Kindt R, Kindt MR. 2019. Package 'BiodiversityR'. Package for community ecology and suitability analysis 2:11-12.

47. Shade A, Stopnisek N. 2019. Abundance-occupancy distributions to prioritize plant core microbiome membership. Current Opinion in Microbiology 49:50-58.

48. Abdelfattah A, Wisniewski M, Schena L, Tack AJ. 2021. Experimental evidence of microbial inheritance in plants and transmission routes from seed to phyllosphere and root. Environmental Microbiology 23:2199-2214.

49. Rezki S, Campion C, Iacomi-Vasilescu B, Preveaux A, Toualbia Y, Bonneau S, Briand M, Laurent E, Hunault G, Simoneau P, Jacques M-A, Barret M. 2016. Differences in stability of seed-associated microbial assemblages in response to invasion by phytopathogenic microorganisms. PeerJ 4:e1923-e1923.

50. Latz MAC, Kerrn MH, Sørensen H, Collinge DB, Jensen B, Brown JKM, Madsen AM, Jørgensen HJL. 2021. Succession of the fungal endophytic microbiome of wheat is dependent on tissue-specific interactions between host genotype and environment. Science of The Total Environment 759:143804.

51. Bender JM, Li F, Adisetiyo H, Lee D, Zabih S, Hung L, Wilkinson TA, Pannaraj PS, She RC, Bard JD. 2018. Quantification of variation and the impact of biomass in targeted 16S rRNA gene sequencing studies. Microbiome 6:1-12.

52. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. 2019. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends Microbiol 27:105-117.

53. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 6:226.

54. Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, Sanders NJ, Cornell HV, Comita LS, Davies KF. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology letters 14:19-28.

55. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R. 2017. Normalization and

microbial differential abundance strategies depend upon data characteristics. Microbiome 5:27.

56. Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ. 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. Proceedings of the National Academy of Sciences 115:E1157-E1165.

57. Song L, Xie K. 2020. Engineering CRISPR/Cas9 to mitigate abundant host contamination for 16S rRNA gene-based amplicon sequencing. Microbiome 8:1-15.

58. Meldau D, Long H, Baldwin I. 2012. A native plant growth promoting bacterium, Bacillus sp. B55, rescues growth performance of an ethylene-insensitive plant genotype in nature. Frontiers in Plant Science 3.

59. Schiltz S, Gaillard I, Pawlicki-Jullian N, Thiombiano B, Mesnard F, Gontier E. 2015. A review: what is the spermosphere and how can it be studied? Journal of Applied Microbiology 119:1467-1481.

60. Parsa S, García-Lemos AM, Castillo K, Ortiz V, López-Lavalle LAB, Braun J, Vega FE. 2016. Fungal endophytes in germinated seeds of the common bean, Phaseolus vulgaris. Fungal biology 120:783-790.

CHAPTER 4: Responses of seed endophytes under drought stress: Field study

Abstract

Plants recruit and select a group of microbial taxa that benefit the plant growth and fitness, and these taxa are transmitted to the next plant generations via seed. There has been growing attention on seed microbiomes and their importance for seed germination and seedling development. However, the study on the response of seed microbiome to abiotic stress, such as drought, is still largely limited. Moreover, the consequences of seed microbiome alteration to the host plant when the plant is exposed to drought are unknown. We assessed the seed endophytes community of common bean cultivated in the field under two different water treatments, with and without rainout shelter. This study is aimed to investigate the responses of seed endophytes to drought stress. The rainout shelter was designed for 50 % water exclusion. The plants were managed under an organic farming system without Nitrogen fertilizer application. Seed bacterial/archaeal community structures were assessed using 16S rRNA gene amplicon sequencing. Analysis of plant productivity showed the differences in plant yield across bean cultivars. Meanwhile, the water treatment has a marginal effect on the plant yield. Given these differences in plant fitness, we expect to observe the differences in seed microbiome across cultivars and the shift of the community when the plant is exposed to drought stress.

Introduction

Drought has been a major obstacle in agriculture and crop production around the world and it has been considered as the most deleterious abiotic stress that leads to reduction of crop production, including for the critical legume crop, common bean (*Phaseolus vulgaris* L.) (1, 2). Water deficits are more pronounced in some warmer and drier regions known for common bean producers (e.g., Latin America and Africa), and common bean cultivation in these areas are

expanding to marginal areas that are more likely to sustain rain-fed conditions (1). In arid areas where the common bean cultivation depends on the rainfall for water supply, drought stress accounts for up to 80% decrease in common bean production (2, 3). Moreover, the adverse effects of drought are exacerbated by global climate changes, such as extreme temperature and irregular rainfall, which contribute to more severe and longer drought periods that threaten global food production and security (2). Therefore, researchers are aimed to improve plant drought stress tolerance through several strategies, such as plant breeding and genome engineering technology. However, the conventional breeding is both time consuming and labor intensive, and the application of plant genetic engineering can discourage consumers (4, 5). Another recent alternative that has received growing attention is harnessing beneficial microbiota associated with the host plant to address the challenges of abiotic stress and develop sustainable agriculture (5-7).

It is clear that plant plays role in recruiting and selecting groups of microbial taxa that benefit plants, for example by the production and secretion of root exudates into the rhizosphere (8-10). The beneficial members of plant microbiome are capable of performing specific functions that are essential for the plant growth and health, such as production of growth promoting phytohormones, nitrogen fixation, phosphate solubilization, and protection against environmental stresses (11-13). Previous studies observed enrichment of plant-associated microbial taxa which possess plant growth promoting activity and stress resistance properties under dessert farming system, suggesting that plant selects microbial community with beneficial traits to survive under drought stress (14). Despite their ability to develop a variety of adaptive and stress responses, plants often need their microbes to mediate protective mechanisms to withstand abiotic stresses (6, 15). It has been reported that plant microbiome responses to

drought stress favors the host plant to deal with the stress, for example drought induces the synthesis of proline and other osmolytes in plant-associated microbiota which are essential for maintaining the integrity of plant cell membranes (16). Another study demonstrated that endophytes inoculation leads to an increase in abscisic acid (ABA) concentration in plant which is positively correlated with an increase in water use efficiency under water deficit condition (17). More recent study confirmed that drought increases the expression of genes related to ABA biosynthesis in endophytes which leads to stomatal closure in plant as an adaptive mechanism to prevent water loss (18).

Microbiome members of seed endophytes are of particular interest because they act as early colonizers of new seedling and a starting point of plant microbiome assembly (19). Moreover, plants may preserve the continuity of beneficial microbes in the seed by transmitting them to progeny through (12, 20, 21). Similar to endophytes in other plant compartments, beneficial seed endophytes colonize and thrive within the seed tissues without causing harm to the host plant, and seed-transmitted endophytes are believed to be more adapted to plants (22). Moreover, because seed is a unique environment, only endophytes with distinct features and competences colonize and survive within seed (23-26). In this aspect, the physiological changes during seed development and maturation influence and select the endophytes community that colonize the seed (23). Plants can heavily depend on seed endophytes for their survival, and studies have investigated the importance of seed endophytes, especially during seed germination and seedling development (21, 27-29). However, little information is available on their roles in plant tolerance to drought stress. Moreover, our knowledge on the impact of drought stress to the seed endophytes of crops, especially common bean, is still largely unknown. Whereas,

understanding the responses of seed endophytes to drought is a fundamental step in developing strategies to promote plant tolerance to drought stress.

In this study, we investigate the drivers of common bean seed endophyte microbial communities and their putative roles in enhancing plant tolerance to drought stress in the field. Specifically, we aimed to (i) understand the effects of drought on seed endophyte microbiome structure and (ii) determine the interactive effects of drought with other factors expected to influence the seed endophyte community structure and composition, such as plant genotype (different bean cultivars), farming site (different geographic locations and latitudes), and farming system (organic and conventional). Rainout shelters in the field were used to apply drought and well-watered treatments to crops grown in Michigan, which is one of the largest US producers and exporters of common bean with 210,000 acres planted and total production of 5.4 million cwt in 2021 (30).

Materials and methods

Plant cultivars

This study used four different common bean cultivars developed by the Dry Bean Breeding and Genetics Program at Michigan State University: Cayenne small red, B18504 black, Rosetta pink, and R99 navy (white) beans. Three first bean cultivars have been tested and adapted to management system in Michigan, and one cultivar, Rosetta, is resistant to heat and drought (**Table 4.1**). Meanwhile, the last cultivar (R99) is a non-nodulating bean cultivar (**Table 4.1**). They have various seed sizes, where Cayenne small red bean has small-medium seed size, B18504 black and R99 navy beans have small seed size, and Rosetta pink bean has medium seed size (**Table 4.1**).

Table 4.1. Description of the common bean cultivars used in this study

Cultivar	Market seed class	Seed size	Attributes	References
Cayenne	Small red bean	Small to medium (36 g/100 seeds)	• High yielding • Resistant to *bean common mosaic virus* (BCMV) and *common bacterial blight* (CBB) • Well adapted to Michigan farming system	(31, 32)
Rosetta	Pink bean	Medium (36 g/100 seeds)	• High yielding • Resistant to common strains of rust and mosaic virus • Well-suited to drought conditions (drought tolerant) • Well adapted to Michigan farming system	(33, 34)
B18504 (Adams)	Black bean	Small (20.9 g/100 seeds)	• High yielding • Resistant to anthracnose • Well adapted to Michigan farming system	(35)
R99	Navy (white) bean	Small (20,2 g/100 seeds)	• Nonnodulation mutant • Adapted to temperate climate (warmer part) of North America	(36)

Field study design

The field experiment was designed to investigate the effect of drought to the four bean cultivars. The experimental fields for this study are in two agricultural sites that represent different latitudes of bean cultivation in Michigan: The Lower Peninsula (in East Lansing, at the MSU Agronomy Farm) and Upper Peninsula (in Chatham, at the Upper Peninsula Research and Extension Center (UPREC) North Farm) of Michigan. A randomized split plot design was used with four replications per plot. There were four cultivars, and each cultivar was grown with or without rainout shelters. Thus, there were eight combinations with four replications or 32 treatment combinations in total. The 10' x 10' rainout shelter was designed for 50% water exclusion; plants under the rainout shelter received 50% less water than those not. The rainout shelters were set in the field when the plants reached the V1 growth stage (first trifoliate). Variables measured were soil chemistry, weekly weather conditions (rainfall precipitation, temperature, soil moisture), agronomic traits (flowering time, days to maturity, plant height), and plant yield (number of harvested plants, total plot biomass, total plot seed weight, seed moisture, 100-seed weight). Precipitation was measured every week with a gutter and barrel system. Rainfall and temperature were recorded and monitored using rain gauges and Hobo temperature loggers (Onset, insert model number). Soil moisture was measured weekly from each subplot using a Field Scout handheld soil moisture probe (manufacturer, model number). Beans were harvested after senescence when they were dried to approximately 18 % of moisture. Biomass yield was calculated by harvesting whole plants from each subplot and weighing them. Seeds were aggregated within a treatment and massed. Statistical analysis for assessing the plant yield differences was conducted by fitting the linear mixed-effects model (LMM) using the lmer() function of the 'lmerTest' package (v3.1.3) (37). Tukey post-hoc test was conducted when the *P*-

value of the effects was less than 0.05 using emmeans() function in 'emmeans package (v1.7.11) (38).

Seed preparation and endophyte microbial DNA extraction

Of the 64 treatments, 63 produced enough seeds to be used for microbiome analysis. The seed numbers varied from 20 to 30 seeds per sample, and ten seeds from each treatment were pooled for DNA extraction. The seeds were weighed and surface sterilized following microbial DNA extraction protocol using PBS soaking procedure as described and developed in our previous study (39). We confirmed the effectiveness of surface sterilization procedure by plating 50 ml of the last rinse water on to Trypticase Soy Agar (TSA) and Potato Dextrose Agar (PDA) and incubated them at 30 ˚C for 2-3 days and 25-26 ˚C for 5 days, respectively. We discarded any seed sample that had microbial growth on these plates. For each "batch" of DNA extraction samples, we included negative and positive controls. The negative control (buffer only) was used to assess microbial contamination and was continued through the entire microbiome profiling process, from DNA extraction to PCR to sequencing. The positive control was an in-house mock community to assess the success of our extraction protocol (39, 40). The DNA extracted from the seed samples were quantified using Qubit™dsDNA BR Assay Kit (ThermoFisher Scientific, Waltham, MA, United States).

PCR amplification and amplicon sequencing

Analysis of seed endophytic bacteria and archaea was performed using PCR amplification of the V4 region of 16S rRNA gene. The universal primer pairs used for PCR amplification were 515f (5'-GTGCCAGCMGCCGCGGTAA-3') and 806r (5'-

GGACTACHVGGGTWTCTAAT-3')(41). The PCR amplification was performed under the following conditions: 94°C for 3 min, followed by 35 cycles of 94°C (45 s), 50°C (60 s), and 72°C (90 s), with a final extension at 72°C (10 min). The amplification was performed in 25 µl mixtures containing 12.5 µl GoTaq®Green Master Mix (Promega, Madison, WI, United States), 0.625 µl of each primer (20 µM), 2 µl of DNA template (1-25 ng per µl), and 9.25 µl nuclease free water. Amplicon library preparation and sequencing were conducted at the Environmental Sample Preparation & Sequencing Facility, Bioscience Division, Argonne National Laboratory using the Illumina MiSeq platform v2 Standard flow cell. The sequencing was performed in a 250 x 250-bp cycles. Additional negative sequencing controls for library preparation were provided by the sequencing facility and included with each sequencing run.

Sequencing analysis and OTU clustering

Bioinformatic analysis of 16S rRNA gene amplicon sequence workflow was performed with QIIME 2 (v2021.4) (42). Demultiplexed paired end raw sequences data were denoised, dereplicated, chimera-removed, and quality filtered using DADA2 plugin (43) implemented on QIIME 2 using '*qiime dada2 denoise-*' command. Before denoising, we assessed the Q-score distribution of our raw sequence data to determine the trimming parameters with FIGARO (Zymo Research (44)). The goal of the FIGARO tool to optimize the trimming of low-quality reads, while maintaining enough overlap for the optimum merging forward and reverse sequences. Operational Taxonomic Unit (OTU) clustering at 99 % of sequence identity threshold was conducted using open reference strategy using *q2-vsearch* plugin (v2021.4.0) (45). In this open reference OTU clustering strategy, all denoised and quality filtered reads first were matched to the reference SILVA database (v.138) (46). Then, reads that did not match to the

reference database were subsequently clustered *de novo*. Finally, closed reference and *de novo* OTUs were combined into a full set of representative sequences.

Taxonomy was assigned through *q2-feature-classifier* plugin (47) using machine-learning based classification method (*classify-sklearn* method) with a Naive Bayesian classifier. Taxonomy assignment was performed at a minimum confidence of 0.8 using pre-trained classifier with SILVA database (v.138) as the reference (46-48). Plant contaminants (chloroplast and mitochondria) and unassigned taxa were removed from the OTU table and the representative sequences using '*qiime taxa filter-table*' and '*qiime taxa filter-seqs*' commands, respectively. Representative sequence alignment was conducted using Multiple Alignment using Fast Fourier Transform (MAFFT) (49). Filtering the potential microbial contaminants from the OTU table was conducted in R (v4.1.2) (50) using the 'microDecon' package (51). Reads were normalized using Cumulative Sum Scaling (CSS) method in metagenomeSeq 'Bioconductor' package in R (52).

Seed-associated microbial community analysis

Ecological analyses of the microbial communities were conducted in R (v4.1.2) (50). Community alpha and beta diversity were calculated on the contaminant-filtered and CSS-normalized OTU table using the vegan package (v2.5-7) (53). Bacterial and archaeal alpha diversity analysis was performed using richness or count of observed OTU and Faith's phylogenetic diversity. Statistical analysis was conducted to investigate the effect of treatment, cultivar, and location on the seed endophyte alpha diversity by fitting the linear mixed-effects model (LMM) using the lmer() function of the 'lmerTest' package (v3.1.3) (37). Treatment, cultivar, and location were treated as fixed factors, and block was treated as a random factor.

Based on the randomized split-plot design, the model used in the study using the lmer function (response variable ~ location × cultivar × treatment + (1 | block/cultivar), data = data). The Tukey post-hoc test was conducted when the *P-value* of the effects was less than 0.05 using emmeans() function in 'emmeans package (v1.7.11) to test which levels were significantly different (38).

The seed endophytes composition and relative abundance were analyzed using the 'Phyloseq' package (v1.38.0) in R (54). Beta diversity analysis was assessed using Jaccard index, which is based on presence-absence (unweighted). As in our previously published analyses, we used an unweighted resemblance to be conservative because most members of the seed microbiome are inactive or dormant (55), thus relative abundance within the seed is not the direct outcome of competitive growth advantages *in situ* (39). The Jaccard distance metric was calculated using vegdist() function in 'vegan'. Principal coordinate analysis (PCoA) plot was used for visualization of the beta diversity analysis. Permutational multivariate analysis of variance (PERMANOVA) (permutations = 999) using the function adonis() from the 'vegan' package and nested.npmanova() from the 'BiodiversityR' package (56) was performed to assess the differences of seed endophyte community structure among treatments, cultivars, and locations. The homogeneity of dispersion (variance) among groups was tested using multivariate analysis using the function betadisper() from the 'vegan' package (53). We performed PERMDISP to test the significant differences in dispersions between groups and Tukey's HSD test to determine which groups differ in relation to the dispersions (variances). We investigated core taxa of seed endophyte community by calculating the occupancy or the proportion of samples in which the taxa are detected. Taxa that were shared and detected in all samples (occupancy = 1) were defined as core taxa (57). We assessed enriched and depleted taxa between

two treatments (with and without rainout shelter) by calculating log_2 fold change in relative abundance using 'DESeq2' package (v1.34.0) (58).

Data and code availability

The computational workflows for sequence processing and ecological statistics are available on GitHub (https://github.com/ShadeLab/PAPER_Bintarti_2021_Bean_Rainoutshelter).

Results and Discussion

Analysis of plant yield showed no differences in plant yield between Lower and Upper Peninsula (LMM, p-value = 0.1). However, we observed differences of plant yield among cultivars in both locations (**Figure 4.1**, LMM, df = 3, F-value = 16.57, p-value = 2.036e-05). The non-nodulating mutant, R99, has the lowest yield among all cultivars, suggesting that plant association with nitrogen-fixing bacteria is essential for plant productivity. We also detected the influence of water treatment to the plant yield (**Figure 4.2**, LMM, df = 1, F-value = 6.17, p-value = 0.02). The effects of the drought on plant yield were more pronounced in the Upper Peninsula than in the Lower Peninsula. in Upper Peninsula compared to Lower Peninsula. Even though we can observe decreased plant yield under rainout shelter treatment for most cultivars, pairwise analyses between treatments within cultivar showed no differences, meaning that the effect of the treatment on the plant was marginal. The result either suggests the resistance of these bean cultivars to drought stress or the stress caused by the water treatment was moderate. Overall, Rosetta has the highest yield among the four cultivars, and this cultivar seems to be least affected by the drought stress. As previously reported, Rosetta pink bean is well adapted to drought

conditions (33). The pink bean is considered the most drought tolerant bean cultivar and is commonly cultivated in the semiarid western states (33).

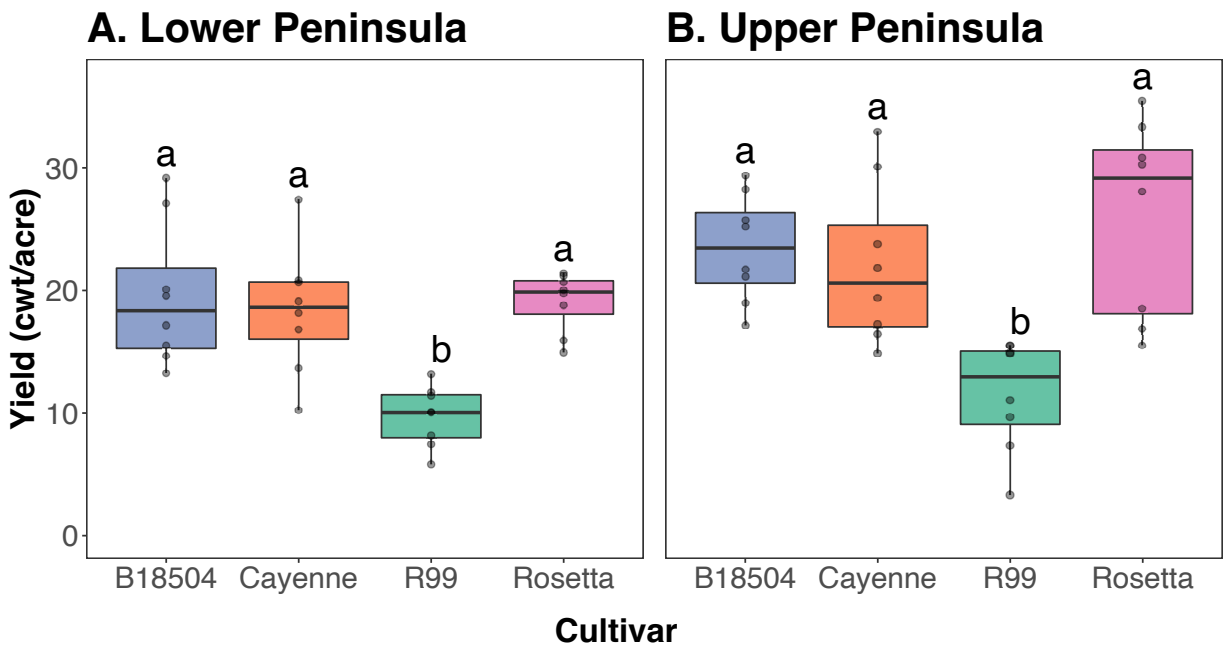

Figure 4.1. Plant yield among cultivars in both locations.

Plant yield (cwt/acre) of for 4 different cultivars planted in (**A**) Lower Peninsula and (**B**) Upper Peninsula. Yield is hundredweight of seeds per acre standardized to 18% moisture content. For each box of the boxplots, circles represent yield calculation per treatment combination. The central horizontal lines represent the mean, the outer horizontal lines of the box represent the 25th and 75th percentiles. Boxes labelled with different letters were significantly different by linear mixed model and post-hoc Tukey's HSD test.

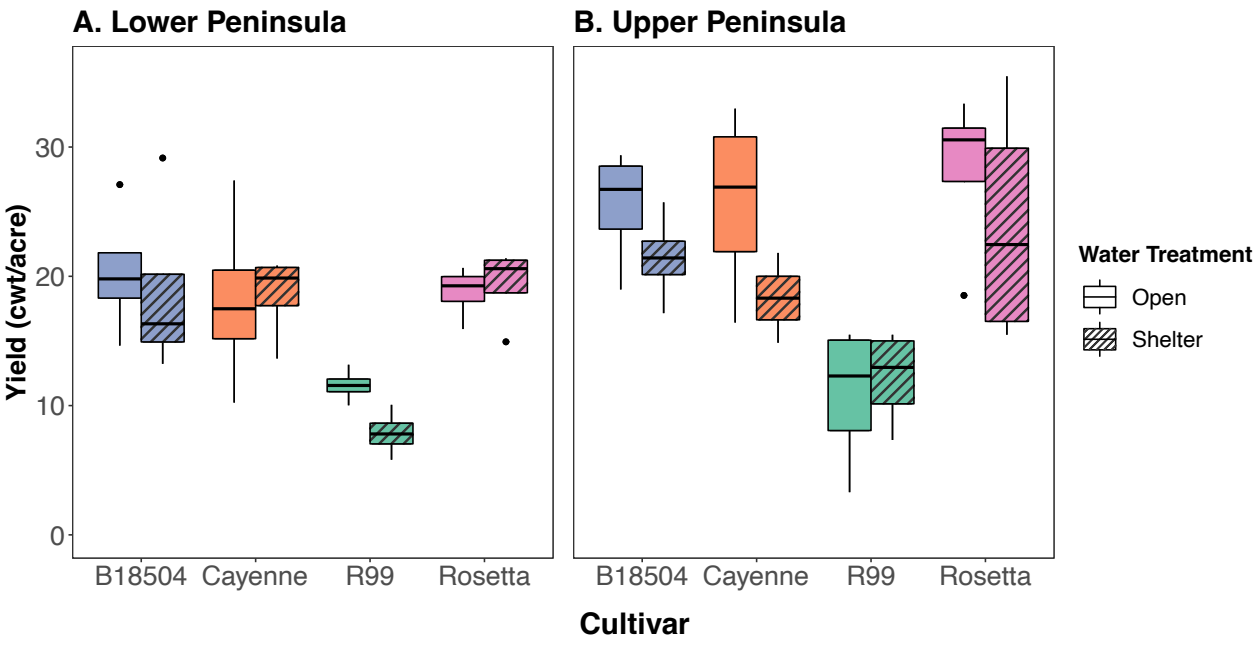

Figure 4.2. Plant yield between treatments within cultivar in both locations.

Effect of rainout shelter treatment to plant's yield (cwt/acre) in two locations (**A**) Lower Peninsula and (**B**) Upper Peninsula. Yield is hundredweight of seeds per acre standardized to 18% moisture content. For each box of the boxplots, the central horizontal lines represent the mean, the outer horizontal lines of the box represent the 25th and 75th percentiles. Pairwise comparison between treatments within the same cultivar was performed using t-test. 'Open' means without rainout shelter and 'Shelter' means with rainout shelter.

Overall, the observed differences in plant yield were due to the difference in plant cultivars. However, water treatment also has some explanatory value for the plant yield. Given these plant yield data, we expect to detect seed microbial community composition and structure differences across different cultivars. Evidence shows that plant genotype shapes the plant microbial communities, specifically the endophytic communities (59, 60). Previous research reported that different from the rhizosphere microbial communities strongly affected by soil edaphic factors, the root endosphere communities are influenced mainly by plant genetics (61). This phenomenon can be explained that endophytes have a very intimate relationship with the host plant, and to colonize and thrive within the plant tissues, endophyte candidates must be able to overcome the host plant's innate immunity (62). The diversity of plant microbiome decreased from the outside plant compartments to inside the plant tissues (60, 61), suggesting microbial selection by the host plant.

Moreover, despite the marginal effect of water treatment on the plant yield, we expect to observe the shift of microbial communities in the seed of treated plants relative to the untreated plants (without rainout shelter). Even though the impact of drought stress was not apparent in the plants, the impact of the stress may be more pronounced in their microbial community structure and composition. Because of the extensive interactions with reciprocal impacts between the plant and its native microbiota, any perturbations that affect the plant may also affect its microbial communities (63, 64). We detected shifts in seed endophyte communities of common bean plants exposed to moderate drought in our previous pilot study (65). Previous studies revealed that drought stress effects are more pronounced on endophytes than rhizosphere microbiome, which is related to the closed interaction between the endophytes with the host plant (60, 66).

Future Directions

Understanding the responses of seed endophytes to drought is a fundamental step in developing strategies to promote plant tolerance to drought stress. Yet, the information on the effect of abiotic stress on plant microbiome or seed endophytes, in particular, is still largely unknown. As bean is an important staple food and crops worldwide, dissecting the plant microbial community interactions in this model crop system will be essential to address critical needs in bean crop production and agricultural sustainability. Specifically, the plant productivity and fitness information across cultivars and drought treatments in this study offer predictions of the endophyte community structure and composition associated with the plant. Future research will be conducted to assess the impact of drought stress on seed endophytes communities. The enriched and depleted taxa will be investigated, and the analysis of crucial members that are responsive to the drought stress can help future research prioritize the particular taxa to develop synthetic communities for application in the field.

APPENDIX

APPENDIX: Contributions to another publication

I have contributed to the following publication during my dissertation work.

> **Bintarti AF**, Kearns PJ, Sulesky-Grieb A, Shade A. Abiotic treatment to common bean plants results in an altered endophytic seed microbiome. *Microbiology Spectrum*: e00210-21 (2022).

My contributions to this works including processing the raw 16S rRNA and ITS1 amplicon sequence data, performing the microbial community analyses, performing the statistical analysis of plant biomass, writing the manuscript, discuss, and revise the manuscript. We observed shifts in the structure and composition of common bean seed endophyte communities of the plants exposed to mild drought and fertilizer treatments relative to the control plants.

REFERENCES

REFERENCES

1. Beebe SE, Rao IM, Blair MW, Acosta-Gallegos JA. 2013. Phenotyping common beans for adaptation to drought. Frontiers in physiology 4:35-35.

2. Losa A, Vorster J, Cominelli E, Sparvoli F, Paolo D, Sala T, Ferrari M, Carbonaro M, Marconi S, Camilli E, Reboul E, Waswa B, Ekesa B, Aragão F, Kunert K. 2022. Drought and heat affect common bean minerals and human diet—What we know and where to go. Food and Energy Security 11:e351.

3. Kazai P, Noulas C, Khah E, Vlachostergios D. 2019. Yield and seed quality parameters of common bean cultivars grown under water and heat stress field conditions. AIMS Agriculture and Food 4:285-302.

4. Backer R, Rokem JS, Ilangumaran G, Lamont J, Praslickova D, Ricci E, Subramanian S, Smith DL. 2018. Plant Growth-Promoting Rhizobacteria: Context, Mechanisms of Action, and Roadmap to Commercialization of Biostimulants for Sustainable Agriculture. Frontiers in Plant Science 9.

5. Ngumbi E, Kloepper J. 2016. Bacterial-mediated drought tolerance: Current and future prospects. Applied Soil Ecology 105:109-125.

6. Meena KK, Sorty AM, Bitla UM, Choudhary K, Gupta P, Pareek A, Singh DP, Prabha R, Sahu PK, Gupta VK, Singh HB, Krishanani KK, Minhas PS. 2017. Abiotic Stress Responses and Microbe-Mediated Mitigation in Plants: The Omics Strategies. Frontiers in Plant Science 8.

7. Coleman-Derr D, Tringe SG. 2014. Building the crops of tomorrow: advantages of symbiont-based approaches to improving abiotic stress tolerance. Frontiers in Microbiology 5.

8. Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V. 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. Proceedings of the National Academy of Sciences 112:E911-E920.

9. Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. Annual Review of Plant Biology 57:233-266.

10. Yin C, Casa Vargas JM, Schlatter DC, Hagerty CH, Hulbert SH, Paulitz TC. 2021. Rhizosphere community selection reveals bacteria associated with reduced root disease. Microbiome 9:86.

11. Egamberdieva D, Wirth SJ, Alqarawi AA, Abd_Allah EF, Hashem A. 2017. Phytohormones and Beneficial Microbes: Essential Components for Plants to Balance Stress and Fitness. Frontiers in Microbiology 8.

12. Jooste M, Roets F, Midgley GF, Oberlander KC, Dreyer LL. 2019. Nitrogen-fixing bacteria and Oxalis – evidence for a vertically inherited bacterial symbiosis. BMC Plant Biology 19:441.

13. Varga T, Hixson KK, Ahkami AH, Sher AW, Barnes ME, Chu RK, Battu AK, Nicora CD, Winkler TE, Reno LR, Fakra SC, Antipova O, Parkinson DY, Hall JR, Doty SL. 2020. Endophyte-Promoted Phosphorus Solubilization in Populus. Frontiers in Plant Science 11.

14. Marasco R, Rolli E, Ettoumi B, Vigani G, Mapelli F, Borin S, Abou-Hadid AF, El-Behairy UA, Sorlini C, Cherif A, Zocchi G, Daffonchio D. 2012. A Drought Resistance-Promoting Microbiome Is Selected by Root System under Desert Farming. PLOS ONE 7:e48479.

15. Poudel M, Mendes R, Costa LAS, Bueno CG, Meng Y, Folimonova SY, Garrett KA, Martins SJ. 2021. The Role of Plant-Associated Bacteria, Fungi, and Viruses in Drought Stress Mitigation. Frontiers in Microbiology 12.

16. Ma Y, Dias MC, Freitas H. 2020. Drought and Salinity Stress Responses and Microbe-Induced Tolerance in Plants. Frontiers in Plant Science 11.

17. Rho H, Van Epps V, Wegley N, Doty SL, Kim S-H. 2018. Salicaceae Endophytes Modulate Stomatal Behavior and Increase Water Use Efficiency in Rice. Frontiers in Plant Science 9.

18. Hereme R, Morales-Navarro S, Ballesteros G, Barrera A, Ramos P, Gundel PE, Molina-Montenegro MA. 2020. Fungal Endophytes Exert Positive Effects on Colobanthus quitensis Under Water Stress but Neutral Under a Projected Climate Change Scenario in Antarctica. Frontiers in Microbiology 11.

19. Shade A, Jacques MA, Barret M. 2017. Ecological patterns of seed microbiome diversity, transmission, and assembly. Curr Opin Microbiol 37:15-22.

20. Bergna A, Cernava T, Rändler M, Grosch R, Zachow C, Berg G. 2018. Tomato Seeds Preferably Transmit Plant Beneficial Endophytes. Phytobiomes Journal 2:183-193.

21. Matsumoto H, Fan X, Wang Y, Kusstatscher P, Duan J, Wu S, Chen S, Qiao K, Wang Y, Ma B, Zhu G, Hashidoko Y, Berg G, Cernava T, Wang M. 2021. Bacterial seed endophyte shapes disease resistance in rice. Nature Plants 7:60-72.

22. White JF, Kingsley KL, Butterworth S, Brindisi L, Gatei JW, Elmore MT, Verma SK, Yao X, Kowalski KP. 2019. Seed-Vectored Microbes: Their Roles in Improving Seedling Fitness and Competitor Plant Suppression, p 3-20. *In* Verma SK, White JJF (ed), Seed Endophytes: Biology and Biotechnology doi:10.1007/978-3-030-10504-4_1. Springer International Publishing, Cham.

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

24. Rodríguez CE, Mitter B, Barret M, Sessitsch A, Compant S. 2018. Commentary: seed bacterial inhabitants and their routes of colonization. Plant and Soil 422:129-134.

25. Mano H, Tanaka F, Watanabe A, Kaga H, Okunishi S, Morisaki H. 2006. Culturable Surface and Endophytic Bacterial Flora of the Maturing Seeds of Rice Plants (Oryza sativa) Cultivated in a Paddy Field. Microbes and Environments 21:86-100.

26. Walitang DI, Kim K, Madhaiyan M, Kim YK, Kang Y, Sa T. 2017. Characterizing endophytic competence and plant growth promotion of bacterial endophytes inhabiting the seed endosphere of Rice. BMC Microbiology 17:209.

27. Hardoim PR, van Overbeek LS, Elsas JDv. 2008. Properties of bacterial endophytes and their proposed role in plant growth. Trends in Microbiology 16:463-471.

28. Verma SK, Kingsley K, Irizarry I, Bergen M, Kharwar RN, White JF, Jr. 2017. Seed-vectored endophytic bacteria modulate development of rice seedlings. J Appl Microbiol 122:1680-1691.

29. Kumar K, Verma A, Pal G, Anubha, White JF, Verma SK. 2021. Seed Endophytic Bacteria of Pearl Millet (Pennisetum glaucum L.) Promote Seedling Development and Defend Against a Fungal Phytopathogen. Frontiers in Microbiology 12.

30. USDA-NASS. 2021. U.S. Department of Agriculture, National Agricultural Statistics Service https://quickstats.nass.usda.gov. Accessed 20 March 2022.

31. Kelly JD, Varner GV, Miklas PN, Cichy KA, Wright EM. 2018. Registration of 'Cayenne' Small Red Bean. Journal of Plant Registrations 12:194-198.

32. Kelly JD, Wright EM, Varner GV, Sprague CL. 2019. 'Cayenne': A new small red bean variety for Michigan [E3405]. East Lansing: Michigan State University, MSU Extension.

33. Kelly JD, Varner GV, Cichy KA, Wright EM. 2012. Registration of 'Rosetta' Pink Bean. Journal of Plant Registrations 6:229-232.

34. Kelly JD, Wright EM, Varner GV, Sprague CL. 2012. 'Rosetta': New pink bean for Michigan [E3185]. East Lansing: Michigan State University, MSU Extension.

35. Kelly JD, Awale HE, Wiersma AT, Wright EM. 2021. Registration of 'Adams' black bean. Journal of Plant Registrations 15:253-259.

36. Park SJ, Buttery BR. 2006. Registration of ineffective nodulation mutant R69 and nonnodulation mutant R99 common bean genetic stocks. Crop Science 46:1415+.

37. Kuznetsova A, Brockhoff PB, Christensen RH. 2017. lmerTest package: tests in linear mixed effects models. Journal of statistical software 82:1-26.

38. Lenth R, Singmann H, Love J, Buerkner P, Herve M. 2019. Package 'emmeans'.

39. Bintarti AF, Sulesky-Grieb A, Stopnisek N, Shade A. 2022. Endophytic Microbiome Variation Among Single Plant Seeds. Phytobiomes Journal 6:45-55.

40. Colovas J, Bintarti AF, Grady KL, Shade A. 2022. Do-it-yourself mock community standard for multi-step assessment of microbiome protocols [Manuscript submitted for publication]. Department of Microbiology and Molecular Genetics. Michigan State University.

41. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences 108:4516-4522.

42. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37:852-857.

43. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13:581-583.

44. Weinstein MM, Prem A, Jin M, Tang S, Bhasin JM. 2019. FIGARO: An efficient and objective tool for optimizing microbiome rRNA gene trimming parameters. bioRxiv doi:10.1101/610394:610394.

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

46. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590-6.

47. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6:90.

48. Robeson MS, II, O'Rourke DR, Kaehler BD, Ziemski M, Dillon MR, Foster JT, Bokulich NA. 2021. RESCRIPt: Reproducible sequence taxonomy reference database management. PLOS Computational Biology 17:e1009581.

49. Katoh K, Misawa K, Kuma Ki, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30:3059-3066.

50. Team RC. 2021. R: A language and environment for statistical computing.

51. McKnight DT, Huerlimann R, Bower DS, Schwarzkopf L, Alford RA, Zenger KR. 2019. microDecon: A highly accurate read-subtraction tool for the post-sequencing removal of contamination in metabarcoding studies. Environmental DNA 1:14-25.

52. Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for microbial marker-gene surveys. Nature methods 10:1200-1202.

53. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P, O'Hara RB, Simpson G, Solymos P, Stevens MHH, Szöcs E, Wagner H. 2020. vegan community ecology package version 2.5-7 November 2020.

54. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.

55. Cope-Selby N, Cookson A, Squance M, Donnison I, Flavell R, Farrar K. 2017. Endophytic bacteria in Miscanthus seed: implications for germination, vertical inheritance of endophytes, plant evolution and breeding. GCB Bioenergy 9:57-77.

56. Kindt R, Kindt MR. 2019. Package 'BiodiversityR'. Package for community ecology and suitability analysis 2:11-12.

57. Shade A, Stopnisek N. 2019. Abundance-occupancy distributions to prioritize plant core microbiome membership. Current Opinion in Microbiology 49:50-58.

58. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15:1-21.

59. Liu J, Abdelfattah A, Wasserman B, Wisniewski M, Droby S, Fazio G, Mazzola M, Wu X. 2022. Contrasting effects of genotype and root size on the fungal and bacterial communities associated with apple rootstocks. Horticulture research 9:uhab013.

60. Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ. 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. Proceedings of the National Academy of Sciences 115:E1157-E1165.

61. Brown SP, Grillo MA, Podowski JC, Heath KD. 2020. Soil origin and plant genotype structure distinct microbiome compartments in the model legume Medicago truncatula. Microbiome 8:139.

62. Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. 2017. Interplay Between Innate Immunity and the Plant Microbiota. Annu Rev Phytopathol 55:565-589.

63. Naylor D, Coleman-Derr D. 2018. Drought Stress and Root-Associated Bacterial Communities. Frontiers in Plant Science 8.

64. Berg G, Grube M, Schloter M, Smalla K. 2014. Unraveling the plant microbiome: looking back and future perspectives. Frontiers in microbiology 5:148-148.

65. Bintarti AF, Kearns PJ, Sulesky-Grieb A, Shade A, Veach A. 2022. Abiotic Treatment to Common Bean Plants Results in an Altered Endophytic Seed Microbiome. Microbiology Spectrum doi:doi:10.1128/spectrum.00210-21:e00210-21.

66. Santos-Medellín C, Liechty Z, Edwards J, Nguyen B, Huang B, Weimer BC, Sundaresan V. 2021. Prolonged drought imparts lasting compositional changes to the rice root microbiome. Nature Plants 7:1065-1077.

CHAPTER 5: Conclusions and Future Directions

Summary

This dissertation examined the driving factors that shape the structure of microbial community associated with two important crops for Michigan's agriculture: perennial apple tree and annual common bean plant. In Chapter 2, we investigated the influences of biogeography on the structure of apple root zone microbiome across 20 sites represent the major apple production orchards in Michigan. Moreover, we also looked at the influences of plant genotype, in this aspect, different scion and rootstock cultivars, on the root zone microbiome. Furthermore, since the root zone is a complex ecological niche and heterogenous environment, we assessed multi-trophic root zone microbiomes, including bacteria/archaea, fungi, nematodes, oligochaetes, and mycorrhizal fungi and investigated their potential intertrophic relationship. We observed that orchard location has stronger impact on the root zone microbiome than rootstock, while we found no evidence of scion as a driving factor of root zone microbiome structure. Further, we detected high diversity and high evenness with many rare core microbiome members. We proposed that the diversity and specific structure are typical of perennial trees. Moreover, we suggested that the community-level functional traits may be more important in determining the structure of the community than the composition. While we expected to observe specific multi-trophic interactions, we did not detect particular patterns of intertrophic interactions that were unique of apple root zone microbiome.

In the next two chapters, we investigated endophyte community associated with common bean seed. Chapter 3 and 4 examined the variability of seed endophytes and their responses to drought stress, respectively, as a fundamental knowledge to better understand their importance and role in promoting plant tolerance to the stress. In Chapter 3, we used individual seeds to evaluate the variability of the seed endophyte community across seeds, pods, and plants. We

observed the highest variability of the seed microbiome at the plant level, which indicates that seed microbiome assessment can be conducted by pooling the seeds from the same plant. Furthermore, the difficulty in assessing the seed microbiome encouraged us to develop a microbial DNA extraction protocol and analysis pipeline to explore the seed endophyte community. These findings provide important information for vertical transmission studies because a single seed is considered a unit that carries not only the plant's genetic information but also a set of microbial inoculums from the parent plant, which eventually colonizes a new seedling. Moreover, assessing the seed microbiome at the seed level allows us to determine the observational unit for future seed microbiome assessment.

In Chapter 4, we investigated the response of seed endophyte community to drought stress. This study was conducted in a field setting with and without rainout shelter representing well-watered and drought condition, respectively. We used different cultivars of common bean and the trials were conducted in two different locations in Michigan that have different latitudes (Lower and Upper Peninsula). We found the water treatment has a marginal effect on the plant yield. Given these differences in plant fitness, we expect to observe the differences in seed microbiome across cultivars and the shift of the community when the plant is exposed to drought stress.

Broadly, these works provide critical information to achieve a basis of knowledge for plant microbiome engineering. They demonstrate different driving factors that structure the plant-associated microbiota. The protocol and microbial community analysis pipeline developed in this work can be applied for seed endophyte community assessment to better understand the vertical transmission of either pathogenic or beneficial microbiota over plant generations. Moreover, these knowledges have important implications for future studies related to

manipulation of the microbial community in order to mitigate biotic stress (soil-borne pathogen attacks) as well as abiotic stress (drought) in crop plants.

Future Directions

The motivations behind harnessing beneficial plant-associated microbiome include increased world population, which leads to increased global demand and consumption of crops for food; global climate change; and increased demand for sustainable agriculture (1). Global climate change will have damaging impacts on commercial agricultural production. Global climate change leads to changes in seasonal precipitation and shifts in temperature, which can cause extreme weather, such as drought, which eventually leads to crop losses (2, 3). Moreover, warmer and rainier weather conditions may also contribute to increased plant pathogen attacks (3-5). In addition, the overuse of chemicals as part of agricultural management practices (e.g., to control phytopathogens or weeds) negatively affects the environment, including humans and animals, leading to an increased demand for sustainable agricultural systems (1). To effectively apply the beneficial microbiota in agronomic settings, we need to understand better the driving factors of the microbial community structure and their response to stresses. Our findings in these works are an essential part of the exploitation of plant microbiome in agriculture and provide a base of knowledge for future works. This section describes every study's specific and broad future directions.

The future direction in assessing the root zone microbiome of apple trees is to conduct comparison study between healthy and unhealthy apple orchards to investigate shifts in the structure of root zone microbial communities. Further, the comparison study can allow us to develop a model or prediction of the structure of the root zone microbiome to determine soil

health (soil health assessment) or provide a prognosis for soil-borne disease occurrence. Moreover, since apple trees are perennial crops, it is expected that there will be dynamic changes in the root zone microbiome due to seasonal variations over the plant lifetime. Thus, future temporal assessment of the dynamics of the root zone microbiome season-to-season will inform more reliable microbial community targets characteristic of perennial trees and potentially engage with pathogens or the environment to repress disease.

This study investigated the variability of the seed microbiome under controlled conditions in the growth chamber. However, it is unclear whether the observations in the growth chamber study are also valid for field study. Hence, a field study under standard management practices will provide us with valuable information that is important for seed microbiome assessment in actual-work settings. We assumed that observed shifts of seed endophyte community and drought-enriched taxa would positively impact the host plant in coping with drought stress. Thus, it would be essential to perform a cultivated-dependent study to characterize those enriched taxa to investigate further their beneficial capabilities in alleviating drought stress. Those taxa may help the host plant coping with the drought stress, for example, by inducing plant stress hormones. Further research using metagenomic and metaproteomic approaches to detect functional genes enriched and expressed during drought stress would be essential. Another potential future study is developing beneficial microbial inoculum that can be tested under controlled conditions and field trials to assess their ability to promote plant tolerance to drought stress.

Plant microbiome research's primary goal is to integrate and apply beneficial plant-associated microbial communities into modern agricultural practices to promote plant growth under a range of environmental conditions, enhance resilience to abiotic stress, and combat

pathogen attacks. Successful integration of plant microbiome into agricultural settings requires large-scale research and careful considerations of the interactions between the host plants, their microbes, the environment, and the management practices. The next significant steps and the emerging challenges in microbiome engineering to support agricultural productivity are developing stable synthetic microbial consortia and establishing an effective and efficient application of microbial consortia in an agricultural environment.

Developing host-microbiome model systems for crop plants is an essential platform for dissecting the mechanisms of plant-microbiome interactions prior to incorporation into the field (6). These studies offer valuable models for plant host-microbe interactions in important agricultural perennial and annual crops. The models provide the microbial community structure (diversity, composition, and interactions) in these particular systems. These works also provide public-accessible open resources such as metadata and sequences repositories as part of the model system. Another crucial aspect in developing a model for plant host-microbe interaction is a standardization protocol, data collection, sample processing, and analysis. One main reason the microbial community associated with above-ground plant compartments (e.g., seed, flower, fruit) receives less attention relative to soil or rhizosphere is the technical difficulties of microbial assessment in those parts of the plants (7). This work offers seed endophytes assessment protocols that are expected to be generally applicable to other crops with similar seed features.

Efforts for integrating plant microbiome into agriculture have been conducted by inoculating individual microbial strains into the field. However, the success of conventional microbial inoculation is highly variable field-to-field or season-to-season (8, 9). The inconsistent performances of microbial inoculants are due to the complexity of indigenous microbial communities and the influence of several factors, including the compatibility with plant hosts

and the environment (10). Research has been focused on developing synthetic microbial communities (SynComs) with desired traits that are incorporated with the critical aspect of multiple interactions between microbes, hosts, and the environment. Defining core microbiota or particular taxa responsive to specific environmental stress is an initial step in developing a synthetic community. Identifying microbial key members can reduce the complexity of the microbial community and guide future research by prioritizing certain groups of microbial taxa and validating their functions (6). Synthetic communities are expected to be more stable when applied in the field than single strains inoculation. Therefore, understanding the biotic and abiotic factors that drive the microbiome structure is critical for generating insights into their stability and resilience to establish robust colonization in particular niches.

One microbial application in agriculture is seed treatment and inoculation. The successful application of synthetic community means that the microbial consortia are capable of colonizing and thriving within the plant. Since early colonization determines successful colonization, it is crucial to introduce the microbial community at the very early stage of plant development (e.g., the seed). One major effort to generate stable and robust microbial inoculants is by combining desired traits for plant fitness with ecological traits which are vital for the community colonization and establishment.

REFERENCES

REFERENCES

1. Saad MM, Eida AA, Hirt H. 2020. Tailoring plant-associated microbial inoculants in agriculture: a roadmap for successful application. Journal of Experimental Botany 71:3878-3901.

2. Kumar S, Kumar M, Wakchaure GC, Bhushan B, Meena KK, Singh AK, Gurumurthy S, Rane J. 2021. Chapter 17 - Microbial management of crop abiotic stresses: Current trends and prospects, p 251-260. *In* Kumar A, Droby S (ed), Microbial Management of Plant Stresses doi:https://doi.org/10.1016/B978-0-323-85193-0.00017-6. Woodhead Publishing.

3. Cavicchioli R, Ripple WJ, Timmis KN, Azam F, Bakken LR, Baylis M, Behrenfeld MJ, Boetius A, Boyd PW, Classen AT, Crowther TW, Danovaro R, Foreman CM, Huisman J, Hutchins DA, Jansson JK, Karl DM, Koskella B, Mark Welch DB, Martiny JBH, Moran MA, Orphan VJ, Reay DS, Remais JV, Rich VI, Singh BK, Stein LY, Stewart FJ, Sullivan MB, van Oppen MJH, Weaver SC, Webb EA, Webster NS. 2019. Scientists' warning to humanity: microorganisms and climate change. Nature Reviews Microbiology 17:569-586.

4. Roy BA, Güsewell S, Harte J. 2004. Response of plant pathogens and herbivores to a warming experiment. Ecology 85:2570-2581.

5. Bebber DP, Ramotowski MA, Gurr SJ. 2013. Crop pests and pathogens move polewards in a warming world. Nature climate change 3:985-988.

6. Busby PE, Soman C, Wagner MR, Friesen ML, Kremer J, Bennett A, Morsy M, Eisen JA, Leach JE, Dangl JL. 2017. Research priorities for harnessing plant microbiomes in sustainable agriculture. PLOS Biology 15:e2001793.

7. Singh BK, Trivedi P. 2017. Microbiome and the future for food and nutrient security. Microbial biotechnology 10:50-53.

8. Glick BR. 2012. Plant growth-promoting bacteria: mechanisms and applications. Scientifica 2012.

9. Ke J, Wang B, Yoshikuni Y. 2021. Microbiome Engineering: Synthetic Biology of Plant-Associated Microbiomes in Sustainable Agriculture. Trends in Biotechnology 39:244-261.

10. Vejan P, Abdullah R, Khadiran T, Ismail S, Nasrulhaq Boyce A. 2016. Role of Plant Growth Promoting Rhizobacteria in Agricultural Sustainability-A Review. Molecules (Basel, Switzerland) 21:573.