12-2023

Effect of Plant Genotype on Plant-Microbe Interactions and Multi-Generation Ecosystem Selection of Microbial Communities Associated with Plant Biomass in Arabidopsis Thaliana

Nachiket Shankar
University of Massachusetts Boston

EFFECT OF PLANT GENOTYPE ON PLANT-MICROBE INTERACTIONS AND MULTI-GENERATION ECOSYSTEM SELECTION OF MICROBIAL COMMUNITIES ASSOCIATED WITH PLANT BIOMASS IN *ARABIDOPSIS THALIANA*

A Dissertation Presented

By

NACHIKET SHANKAR

Submitted to the Office of Graduate Studies,

University of Massachusetts Boston,

In fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2023

Molecular Cellular and Organismal Biology Program

EFFECT OF PLANT GENOTYPE ON PLANT-MICROBE INTERACTIONS AND MULTI- GENERATION ECOSYSTEM SELECTION OF MICROBIAL COMMUNITIES ASSOCIATED WITH PLANT BIOMASS IN *ARABIDOPSIS THALIANA*

A Dissertation Presented

By

NACHIKET SHANKAR

Rick Kesseli, Professor
Chairperson of Committee

Michael Shiaris, Professor Emeritus
Member

Douglas Woodhams, Professor
Member

Karen Ricciardi, Associate Professor
Member

Linda Huang, Program Director
Molecular, Cellular, and Organismal Biology Program

Rachel Skvirsky, Chairperson
Biology Department

ABSTRACT

EFFECT OF PLANT GENOTYPE ON PLANT-MICROBE INTERACTIONS AND MULTI- GENERATION ECOSYSTEM SELECTION OF MICROBIAL COMMUNITIES ASSOCIATED WITH PLANT BIOMASS IN *ARABIDOPSIS THALIANA*

December 2023

Nachiket Shankar, B.E., MS Ramaiah Institute Of Technology
PhD, University of Massachusetts Boston

Directed By Professor Rick Kesseli

The microbiome's role in shaping host phenotypes is a critical area of investigation, with implications for ecology, evolution, and host health. Dynamic plant-microbe interactions are influenced by factors like soil type, environment, and genotype. Understanding their impact on microbial communities is key for tailored plant benefits. An artificial ecosystem selection experiment was done for eight generations with *Arabidopsis thaliana Ler* and Cvi. This revealed distinct microbial communities shaped by genotypes and biomass treatments. Initially, environment dominated, but over time, genotype and biomass gained influence, explaining ~40% of the variation. Moreover, genotype-specific rhizobacterial associations were observed, enhancing understanding of community dynamics and genetics, with potential for agricultural applications.

Plant genes interact with microorganisms, fostering beneficial or antagonistic relationships. This control affects microbe abundance, aiding nutrient uptake, disease defense, and stress tolerance. Despite a century of research, our understanding of these genetic mechanisms remains limited. Our study focused on using near-isogenic lines (NILs) obtained from crossing *Arabidopsis thaliana* L*er* and Cvi to narrow down a plant genomic region previously identified in the lab. This region spans approximately 3.75 Mbp on chromosome 1 and houses around 995 genes. Employing fine-mapping with near-isogenic lines and metagenomic data, we confirm a small yet significant genotype impact on microbial community structure, identifying genotype-specific microbial taxa abundance. Our work reduces the candidate region to 418 genes, advancing insight into the genetic control imposed by *Arabidopsis thaliana* on the microbiome.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Rick Kesseli, for being an inspirational teacher, scientist, and guide over the past eight years. I feel honored to have had you as my advisor and am truly grateful for the research opportunity.

My sincere appreciation goes to my committee members, Doug, Karen, Michael, and Rick, for their guidance, support, valuable insight, and advice. I would also like to extend my thanks to my collaborators, Prateek Shetty and Maroti Gergely, who provided invaluable assistance during my research.

A special acknowledgment goes to James Allen from the greenhouse, without whom I would not have been able to grow the thousands of plants required for my experiments. I also express my gratitude to the Biology Department Staff, Alexa MacPherson, Maria Mahoney, Alana L Boyle, Rebecca Lawless, and Sarah Yellamaty for all their help.

I want to convey my appreciation to all past and present members of the Kesseli lab, including Stuart Morey, Lisa Mazumder, Anastasia Mozharova, Emmi Kurosawa, Dina Tsirelson, and Trudi Gulick, for their unwavering support and friendship.

Lastly, I want to extend my heartfelt thanks to my family—my parents, Ambika, Prasan, Mahesh—and my loving wife, Sneha, for their constant support, encouragement, and love over the years. Thank you all for always being there.

PREFACE

This thesis is divided into three chapters, and each chapter deals with plant-microbe interactions in *Arabidopsis thaliana*. Chapter 1 gives a broad overview of plant-microbe interactions and encompasses the scope and importance of study in this field. This chapter builds up to the two main studies addressed in the thesis and provides the basis for the dissertation.

Chapter 2, The role of microbial community assembly on the plant is highlighted. I utilized, two common inbred genotypes of *Arabidopsis thaliana*, Cvi and L*er*, to carry out generational ecosystem selection for a total of eight generations. The composition of the resulting microbial community was shaped by a complex interplay between environmental factors, genotypes, and biomass selection treatment (high or low biomass selection). In the initial phases of the experiment, the environment played a dominant role in shaping the microbial community composition, while the genotype and biomass treatment had modest but significant impacts. Over time, the plant genotype and selection strategy gained more influence, ultimately leading to the plant genotype becoming the primary determinant of the microbial community composition. This study enhances our understanding of both the microbial community dynamics and plant genetics and can be applied to agricultural settings to improve plant growth and in soil restoration.

Chapter 3 aims to narrow down the candidate regions involved in plant genetic control of the microbiome in *Arabidopsis thaliana* L*er* based on a previously identified quantitative trait loci (QTL) region. The identified QTL region spans approximately 3.75 Mbp at the beginning of chromosome 1 of *Arabidopsis thaliana* L*er* consisting of ~995

genes. A highly replicated (12 replicates) fine-mapping approach using near-isogenic lines (NILs) is employed to analyze 16S rRNA metagenomic data coupled with statistical analyses. The study confirmed the findings from previous studies that have shown a significant but small effect of plant genotype on the microbial community structure. In addition, we found evidence of genotype-specific differential abundance of specific microbial taxa. Ultimately, we narrowed down the candidate region from ~995 genes to 418 genes. This study is a step forward in deciphering the mechanisms behind plant genetic control in *Arabidopsis thaliana*.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS vi

PREFACE vii

LIST OF TABLES xi

LIST OF FIGURES xii

LIST OF SUPPLEMENTARY FIGURES AND TABLES xiv

CHAPTER PAGE

1. CHAPTER 1 INTRODUCTION 1

Introduction 1

The “holobiont” concept 1

Impacts on plant nutrition, health, and abiotic stress 2
Nutrient uptake and acquisition 3
Disease resistance 6
Abiotic stress 8

Harnessing the power of a microbial consortia 9

Microbial community assembly 11

Plant genetic control 12

Goals of this dissertation 13
Part 1 13
Part 2 14

Literature Cited Chapter 1 17

2. CHAPTER 2 MULTI-GENERATION ECOSYSTEM SELECTION OF RHIZOSPHERE MICROBIAL COMMUNITIES ASSOCIATED WITH PLANT BIOMASS IN ARABIDOPSIS THALIANA 27

Abstract 27

Introduction 28

Material and Methods 31
Multi-generation selection of soil ecosystem 31
DNA extraction and 16S rRNA amplicon library prep 33
Sequence Data Analysis & Statistics 33

CHAPTER PAGE

Results 36
Plant Biomass 36
Microbial community Composition 37
Alpha Diversity 38
Beta Diversity 38
Sloan Neutral Model 40
Differential Abundance 40

Discussion 42

Supplementary figures and tables 2 57

Literature Cited Chapter 2 62

3. CHAPTER 3 CHARACTERIZATION OF RHIZOSPHERE MICROBIOMES ASSOCIATED WITH NEAR-ISOGENIC LINES DERIVED FROM ARABIDOPSIS THALIANA CVI AND LER USING 16S rRNA METAGENOMIC SEQUENCING 67

Abstract 67

Introduction 69

Material and Methods 71
Plant seeds, growth conditions, and soil 71
DNA extraction and 16s rRNA gene library prep 72
Sequence Data Processing 73
Statistical Analyses 73
Data mining and Gene Ontology 74

Results 75
Microbial community Composition 75
Alpha Diversity 76
Beta Diversity 76
Differential Abundance and Candidate Regions 77

Discussion 79

Literature Cited Chapter 3 104

4. CHAPTER 4 CONCLUSIONS 110

CONCLUSIONS 110

BIOGRAPHICAL SKETCH OF THE AUTHOR 116

LIST OF TABLES

Table Page

Table 3-1 Genomic markers and number of genes for NILs on chromosome 1..... 83

LIST OF FIGURES

Figure Page

1.1 An overview of plant-microbe interactions 16

2.1 Results of the ecosystem selection on above-ground plant biomass 48

2.2 Deviations from the mean biomass 49

2.3 Taxonomic distribution - Phylum level 50

2.4 Taxonomic distribution - Class level 52

2.5 Alpha diversity 52

2.6 Principle Coordinate Analysis 53

2.7 R-squared value obtained from the Adonis2 test 54

2.8 Burns neutral model 55

2.9 Differential abundance at family level 56

3.1 Cvi, Ler and Near isogenic lines on chromosome 1 84

3.2 Rhizotube apparatus 85

3.3 Above ground biomass of Cvi, Ler and NILs 86

3.4 Growth stage at day 32 for Cvi, Ler and NILs 87

3.5 Taxonomic distribution at Phylum level 88

3.6 Taxonomic distribution at Class level 89

3.7 Alpha diversity of Cvi, Ler and NILs 90

3.8 Principle coordinate analysis for weighted and unweighted unifrac 91

3.9 R-squared value obtained from Pairwise PERMANOVA 92

3.10 Differentially abundant OTUs on chromosome 1 93

Figure Page

3.11 Heatmap of differentially abundant OTUs between N1-2.1 and N1-2.5 94

3.12 Heatmap of differentially abundant OTUs between N1-2.8 and N1-3 95

3.13 Transgressive segregation of microbes between NILs and parents 96

LIST OF SUPPLEMENTARY FIGURES AND TABLES

Figure Page

Supplementary Figure S2. 1 Rhizotube apparatus 57

Table Page

Supplementary Table 2. 1 Statistical comparisons of biomass for Ler samples 58

Supplementary Table 2. 2 Statistical comparisons of biomass for Cvi samples 58

Supplementary Table 2. 3 ANOVA for Observed alpha diversity index 59

Supplementary Table 2. 4 ANOVA for Chao1 alpha diversity index 59

Supplementary Table 2. 5 ANOVA for Shannon alpha diversity index 60

Supplementary Table 2. 6 ANOVA for Simpson alpha diversity index 60

Supplementary Table 2. 7 Linear mixed effect model for Shannon diversity index 61

Supplementary Table 2. 8 Linear mixed effect model for Simpson diversity index 61

Supplementary Table S3. 1 Statistical comparison of biomass of NILs, Cvi and Ler 97

Supplementary Table S3. 2 Summary of ANOVA for biomass 97

Supplementary Table S3. 3 Statistical comparison for observed number of OTUs 98

Supplementary Table S3. 4 Statistical comparisons for Chao1 index 99

Supplementary Table S3. 5 Statistical comparisons for Shannon index 100

Supplementary Table S3. 6 Statistical comparisons for Simpson index 101

Supplementary Table S3. 7 PERMANOVA for unweighted unifrac distances 102

Supplementary Table S3. 8 PERMANOVA for weighted unifrac distances 102

Table Page

Supplementary Table S3. 9 Comprehensive results from differential abundance analyses between NILs 103

CHAPTER 1

REVIEW OF PLANT-MICROBE INTERACTIONS IN THE CONTEXT OF MICROBIAL COMMUNITY ASSEMBLY, PLANT GENETIC CONTROL, AND ECOSYSTEM SELECTION

Introduction

Plant hosts harbor a diverse group of microorganisms residing in the region surrounding the roots (rhizosphere) and extending up to the leaves (phyllosphere), collectively known as the plant microbiome. These microorganisms can establish complex relationships with plants and provide a variety of fitness benefits to the plant, including but not limited to, disease resistance (1–3), growth promotion (4,5), nutrient uptake (6,7), and stress tolerance (8–10). The practical implications of plant-microbe interactions on agriculture, and the ecology and evolution of plants, have sparked a keen interest in understanding the factors that influence and shape plant-associated microbiota. The emphasis of this chapter is on the biotic and abiotic factors that influence the rhizosphere microbiome (RM) and the advantages and detrimental impacts the RM has on plant growth, disease resistance, and nutrient acquisition. Further, we explore the impact of plant genetic control on the RM and the factors influencing microbial community assembly.

The "holobiont" concept

The conventional understanding of a host phenotype is centered around the interplay between genotype and environment, resulting in observable physical traits or growth patterns. However, in the past two decades, research on the microbiome has emphasized the significance

of the host's microbial community as an essential component of the plant's environment and of this equation. The role of the microbiome in shaping the host phenotype has emerged as a critical area of investigation, for diverse fields such as ecology, and host health.

The term "holobiont" was initially conceptualized as the plant host and its associated symbionts (11). However, today the holobiont has expanded to include the entire microbiome as a unit of selection (12). The plant microbiome represents a rich source of functional diversity that is not encoded within the host genome. The interactions between the plant host and its microbiome are dynamic and reciprocal, with the plant shaping its immediate environment through exuding specific metabolites, thereby promoting the growth of specific microbial taxa, while the microbiome in turn influences plant health and growth (Fig. 1.1).

Impacts on plant nutrition, health, and abiotic stress

Plant-microbe interactions encompass a spectrum of both beneficial and harmful outcomes that play pivotal roles in shaping ecosystems and agriculture. Plant pathogens can lead to diseases, causing reduced crop yields and economic losses and parasitic plants exploit host resources, impacting plant fitness and community dynamics. In recent years research has been directed toward harnessing the beneficial aspects of these interactions. The sections below review the types of interactions, their underlying mechanisms, and strategies that harness the benefits of the plant microbiome.

The importance of the potential beneficial effects of plant-associated microbiota has been recognized in the fields of agriculture, ecology, and evolution. For instance, nitrogen-fixing bacteria, like *Rhizobium* species, convert atmospheric nitrogen into a usable form for

plants, contributing to enhanced nitrogen availability. Regarding plant health, some members of the microbiome act as a defensive shield, providing protection against pathogens. For example, *Bacillus* species, produce antimicrobial compounds, effectively suppressing harmful pathogens (13,14) and *Trichoderma*, a genus of fungi, is known for its biocontrol abilities against plant diseases (15,16). When it comes to abiotic stress, the microbiome can play a vital role in enhancing plant resilience (17). For instance, plant growth-promoting rhizobacteria (PGPR) can improve plant tolerance to drought stress by producing stress-related hormones and promoting root growth (18). As a result, extensive research has been done to better understand the role of the microbiome in the acquisition of essential nutrients (nitrogen, phosphorus, and iron), conferring resistance to diseases and mitigating abiotic stress. The section below delves deeper into each of these aspects.

Nutrient uptake and acquisition

The microbiome plays a crucial role in plant nutrient acquisition by enhancing nutrient availability and uptake. Beneficial microorganisms in the rhizosphere and plant tissues facilitate nutrient solubilization, fixation, and cycling, making essential nutrients more accessible to plants. Additionally, some microbes form symbiotic associations, such as mycorrhizal fungi and nitrogen-fixing bacteria, which directly contribute to plant nutrient uptake and utilization. Overall, the microbiome acts as a vital partner for plants, promoting their nutrient acquisition and supporting their growth and health.

Nitrogen

Nitrogen is an essential macronutrient for plant growth and productivity. Nitrogen-fixing bacteria or "Diazotrophs" are capable of fixing nitrogen by converting atmospheric nitrogen to ammonia (7). More importantly, biological nitrogen fixation (BNF) is a trait exclusive to prokaryotes. Depending on the type of diazotroph, BNF can occur in soil and on plant roots (non-symbiotic or associative) or within plant root nodules (symbiotic).

Associative diazotrophs are mostly present in the soil or reside on the surface of roots and do not form nodules. They are omnipresent and genetically diverse, belonging to different genera such as *Axoarcus*, *Axospirilium*, *Burkholderia*, *Citrobacter*, among others (19). They fix very low amounts of nitrogen in the soil (1–20 kg N ha^{-1} $year^{-1}$) (20).

Symbiotic diazotrophs are typically involved in nitrogen fixation within specialized structures on plant roots called nodules. *Rhizobia* and *Frankia* are among the most widely researched diazotrophs in root nodulation. More than 150 kg N ha^{-1} of nitrogen is reported to be fixed by these kinds of associations, which is very high and comparable to nitrogen added by fertilizers (21). *Rhizobia* are the common symbionts of leguminous plants and often form strong mutualistic relationships (22). Some examples of identified nitrogen-fixing bacteria include *Rhizobium*, *Bradyrhizobium*, *Burkholderia*, *Achromobacter*, and *Gluconacetobacter diazotrophicus* (23–27). *Frankia* on the other hand associates with non-legumes. *Frankia* is a filamentous, gram-positive *Actinomycetes* and can symbiotically associate with about 220 species of dicots (28). However, the use of nitrogen-fixing bacteria in agriculture is still limited to certain strains and crops, such as *Gluconacetobacter diazotrophicus* in sugarcane and *Rhizobium leguminosarum* in leguminous plants (29).

Phosphate

Arbuscular mycorrhizal (AM) and Ectomycorrhizal (EM) fungi are a type of beneficial soil fungi that establish a symbiotic relationship with the roots of the majority of plants (30). Their fine and extensive hyphae allow them to absorb nutrients, such as phosphorous, that would otherwise be inaccessible to the plant. The fungus then passes the nutrients to the plant in exchange for sugars (31–35). However, not all plants can form a symbiotic relationship with AM or EM fungi, and non-mycorrhizal plants have alternative adaptations that allow colonization by other fungi, such as *Colletotrichum tofieldiae* (36). Recent research has demonstrated that in conditions of low phosphorus, suppressing the plant's immune system leads to the development of an inorganic-phosphate assimilating fungal microbiome that enhances the growth of non-mycorrhizal plants in phosphorus-deficient soils (9,37).

Iron

Iron is an essential nutrient for plants, but it is often found in soil in forms that are not readily available to plants. Iron is involved in many important processes in plants, including producing chlorophyll, maintaining cell membrane stability, and regulating oxidative stress. Microbial siderophores are small molecules produced by microbes and can chelate iron, making it more available to plants (38). By chelating iron, microbial siderophores can increase the uptake of iron by plant roots (39–41). By increasing the availability of iron to the plant, microbial siderophores can improve the growth and health of plants, especially in soils that are low in iron (42–45). In the face of iron limitation, plants have evolved two methods of acquiring this nutrient. They either make inorganic iron more soluble in the area surrounding

the roots (Strategy I) or release phytosiderophores which are then transported into the root tissue through a specific system (Strategy II) (46). Some microbial siderophores can outcompete pathogens for iron, thus reducing pathogen growth and potential harm to the plant (47). For example, Coumarins, plant-derived secondary metabolites, shape the root bacterial community of *Arabidopsis thaliana* by promoting iron mobilization and suppressing a common *Pseudomonas* species that compete with the plant for iron through reactive oxygen species (48).

Disease resistance

The idea that bacteria can protect plants against pathogens is referred to as "bacterial biological control" or "biocontrol". This concept has been around for a long time. In 1995, Ryan et al. showed that *Pseudomonas spp.* is effective against take-all disease, a soil-borne disease affecting a wide range of crops including wheat, barley, and carrots (49). There are several mechanisms by which microorganisms can ward off plant pathogens. These include but are not limited to competition for nutrients and space, antibiotic production, induced plant systemic acquired resistance, and production of quorum sensing (QS) and volatile organic compounds (VOCs).

Microorganisms that are beneficial to plants can compete with pathogens for essential nutrients, trace elements, and space, effectively reducing the growth and colonization of the pathogen. For example, the presence of *Pseudomonas spp*. has been shown to reduce the growth of microorganisms that cause plant diseases and compete with them for resources (1–3). Rhizosphere fungi are also known to produce several secondary metabolites that have

antibiotic properties, some of the most well-known ones are *Aspergillus nidulans* which produce penicillin-like β-lactams and *Cephalosporium acremonium* that produce cephalosporins (50,51). Some beneficial microorganisms produce antibiotics that can inhibit the growth of pathogens for example, *Pseudomonas* strains produce 2,4-Diacetylphloroglucinol (DAPG) which are toxic to plant pathogenic fungi (52–54).

Beneficial microorganisms can induce the plant's systemic acquired resistance (SAR), a mechanism that provides long-lasting protection against a wide range of pathogens. For example, mycorrhizal fungi have been shown to suppress SAR, which is a defense response of plants to pathogen infection. SAR is mediated by the salicylic acid (SA) signaling pathway, which results in the production of phytoalexins, pathogenesis-related (PR) proteins, and other defense-related compounds (55–58). The presence of the hormone salicylic acid (SA) also influences the composition of the root microbiome, as plants with altered SA signaling have been shown to have different relative abundances of specific bacterial families (enriched in *Bacteroidetes* and *Actinobacteria* and depleted in *Betaproteobacteria*) compared to wild-type plants. These results indicate that different bacterial strains can respond differently to SA, thereby affecting the microbial community structure of the root (59).

Pseudomonas syringae has been shown to manipulate the jasmonic acid (JA) pathway to successfully establish itself in the plant. For example, *P. syringae* has been demonstrated to secrete effector proteins that can interfere with the JA pathway, blocking the activation of plant defense responses and allowing the bacterium to establish itself in the plant (60). It is also important to highlight that studies have also demonstrated that *Arabidopsis* under pathogenic

attack modulate the exudates released through their roots, leading to an enrichment of *Microbacterium*, *Stenotrophomonas*, and a *Xanthomonas* sp. in the rhizosphere (61).

Alternative mechanisms of plant disease resistance such as VOCs have been receiving more attention over the past years. VOCs play a significant role in plant disease resistance. When plants are exposed to pathogens or pests, they often produce VOCs as part of their defense mechanisms (62). These emitted VOCs act as signaling molecules that help neighboring plants detect the presence of potential threats. Some were shown to modulate plant growth and mediate the intricate dialogues between microorganisms and plants (63).

Abiotic stress

The rhizosphere microbiome can play a critical role in mitigating the effects of abiotic stress such as drought, extreme temperatures, salinity, and heavy metal toxicity that can harm plant growth and survival (64). Plant growth-promoting rhizobacteria (PGPR) with ACC (1-aminocyclopropane-1-carboxylic acid) deaminase activity can help reduce the levels of ethylene in stressed plants, potentially improving their ability to survive and recover from water stress (65,66). PGPR can also help plants cope with salt stress, for example, wheat plants grown in saline soils showed a significant number of bacterial isolates with plant growth-promoting traits and were also tolerant to high levels of NaCl (8%) (67–69). The bacteria *Burkholderia phytofirmans* PsJN has been shown to enhance root growth and physiological activity in grapevine plants, even at temperatures as low as 4°C. It has also been shown that the combination of *Bradyrhizobium japonicum* and *Serratia proteamaculans* was found to

stimulate soybean growth, even at 15°C, where soybean nodule infection and nitrogen fixation are usually inhibited (51,70).

Harnessing the power of a microbial consortia

Traditionally, single microbial isolates were selected based on their ability to directly express functions of interest (71), such as nutrient mineralization, nitrogen fixation, or pathogen suppression (53). Microbial inoculations can also modulate plant responses to abiotic stress such as salinity. While both bacteria and fungi improve plant tolerance to salt stress, fungi play a role of greater importance under abiotic stress (72). Glycophytes, or salt-sensitive plants, depend more on fungal symbiosis to improve nutrient uptake and plant biomass, than halophytes under salt stress (73). Although imported microbes can have immediate impacts, they often fail to establish in required densities, making these effects transitory (74) and in some cases, inoculants are outcompeted by native soil microbes. Microbial inoculations are more successful in reproducing functions of interest if they are introduced as a community or a group of species. This is because diverse inoculants can occupy a wider resource niche, allowing them to compete with native microbes more effectively. For instance, plant inoculation with a consortium of *Pseudomonas* species showed that increased plant biomass and decreased disease frequency were correlated with an increase in the species richness of the inoculation (6,75). The evolutionary history of the applied inoculant also influences the potential success under stress conditions. Salt-stressed mangroves inoculated with endophytes from high-salinity environments showed enhanced tolerance, while endophytes from freshwater environments did not ameliorate the effects of salt stress (76). Rice inoculated with

fungi isolated from coastal habitats showed improved salt tolerance, while rice inoculated with isolates from agricultural soils conferred no tolerance (77). It is also important to highlight that improvement in plant phenotype conferred by microbial inoculations can be transferred across different plant species. For instance, tomato and cucumber plants showed improved drought and salt tolerance when inoculated with fungi isolated from the roots of desert plants that had been exposed to similar stressors (78). Halotolerant bacterial consortia isolated from Avocado trees improved the salt tolerance of Avocado saplings (79) and wheat (80). In fact, researchers have shown it is possible to engineer an artificial synthetic community of five bacterial strains isolated from the rhizosphere of a desert plant to improve salt tolerance in tomatoes (81).

The above-mentioned studies involve direct isolation and transfer of microbes to elicit improved plant tolerance to abiotic stress. Indirect selection of microbial consortia capable of inducing specific traits of interest can be carried out through host-assisted microbiome selection. It also allows for the selection of microbes as an interdependent network rather than as individual isolates allowing even greater persistence in the environment. Host-assisted microbiome selection is a method that involves the transfer of microbiomes with desired effects on a host from one generation to the next. This is done by identifying individuals that display the strongest desired phenotype and using their microbiomes for the next generation. This process helps to perpetuate beneficial microbiomes and improve the host's fitness. Only four studies have used this approach, and all have been carried out in plant-soil systems. These approaches have led to the selective generation of microbiomes that can alter the biomass and flowering time in *Arabidopsis thaliana* (82,83), to improve drought tolerance in wheat (84) and improve salt tolerance in *Brachypodium distachyon* (85)

Microbial community assembly

The rhizosphere microbial community diverges from, yet originates primarily from, the bulk soil. (86–88). A multitude of biotic and abiotic factors influence the rhizosphere microbiome, including but not limited to plant root exudates (89–91), soil type (92–96), environment (97), and various aspects of the plant, including species (98), genotype (99–101) and developmental stage (102,103). Soil type has been shown to have a major effect on the microbial community (104–106), while the impact of plant species, genotype, and developmental stage is relatively less pronounced.

In order to respond to plant cues, rhizosphere microbes that have a higher number of genes that are involved in – chemotaxis, flagella assembly, biofilm formation, and transport – have a selective advantage (107–110). Despite the many factors that influence and shape the rhizosphere microbiome throughout the life cycle of a plant, there is a small fraction of the microbiome that remains present at high relative abundances throughout all developmental stages (111–114).

Previous studies in *Arabidopsis thaliana* have demonstrated the core microbiome as early as 2012 in two different publications (115,116). Both arrived at similar results. The environment played a greater role in offering a reservoir of microbes for the plant genotypes to select from. The environment was a stronger predictor for microbiome structure compared to the genotype. The microbial diversity increased as you moved away from the endophytic compartment and toward the bulk soil. Importantly, microbes found in the bulk soil were not enriched in the endophytic compartment and vice-versa. The microbiome of greenhouse-grown

plants was similar to that of the field-grown plants specifically at different plant compartments – rhizosphere, woody stem, and endophytic compartment, indicating an active role of the plant host in creating and maintaining an environment where certain microbes have improved fitness. There are certain phyla enriched in each compartment and zone of the plant. The study found that several taxa (*Acidobacteria, Verrucomicrobia, Gemmatimonadetes* and *Proteobacterial* families) that were present in both the rhizosphere and the bulk soil, were depleted in the endophytic compartment, which was shown to be dominated by *Actinobacteria*, *Proteobacteria,* and *Firmicutes*.

A follow-up study confirmed that variation in microbiome communities between different *Arabidopsis* varieties and sister species was largely quantitative with host phylogenetic distance not explaining microbiome community variation (117). Despite having a large divergence between different host species, the diversity of the microbes interacting with the plant host at the root endophytic zone was largely consistent and reproducible.

Plant genetic control

By understanding the genetic factors that influence the composition and functioning of the plant-associated microbiome, we can gain insights into the complex interactions between plants and microorganisms. This knowledge enables us to identify key genes that govern beneficial microbial associations. As highlighted in the previous section, most studies have shown a significant, but weak effect of genotype on the rhizosphere microbiome (117–119). Plant roots have the ability to modulate the physiology and biochemistry in the rhizosphere. Root exudates, released by the roots into the nearby soil (the rhizosphere), mediate microbial

community assembly constitute an essential part of this response (120–122). Identifying mechanisms and plant genes associated with the assembly of specific microbial taxa is essential for devising innovative plant breeding strategies that harness beneficial microorganisms to promote plant growth and health. Some studies have suggested the participation of genes associated with plant defense and cross-membrane transport of organic compounds, like ATP-binding cassette (ABC) transporter and secondary metabolites, in influencing the rhizosphere community (60).

Goals of this dissertation

The literature discussed above emphasizes the importance and practical applications of plant-microbe interactions in agriculture, ecology, and evolution. It is apparent that gaining a deep understanding of plant-microbe interactions involves studying processes driven by both microbe-microbe interactions as well as by plant genetic control. This dissertation addresses and furthers the knowledge in both these key areas of research in the field.

Part 1

The second chapter of this dissertation aimed to understand the dynamics of microbe-mediated change of host phenotype through a multi-generation artificial ecosystem selection design.

Artificial ecosystem selection was carried out in *Arabidopsis thaliana* L*er* and Cvi with respect to the phenotypic trait, above-ground plant biomass (selecting for two contrasting

phenotypes of high and low biomass) as previously described by Swenson *et. al.* (82). The selected ecosystems serve as "parents" for the next "generation of the soil", and over multiple generations, the phenotypic trait should shift in the direction of selection i.e., result in either lower or higher biomass of the plant. The root-associated microbial communities were sequenced in each generation and analyzed.

Plant traits such as drought tolerance, flowering time, plant biomass, and salt tolerance can all be modulated by either generational selection or inoculating the plant with microbes sympatric to the drought or salt environments. Mueller *et. al.*, (85) used generation selection to select for microbes that could improve plant seed production by up to 205%. Soil inoculum obtained from drought-exposed soils improved *Arabidopsis* biomass under drought conditions. Older studies such as Swenson et al. (82) have shown similar results without necessarily exploring the microbiome. This study aims to identify the driving factors that determine the microbial community assembly and, in turn, influence plant biomass through longitudinal analysis of next-generation sequencing 16s rRNA metagenomic data.

Part 2

The third chapter of this study addresses the role of plant genetic control on the rhizosphere bacteria community by using a traditional mapping approach with near-isogenic lines (NILs) and *Arabidopsis thalian* L*er* and Cvi. Several studies have attempted to understand the underlying genetic mechanisms giving rise to a specific rhizosphere bacterial community (83,123–128). There are a few studies that have indicated the involvement of genes related to plant defense and cross-membrane transport of organic substances such as ABC

transporters and secondary metabolites in shaping the rhizosphere community (60,129–132). However, despite over a hundred years of research in this area, the underlying plant genetic mechanisms in structuring the rhizosphere are still poorly understood.

Traditional approaches used to understand the effect of plant genetics on specific microbes such as quantitative trait loci (QTL) mapping using recombinant inbred lines (RILs), near-isogenic lines (NILs), or existing mutant lines can narrow down large candidate chromosomal regions conferring respective traits in their respective model systems (133–135). This study aims to narrow down the candidate regions in *Arabidopsis thaliana* L*er* based QTL region previously identified by a lab member, Stuart Morey (Boston, USA). The QTL region identified spans about 3.75 Mbp. at the end of chromosome 1 of *Arabidopsis thaliana* L*er*. This is a large genomic region consisting of more than ~400 genes. To explore the plant genetic effects of NILs derived from a cross between *Arabidopsis thaliana* L*er* and Cvi parents on the microbiome, I performed a highly replicated study involving 12 biological replicates and analyzed it using 16S rRNA metagenomic sequencing.

Figures Chapter 1

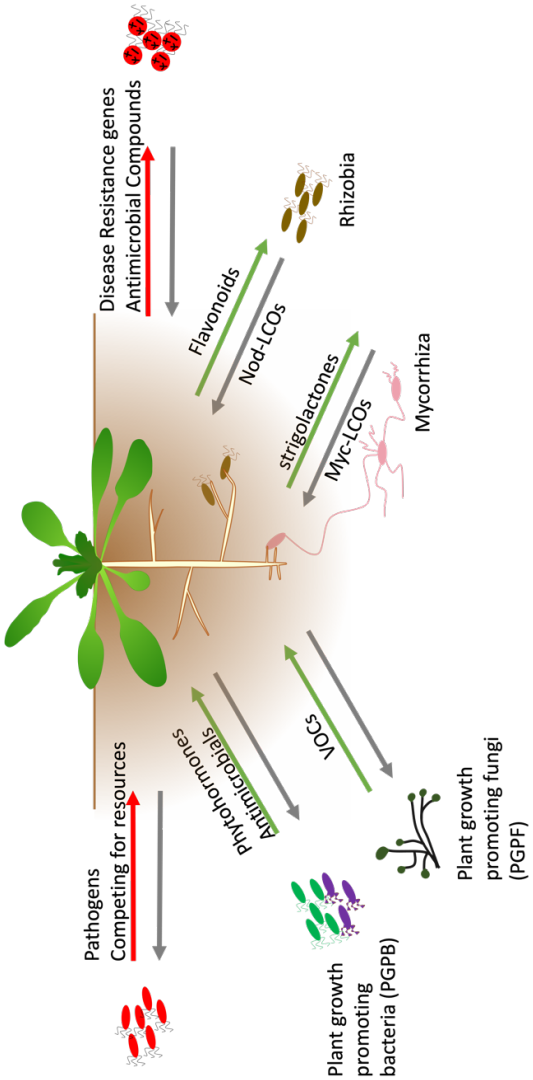


Figure 1.1. An overview of plant-microbe interactions. The left side of the figure indicates some of the well-known microbial directed interactions. The right-side of the figure illustrates some of the well-known plant directed interactions.

Literature Cited Chapter 1

1. Santhanam R, Luu VT, Weinhold A, Goldberg J, Oh Y, Baldwin IT. Native root-associated bacteria rescue a plant from a sudden-wilt disease that emerged during continuous cropping. Proc Natl Acad Sci U S A. 2015;112(36).
2. Trivedi P, Delgado-Baquerizo M, Trivedi C, Hamonts K, Anderson IC, Singh BK. Keystone microbial taxa regulate the invasion of a fungal pathogen in agro-ecosystems. Soil Biol Biochem. 2017;111.
3. Mendes R, Kruijt M, De Bruijn I, Dekkers E, Van Der Voort M, Schneider JHM, et al. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. Science (1979). 2011;332(6033).
4. Ahkami AH, Allen White R, Handakumbura PP, Jansson C. Rhizosphere engineering: Enhancing sustainable plant ecosystem productivity. Vol. 3, Rhizosphere. Elsevier B.V.; 2017. p. 233–43.
5. Gouda S, Kerry RG, Das G, Paramithiotis S, Shin HS, Patra JK. Revitalization of plant growth promoting rhizobacteria for sustainable development in agriculture. Vol. 206, Microbiological Research. 2018.
6. Hu J, Wei Z, Weidner S, Friman VP, Xu YC, Shen QR, et al. Probiotic Pseudomonas communities enhance plant growth and nutrient assimilation via diversity-mediated ecosystem functioning. Soil Biol Biochem. 2017 Oct 1;113:122–9.
7. Dobbelaere S, Vanderleyden J, Okon Y. Plant growth-promoting effects of diazotrophs in the rhizosphere. Vol. 22, Critical Reviews in Plant Sciences. 2003.
8. Vurukonda SSKP, Vardharajula S, Shrivastava M, SkZ A. Enhancement of drought stress tolerance in crops by plant growth promoting rhizobacteria. Vol. 184, Microbiological Research. Elsevier GmbH; 2016. p. 13–24.
9. Castrillo G, Pereira PJ, Teixeira L, Paredes SH, Law TF, De Lorenzo L, et al. Root microbiota drive direct integration of phosphate stress and immunity [Internet]. Vol. 543. 2017. Available from: http://www.nature.com/authors/editorial_policies/license.html#terms
10. Tewari S, Arora NK. Soybean Production Under Flooding Stress and Its Mitigation Using Plant Growth-Promoting Microbes. In: Environmental Stresses in Soybean Production: Soybean Production. 2016.
11. Margulis L. Symbiogenesis and symbionticism. In: Margulis L, Fester R, editors. Symbiosis As a Source of Evolutionary Innovation: Speciation and Morphogenesis. MIT Press; Cambridge, MA. 1991;
12. Zilber-Rosenberg I, Rosenberg E. Role of microorganisms in the evolution of animals and plants: The hologenome theory of evolution. Vol. 32, FEMS Microbiology Reviews. 2008. p. 723–35.
13. Radhakrishnan R, Hashem A, Abd Allah EF. Bacillus: A biological tool for crop improvement through bio-molecular changes in adverse environments. Front Physiol. 2017 Sep 6;8(SEP).

14. de Maagd R. Bacillus thuringiensis-based Products for Insect Pest Control. In: Lugtenberg B, editor. Principles of Plant-Microbe Interactions - Microbes for Sustainable Agriculture [Internet]. Cham: Springer International Publishing; 2015 [cited 2014 Dec 8]. p. 185–92. Available from: http://link.springer.com/10.1007/978-3-319-08575-3
15. Vinale F, Sivasithamparam K, Ghisalberti EL, Marra R, Woo SL, Lorito M. Trichoderma-plant-pathogen interactions. Soil Biol Biochem. 2008;40(1):1–10.
16. Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, et al. Trichoderma: The genomics of opportunistic success. Vol. 9, Nature Reviews Microbiology. 2011. p. 749–59.
17. Khan N, Ali S, Shahid MA, Mustafa A, Sayyed RZ, Curá JA. Insights into the interactions among roots, rhizosphere, and rhizobacteria for improving plant growth and tolerance to abiotic stresses: A review. Vol. 10, Cells. MDPI; 2021.
18. Marasco R, Rolli E, Ettoumi B, Vigani G, Mapelli F, Borin S, et al. A Drought Resistance-Promoting Microbiome Is Selected by Root System under Desert Farming. PLoS One. 2012;7(10).
19. Santi C, Bogusz D, Franche C. Biological nitrogen fixation in non-legume plants. Vol. 111, Annals of Botany. 2013. p. 743–67.
20. Cleveland CC, Townsend AR, Schimel DS, Fisher H, Howarth RW, Hedin LO, et al. Global patterns of terrestrial biological nitrogen (N 2) fixation in natural ecosystems . Global Biogeochem Cycles. 1999 Jun;13(2):623–45.
21. Reed HE, Martiny JBH. Testing the functional significance of microbial composition in natural communities. FEMS Microbiol Ecol. 2007;62(2):161–70.
22. Franche C, Lindström K, Elmerich C. Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. Vol. 321, Plant and Soil. 2009. p. 35–59.
23. Guimarães AA, Jaramillo PMD, Nóbrega RSA, Florentino LA, Silva KB, de Souza Moreira FM. Genetic and symbiotic diversity of nitrogen-fixing bacteria isolated from agricultural soils in the western amazon by using cowpea as the trap plant. Appl Environ Microbiol. 2012 Sep;78(18):6726–33.
24. Zehr JP, Jenkins BD, Short SM, Steward GF. Nitrogenase gene diversity and microbial community structure: A cross-system comparison. Vol. 5, Environmental Microbiology. 2003. p. 539–54.
25. Gaby JC, Buckley DH. A global census of nitrogenase diversity. Environ Microbiol. 2011 Jul;13(7):1790–9.
26. Reis VM, Teixeira KR dos S. Nitrogen fixing bacteria in the family Acetobacteraceae and their role in agriculture. Vol. 55, Journal of Basic Microbiology. Wiley-VCH Verlag; 2015. p. 931–49.
27. Reis VM, Teixeira KR dos S. Nitrogen fixing bacteria in the family Acetobacteraceae and their role in agriculture. Vol. 55, Journal of Basic Microbiology. Wiley-VCH Verlag; 2015. p. 931–49.
28. Granqvist E, Sun J, Op den Camp R, Pujic P, Hill L, Normand P, et al. Bacterial-induced calcium oscillations are common to nitrogen-fixing associations of nodulating legumes and nonlegumes. New Phytologist. 2015 Aug 1;207(3):551–8.

29. dos Reis M, Wernisch L, Savva R. Unexpected correlations between gene expression and codon usage bias from microarray data for the whole Escherichia coli K-12 genome. Nucleic Acids Res [Internet]. 2003 Dec 1 [cited 2014 Jan 14];31(23):6976–85. Available from: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkg897
30. Averill C, Bhatnagar JM, Dietze MC, Pearse WD, Kivlin SN. Global imprint of mycorrhizal fungi on whole-plant nutrient economics. Proc Natl Acad Sci U S A. 2019 Nov 12;116(46):23163–8.
31. Richardson AE, Barea JM, McNeill AM, Prigent-Combaret C. Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms. Vol. 321, Plant and Soil. 2009. p. 305–39.
32. Adeleke RA, Cloete TE, Bertrand A, Khasa DP. Iron ore weathering potentials of ectomycorrhizal plants. Mycorrhiza. 2012 Oct;22(7):535–44.
33. Hawkins HJ, Johansen A, George E. Uptake and transport of organic and inorganic nitrogen by arbuscular mycorrhizal fungi. Vol. 226, Plant and Soil. 2000.
34. Gianinazzi S, Gollotte A, Binet MN, van Tuinen D, Redecker D, Wipf D. Agroecology: The key role of arbuscular mycorrhizas in ecosystem services. Vol. 20, Mycorrhiza. 2010. p. 519–30.
35. Johnson NC, Graham JH. The continuum concept remains a useful framework for studying mycorrhizal functioning. Vol. 363, Plant and Soil. Kluwer Academic Publishers; 2013. p. 411–9.
36. Hacquard S, Garrido-Oter R, González A, Spaepen S, Ackermann G, Lebeis S, et al. Microbiota and host nutrition across plant and animal kingdoms [Internet]. Vol. 17, Cell Host and Microbe. 2015 [cited 2017 Mar 1]. p. 603–16. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1931312815001675
37. Almario J, Jeena G, Wunder J, Langen G, Zuccaro A, Coupland G, et al. Root-associated fungal microbiota of nonmycorrhizal Arabis alpina and its contribution to plant phosphorus nutrition. Proc Natl Acad Sci U S A. 2017 Oct 31;114(44):E9403–12.
38. Marschner P. Marschner's Mineral Nutrition of Higher Plants: Third Edition. Marschner's Mineral Nutrition of Higher Plants: Third Edition. 2011.
39. Crowley DE. Microbial Siderophores in the plant rhizosphere. In: Iron Nutrition in Plants and Rhizospheric Microorganisms. 2006.
40. Vansuyt G, Robin A, Briat JF, Curie C, Lemanceau P. Iron Acquisition from Fe-Pyoverdine by Arabidopsis thaliana. / 441 MPMI. 2007;20(4):441–7.
41. Lemanceausoils P, Bauer P, Kraemer S, Briat JF. Iron dynamics in the rhizosphere as a case study for analyzing interactions between soils, plants and microbes. Vol. 321, Plant and Soil. 2009. p. 513–35.
42. Buckling A, Harrison F, Vos M, Brockhurst MA, Gardner A, West SA, et al. Siderophore-mediated cooperation and virulence in Pseudomonas aeruginosa. Vol. 62, FEMS Microbiology Ecology. 2007. p. 135–41.
43. Lindsay WL, Schwab AP. The Chemistry of Iron in Soils and its Availability to Plants. J Plant Nutr. 1982;5(4–7).
44. Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial iron homeostasis. Vol. 27, FEMS Microbiology Reviews. Elsevier; 2003. p. 215–37.

45. Hider RC, Kong X. Chemistry and biology of siderophores. Vol. 27, Natural Product Reports. 2010.
46. Walker EL, Connolly EL. Time to pump iron: iron-deficiency-signaling mechanisms of higher plants. Vol. 11, Current Opinion in Plant Biology. 2008. p. 530–5.
47. Gu S, Wei Z, Shao Z, Friman VP, Cao K, Yang T, et al. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. Nat Microbiol. 2020 Aug 1;5(8):1002–10.
48. Voges MJEEE, Bai Y, Schulze-Lefert P, Sattely ES. Plant-derived coumarins shape the composition of an Arabidopsis synthetic root microbiome. Proc Natl Acad Sci U S A. 2019 Jun 18;116(25):12558–65.
49. Ryan CA, Lamb CJ, Jagendorf AT, Kolattukudy PE, Cook RJ, Thomashow LS, et al. "Self-Defense by Plants: Induction and Signalling Pathways," organized by Molecular mechanisms of defense by rhizobacteria against root disease [Internet]. Vol. 92. 1995. Available from: https://www.pnas.org
50. Hoffmeister D, Keller NP. Natural products of filamentous fungi: Enzymes, genes, and their regulation. Vol. 24, Natural Product Reports. 2007. p. 393–416.
51. Ait Barka E, Nowak J, Clément C. Enhancement of chilling resistance of inoculated grapevine plantlets with a plant growth-promoting rhizobacterium, Burkholderia phytofirmans strain PsJN. Appl Environ Microbiol. 2006 Nov;72(11):7246–52.
52. Haas D, Défago G. Biological control of soil-borne pathogens by fluorescent pseudomonads. Vol. 3, Nature Reviews Microbiology. 2005. p. 307–19.
53. Lugtenberg B, Kamilova F. Plant-growth-promoting rhizobacteria. Vol. 63, Annual Review of Microbiology. 2009. p. 541–56.
54. Raaijmakers JM, Mazzola M. Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. Vol. 50, Annual Review of Phytopathology. 2012. p. 403–24.
55. De Meyer G, Höfte M. Biological Control Salicylic Acid Produced by the Rhizobacterium Pseudomonas aeruginosa 7NSK2 Induces Resistance to Leaf Infection by Botrytis cinerea on Bean. 1997.
56. Audenaert K, Pattery T, Cornelis P, Höfte M. Induction of Systemic Resistance to Botrytis cinerea in Tomato by Pseudomonas aeruginosa 7NSK2: Role of Salicylic Acid, Pyochelin, and Pyocyanin. Vol. 15. 2002.
57. Barriuso J, Ramos Solano B, Gutiérrez Mañero FJ. Protection against pathogen and salt stress by four plant growth-promoting rhizobacteria isolated from Pinus sp. on Arabidopsis thaliana. Phytopathology. 2008 Jun;98(6):666–72.
58. van de Mortel JE, de Vos RCH, Dekkers E, Pineda A, Guillod L, Bouwmeester K, et al. Metabolic and transcriptomic changes induced in arabidopsis by the rhizobacterium Pseudomonas fluorescens SS101. Plant Physiol. 2012;160(4):2173–88.
59. Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M, et al. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. Science (1979) [Internet]. 2015;349(6250):860–4. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.aaa8764

60. Millet YA, Danna CH, Clay NK, Songnuan W, Simon MD, Werck-Reichhart D, et al. Innate immune responses activated in Arabidopsis roots by microbe-associated molecular patterns. Plant Cell. 2010;22(3):973–90.
61. Berendsen RL, Vismans G, Yu K, Song Y, De Jonge R, Burgman WP, et al. Disease-induced assemblage of a plant-beneficial bacterial consortium. ISME Journal. 2018;12(6).
62. Bailly A, Weisskopf L. The modulating effect of bacterial volatiles on plant growth current knowledge and future challenges. Vol. 7, Plant Signaling and Behavior. Landes Bioscience; 2012.
63. Effmert U, Kalderás J, Warnke R, Piechulla B. Volatile Mediated Interactions Between Bacteria and Fungi in the Soil. Vol. 38, Journal of Chemical Ecology. 2012. p. 665–703.
64. Hartman K, Tringe SG. Interactions between plants and soil shaping the root microbiome under abiotic stress. Biochemical Journal. 2019 Oct 11;476(19):2705–24.
65. Grichko VP, Glick BR. Amelioration of flooding stress by ACC deaminase-containing plant growth-promoting bacteria. 2001.
66. Tewari S, Arora NK. Soybean Production Under Flooding Stress and Its Mitigation Using Plant Growth-Promoting Microbes. In: Environmental Stresses in Soybean Production: Soybean Production. Elsevier Inc.; 2016. p. 23–40.
67. Ansari FA, Ahmad I, Pichtel J. Growth stimulation and alleviation of salinity stress to wheat by the biofilm forming Bacillus pumilus strain FAB10. Applied Soil Ecology. 2019 Nov 1;143:45–54.
68. Upadhyay SK, Singh DP, Saikia R. Genetic Diversity of Plant Growth Promoting Rhizobacteria Isolated from Rhizospheric Soil of Wheat under Saline Condition. Curr Microbiol. 2009 Nov;59(5):489–96.
69. Tahir M, Ahmad I, Shahid M, Shah GM, Farooq ABU, Akram M, et al. Regulation of antioxidant production, ion uptake and productivity in potato (Solanum tuberosum L.) plant inoculated with growth promoting salt tolerant Bacillus strains. Ecotoxicol Environ Saf. 2019 Aug 30;178:33–42.
70. Zhang F, Lynch DH, Smith DL. Impact of low root temperatures in soybean [Glycine max. (L.) Merr.] on nodulation and nitrogen fixation. Environ Exp Bot. 1995;35(3).
71. Compant S, Van Der Heijden MGA, Sessitsch A. Climate change effects on beneficial plant-microorganism interactions. FEMS Microbiol Ecol. 2010;73(2):197–214.
72. Porter SS, Bantay R, Friel CA, Garoutte A, Gdanetz K, Ibarreta K, et al. Beneficial microbes ameliorate abiotic and biotic sources of stress on plants. Vol. 34, Functional Ecology. Blackwell Publishing Ltd; 2020. p. 2075–86.
73. Pan J, Peng F, Tedeschi A, Xue X, Wang T, Liao J, et al. Do halophytes and glycophytes differ in their interactions with arbuscular mycorrhizal fungi under salt stress? A meta-analysis. Bot Stud. 2020 Dec 1;61(1).
74. Vandeputte OM, Kiendrebeogo M, Rasamiravaka T, Stévigny C, Duez P, Rajaonson S, et al. The flavanone naringenin reduces the production of quorum sensing-controlled virulence factors in pseudomonas aeruginosa PAO1. Microbiology (N Y). 2011 Jul;157(7):2120–32.

75. Liu K, Hu H, Wang W, Zhang X. Genetic engineering of Pseudomonas chlororaphis GP72 for the enhanced production of 2-Hydroxyphenazine. Microb Cell Fact. 2016 Jul 28;15(1).
76. Subedi SC, Allen P, Vidales R, Sternberg L, Ross M, Afkhami ME. Salinity legacy: Foliar microbiome's history affects mutualist-conferred salinity tolerance. Ecology. 2022 Jun 1;103(6).
77. Rodriguez R, Redman R. More than 400 million years of evolution and some plants still can't make it on their own: Plant stress tolerance via fungal symbiosis. In: Journal of Experimental Botany. 2008. p. 1109–14.
78. Hosseyni Moghaddam MS, Safaie N, Soltani J, Hagh-Doust N. Desert-adapted fungal endophytes induce salinity and drought stress resistance in model crops. Plant Physiology and Biochemistry. 2021 Mar 1;160:225–38.
79. Barra PJ, Inostroza NG, Mora ML, Crowley DE, Jorquera MA. Bacterial consortia inoculation mitigates the water shortage and salt stress in an avocado (Persea americana Mill.) nursery. Applied Soil Ecology. 2017 Mar 1;111:39–47.
80. Barra PJ, Inostroza NG, Acuña JJ, Mora ML, Crowley DE, Jorquera MA. Formulation of bacterial consortia from avocado (Persea americana Mill.) and their effect on growth, biomass and superoxide dismutase activity of wheat seedlings under salt stress. Applied Soil Ecology. 2016 Jun 1;102:80–91.
81. Schmitz L, Yan Z, Schneijderberg M, de Roij M, Pijnenburg R, Zheng Q, et al. Synthetic bacterial community derived from a desert rhizosphere confers salt stress resilience to tomato in the presence of a soil microbiome. ISME Journal. 2022 Aug 1;16(8):1907–20.
82. Swenson W, Wilson DS, Elias R. Artificial ecosystem selection. Proc Natl Acad Sci U S A [Internet]. 2000 Aug 1 [cited 2015 Oct 8];97(16):9110–4. Available from: http://www.pnas.org/content/97/16/9110.full
83. Panke-buisse K, Poole AC, Goodrich JK, Ley RE, Kao-kniffin J. Selection on soil microbiomes reveals reproducible impacts on plant function. ISME J [Internet]. 2014;1–10. Available from: http://dx.doi.org/10.1038/ismej.2014.196
84. Jochum MD, McWilliams KL, Pierson EA, Jo YK. Host-mediated microbiome engineering (HMME) of drought tolerance in the wheat rhizosphere. PLoS One. 2019 Dec 1;14(12).
85. Mueller UG, Sachs JL. Engineering Microbiomes to Improve Plant and Animal Health. Trends Microbiol [Internet]. 2015 Oct [cited 2017 Mar 1];23(10):606–17. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0966842X15001729
86. Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. Trends Plant Sci [Internet]. 2012;17(8):478–86. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1360138512000799\npapers2://publication/doi/10.1016/j.tplants.2012.04.001
87. Hein JW, Wolfe G V., Blee KA. Comparison of rhizosphere bacterial communities in Arabidopsis thaliana mutants for systemic acquired resistance. Microb Ecol. 2008 Feb;55(2):333–43.
88. Sasse J, Martinoia E, Northen T. Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? Vol. 23, Trends in Plant Science. Elsevier Ltd; 2018. p. 25–41.

89. Dennis PG, Miller AJ, Hirsch PR. Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? Vol. 72, FEMS Microbiology Ecology. 2010. p. 313–27.
90. Hartmann A, Schmid M, van Tuinen D, Berg G. Plant-driven selection of microbes. Vol. 321, Plant and Soil. 2009. p. 235–57.
91. Jones DL, Nguyen C, Finlay RD. Carbon flow in the rhizosphere: Carbon trading at the soil-root interface. Vol. 321, Plant and Soil. 2009. p. 5–33.
92. Chaparro JM, Badri D V, Vivanco JM. Rhizosphere microbiome assemblage is affected by plant development. ISME J [Internet]. 2014;8(4):790–803. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24196324
93. Marschner P, Marschner P, Crowley D, Crowley D, Yang CH, Yang CH, et al. Development of specific rhizosphere bacterial communities in relation to plant species, nutrition and soil type. Plant Soil. 2004;261(1–2):199–208.
94. Garbeva P, van Veen J a, van Elsas JD. Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. Annu Rev Phytopathol. 2004;42(29):243–70.
95. Tkacz A, Cheema J, Chandra G, Grant A, Poole PS. Stability and succession of the rhizosphere microbiota depends upon plant type and soil composition. ISME J [Internet]. 2015;1–11. Available from: http://dx.doi.org/10.1038/ismej.2015.41
96. Berg G, Smalla K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. FEMS Microbiol Ecol. 2009;68(1):1–13.
97. Jin Y, Xia J, Pan Z, Yang J, Wang W, Fu Z. Polystyrene microplastics induce microbiota dysbiosis and inflammation in the gut of adult zebrafish. Environmental Pollution. 2018 Apr 1;235:322–9.
98. Barbosa Lima A, Cannavan FS, Navarrete AA, Teixeira WG, Kuramae EE, Tsai SM. Amazonian Dark Earth and Plant Species from the Amazon Region Contribute to Shape Rhizosphere Bacterial Communities. Microb Ecol. 2015 May 1;69(4):855–66.
99. Marques JM, da Silva TF, Vollu RE, Blank AF, Ding GC, Seldin L, et al. Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown sweet potato plants. FEMS Microbiol Ecol. 2014;88(2):424–35.
100. Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA. A Synthetic Community Approach Reveals Plant Genotypes Affecting the Phyllosphere Microbiota. PLoS Genet. 2014;10(4).
101. Zancarini A, Mougel C, Voisin AS, Prudent M, Salon C, Munier-Jolain N. Soil Nitrogen Availability and Plant Genotype Modify the Nutrition Strategies of M. truncatula and the Associated Rhizosphere Microbial Communities. PLoS One. 2012;7(10).
102. Yuan J, Chaparro JM, Manter DK, Zhang R, Vivanco JM, Shen Q. Roots from distinct plant developmental stages are capable of rapidly selecting their own microbiome without the influence of environmental and soil edaphic factors. Soil Biol Biochem. 2015;89(AUGUST):206–9.

103. Houlden A, Timms-Wilson TM, Day MJ, Bailey MJ. Influence of plant developmental stage on microbial community structure and activity in the rhizosphere of three field crops. FEMS Microbiol Ecol. 2008;65(2):193–201.
104. Pieterse CMJ, de Jonge R, Berendsen RL. The Soil-Borne Supremacy. Vol. 21, Trends in Plant Science. Elsevier Ltd; 2016. p. 171–3.
105. Edwards KJ, Bach W, McCollom TM. Geomicrobiology in oceanography: Microbe-mineral interactions at and below the seafloor. Vol. 13, Trends in Microbiology. 2005. p. 449–56.
106. Philippot L, Raaijmakers JM, Lemanceau P, Van Der Putten WH. Going back to the roots: The microbial ecology of the rhizosphere. Vol. 11, Nature Reviews Microbiology. 2013. p. 789–99.
107. Xu J, Zhang Y, Zhang P, Trivedi P, Riera N, Wang Y, et al. The structure and function of the global citrus rhizosphere microbiome. Nat Commun. 2018 Dec 1;9(1).
108. Zhang Y, Xu J, Riera N, Jin T, Li J, Wang N. Huanglongbing impairs the rhizosphere-to-rhizoplane enrichment process of the citrus root-associated microbiome. Microbiome. 2017 Aug 10;5(1):97.
109. Levy A, Salas Gonzalez I, Mittelviefhaus M, Clingenpeel S, Herrera Paredes S, Miao J, et al. Genomic features of bacterial adaptation to plants. Nat Genet. 2018 Jan 1;50(1):138–50.
110. Knief C, Delmotte N, Chaffron S, Stark M, Innerebner G, Wassmann R, et al. Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. ISME Journal. 2012 Jul;6(7):1378–90.
111. De Souza RSC, Okura VK, Armanhi JSL, Jorrín B, Lozano N, Da Silva MJ, et al. Unlocking the bacterial and fungal communities assemblages of sugarcane microbiome. Sci Rep. 2016 Jun 30;6.
112. Edwards JA, Santos-Medellín CM, Liechty ZS, Nguyen B, Lurie E, Eason S, et al. Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. PLoS Biol. 2018 Feb 23;16(2).
113. Cregger MA, Veach AM, Yang ZK, Crouch MJ, Vilgalys R, Tuskan GA, et al. The Populus holobiont: Dissecting the effects of plant niches and genotype on the microbiome. Microbiome. 2018;6(1).
114. Wei Z, Gu Y, Friman VP, Kowalchuk GA, Xu Y, Shen Q, et al. Initial soil microbiome composition and functioning predetermine future plant health [Internet]. Vol. 5, Sci. Adv. 2019. Available from: https://www.science.org
115. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature [Internet]. 2012;488(7409):86–90. Available from: http://dx.doi.org/10.1038/nature11237
116. Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, et al. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature [Internet]. 2012 Aug 1 [cited 2016 Oct 6];488(7409):91–5. Available from: http://www.nature.com/doifinder/10.1038/nature11336

117. Schlaeppi K, Dombrowski N, Oter RG, Ver Loren Van Themaat E, Schulze-Lefert P. Quantitative divergence of the bacterial root microbiota in Arabidopsis thaliana relatives. Proc Natl Acad Sci U S A. 2014;111(2):585–92.
118. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature [Internet]. 2013 Aug 1 [cited 2016 Oct 6];488(7409):86–90. Available from: http://www.nature.com/doifinder/10.1038/nature11237
119. Coleman-Derr D, Desgarennes D, Fonseca-Garcia C, Gross S, Clingenpeel S, Woyke T, et al. Biogeography and cultivation affect microbiome composition in the drought-adapted plant Subgenus Agave. in Review. 2015;798–811.
120. Miller SB, Heuberger AL, Broeckling CD, Jahn CE. Non-targeted metabolomics reveals sorghum rhizosphere-associated exudates are influenced by the belowground interaction of substrate and sorghum genotype. Int J Mol Sci. 2019 Jan 2;20(2).
121. Micallef SA, Shiaris MP, Colón-Carmona A. Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. J Exp Bot. 2009 Apr;60(6):1729–42.
122. Narula N, Kothe E, Behl RK. Role of root exudates in plant-microbe interactions. Journal of Applied Botany and Food Quality. 2009;82(2):122–30.
123. De Bruijn FJ, Szabados L, Schell J. Chimeric genes and transgenic plants are used to study the regulation of genes involved in symbiotic plant-microbe interactions (nodulin genes). Dev Genet [Internet]. 1990;11(3):182–96. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2279354
124. Read D. plant-microbe mutualisms and community structure. Biodiversity and ecosystem function [Internet]. 1993;181–209. Available from: http://link.springer.com/chapter/10.1007/978-3-642-58001-7_9
125. Siqueira JO, Nair MG, Hammerschmidt R, Safir GR, Putnam AR. Significance of phenolic compounds in plant-soil-microbial systems. Vol. 10, Critical Reviews in Plant Sciences. 1991. p. 63–121.
126. Quigley NB, Roberts DM, Ownley BH. Molecular plant—microbe interactions. Vol. 1, Trends in Plant Science. 1996. p. 366–8.
127. Veronese P, Ruiz MT, Coca MA, Hernandez-Lopez A, Lee H, Ibeas JI, et al. In defense against pathogens. Both plant sentinels and foot soldiers need to know the enemy. Plant Physiol. 2003;131(4):1580–90.
128. Mark GL, Dow JM, Kiely PD, Higgins H, Haynes J, Baysse C, et al. Transcriptome profiling of bacterial responses to root exudates identifies genes involved in microbe – plant interactions. Plating. 2005;(Track II).
129. Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M, et al. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. Science (1979). 2015 Aug 21;349(6250):860–4.
130. Adams-Phillips L, Wan J, Tan X, Dunning FM, Meyers BC, Michelmore RW, et al. Discovery of ADP-ribosylation and other plant defense pathway elements through expression profiling of four different Arabidopsis-Pseudomonas R-avr interactions. Mol Plant Microbe Interact. 2008;21(5):646–57.

131. Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, et al. Plant ABC Transporters. The Arabidopsis book / American Society of Plant Biologists [Internet]. 2011;9:e0153. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3268509&tool=pmcentrez&rendertype=abstract
132. Yazaki K, Sugiyama A, Morita M, Shitan N. Secondary transport as an efficient membrane transport mechanism for plant secondary metabolites. Phytochemistry Reviews. 2008;7(3):513–24.
133. Balint-Kurti P, Simmons SJ, Blum JE, Ballaré CL, Stapleton AE. Maize leaf epiphytic bacteria diversity patterns are genetically correlated with resistance to fungal pathogen infection. Mol Plant Microbe Interact [Internet]. 2010 Apr 1 [cited 2015 Dec 14];23(4):473–84. Available from: http://apsjournals.apsnet.org/doi/abs/10.1094/MPMI-23-4-0473
134. Smith KP, Handelsman J, Goodman RM. Genetic basis in plants for interactions with disease-suppressive bacteria. Proc Natl Acad Sci U S A [Internet]. 1999;96(9):4786–90. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=21769&tool=pmcentrez&rendertype=abstract
135. Li P, Wind JJ, Shi X, Zhang H, Hanson J, Smeekens SC, et al. Fructose sensitivity is suppressed in Arabidopsis by the transcription factor ANAC089 lacking the membrane-bound domain. Proc Natl Acad Sci U S A [Internet]. 2011;108(8):3436–41. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3044370&tool=pmcentrez&rendertype=abstract

CHAPTER 2

MULTI-GENERATION ECOSYSTEM SELECTION OF RHIZOSPHERE MICROBIAL COMMUNITIES ASSOCIATED WITH PLANT BIOMASS IN *ARABIDOPSIS THALIANA*

Nachiket Shankar[1*], Prateek Shetty[2], Rick Kesseli[1]

[1]Department of Biology, University of Massachusetts, Boston, MA 02125,

[2]Institute of Plant Biology, Hungarian Academy of Sciences, Biological Research Centre, 6726 Szeged, Hungary

Abstract

The role of the microbiome in shaping the host phenotype has emerged as a critical area of investigation, with implications in ecology, evolution, and host health. The complex and dynamic interactions involving plants and their diverse rhizosphere microbial communities are influenced by a multitude of factors, including but not limited to soil type, environment, and plant genotype. Understanding the impact of these factors on microbial community assembly is key to yielding host-specific and robust benefits for plants, yet remains challenging. Here we ran an artificial ecosystem selection experiment, over eight generations, in *Arabidopsis thaliana* L*er* and Cvi to select soil microbiomes associated with higher or lower biomass of the host. This resulted in divergent microbial communities, shaped by a complex interplay between random environmental variations, plant genotypes, and biomass selection pressures. In the initial phases of the experiment, the genotype and the biomass selection treatment have modest but significant impacts. Over time, the plant genotype and biomass treatments gain more influence, explaining ~40% of the variation in the microbial community composition. Furthermore, a genotype-specific association of a plant growth-promoting

rhizobacterial taxon, Labraceae with L*er* and Rhizobiaceae with Cvi, is observed under selection for high biomass. This study enhances our understanding of both microbial community dynamics and plant genetics and can be applied to agricultural settings to improve plant growth and soil restoration.

Introduction

The conventional understanding of the host phenotype involves genetics and environment shaping observable traits. Yet, the last two decades have underscored the microbiome's significance in shaping the host phenotype, driven by extensive research on its role in ecology, evolution, and host health. The plant microbiome represents a rich source of functional diversity that is not encoded within the host genome. The interactions between the plant host and its microbiome are dynamic and reciprocal, with the plant shaping its immediate environment by exuding specific metabolites, thereby promoting the growth of specific microbial taxa, while the microbiome in turn influences plant health and growth (1). A multitude of biotic and abiotic factors influence the dynamic nature of the microbiome, including but not limited to plant root exudates (2–4), soil type (5–9), environment (10), and various aspects of the plant, including species (11), genotype (4,12–15) and developmental stage (16,17).

Previous studies in *Arabidopsis thaliana* have shown that the microbial diversity of the soil reduced with proximity to the endophytic compartment, and microbes found in the bulk soil were not enriched in the endophytic compartment and vice-versa (18–20). Furthermore, the microbiome of greenhouse-grown plants was similar to that of the field-grown plants, in

the different plant compartments – rhizosphere, woody stem, and endophytic compartment, indicating an active role of the plant host in creating and maintaining an environment where certain microbes have improved fitness (21). The soil environment was a stronger predictor of bulk soil and rhizosphere microbiome structure, though plant genotype had a weak influence in some studies (20,22,23). A follow-up study confirmed that variation in microbiome communities depends more on the environment than on different *Arabidopsis* varieties and sister species. Though there were some differences between different host species, the core microbiome interacting with the plant host at the root endophytic zone was largely consistent and reproducible (24).

Unfortunately, most prior studies were conducted in a single plant growth cycle, precluding an analysis of the temporal stability of the plant-microbiome association. Observations from plant-soil feedback studies highlight that plant-microbiome interactions can be altered by successive growth cycles of a given plant in the same soil (25,26). Prior studies have employed plant-mediated selection on the soil ecosystem over multiple generations, and resulted in consistent effects on plant characteristics such as biomass, flowering time, and germination (27,28). By selecting for soils where hosts show the desired phenotype, such as high biomass or altered flowering time, it is possible to enrich microbes that modulate host traits. Additionally, studies showed the plant's response to abiotic stress, such as drought and salt stress, could be influenced in two ways. First, through generational selection, where plants with desirable stress responses were chosen over multiple generations (29–31). Second, by introducing beneficial microbes that were associated with plants and had previously lived in similar drought or salt conditions (32). Mueller et al., (29) used generational selection to create

microbial communities that could improve plant seed production by up to 205%. Soil inoculum obtained from drought-exposed soils improved wheat biomass under drought conditions (32). Older studies, such as Swenson et al. (28), also have shown similar results under optimal growing conditions without necessarily exploring the microbiome. A recent study conducted on the rhizosphere microbiome of wild and domesticated tomato plants over multiple generations demonstrated an escalating influence of host genotype on the microbiome community (33). A separate study on the phyllosphere microbiome of various tomato genotypes indicated a declining impact of host genotype across successive generations (34).

Despite several earlier efforts, the development of a robust and host-specific microbial community with long-lasting beneficial effects on the plant host remains a challenge.

Using an artificial ecosystem selection experiment, we sought to advance our understanding of how plant genotype, environment, and biomass selection treatment (henceforth referred to as biomass treatment) impact the assembly of host microbial communities. Treatments selecting for the microbial communities of high and low biomass plants did affect the growth of plants in subsequent generations, though unknown environmental factors had substantial effects throughout the experiment. In the first few generations of the experiment, the genotype and biomass treatments played a modest role in shaping the microbial community. Over time, the plant genotype and biomass treatments had increasing influence, explaining ~40% of the difference in the microbial community composition. Moreover, we observe an enrichment of known plant growth-promoting rhizobacteria (PGPR) in the high biomass treatment, indicating potentially beneficial host-specific interactions.

Material and Methods

Multi-generation selection of soil ecosystem

Arabidopsis thaliana Cvi and L*er* accessions were cultivated using custom-made "rhizotubes," (Stuart Morey, Univ. of Massachusetts, Boston, unpublished). The tubes were equipped with a black polyethylene sleeve, which effectively blocked light penetration. This design also facilitated the convenient removal of the plant from the pot, granting full access to the root system. (Supplementary Figure S2.1).

The potting soil used was PRO-MIX PGX, a commercial mixture comprising 80-90% sphagnum peat moss and small quantities of perlite. It was autoclaved twice for 40 minutes with a 48-hour interval between each sterilization. The potting soil was then sifted through a 3mm sieve and combined with field soil in a 6:1 ratio respectively. The field soil obtained from the Center for Agricultural Research in Waltham, Massachusetts, was composed of 44% sand, 49% silt, and 7% clay representative of agricultural and grassland ecosystems.

The mixture was homogenized using a custom-made cement mixer and attached with sterile bins to avoid cross-contamination. Distilled water was gradually added to achieve a final ratio of 1:2 (water to soil). The soil was incubated at room temperature for two days following inoculation before planting the seeds. All seeds used in the experiment were obtained from a single parent plant. Before planting, the seeds were treated with a solution of 50% bleach vol/vol and a drop of tween 20 for 10 minutes and rinsed ten times with sterile distilled water. We placed 3-4 seeds in the center of each pot and kept only one seedling per pot after emergence. All plants were grown at 22^{o}C day/ 18^{o}C night with a 12/12 hour day

night cycle in a controlled growth chamber. Relative humidity ranged from 35-60% and the light intensity 96uE. Fertilizer was not applied to plants.

In the first generation of the experiment, 100 plants of each accession (Cvi and L*er*) were grown separately in individual pots. The plants were arranged in a randomized block design to avoid batch effects. All plants were harvested 35 days after germination. The above-ground portion of the plant, the rhizosphere, and the bulk soil from each rhizotube were separated, ensuring there was no cross-contamination (see Supplementary Figure S2.1). The above-ground part of the plant, stem, and leaves were dried at 70^{o}C for 4 days. All the plants were weighed individually on a closed weighing scale accurate to 1 mg. The root-soil complex (comprising the rhizosphere and the endosphere, henceforth referred to as the rhizosphere) was obtained by shaking the excess soil off the root and placing the root and remaining attached soil in a sterile 5 ml tube. Bulk soil samples were moved to a sterile Ziplock bag. Tubes and bags were immediately transferred to dry ice and then stored at -80^{o}C for DNA analysis.

The inoculation process for the next generation of soil involved using the bulk soil of the top five and bottom five plants by weight, resulting in two different biomass treatments for each genotype: high biomass L*er*, low biomass L*er*, high biomass Cvi, and low biomass Cvi. Starting from the second generation of the experiment, 50 plants are grown on each of the four soil inoculants to maintain a manageable experiment size. This process is repeated for eight generations. There was substantial stochastic variation in above-ground plant biomass and during generation 6 all plants died for unknown reasons. That experiment was therefore restarted using spare soil from generation 5. Plants in uninoculated sterile potting

soil control were grown from generation 3 onwards to act as a reference for random environmental variation (REV).

DNA extraction and 16S rRNA amplicon library prep

Microbial DNA was isolated from the frozen rhizosphere samples using the Machery-Nagel Nucleospin Soil DNA extraction kit (MACHEREY-NAGEL Inc., Allentown, PA, USA). Approximately 0.1g of rhizosphere soil sample was used for DNA isolation. All samples were diluted to 5 ng ul-1 with PCR-grade water. 16S rRNA gene was amplified from the isolated DNA samples in triplicate in 96-well PCR plates. The PCR primers used were for the 16S rRNA V4 region, 515F (5'-GTGYCAGCMGCCGCGGTAA-3') (ref. 35) and 806R (5'-GGACTACNVGGGTWTCTAAT-3') (ref. 36) for downstream paired-end Illumina (Illumina, Inc., San Diego, CA, USA) barcoded sequencing (37). The PCR cycling conditions were as follows: 94^{o}C for 3 min; 25 cycles of 94^{o}C for 45 s, 50^{o}C for 60 s, 72^{o}C for 90 s, and final elongation at 72^{o}C for 10 min. The triplicate amplified samples were pooled and then purified and normalized using the SequalPrep™ Normalization Plate Kit (Invitrogen Corporation, Carlsbad, Canada). Finally, multiplexed paired-end sequencing was carried out in the Illumina MiSeq platform using earth microbiome project (EMP) primers (38).

Sequence Data Analysis & Statistics

The paired-end sequences obtained from the Illumina MiSeq were demultiplexed with QIIME 2 and converted into individual sequence fastq files for each sample (37,39). The rest

of the sequence processing was carried out in R (Version 4.2.2) using the DADA2 package (Version 1.24.0) (ref. 40). The reads were processed in R using the following command in DADA2; `filterAndTrim(fnFs, filtFs, fnRs, filtRs, truncLen=c(145,145), minLen=50, maxN=0, maxEE=c(2,2), truncQ=2, rm.phix=TRUE, compress=TRUE, multithread=TRUE). `

De novo OTU (Operational Taxonomic Units) picking was performed using DADA2 (37) which resolves amplicon sequence variants (ASVs) to single nucleotide differences. Chimeras were removed. The taxonomy was assigned using a 100% cut-off for species-level identification with the (2022) GTDB 16S rRNA reference database (41). The phylogenetic tree was constructed using IQTree (42), with 1000 ultrafast bootstraps and Modelfinder to identify the best model. In this study, the best-fit model was SYM+R10 chosen according to BIC. The data comprising the OTU table, phylogenetic tree, taxonomy table, and sample meta-data were then parsed through R studio using the phyloseq package (43). Here, it was processed to remove: OTUs identified as mitochondrial or chloroplast, which had less than five reads across all samples, and samples that had fewer than 8000 reads. The packages phytools, phyloseq, microbiome, and ggplot2 were used for further data analysis.

Diversity Analyses

Alpha diversity was computed using four metrics, Observed, Shannon, Chao1, and Simpson diversity indices. Beta diversity was calculated using weighted and unweighted UniFrac, and subsequently Principle Coordinate Analysis (PCoA) plots were constructed using the first two axes that explain the most variance. Linear mixed effect models using,

lmerTest package in R were used to determine changes in alpha and beta diversity over time (lmer (DiveristyMetric, Genotype*Generation*Biomass Treatment, (1|Generation)), with generation as a fixed and random effect. Pairwise comparisons were determined with *emmeans* and adjusted p-value obtained with Tukey.

Adonis

PERMANOVA was carried out for each generation using the adonis2 test. Both weighted and unweighted UniFrac distance matrices were used to account for the abundance, presence/absence, and phylogeny of the microbiome. The following model was used `adonis2(distance.matrix~Biomass Treatment+Genotype, data = meta, permutations=999)`.

Neutral Model

To determine the impact of neutral processes like drift and dispersal or by deterministic selective forces such as plant genotype and biomass treatments on microbial community assembly, we carried out the analysis as described in Burns et al. (44), which fits the data to the neutral model for Prokaryotes from Sloan et al. (45). The following design was used, sncm.fit (spp=generation(n)_Biomass Treatment, pool = generation(n), stats=T), where n is the nth generation and 'Biomass Treatment' is high or low biomass samples.

Differential Abundance

There are several challenges in estimating differentially abundant (DA) taxa in microbiome data, these include high variability in abundance, zero-inflated data, and the

compositional nature of the data. Nearing et al. (46) compared several methods across 38 datasets and found that different methods often identified different sets of DA taxa. The DA taxa are estimated using ANCOMBC2 (47), a conservative and robust approach. ANCOMBC2 incorporates bias correction, effectively addressing sampling-specific and sequencing biases present in the data. This feature ensures that the analysis is not skewed by any systematic errors introduced during the sampling or sequencing process. It also conducts a sensitivity analysis for the pseudo-count addition to assess the impact of different pseudo-count values on zero counts for each taxon.The analysis was run on L*er* and Cvi samples, from OTU to Family taxonomic levels with (fixed effect = (Biomass Treatment + Genotype + Generation), group = Biomass Treatment).

Results

Plant Biomass

The ecosystem selection experiment was carried out for the trait plant biomass on two different accessions of *Arabidopsis thaliana,* L*er* and Cvi for eight generations. The two selected treatments were high and low above-ground biomass. The absolute values for the mean biomass of plants (n=50) from each genotype and treatment changed substantially through the course of eight generations of the experiment. Despite the stochastic fluctuations in biomass caused by random environmental variation (REV), the high biomass selected lines were always the same or greater than those of the low selected biomass for both genotypes (Figures 2.1a and 2.1b). Due to the significant drop in biomass between generations 1 and 2,

a sterile potting soil uninoculated control was planted in all subsequent generations to serve as a reference for REV. The uninoculated control does show that the drastic drop in biomass after generations 1 and 2, was not a result of the inoculated soil but other largely unknown factors and growing conditions. Above-ground biomass in terms of deviations from the mean removes some of these REVs and gives a clearer representation of the difference in phenotype seen in every generation (Figures 2.2a and 2.2b). Significant differences between the low and high biomass treatments become apparent from generation 4 onwards. Plants of all treatments in generation 6 died resulting in a sharp dip, which necessitated repeating that generation (described in the Materials and Methods).

Microbial community Composition

After preprocessing in DADA 2, the Illumina MiSeq (SY-410-1003) generated a total of 1,897,732 reads, with an average of 12822 reads per sample. Singletons and chimeras were removed during pre-processing. The most well-represented phyla were Proteobacteria (36.2%), Bacteroidetes (10.3%), Planctomycetes (7.8%), and Actinobacteria (7.6%) (Figure 2.3). The most well-represented Classes identified in the dataset were Alphaproteobacteria (24.7%), Gammaproteobacteria (11.4%), Bacteroidia *(10.1%)*, and Verrucomicrobiae (6.2%) (Figure 2.4). The relative abundance is typical of microbiome data, heavily weighted to a few abundant groups with a long tail of rare taxa.

Alpha Diversity

We used four metrics to assess alpha diversity (Figure 2.5). Observed ($F_{7,160}$ =333.054, P<0.001) and Chao1 ($F_{7,160}$ =195.54, P<0.001) estimate species richness, and both showed a statistically significant decrease in OTUs (1477 to ~580) and richness across generations. The Shannon diversity index ($F_{7,160}$ =82.23, P<0.001), which is more sensitive to the difference in abundance, also showed this steady decrease across generations. The Simpson diversity ($F_{7,160}$ =14.46, P<0.001), which is more sensitive to evenness, exhibits fluctuations across successive generations, albeit without any distinct trend. In addition, a linear mixed-effect model shows significant interactions between genotype:biomass treatment, genotype:generation and genotype:biomass treatment:generation, for both Shannon and Simpson diversity indices (Supplementary Tables 2.7 and 2.8). No significant interaction terms were found for Observed and Chao1, which do not consider the abundance of OTUs in their calculations.

Beta Diversity

Principal coordinate analyses (PCoA) of the weighted UniFrac distances were plotted with the first two axes that captured the most variance in the data, from 73% in generation 1 to greater than 90% in generation 8 (Figure 2.6). UniFrac distance considers the evolutionary relationships between taxa, which makes it more biologically meaningful than other distance metrics that do not consider the evolutionary history. Weighted UniFrac is more sensitive to differences in the abundance of taxa. The microbiomes associated with the treatments and

genotypes are very similar in generation 1, but begin to diverge by generation 2. By generation 3, the microbiomes of the L*er* and Cvi genotypes have diverged and by generation 5 the high and low biomass treatments appear to have further distinguished the microbial communities. The divergence of the microbial community in generations 4 and 5 caused by the high/low biomass treatments also aligns with the onset of notable differences in the above-ground plant biomass within these generations.

The complex interplay between genotype and biomass treatments was modeled using a PERMANOVA test with both weighted and unweighted UniFrac distances to account for differences in abundance and presence-absence of OTUs respectively. The resulting R^2 values for genotype and biomass treatment and residuals as a proxy for REVs were plotted (Figure 2.7a and 2.7b). For both weighted and unweighted UniFrac, the variance in microbial community explained by both genotype and biomass treatment increased significantly over the eight generations. Despite observing a decrease in the variance accounted for by the residuals, which could potentially signify a decline in the influence of REVs, they accounted for ~50 percent of the variability within the microbial community in generation 8. For all metrics, weighted UniFrac exhibits more fluctuations compared to the unweighted UniFrac. Weighted UniFrac analyses found that genotype ($F_{1,39}$=1.55, P<0.05) and biomass treatment ($F_{1,39}$=1.33, P<0.001) explained 5.7% and 13.7%, respectively, of the dissimilarity between microbiomes in generation 1. By generation 8, genotype ($F_{1,14}$=5.14, P<0.05) and biomass treatment ($F_{1,14}$=5.48, P<0.001) explain 24% and 22%, respectively, of the dissimilarity between microbiomes (Supplementary Tables 2.9 to 2.10). For unweighted UniFrac, the influence of genotype on the microbial community increased from 3.3% in generation 1

($F_{1,39}$=1.33, P<0.01) to 26% in generation 8 ($F_{1,39}$=5.13, P<0.001), and for biomass increased from 3.8% in generation 1 ($F_{1,14}$=1.55 ,P<.001) to 13% in generation 8 ($F_{1,14}$=2.58, P<0.05) (Supplementary Tables 2.11 and 2.12).

Sloan Neutral Model

Many forces can alter microbial community structure and dynamics and in our study these can be divided into selective forces such as the plant genotype being colonized and the biomass selection process, or neutral, stochastic forces innate in any experiment and environment. The Sloan neutral model was fit to the data to assess the importance of selective versus neutral drivers of change in the experiment. If neutral processes were the driving force in the microbial community assembly, the null hypothesis would be that all generations of the experiment would fit the model equally well. The model fit to neutrality represented by the R^2 value decreases from 0.795 in generation 1 to 0.58 in generation 8 demonstrating the increasing importance of selective drivers through the course of the experiment (Figure 2.8).

Differential Abundance

Differentially abundant (DA) taxa that distinguished between the high and low biomass treatments were determined using ANCOMBC2 at multiple taxonomic levels, ranging from the OTU to the Class. Taxa lacking classification at the Family level were denoted by the subsequent identifiable taxonomic tier. The OTUs that were present in higher (red) and lower (blue) abundance in the high biomass treatment are identified (Figure 2.9).

Among the DA taxa in this study, several are known to benefit plants. The Class Bacilli promote plant growth (48). The Class Gemmatimonadetes are positively associated with vegetation restoration, plant richness, and soil nutrients (49). Cytophagacea are chemoorganotrophs and important for remineralizing organic materials into micronutrients. They could support both mycelial growth and plant nutrition (50).

Enterobacteriaceae, Paenibacillaceae, and JACDCH01 were all present in higher abundance in the high biomass treatments of both L*er* and Cvi. Most Paenibacillaceae members predominantly inhabited soil, frequently in close association with plant roots (51). These rhizobacteria were known to play a significant role in enhancing plant growth and possessing potential applications in agriculture and *Enterobacteriaceae* had been reported to enhance plant growth (52–54).

Additionally, Order_NS11-12g, BJHT01, JACDCH01, and UBA6156 were more abundant in Cvi under high biomass conditions and less abundant in L*er* under high biomass conditions. The majority of these were uncultivated or candidate taxa. Two alternative families in the Order Rhizobiales appeared differentially associated with the two plant genotypes: Labraceae with L*er* and Rhizobiaceae with Cvi (55). These could indicate genotype-specific interactions with different members of the microbial community.

Discussion

We applied artificial ecosystem selection for eight generations, in *Arabidopsis thaliana* L*er* and Cvi to select soil microbiomes associated with higher or lower biomass of the host. In contrast to some previous studies (27,33), we did not fertilize plants thus maintaining nutrient limitation, and thereby promoting the plant's interaction with the microbiome (56). Over the course of eight generations, a response to the selection, most apparent after generations 4 and 5, was observed both in the microbiome characteristics that shifted gradually and in the plant biomass. Microbiome selection noticeably influenced plant biomass despite large phenotypic variation from one generation to the next. Stochasticity in growth through generations may be attributed to random environmental variations (REVs) as indicated by Swenson et al. (28). This was observed in the uninoculated sterile potting soil reference that showed similar patterns of variability as the inoculated treatments across generations (Figure 2.1). Ecological variability among generations even under controlled growth chamber conditions is a common characteristic of studies (57) and may be due to minute fluctuations in growth chamber conditions or batch and age of potting soil.

We observed a gradual decline in microbial species richness during the initial generations of the experiment. This was evident from the observed OTU counts, as well as the Chao1 and Shannon diversity indices. This pattern is consistent with findings from other studies (32–34), highlighting the impact of selection pressures as the microbial communities adapted to the host plant environment. However, in our study which continued for about twice the number of generations of these earlier studies, we observed a stabilization of richness and alpha diversity in the later generations.

To further understand the complex interplay between the forces of directed selection (plant genotype and biomass treatment), we did a PCoA. Over successive generations, the results demonstrate a strengthening effect of plant genotypes and biomass treatment on the microbial community (Figure 2.6). This result was modeled using PERMANOVA , for both weighted and unweighted UniFrac distances. Both weighted and unweighted UniFrac show marked increases in the proportion of variance explained by genotype and biomass treatment, with unweighted UniFrac exhibiting larger increases from generation to generation (Figure 2.7). Both metrics consider phylogenetic relatedness. Weighted UniFrac considers the absolute abundance of OTUs and is generally used to study changes in microbial community structure. Unweighted UniFrac considers only the presence/absence of OTUs and is generally used to study changes in microbial community composition. The results suggest that variance in microbiome composition as explained by genotype consistently increases, whereas variance in microbiome structure as explained by genotype, fluctuates more with changes in abundant taxa.

Here we would like to highlight that pronounced shifts in biomass between the high/low biomass treatments, alpha diversity, and beta diversity all occur around generations 4 and 5 (Figure 2.1). This marks the point at which the changes in alpha diversity stabilize (Figure 2.5) and the divergence between microbial communities by high/low biomass treatments in the PCoA plots becomes more pronounced (Figure 2.6). It is possible that over the first four to five generations, the starting microbial community underwent a period of restructuring before it stabilized and formed four distinct communities under selection by high/low biomass treatments and by plant genotype. We propose a complex interaction between plants and their

associated microbiome ensued, where differences in root exudation patterns between the two genotypes presumably established associations with microbes in the early generations. Simultaneously, during this period, the biomass treatment likely promoted host-specific microbe-mediated interactions that modulate plant biomass. This restructuring of the microbial community is driven by both structural (abundance of OTUs) and compositional (presence/absence of OTUs) effects in the microbiome (Figure 2.7). Differences in the abundance of taxa driven by the biomass and genotype selection pressures, was a larger contributor to the generation to generation variation during the experiment.

Changes in the assembly of microbial communities can arise from either selective pressure, like the plant genotype or biomass treatment observed in this experiment, or stochastic processes, like minute changes in growth chamber humidity or microenvironment. To gain deeper insights into the influence of selection on microbial community assembly, a neutral model was fit to the data (44,45). Interestingly, this revealed a progressive decline in fit to neutrality, indicated by decreasing R^2 values and increasing AIC values (Figure 2.8). This suggests an increasing influence of selective forces such as biomass treatment and genotype over multiple generations of ecosystem selection. This was also previously observed by Morella et al. (34).

Previous studies in *Arabidopsis thaliana* have often detected only a weak genotypic effect on the rhizosphere microbiome (23,24,58). However, a majority of these studies are focused on a single growth cycle. Insights from plant-soil feedback research emphasize that the interactions between plants and their microbiomes can be modified by consecutive growth cycles of the same plant species in its respective soil (25,26). Recent research has presented

conflicting findings on the influence of host genotype (33,34,59). The different experimental designs, communities being assessed and the length of these experiments, alongside high stochastic environmental variation that is clearly associated with these microbial studies likely account for the conflicting findings. In our investigation, we find that during the initial stages, genotype and biomass treatments have modest but significant impacts. Over time, the plant genotype and biomass treatments have an increasing influence. Together these selective forces explained ~40% of the variation in the microbial community composition in these later generations.

A key aim of microbiome engineering has been to develop host-specific microbial communities that impart lasting beneficial effects on the plant host (27,29,32,60). Here we show the enrichment of some common taxa in high biomass treatments but also some genotype specific changes. Despite starting with the same soil, only three common families were enriched in the high biomass treatment samples of both the L*er* and Cvi genotypes. In addition, well-known plant growth promoting rhizobacteria were enriched in both L*er* and Cvi high biomass treatment samples, however possible genotype-specific interactions are involved. In the Order Rhizobiales, Labraceae was enriched in L*er,* while Rhizobiaceae was enriched in Cvi (55). Moreover, in high biomass conditions, Cvi had increased levels of the taxa Order_NS11-12g, BJHT01, JACDCH01, and UBA6156, while L*er* had decreased levels of these. These data suggest genotype specific plant genetic control. Multiple studies have shown variations in the structure of the rhizosphere microbiome, even within closely related plant genotypes, highlighting the importance of genotype-specific root exudates in forming associations with the corresponding microbiome (61–65). Further exploration must prioritize

highlighting the mechanisms that uphold these interactions through the utilization of comparative metagenome and metatranscriptome analyses. This will play a part in decoding chemical crosstalk, which could eventually foster interactions between plants and microbes, leading to an overall enhancement in plant fitness.

The relationship between plants and microorganisms in the rhizosphere involves complex and diverse interactions, which influence crucial ecological and physiological processes. These interactions can be mutualistic, competitive, or antagonistic in nature. Many biotic and abiotic factors act in concert to influence the dynamic rhizosphere microbiome. Gaining insights into how these factors influence the assembly of microbial communities is crucial for obtaining targeted and long-lasting advantages for plants. Nevertheless, establishing a strong and host-specific microbial community that consistently provides beneficial effects continues to be a persistent challenge. We have shown that, despite stochastic fluctuation due to REVs, it is possible to select for microbial communities that impact biomass in a host-specific manner within four generations, addressing this key challenge in microbial community engineering (32,60). The rhizosphere microbiome that evolves under plant-mediated selection offers better chances of survivability and efficacy when applied as inoculum to the plant (66,67). This study enhances our understanding of the temporal dynamics involved in microbial assembly in the rhizosphere and has implications for sustainable agriculture, evolution, and ecology.

Author Contributions

N.S. Conceptualization, methodology, data analysis and visualization, data collection, writing-original draft, writing-reviewing and editing. P.S. Methodology, data analysis, writing-reviewing, and editing. R.K. resources, funding, writing-reviewing, and editing.

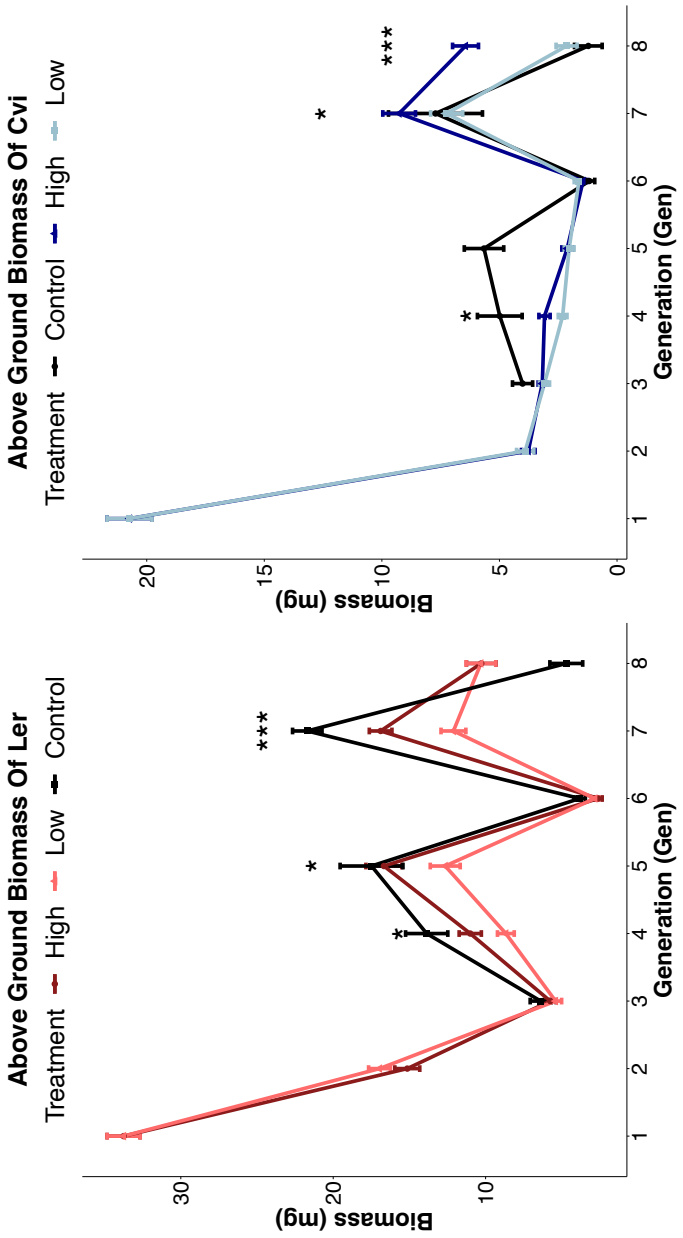


Figure 2.1. Results of the ecosystem selection on above-ground plant biomass. Genotypes L*er* (a) and Cvi (b) of Arabidopsis thaliana are represented. Dark lines represent the high biomass treatment, and light lines represent the low biomass treatment. Each point represents the mean of n=50 microcosms. The mean biomass of the selection lines differed for several generations of both genotypes (t.test p-value * < 0.05; **<0.01;*** <0.001).

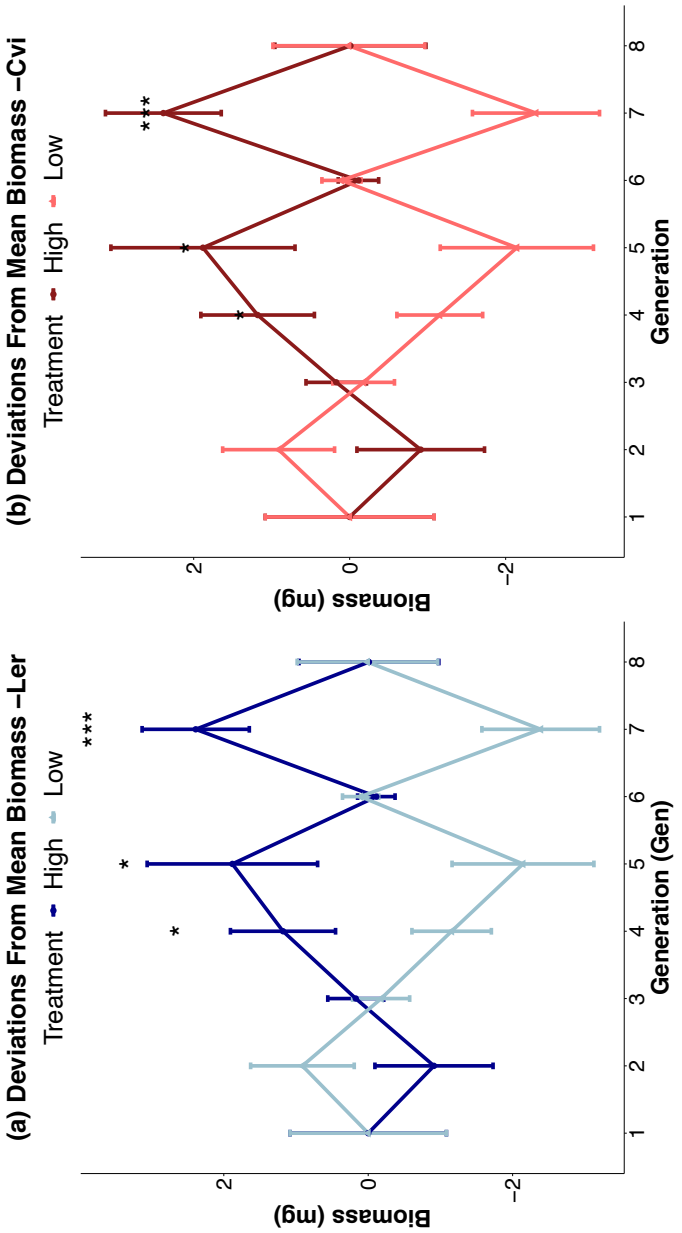


Figure 2.2 Deviations from mean biomass. By plotting deviations from the mean, the influence of REVs on above-ground biomass is mitigated for genotypes L*er* (a) and Cvi (b). The dark lines denote high biomass treatments, while the light lines represent low biomass treatments. Each data point is the average of n=50 microcosms. (t.test p-value * < 0.05;**<0.01; *** <0.001).

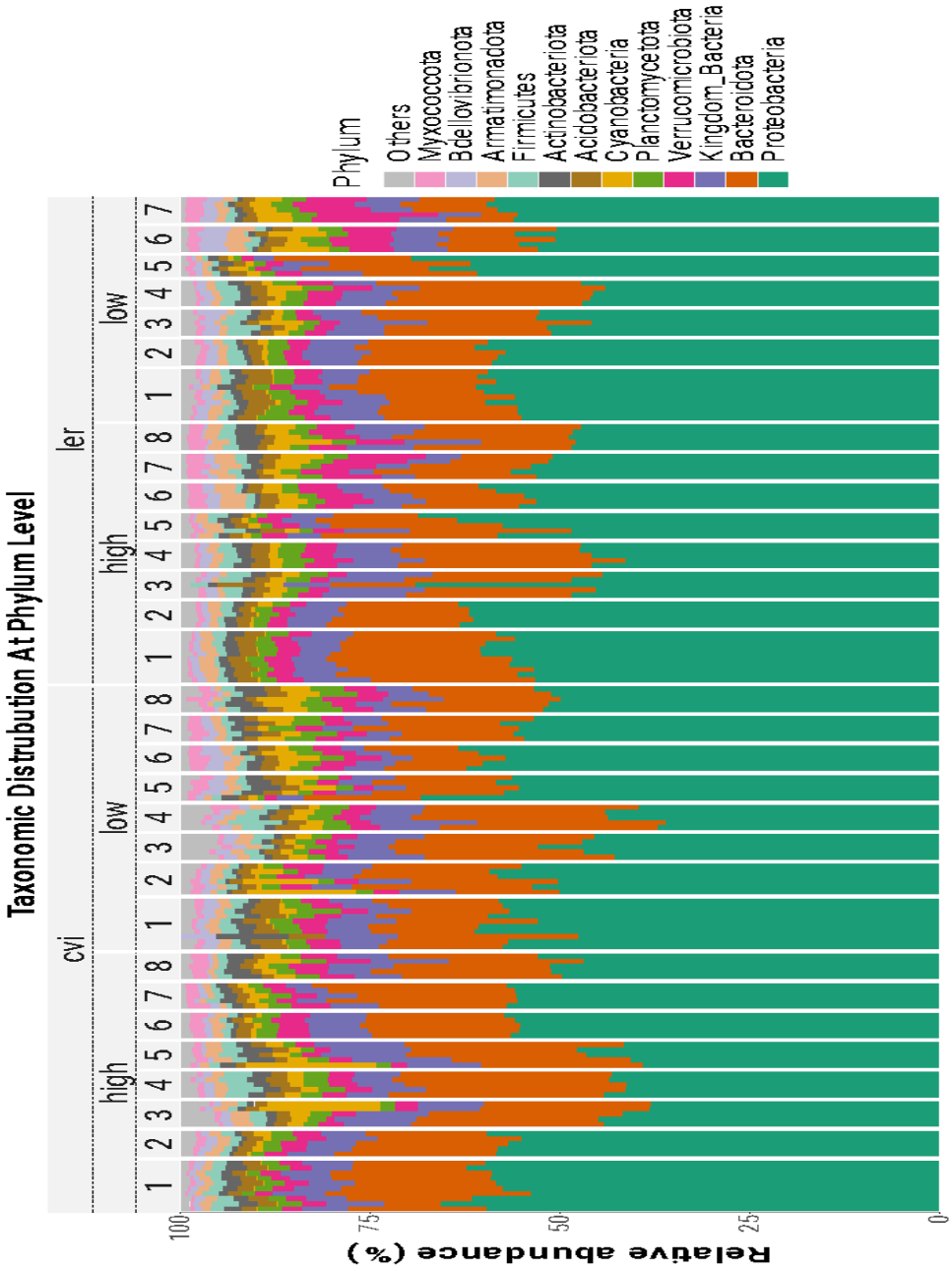


Figure 2.3. Taxonomic distribution - Phylum level. Taxonomic distribution of the microbial community represented in terms of the relative abundance at the Phylum level. The plot shows eight generations (1-8) for the high and low biomass treatments of both the Cvi (left) and L*er* (right) genotypes of *Arabidopsis thaliana*. The microbial community is dominated by *Proteobacteria*, *Bacteroidetes*, *Verrucomicrobia, and Planctomycetes* which together comprise nearly 73% of the bacterial community.

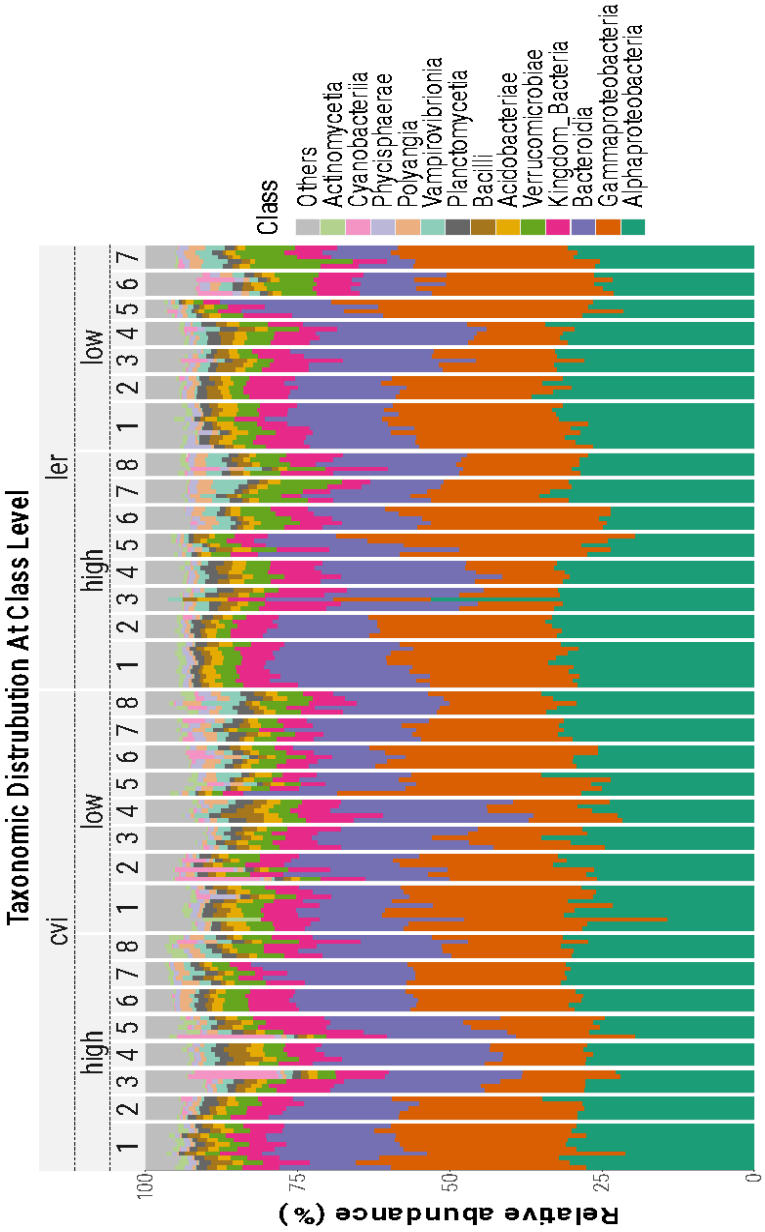


Figure 2.4. Taxonomic distribution - Class level. Taxonomic distribution of the microbial community is represented in terms of the relative abundance at the Class level. The plot shows eight generations (1-8) for the high and low biomass treatments of both the Cvi (left) and L*er* (right) genotypes of *Arabidopsis thaliana*. Ten samples were assayed for generation 1 experiments while five were done for all subsequent generations. The most represented classes across all samples were *Alphaproteobacteria* (24.7%), *Gammaproteobacteria* (11.4%), *Bacteroidia* (10.1%), and *Verrucomicrobiae* (6.2%).

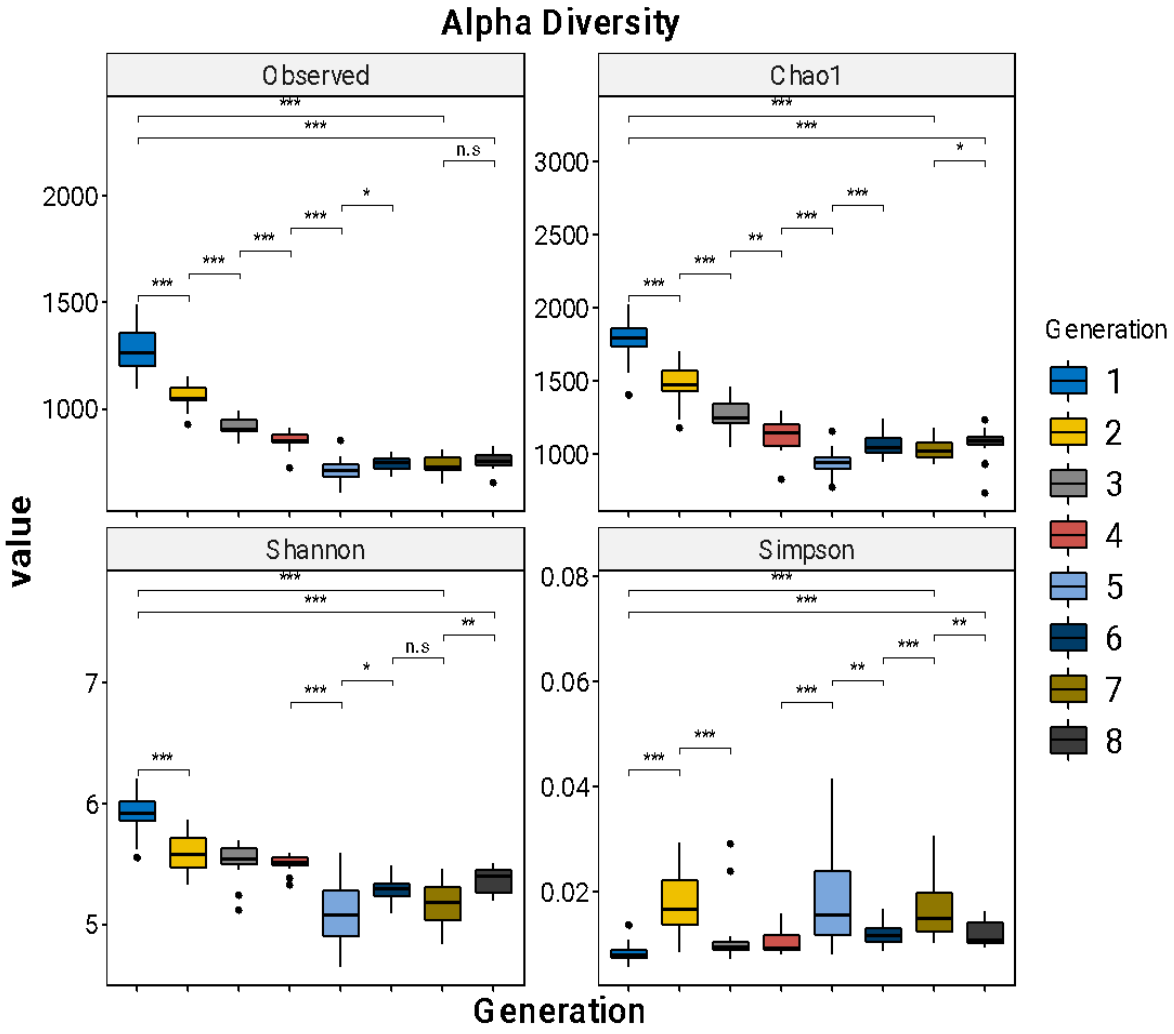


Figure 2.5. Alpha diversity. Over the course of eight generations, four Alpha diversity indices were computed, including Observed, Chao1 estimate, Shannon, and Simpson indices. Results indicate a significant decline in both Observed and Chao1 estimate species richness. Additionally, the Shannon index was more responsive to species richness and exhibited a noteworthy decrease from generation 1 to 8. In contrast, the Simpson index was more sensitive to evenness and demonstrated fluctuations across successive generations without any distinct trend. (wilcox test *<0.05, **<0.01, *** <0.001).

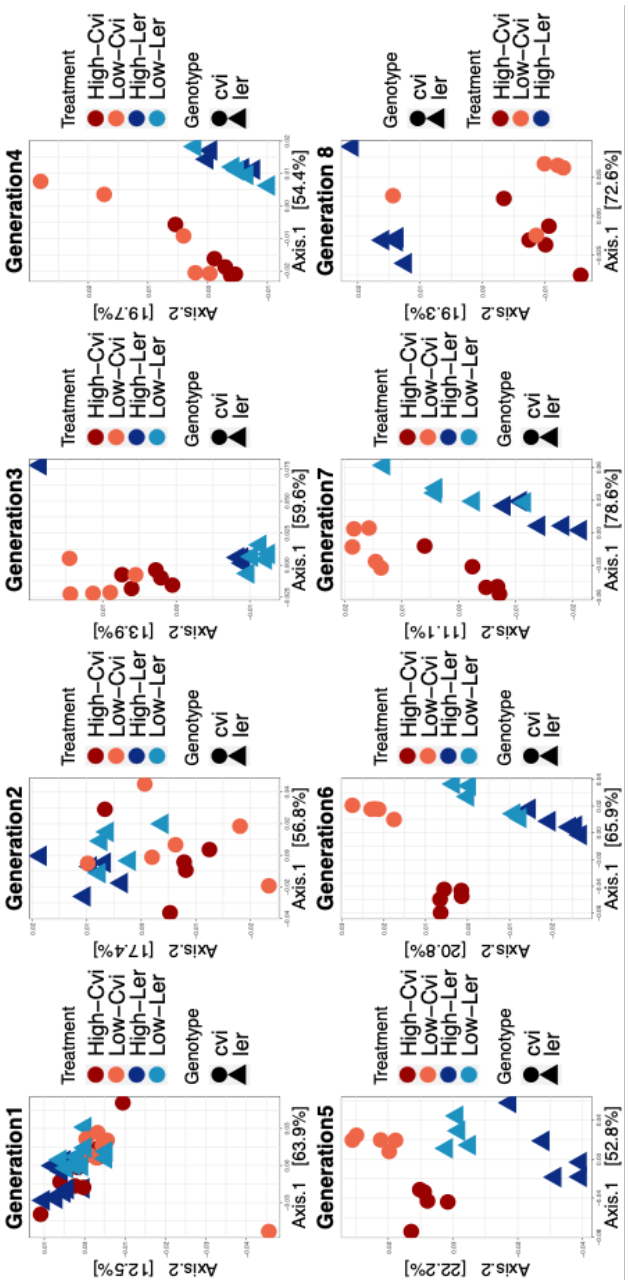


Figure 2.6. Principle Coordinate Analysis (PCoA). PCoA using Unweighted UniFrac was performed for generations 1 through 8. Each point represents an individual sample with microbial communities defined of Cvi represented by red/orange circles and L*er* by dark/light blue triangles. The first two coordinate axes plotted accounted highest variation in the data, ranging from 13% in generation 1 to approximately 52% in generation 8. While there was no discernible clustering in generation 1, both genetic lines and high/low biomass treatments exhibited greater clustering over the course of the experiment.

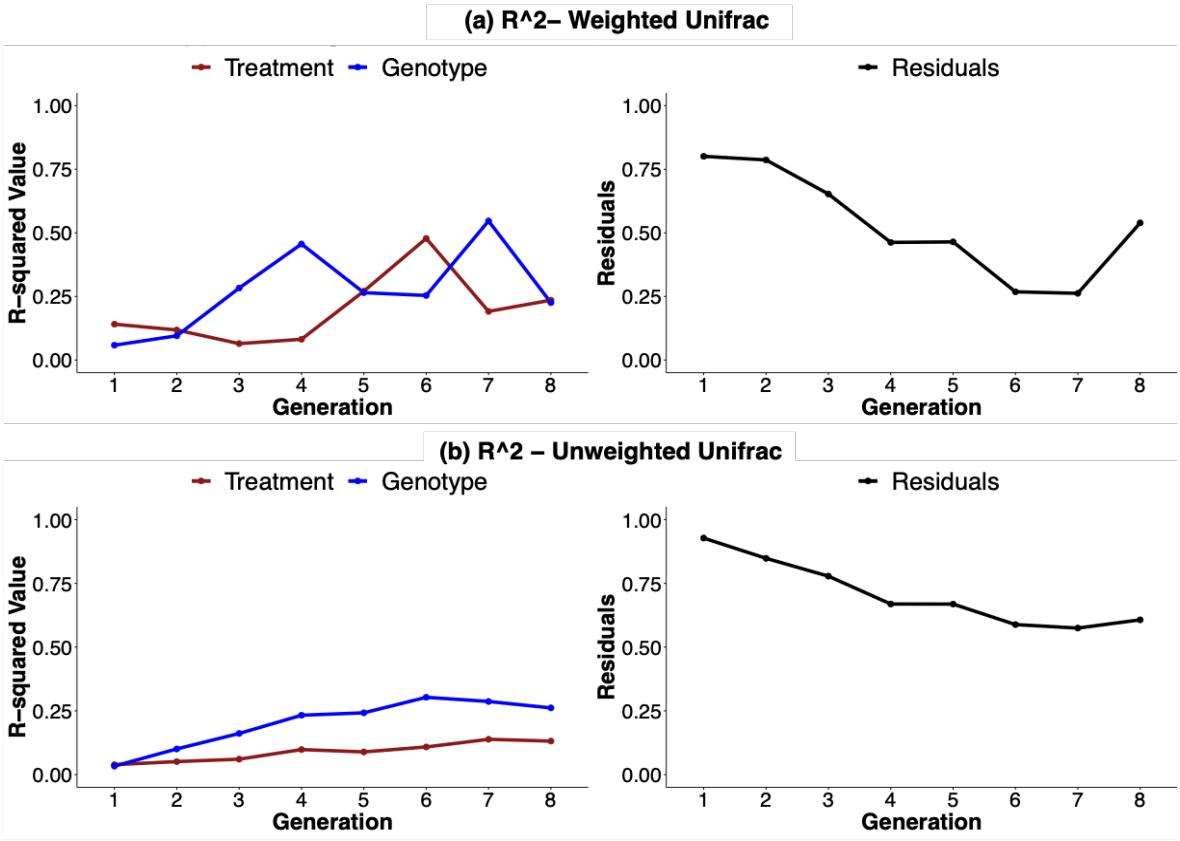


Figure 2.7. The R^2 value obtained from PERMANOVA using adonis2 test. This was done for generations 1 through 8 for (a) weighted and (b) unweighted-UniFrac distances. The resulting R^2 and variance explained by residuals values for the biomass treatment (brown), genotype (blue) and residuals (black) are plotted. Residuals are plotted as a proxy for random environmental variations (REVs). There is a marked increase in the influence of both genotype and biomass treatment on differences in the microbial community from generations 1 through 8. Despite observing a decrease in the variance accounted for by the residuals, which could potentially signify a decline in the influence of REVs, the model elucidates more than 50 percent of the variability within the microbial community in the 8th generation. These differences are more pronounced in the weighted vs unweighted UniFrac. Model (distance.matrix~biomass treatment +genotype).

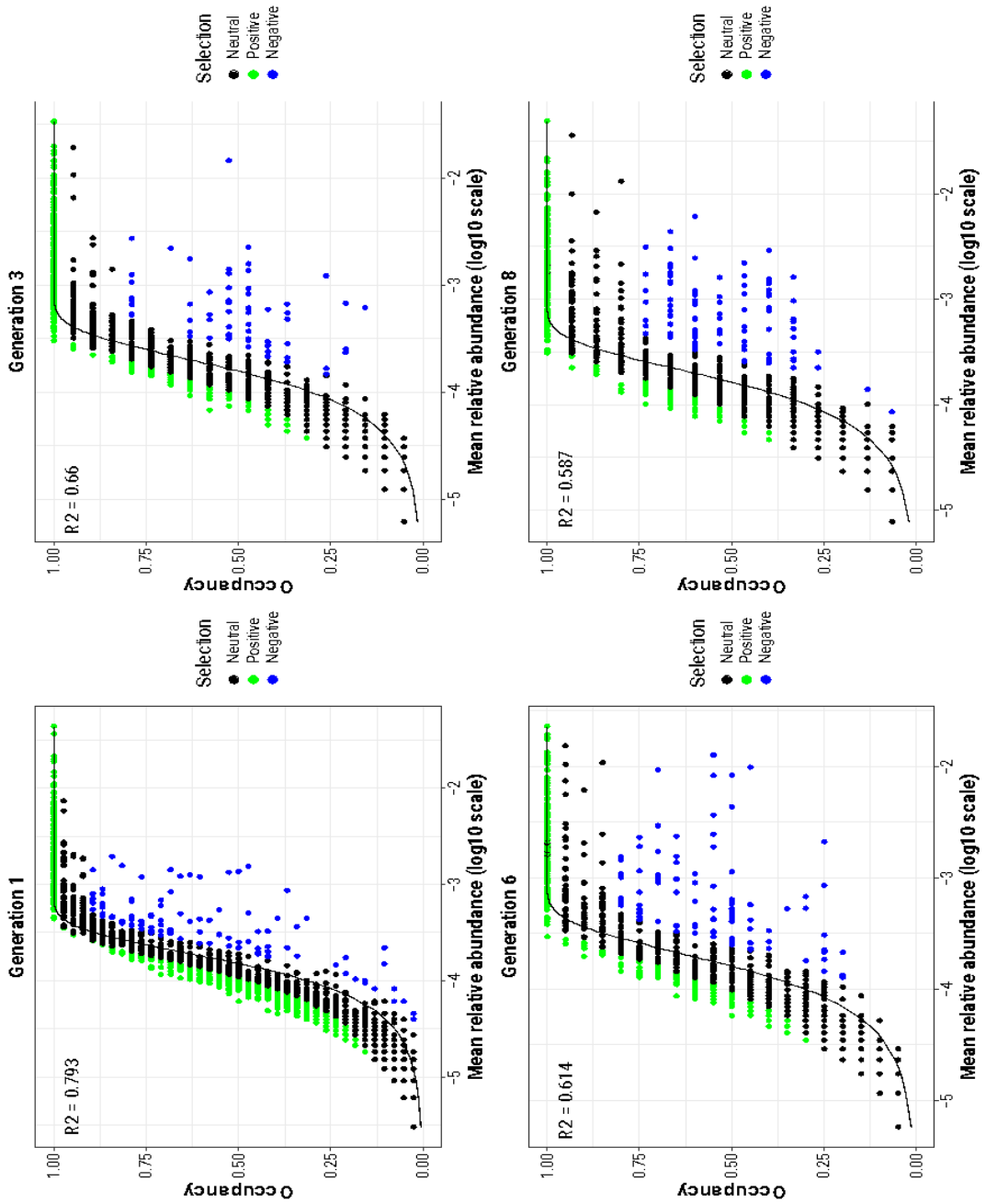


Figure 2.8. Burn's neutral model. The occupancy (prevalence of OTU in samples) plotted against log(10) abundance is depicted using the Burns model. OTUs that are neutral, i.e., not selected for or against, are indicated in black. Green and blue-colored OTUs signify positive and negative selection, respectively. This plot highlights the presence of numerous microbes that undergo selection throughout the experiment.exhibited greater clustering over the course of the experiment.

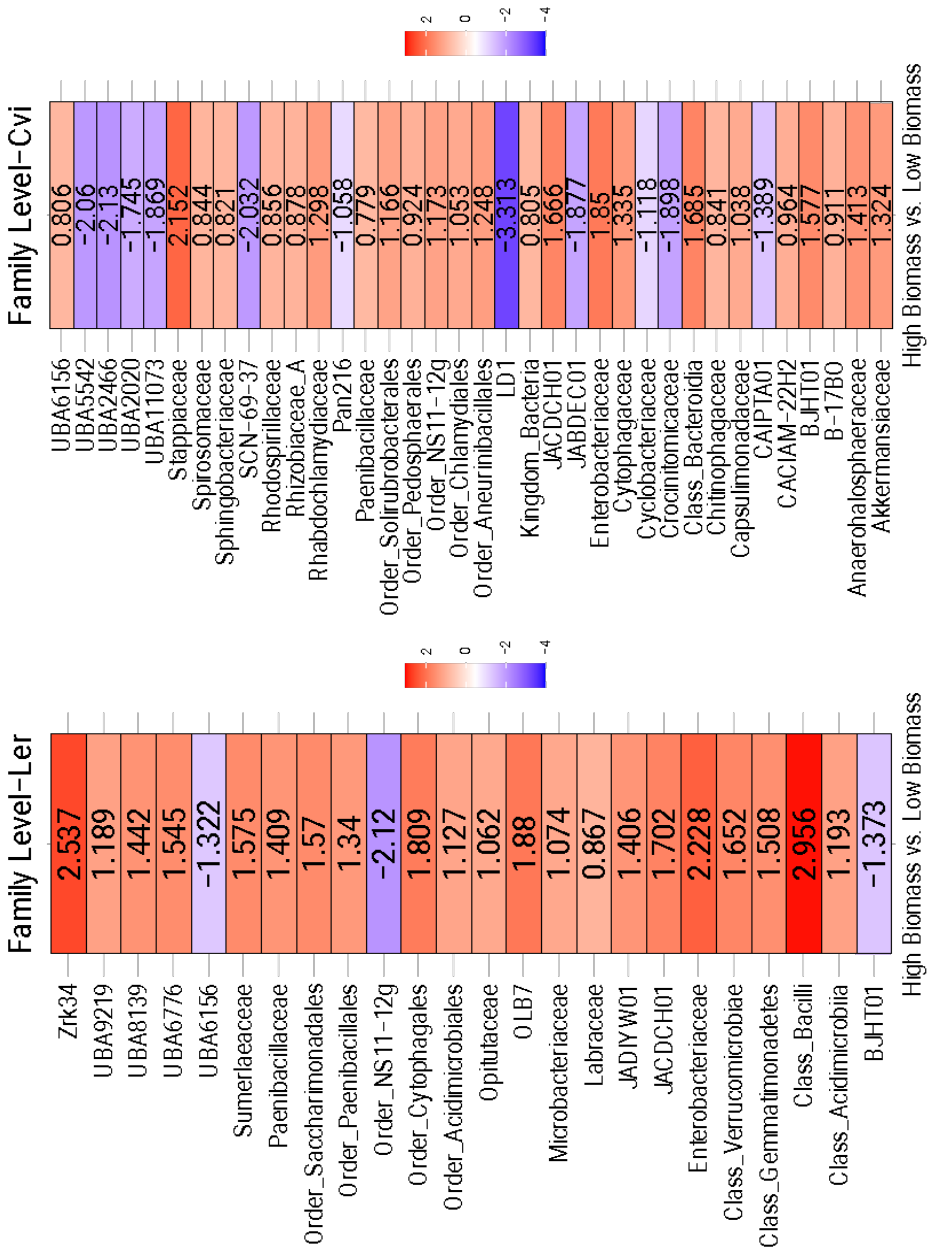
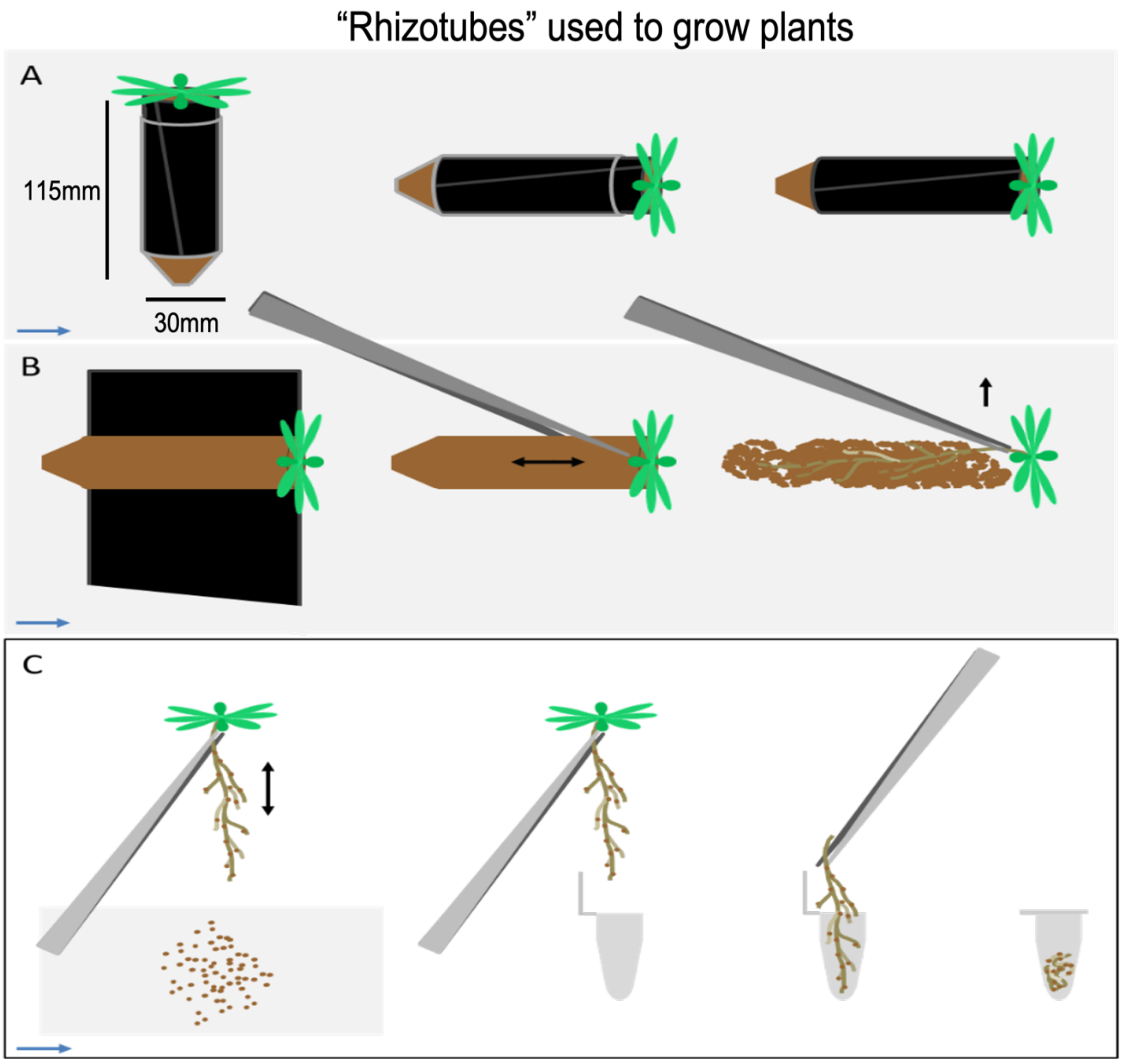


Figure 2.9. Differential abundance at family level. ANCOMBC2 results are illustrated as log fold changes at the family level for high biomass vs low biomass samples. Each accession listed in the table represents a single OTU in either the L*er* or Cvi microbial community identified to the family level if possible. The scale indicates log fold change, red and blue for positive and negative fold change in the high biomass samples respectively. Well-known plant growth-promoting bacteria, such as *Paenibacillaceae, Bacilli, Labraceae, Rhizobiaceae* and *Bdellovibrio* are present in higher abundance. A vast majority of the other taxa identified are uncultivated and have little to no information on them.

SUPPLEMENTARY FIGURES AND TABLES 2



Supplementary Figure S2.1. Rhizotube apparatus. The purpose of this illustration is to demonstrate how the custom-made rhizotubes facilitate the collection of plant rhizosphere. The black insert within the tube is unwrapped to extract the rhizosphere, which is then separated from the bulk soil by shaking the roots. After removing the above-ground portion of the plant, the root-soil complex (comprising the rhizosphere and the endosphere) was obtained by shaking off the excess soil and then placed in a sterile 5ml tube and bulk soil samples were moved to a sterile Ziplock bag. Both were immediately transferred to dry ice and then stored at -80°C for DNA analysis.

Generation	group1	group2	p	p.signif
Gen4	High	Low	1.222020e-02	*
Gen5	High	Low	1.032381e-02	*
Gen7	High	Low	3.562072e-05	****

Supplementary Table 2.1 Statistical comparisons of biomass for L*er* samples. (t.test p-value * < 0.05; **<0.01; *** <0.001).

Generation	group1	group2	p	p.signif
Gen4	High	Low	1.275260e-02	*
Gen7	High	Low	4.123576e-02	*
Gen8	High	Low	6.758467e-08	****

Supplementary Table 2.2 Statistical comparisons of biomass for Cvi samples. (t.test p-value * < 0.05; **<0.01; *** <0.001).

```
             Df Sum Sq Mean Sq F value Pr(>F)
Generation    7 158.13  22.590 333.054 <2e-16 ***
Lineage       1   0.01   0.009   0.130  0.719
Regime        1   0.01   0.011   0.169  0.681
Residuals   160  10.85   0.068
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

Supplementary Table 2.3. ANOVA for Observed alpha diversity index. This shows a significant influence of generation on alpha diversity. Lineage is genotype and Regime is biomass treatment. Design (Observed ~ Generation + Lineage + Regime) Df: degrees of freedom; SumOfSqs : sum of squares; Mean sq: mean squares; F: F statistic; Pr(>F): p-value.

```
             Df Sum Sq Mean Sq F value Pr(>F)
Generation    7 151.27  21.610 195.548 <2e-16 ***
Lineage       1   0.00   0.000   0.000  0.995
Regime        1   0.05   0.046   0.416  0.520
Residuals   160  17.68   0.111
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

Supplementary Table 2.4. ANOVA for Chao1 alpha diversity index. This shows a significant influence of generation on alpha diversity. Lineage is genotype and Regime is biomass treatment. Design (Observed ~ Generation + Lineage + Regime) Df: degrees of freedom; SumOfSqs : sum of squares; Mean sq: mean squares; F: F statistic; Pr(>F): p-value.

```
             Df Sum Sq Mean Sq F value Pr(>F)
Generation    7 132.11  18.873  82.238 <2e-16 ***
Lineage       1   0.01   0.009   0.040  0.841
Regime        1   0.16   0.160   0.695  0.406
Residuals   160  36.72   0.229
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

Supplementary Table 2.5. ANOVA for Shannon alpha diversity index. This shows a significant influence of generation on alpha diversity. Lineage is genotype and Regime is biomass treatment. Design (Observed ~ Generation + Lineage + Regime) Df: degrees of freedom; SumOfSqs : sum of squares; Mean sq: mean squares; F: F statistic; Pr(>F): p-value.

```
             Df Sum Sq Mean Sq F value   Pr(>F)
Generation    7  65.46   9.351   14.46 1.66e-14 ***
Lineage       1   0.00   0.000    0.00    0.998
Regime        1   0.07   0.071    0.11    0.740
Residuals   160 103.47   0.647
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

Supplementary Table 2.6. ANOVA for Simpson alpha diversity index. This shows a significant influence of generation on alpha diversity. Lineage is genotype and Regime is biomass treatment. Design (Observed ~ Generation + Lineage + Regime) Df: degrees of freedom; SumOfSqs : sum of squares; Mean sq: mean squares; F: F statistic; Pr(>F): p-value.

```
Type III Analysis of Variance Table with Satterthwaite's method
                          Sum Sq Mean Sq NumDF DenDF F value   Pr(>F)
Lineage                   0.0035  0.0035     1   139  0.0065 0.936050
Regime                    0.0945  0.0945     1   139  0.1770 0.674654
Generation                1.4097  0.2014     7   139  0.3771 0.914458
Lineage:Regime            4.6552  4.6552     1   139  8.7162 0.003704 **
Lineage:Generation       11.2206  1.6029     7   139  3.0013 0.005765 **
Regime:Generation         2.8505  0.4072     7   139  0.7624 0.619688
Lineage:Regime:Generation 9.7397  1.6233     6   139  3.0393 0.007990 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Supplementary Table 2.7. Linear mixed effect model for Shannon diversity index. This shows a significant effect of the generation and interaction term with genotype (“Lineage”) and biomass treatment (“Regime”). Design lmer (DiveristyMetric, Genotype*Generation*Regime, (1|Generation)) Sum Sq: sum of squares; Mean sq : mean squares; NumDF: numerator degrees of freedom; DenDF: denominator degrees of freedom; F value: F statistic; Pr(>F): p-value.

Supplementary Table 2. 7 Linear mixed effect model for Shannon diversity index

```
Type III Analysis of Variance Table with Satterthwaite's method
                          Sum Sq Mean Sq NumDF DenDF F value   Pr(>F)
Lineage                   0.0318  0.0318     1   139  0.0585 0.809224
Regime                    0.0580  0.0580     1   139  0.1068 0.744271
Generation                4.1433  0.5919     7   139  1.0900 0.372960
Lineage:Regime            4.1906  4.1906     1   139  7.7171 0.006226 **
Lineage:Generation       10.3724  1.4818     7   139  2.7287 0.011039 *
Regime:Generation         3.1643  0.4520     7   139  0.8324 0.562098
Lineage:Regime:Generation 10.2760  1.7127     6   139  3.1539 0.006250 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Supplementary Table 2.8. Linear mixed effect model for Simpson diversity index. This shows significant effect of the interaction term with genotype (“Lineage”) and biomass treatment (“Regime”). Design lmer (DiveristyMetric, Genotype*Generation*Regime, (1|Generation)) Sum Sq: sum of squares; Mean sq : mean squares; NumDF: numerator degrees of freedom; DenDF: denominator degrees of freedom; F value: F statistic; Pr(>F): p-value.

Literature Cited Chapter 2

1. Gouda S, Kerry RG, Das G, Paramithiotis S, Shin HS, Patra JK. Revitalization of plant growth promoting rhizobacteria for sustainable development in agriculture. Microbiol Res. 2018;206.
2. Jones DL, Nguyen C, Finlay RD. Carbon flow in the rhizosphere: Carbon trading at the soil-root interface. Plant Soil. 2009;321(1–2):5–33.
3. Dennis PG, Miller AJ, Hirsch PR. Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? FEMS Microbiol Ecol. 2010;72(3):313–27.
4. Hartmann A, Schmid M, van Tuinen D, Berg G. Plant-driven selection of microbes. Plant Soil. 2009;321(1–2):235–57.
5. Chaparro JM, Sheflin AM, Manter DK, Vivanco JM. Manipulating the soil microbiome to increase soil health and plant fertility. Biol Fertil Soils. 2012;48(5):489–99.
6. Marschner P, Marschner P, Crowley D, Crowley D, Yang CH, Yang CH, et al. Development of specific rhizosphere bacterial communities in relation to plant species, nutrition and soil type. Plant Soil. 2004;261(1–2):199–208.
7. Garbeva P, van Veen J a, van Elsas JD. Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. Annu Rev Phytopathol. 2004;42(29):243–70.
8. Tkacz A, Cheema J, Chandra G, Grant A, Poole PS. Stability and succession of the rhizosphere microbiota depends upon plant type and soil composition. ISME J. 2015;1–11.
9. Berg G, Smalla K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. FEMS Microbiol Ecol. 2009;68(1):1–13.
10. Jin H, Yang X, Liu R, Yan Z, Li X, Li X, et al. Bacterial community structure associated with the rhizosphere soils and roots of Stellera chamaejasme L. along a Tibetan elevation gradient. Ann Microbiol. 2018 May 1;68(5):273–86.
11. Barbosa Lima A, Cannavan FS, Navarrete AA, Teixeira WG, Kuramae EE, Tsai SM. Amazonian Dark Earth and Plant Species from the Amazon Region Contribute to Shape Rhizosphere Bacterial Communities. Microb Ecol. 2015 May 1;69(4):855–66.
12. Marques JM, da Silva TF, Vollu RE, Blank AF, Ding GC, Seldin L, et al. Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown sweet potato plants. FEMS Microbiol Ecol. 2014;88(2):424–35.
13. Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA. A Synthetic Community Approach Reveals Plant Genotypes Affecting the Phyllosphere Microbiota. PLoS Genet. 2014;10(4).
14. Zancarini A, Mougel C, Voisin AS, Prudent M, Salon C, Munier-Jolain N. Soil Nitrogen Availability and Plant Genotype Modify the Nutrition Strategies of M. truncatula and the Associated Rhizosphere Microbial Communities. PLoS ONE. 2012;7(10).
15. Wei S, Liu B, Ni K, Ma L, Shi Y, Leng Y, et al. Rhizosphere Microbial Community Shows a Greater Response Than Soil Properties to Tea (Camellia sinensis L.) Cultivars. Agronomy. 2023 Jan 1;13(1).

16. Yuan J, Chaparro JM, Manter DK, Zhang R, Vivanco JM, Shen Q. Roots from distinct plant developmental stages are capable of rapidly selecting their own microbiome without the influence of environmental and soil edaphic factors. Soil Biol Biochem. 2015;89(AUGUST):206–9.
17. Houlden A, Timms-Wilson TM, Day MJ, Bailey MJ. Influence of plant developmental stage on microbial community structure and activity in the rhizosphere of three field crops. FEMS Microbiol Ecol. 2008;65(2):193–201.
18. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature. 2013 Aug 1;488(7409):86–90.
19. Bulgarelli D, Rott M, Schlaeppi K, Ver E, Themaat LV, Ahmadinejad N, et al. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature. 2012;488(7409):91–5.
20. Coleman-Derr D, Desgarennes D, Fonseca-Garcia C, Gross S, Clingenpeel S, Woyke T, et al. Plant compartment and biogeography affect microbiome composition in cultivated and native Agave species. New Phytol. 2016 Jan;209(2):798–811.
21. Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, et al. Structure, variation, and assembly of the root-associated microbiomes of rice. Proc Natl Acad Sci U S A. 2015 Feb 24;112(8):E911–20.
22. Wagner MR, Lundberg DS, Del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat Commun. 2016 Jul 12;7.
23. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature. 2012;488(7409):86–90.
24. Schlaeppi K, Dombrowski N, Oter RG, Ver Loren Van Themaat E, Schulze-Lefert P. Quantitative divergence of the bacterial root microbiota in Arabidopsis thaliana relatives. Proc Natl Acad Sci U S A. 2014;111(2):585–92.
25. Hu L, Robert CAM, Cadot S, Zhang X, Ye M, Li B, et al. Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. Nat Commun. 2018 Dec 1;9(1).
26. Van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T, et al. Plant-soil feedbacks: The past, the present and future challenges. J Ecol. 2013 Mar;101(2):265–76.
27. Panke-buisse K, Poole AC, Goodrich JK, Ley RE, Kao-kniffin J. Selection on soil microbiomes reveals reproducible impacts on plant function. ISME J. 2014;1–10.
28. Swenson W, Wilson DS, Elias R. Artificial ecosystem selection. Proc Natl Acad Sci U S A. 2000 Aug 1;97(16):9110–4.
29. Mueller UG, Juenger TE, Kardish MR, Carlson AL, Burns KM, Edwards JA, et al. Artificial Selection on Microbiomes To Breed Microbiomes That Confer Salt Tolerance to Plants METHODS AND PROTOCOLS [Internet]. 2021. Available from: https://journals.asm.org/journal/msystems
30. Lau JA, Lennon JT. Rapid responses of soil microorganisms improve plant fitness in novel environments. Proc Natl Acad Sci. 2012;109(35):14058–62.

31. terHorst CP, Lennon JT, Lau JA. The relative importance of rapid evolution for plant-microbe interactions depends on ecological context. Proc R Soc B Biol Sci. 2014 Jun 22;281(1785):20140028.
32. Jochum MD, McWilliams KL, Pierson EA, Jo YK. Host-mediated microbiome engineering (HMME) of drought tolerance in the wheat rhizosphere. PLoS ONE. 2019 Dec 1;14(12).
33. Cordovez V, Rotoni C, Dini-Andreote F, Oyserman B, Carrión VJ, Raaijmakers JM. Successive plant growth amplifies genotype-specific assembly of the tomato rhizosphere microbiome. Sci Total Environ. 2021 Jun 10;772.
34. Norma M. Morella, Francis Cheng-Hsuan Weng, Pierre M. Joubert, C. Jessica E. Metcalf, Steven Lindow, Britt Koskella. Successive passaging of a plant-associated microbiome reveals robust habitat and host genotype-dependent selection. PNAS. 2020 Jan 14;117(2).
35. Parada AE, Needham DM, Fuhrman JA. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol. 2016 May 1;18(5):1403–14.
36. Apprill A, Mcnally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat Microb Ecol. 2015 Jun 4;75(2):129–37.
37. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012 Aug;6(8):1621–4.
38. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth's multiscale microbial diversity. Nature. 2017 Nov 1;551(7681):457.
39. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019 Aug 1;37(8):852–7.
40. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jun 29;13(7):581–3.
41. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinforma Oxf Engl. 2022 Nov 30;38(23):5315–6.
42. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020 May 1;37(5):1530–4.
43. Bunn AG. A dendrochronology program library in R (dplR). Dendrochronologia. 2008 Oct 1;26(2):115–24.
44. Burns AR, Stephens WZ, Stagaman K, Wong S, Rawls JF, Guillemin K, et al. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. ISME J. 2016;10(3):655–64.
45. Sloan WT, Woodcock S, Lunn M, Head IM, Curtis TP. Modeling taxa-abundance distributions in microbial communities using environmental sequence data. In: Microbial Ecology. 2007. p. 443–55.

46. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. Nat Commun. 2022 Dec 1;13(1).
47. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nat Commun. 2020 Dec 1;11(1).
48. Radhakrishnan R, Hashem A, Abd Allah EF. Bacillus: A biological tool for crop improvement through bio-molecular changes in adverse environments. Front Physiol. 2017 Sep 6;8(SEP).
49. Mujakić I, Piwosz K, Koblížek M. Phylum Gemmatimonadota and Its Role in the Environment. Microorganisms. 2022 Jan 1;10(1).
50. Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E. The Prokaryotes: A Handbook on the Biology of Bacteria: Archaea. Bacteria: Firmicutes, Actinomycetes. The Prokaryotes: A Handbook on the Biology of Bacteria. 2006.
51. Grady EN, MacDonald J, Liu L, Richman A, Yuan ZC. Current knowledge and perspectives of Paenibacillus: A review. Microb Cell Factories. 2016 Dec 1;15(1).
52. Ramesh A, Sharma SK, Sharma MP, Yadav N, Joshi OP. Plant Growth-Promoting Traits in Enterobacter cloacae subsp. dissolvens MDSR9 Isolated from Soybean Rhizosphere and its Impact on Growth and Nutrition of Soybean and Wheat Upon Inoculation. Agric Res. 2014;3(1):53–66.
53. Kämpfer P, Ruppel S, Remus R. Enterobacter radicincitans sp. nov., a plant growth promoting species of the family Enterobacteriaceae. Syst Appl Microbiol. 2005 Apr 22;28(3):213–21.
54. Khalifa AYZ, Alsyeeh AM, Almalki MA, Saleh FA. Characterization of the plant growth promoting bacterium, Enterobacter cloacae MSR1, isolated from roots of non-nodulating Medicago sativa. Saudi J Biol Sci. 2016;23(1):79–86.
55. Harkes P, van Steenbrugge JJM, van den Elsen SJJ, Suleiman AKA, de Haan JJ, Holterman MHM, et al. Shifts in the Active Rhizobiome Paralleling Low Meloidogyne chitwoodi Densities in Fields Under Prolonged Organic Soil Management. Front Plant Sci. 2020 Jan 10;10.
56. Meier MA, Xu G, Lopez-Guerrero MG, Li G, Smith C, Sigmon B, et al. Association analyses of host genetics, root-colonizing microbes, and plant phenotypes under different nitrogen conditions in maize. eLife. 2022;11.
57. Porter AS, Evans-Fitz.Gerald C, McElwain JC, Yiotis C, Elliott-Kingston C. How well do you know your growth chambers? Testing for chamber effect using plant traits. Plant Methods. 2015;11(1).
58. Bulgarelli D, Garrido-Oter R, Münch PC, Weiman A, Dröge J, Pan Y, et al. Structure and function of the bacterial root microbiota in wild and domesticated barley. Cell Host Microbe. 2015;17(3):392–403.
59. Brown SP, Grillo MA, Podowski JC, Heath KD. Soil origin and plant genotype structure distinct microbiome compartments in the model legume Medicago truncatula. Microbiome. 2020 Sep 28;8(1).
60. Basu A, Prasad P, Das SN, Kalam S, Sayyed RZ, Reddy MS, et al. Plant Growth Promoting Rhizobacteria (PGPR) as Green Bioinoculants: Recent Developments, Constraints, and Prospects. Sustainability. 2021 Jan 22;13(3):1140.

61. Micallef S, Colón-Carmona A. Genetic and Developmental Control of Rhizosphere Bacterial Communities. In: de Bruijn FJ, editor. Molecular Microbial Ecology of the Rhizosphere [Internet]. 1st ed. Wiley; 2013 [cited 2023 Aug 20]. p. 257–63. Available from: https://onlinelibrary.wiley.com/doi/10.1002/9781118297674.ch24
62. Micallef SA, Channer S, Shiaris MP, Colón-Carmona A. Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. Plant Signal Behav. 2009 Aug;4(8):777–80.
63. Lee S, I. Sergeeva L, Vreugdenhil D. Natural variation of hormone levels in Arabidopsis roots and correlations with complex root architecture: Natural variation of phytohormones in Arabidopsis roots. J Integr Plant Biol. 2018 Apr;60(4):292–309.
64. Matsushima C, Shenton M, Kitahara A, Wasaki J, Oikawa A, Cheng W, et al. Multiple analysis of root exudates and microbiome in rice (Oryza sativa) under low P conditions. Arch Microbiol. 2021 Nov 1;203(9):5599–611.
65. Huang XF, Chaparro JM, Reardon KF, Zhang R, Shen Q, Vivanco JM. Rhizosphere interactions: Root exudates, microbes, and microbial communities1. Botany. 2014;92(4):267–75.
66. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. Plant–microbiome interactions: from community assembly to plant health. Nat Rev Microbiol. 2020;18(11).
67. Gu S, Wei Z, Shao Z, Friman VP, Cao K, Yang T, et al. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. Nat Microbiol. 2020 Aug 1;5(8):1002–10.

CHAPTER 3

CHARACTERIZATION OF RHIZOSPHERE MICROBIOMES ASSOCIATED WITH NEAR-ISOGENIC LINES DERIVED FROM *ARABIDOPSIS THALIANA* CVI AND L*ER* USING 16S rRNA METAGENOMIC SEQUENCING

Nachiket Shankar[1*], Stuart Morey[1*], Julia Rosenwald[2], Miraj J Rawal[1], Rick Kesseli[1]

[1]Department of Biology, University of Massachusetts, Boston, MA 02125,

[2]Departement of Biology, Franklin and Marshall College, PA 17603

Abstract

Plant genetic control of the microbiome refers to the influence exerted by the plant's genetic makeup on the composition and functioning of the microbial community associated with its roots and aerial parts. Plants release root exudates, into the soil, shaping the microbial communities in the rhizosphere. In particular, the product of genes in the plant can interact with specific microorganisms, leading to symbiotic or antagonistic relationships. Plant genetic control plays a crucial role in determining the abundance and diversity of beneficial microorganisms, which contribute to nutrient acquisition, disease resistance, and stress tolerance in plants. However, despite over a hundred years of research in this area, the underlying plant genetic mechanisms in structuring the rhizosphere are still poorly understood. This study aims to narrow down the candidate regions involved in plant genetic control of the microbiome in *Arabidopsis thaliana* L*er* based on a previously identified quantitative trait loci (QTL) region, Stuart Morey (Boston, USA). The QTL region identified spans of about 3.75 Mbp at the beginning of chromosome 1 of *Arabidopsis thaliana* L*er* consisting of around ~995 genes. A highly replicated (12 replicates) fine-mapping approach

using near-isogenic lines (NILs) is employed to analyze 16S rRNA metagenomic data coupled with statistical analyses. The study confirms the findings from previous studies that have shown a significant but small effect of plant genotype on the rhizosphere microbial community structure. In addition, we find evidence of genotype-specific differential abundance of specific microbial taxa. Finally, we narrow down the candidate region from ~995 genes to 418 genes. This study is a step forward in deciphering the mechanisms behind plant genetic control of the microbiome in *Arabidopsis thaliana*.

Introduction

Plant roots consistently release substances into the rhizosphere, a phenomenon referred to as exudation or rhizodeposition. This release can range from less than 10% to as much as 44% of the plant's net carbon assimilation (1–4). Root exudates include compounds such as amino acids, sugars, secondary metabolites, resins, and enzymes that may be transported across the cell membrane or released from root border cells (5). The plant root exudates differ across plant species and plant developmental stages. Evidence suggests that these differences may underpin plant modulation of rhizosphere bacterial communities at a species- or even cultivar-specific manner (4,6–12).

Several studies have attempted to further understand the underlying genetic mechanisms governing the release of specific plant exudates and the resulting colonization of specific rhizosphere bacteria (13–19). There are a few studies that have indicated the involvement of genes related to plant defense and cross-membrane transport of organic substances such as ABC transporters and secondary metabolites in shaping the rhizosphere community (20–24). However, despite decades of research in this area, the underlying plant genetic mechanisms that structure the rhizosphere microbiome are still not very well understood. The traditional approaches used to identify and understand the genetic bases of traits including plant traits that interact with microbes, include quantitative trait loci (QTL) mapping using recombinant inbred lines (RILs) or near-isogenic lines (NILs) to narrow down large chromosomal regions in these model systems (25–27). Recent advances in sequencing have sparked increasing research interest in employing quantitative approaches using RILs for mapping the microbiome as a phenotype, encompassing both the phyllosphere and rhizosphere (28–31). However, the mapping of microbiome-associated QTL remains a

sparsely explored domain, and this is an upcoming area of research, with only one previous study in *Arabidopsis thaliana* (28,31).

This study leverages NILs to validate and narrow down a previously identified microbiome-associated QTL region (Stuart Morey Boston, USA, unpublished data). NILs are a set of genetically similar plant lines or varieties that differ only at a specific targeted gene or genomic region and are commonly used to investigate the effects of a particular gene or genomic region on various traits while minimizing genetic background noise. The NILs used in this study were developed by Keurentjes et al. (32) with the parents L*er* and Cvi. These NILs are nearly identical to L*er* but with a small region of Cvi on the top of chromosome 1 (Figure 3.1). Chromosome 1 in *Arabidopsis thaliana* has been implicated in plant functions such as plant growth (33), development (34), disease resistance (35), and recruitment of beneficial rhizobacteria (36). The NILs, N1-2.1, N1-2.5, N1-2.8, and N1-3, localized on the top of chromosome 1 correspond to a previously identified QTL that affected the microbial community (Stuart Morey Boston, USA, unpublished data).

In this study, we conducted a highly replicated study (12 biological replicates) to examine plant genetic effects of NILs obtained from a cross between parents L*er* and Cvi on the microbiome using 16S rRNA metagenomic sequencing data. The results confirm the findings from previous studies that have shown a significant but small effect of plant genotype on the microbial community structure (37,38). We find several taxa that are differentially associated with the NILs including *Bacillus flexus,* known plant growth-promoting bacteria (39–41). We also identify some candidate genes in the plants that have been previously implicated to impact microbial taxa and communities in the soil including *Bacillus* spp (42).

Material and Methods

Plant seeds, growth conditions, and soil

The *Arabidopsis thaliana* accessions grown in this experiment were L*er*, Cvi, N1-2.1, N1-2.5, N1-2.8 and N1-3, which are NILs with L*er* with small fragments of Cvi at the end of Chromosome 1. The potting soil used was PRO-MIX PGX, a commercial mixture comprising 80-90% sphagnum peat moss and small quantities of perlite. It was autoclaved twice for 40 minutes with a 48-hour interval between each sterilization. The potting soil was then sifted through a 3mm sieve and combined with field soil in a 3.5:1 ratio respectively. The field soil obtained from the Center for Agricultural Research in Waltham, Massachusetts, is composed of 44% sand, 49% silt, and 7% clay, representative of agricultural and grassland ecosystems. The mixture was homogenized using a cement mixer sterilized with 70%. ethanol. Distilled water was gradually added to achieve a final ratio of 1:2 (water to soil). Before planting, all seeds were sterilized for 8 minutes in a solution of 50% volume/volume bleach with one drop of Tween^{TM} 20 and rinsed 8 times with sterile distilled water. Seeds were then cold stratified at 4^{o}C for 2 days and planted directly onto the soil-mix. Twenty-four biological replicates of each accession were planted in separate pots, as well as 24 replicates each of unplanted uninoculated and inoculated potting soil. The entire setup with planted and unplanted pots was arranged in a completely randomized design to minimize any batch effects. All plants were grown at 22^{o}C day/ 18^{o}C night with 12/12-hour day night cycle in a controlled growth chamber. Relative humidity ranged from 35-60 and the light intensity 96uE.

Twelve replicates each were harvested 32 days (early) and 39 days (late) after planting. The above-ground portion of the plant, the rhizosphere, and the bulk soil were

separated from each other for each rhizotube ensuring there was no cross-contamination (Figure 3.2). The above-ground parts of the plant, stem, and leaves were dried at 70^{o}C for 4 days. All the plants were weighed individually on a closed weighing scale accurate to 1mg. The root-soil complex (comprising the rhizosphere and the endosphere, henceforth referred to as the rhizosphere) was obtained by shaking the excess soil off the root and placing the root and the remaining attached soil in a sterile 5 ml tube. Bulk soil samples were moved to a sterile Ziplock bag. Tubes and bags were immediately transferred to dry ice and then stored at -80^{o}C for DNA analysis.

DNA extraction and 16s rRNA gene library prep

Microbial DNA is isolated from the frozen rhizosphere sample using the Machery-Nagel Nucleospin Soil DNA extraction kit (MACHEREY-NAGEL Inc., Allentown, PA, USA). Approximately 0.1g of rhizosphere soil sample was used for DNA isolation. All samples were diluted to 5 ng ul^{-1} with PCR-grade water. 16S rRNA gene was amplified from the isolated DNA samples in triplicate in 96-well PCR plates. The PCR primers used were for the 16S rRNA V4 region, 515F (5'-GTGYCAGCMGCCGCGGTAA-3') (ref. 36) and 806R (5'-GGACTACNVGGGTWTCTAAT-3') (ref. 37) for downstream paired-end Illumina (Illumina, Inc., San Diego, CA, USA) barcoded sequencing (45). The PCR cycling conditions used were as follows: 94^{o}C for 3 min; 25 cycles of 94^{o}C for 45 s, 50^{o}C for 60 s, 72^{o}C for 90 s, and final elongation at 72^{o}C for 10 min. The triplicate amplified samples were pooled and then purified and normalized using the SequalPrep™ Normalization Plate Kit (Invitrogen Corporation, Carlsbad, Canada). Finally, multiplexed paired-end sequencing

was carried out in the Illumina MiSeq platform using earth microbiome primers (46).

Sequence Data Processing

The paired-end sequences obtained from the Illumina Miseq were demultiplexed with QIIME 2 and converted into individual sequence fastq files for each sample (45,47). The rest of the sequence processing was carried out in R using the DADA2 package. The reads were processed in R using the following command in DADA2

`filterAndTrim(fnFs, filtFs, fnRs, filtRs, truncLen=c(145,145), minLen=50, maxN=0, maxEE=c(2,2), truncQ=2, rm.phix=TRUE, compress=TRUE, multithread=TRUE)`

De novo OTU (Operational Taxonomic Units) picking was performed using DADA2 (48). It has been shown to resolve amplicon sequence variants (ASVs) up to single nucleotide differences. Chimeras were removed. The taxonomy was assigned using a 100% cut-off for species-level identification by using the most recently updated (2022) GTDB reference database (49). The phylogenetic tree was constructed using IQTree (50), with 1000 ultrafast bootstraps and Modelfinder to identify the best model, in this case, the best-fit model was SYM+R10 chosen according to BIC. The data comprising the OTU table, phylogenetic tree, taxonomy table, and sample meta-data were then parsed through R studio using the phyloseq package (51).

Statistical Analyses

Alpha diversity was computed using four metrics, Observed, Chao1, Shannon, and Simpson diversity indices. Beta diversity was calculated using weighted UniFrac, unweighted UniFrac and Bray-Curtis metrics, and Principle Coordinate Analysis plots were constructed using the first two axes that explain the most variance. PERMANOVA was

carried out for the dataset using the adonis2 test. Both weighted and unweighted UniFrac distance matrices were used to account for the abundance, presence/absence, and phylogeny of the microbiome. The following model was used to assess the effects of genotype and biomass on microbial community `adonis2(distance.matrix~Genotype+Biomass, data = meta, permutations=999)`.

Differentially abundant taxa were estimated using the DeSeq2 package with default parameters in R (52). Starting from the end of Table 3.1. (NIL 1-2.1), each NIL was tested against the one directly above it to determine differentially abundant OTUs. The strength of this method lies in the sample size, which is ideal for an experiment of this scale with 12 replicates per treatment.

Data mining and Gene Ontology

The QTL region that was previously identified corresponds to the top of chromosome 1 *Arabidopsis thaliana*. Based on the genome annotation available on The Arabidopsis Information Resource (TAIR), all genes from this region were identified (53). The gene list was utilized to extract information on SNPs for each gene differentiating L*er* and Cvi from the 1001 genomes API using custom scripts (54). The impact of a particular SNP for an individual gene was classified as having a "low," "moderate," or "high" impact, depending on the nature of the mutation. The gene ontology (GO) annotations and gene set enrichment data for moderate and high impact SNPs were obtained using g:Profiler.

Results

Biomass and Growth Stage

There were large differences in the biomass among the genetic lines (Figure 3.3). The NILs had significantly lower biomass than either of the two parents. ANOVA was carried out to test the influence of genotype and harvest date on biomass. The genotype was a strong predictor of biomass, and the harvest date was a weaker predictor of biomass (Supplementary Table 3.1 and Table 3.2). The growth stage (Vegetative, Budding, or Bolting) at day 32 is depicted across all samples (Figure 3.4). The results showed that there is a substantial number of plants in the vegetative stage of development for both Cvi and L*er*, whereas a majority of NILs have proceeded into the reproductive phase, displaying signs of either budding or bolting.

Microbial community Composition

After pre-processing in DADA 2, the Illumina Miseq (SY-410-1003) generated a total of 3,453,277 reads, with an average of 13,231 reads per sample. Singletons and chimeras were removed during pre-processing. The phyla that were most represented were Proteobacteria (41.2%), Bacteroidetes (15.5%), Acidobacteria (7.5%), and Verrucomicrobia (7.4%) (Figure 3.5). Among the Classes identified in the dataset, the most represented were Alphaproteobacteria (21.4%), Saprospirae (10.1%), Deltaproteobacteria *(7.5%)*, and Gammaproteobacteria (6.4%) (Figure 3.6). The relative abundance in the dataset is typical of the microbiome and heavily weighted to a few abundant groups with a long tail of rare taxa.

At these taxonomic scales, there were no major difference among the microbial communities associated with the six genotypes of this study.

Alpha Diversity

Four metrics for alpha diversity were used to characterize the microbial communities (Figure 3.7). Observed and Chao1 which represent the total number of species in the dataset revealed similar levels of species richness among the four NILs and L*er* but a significantly lower level for Cvi. Some of these differences in richness between Cvi and N1-2.5 and N1-2.8 are lost in the Shannon analysis which weighs abundant taxa more heavily (Supplementary Tables 3.3 and 3.4). In contrast, the Simpson index which is more sensitive to evenness revealed the community associated with L*er* to be significantly lower than the others and N1-2.5 and N1-2.8 have the highest values (Supplementary Tables 3.5).

Beta Diversity

Principle Coordinate Analysis (PCoA) using weighted and unweighted UniFrac were plotted to gauge whether microbial communities are influenced by the genotype of the host plant (Figure 3.8). Weighted UniFrac considers the relative abundances of shared species between samples. Unweighted UniFrac accounts for changes in the presence or absence of species between samples. The first two coordinate axes accounted for the highest variation in the data, ranging from 11.4% for unweighted UniFrac and 36% for weighted UniFrac. As expected, microbial communities of Cvi were differentiated from those of L*er* and its NILs and this was most evident in the unweighted UniFrac plot. However, the variance accounted for by the first two axes in both plots is low.

PERMANOVA was done using adonis2 test for the data using both weighted and unweighted UniFrac distances to estimate the effect of genotype and biomass on the microbial community (Supplementary Table 3.2 and 3.3). For both weighted and unweighted UniFrac analyses, there were small, but significant impacts of both genotype and biomass on the microbial communities. To further analyze the variance explained in the microbial community associated with the plant genotypes, a pairwise PERMANOVA was conducted using *pairwiseAdonis* package in R (Figure 3.9). The weighted and unweighted UniFrac revealed similar patterns, with more pronounced effects in the weighted UniFrac, which considers absolute abundance. As expected, prominent differences were detected between the microbial community of Cvi and those of L*er* and its NILs. Greater disparities existed between the NILs and Cvi in contrast to the distinctions between Cvi and L*er*, suggesting a possible intricate genetic interplay involving the genes in the upper region of chromosome 1 from Cvi in the context of the L*er* background. Comparatively, the dissimilarities among the NILs were of lesser magnitude, and no noteworthy distinctions emerged between the microbiomes of N1-2.5 and N1-2.8. Furthermore, the weighted UniFrac revealed the microbiome of L*er* to be more similar to the N1-2.1 microbiome than the N1-2.5, N1-2.8 and N1-3 microbiomes.

Differential Abundance and Candidate Regions

The NILs were compared to determine differentially abundant (DA) OTUs to narrow down the region of interest on chromosome 1 (Figure 3.10). The comparison between NIL 1-2.5 Vs NIL 1-2.1 and NIL 1-3 Vs 1-2.8 yielded 15 and 36 DA OTUs respectively. A heat

map of these results is plotted in Figure 3.11 and Figure 3.12. Taxa from the families Chitinophagaceae, Sphingobacteriaceae, Pseudomonadaceae, and Bdellovibrio were present in higher abundance in the N1-2.5 compared to N1-2.1 samples and, Sphingobacteriales , Chthoniobacterales, and Burkholderiales were present in higher abundance in N1-3 compared to N1-2.8. *Bacillus flexus*, *Alkanibacter difficillis* and a bacteria from Family Erythrobacteraceae exhibited greater abundance in the Cvi and L*er* parents when contrasted with the NILs N1-2.5, N1-2.8, and N1-3 (Figure 3.13). However, no significant DA was found for N-1.2.1. This suggests that the interaction of the Cvi region between N1-2.1 and N1-2.5 with the L*er* background plays a role in the differential abundance of these taxa. This section is highlighted in orange as a potential candidate region (Figure 3.13). Results of differentially abundant taxa can be found in supplementary materials (Supplementary Table 3.4).

The genomic regions spanning N1-2.1 to N1-2.5 and N1-2.8 to N1-3 exhibit the highest number of DA taxa. These sections are the prime contenders for containing genes linked to the modulation of the rhizosphere microbiome. The SNPs differentiating Cvi and L*er* genes were determined for these regions, and subsequent gene enrichment analyses focused on SNPs with moderate and high impacts. Among these, 90 genes exhibited moderate-impact SNPs, while 6 genes displayed high-impact SNPs within the N1-2.1 to N1-2.5 region. This region housed the genes *EDR1* and *RTM1*, previously recognized for their involvement in *Arabidopsis thaliana* pathogen resistance (55), which displayed moderate and high-impact SNPs, respectively, between Cvi and L*er*. Additionally, a high-impact SNP in *AAO4* (AT1G04850), a gene linked to benzaldehyde dehydrogenase activity, was previously found to be associated with *Bacillus aryabhattai* (42). The region between N1-2.8 to N1-3

had 137 genes with moderate-impact SNPs and 18 with high-impact SNPs. GO annotation data indicated that this region contains genes associated with, but not limited to, secondary metabolite production, transmembrane transport, and the regulation of defense response. Comprehensive data on SNPs and GO annotations can be found in supplementary data.

Discussion

The recruitment of microbes in the rhizosphere is a complex process and understanding the mechanism of plant genetic control of the microbiome is a prime area of investigation (38,56–59). In this study, we examined the effects of plant genotype on the microbiome and reduced the size of down a previously identified QTL region. The top of chromosome 1 has been previously found to be associated with plant growth and development (33,34,60). The RILs of the L*er* x Cvi cross have been shown to exhibit transgressive segregation, where the phenotypic range of a trait in the offspring, extends beyond the range observed in the parental lines. This phenomenon can occur due to the combination of favorable alleles from both parents. Earlier research indicated that the upper region of chromosome 1 in Cvi was associated with early flowering time (61–64). The *CRY2* (AT1G04400) gene found in the N1-2.1 region has been demonstrated to promote flowering through a single amino acid substitution in the Cvi allele, resulting in accelerated developmental cycling and early flowering (65). The process of early flowering can result in reduced biomass due to the redistribution of resources and energy from vegetative growth to reproductive growth (66,67), as evident from the significant biomass and developmental stage disparities between the NILs and parental lines (Figure 3.3 and Figure 3.4).

Previous studies show conflicting evidence on genotype-specific mediated microbiome assembly, with some findings suggesting a weak effect and some suggesting a

stronger effect (68–71). In this study, we find, there are significant influences of the plant genotype on species richness (Observed, Chao1 and Shannon statistics). Specifically, the microbial communities of Cvi are substantially less rich than those from L*er* and the NILs derived from L*er*. Interestingly, while the richness of the NILs and L*er* are all similar, the evenness patterns revealed with the Simpson statistic are not as clear, possibly indicating that the rearranged genomes of the NILs are altering the abundance of microbes and disrupting the typical host-microbe interactions of the parental lines (Figure 3.7). The Beta Diversity as revealed in the PCoA plots (Figure 3.8) does show that the Cvi and L*er* associated microbial communities are different, and that the NILs which are genetically similar to L*er*, have similar though more diverse communities than those of L*er*. Further analyses testing the influence of genotype on the microbiome, using PERMANOVA (Figure 3.9), showed that genotype has a significant but small influence on the microbial community structure supporting the landmark paper on the core microbiome of *Arabidopsis thaliana* as well as subsequent studies (38,70). The outcomes, derived from both weighted and unweighted UniFrac exhibit consistent trends, albeit with more prominent effects evident in the context of the weighted UniFrac. This may suggest that changes in the abundance of specific taxa are being influenced by variations in root exudation patterns, microbe-microbe interactions, or other subtle changes in the microenvironment. Additionally, the results from pairwise PERMANOVA underscore the importance of the host genetics in the establishment of rhizosphere microbial communities. The prominent differences between the microbial communities of Cvi when contrasted to those of L*er* and the NILs, supports this contention. The distinctions among communities associated with the NILs were less pronounced as

expected. The genetically similar L*er* and N1-2.1 lines generated very similar microbial communities as did the N1-2.5 and N1-2.8 lines.

Further analyses of the genotype's influence on the microbiome, were done with differential abundance testing using DeSeq2. The results identified 15 DA OTUs when comparing the microbiomes of N1-2.1 and N 1-2.5 and 36 DA OTUs with the N1-2.8 and N1-3 comparison (Figure 3.11 and 3.12). Taxa belonging to the Chitinophagaceae, Sphingobacteriaceae, Pseudomonadaceae, and Bdellovibrio families exhibited increased abundance in N1-2.5 compared to N1-2.1 microbiomes. Similarly, Sphingobacteriales, Chthoniobacterales, and Burkholderiales were more abundant in N1-3 when compared to N1-2.8.

Subsequent to the characterizing of DA taxa, we identified SNPs in the genes of these regions that differ for the parental Cvi and L*er* lines. Further, we focused on SNPs that would have significant impacts on gene products. Within the N1-2.1 to N1-2.5 region, we identified moderate and high-impact SNPs in genes such as *EDR1* and *RTM1*, which have been previously recognized for their role in *Arabidopsis thaliana* pathogen resistance (55). Another high-impact SNP identified, *AAO4* (AT1G04850), encodes a protein linked to benzaldehyde dehydrogenase activity and associated with *Bacillus aryabhattai* (42). Additionally, the region spanning N1-2.8 to N1-3 harbored genes associated with diverse functions including secondary metabolite production, transmembrane transport, and defense response regulation, which are important in the modulation of microbial communities (72,73). A noteworthy finding from the DA analyses is that L*er* and Cvi exhibited transgressive segregation, with the recombinant NILs having significantly lower abundance of some microbes, including *Bacillus flexus*, *Akanibacter difficillis*, and an OTU from the

Family Erythrobacteraceae, then either of the parental lines (Figure 3.13). This could potentially be attributed to the interaction of alleles from genes of Cvi in the N1-2.5 region, such as *AAO4*, with the L*er* background.

Although there is a low degree of species-level resolution of 16S rRNA data using the V4 region, it is important to note that these are significant results, as this was a very well-replicated study and does indicate that there is host-specific genotypic control of different microbial taxa in *Arabidopsis thaliana.* In conclusion, this study gives us more insight into the specific genetic effects of NILs at the top of chromosome 1, on plant-microbe interactions and reinforces the application of conventional mapping methods in narrowing down the potential regions responsible for plant-microbe interactions. Further studies to understand the mechanisms of plant genetic control on the microbiome, entail fine mapping coupled with plant gene expression and complete metagenomic sequencing.

Author Contributions

N.S. Conceptualization, methodology, data collection, data analysis and visualization, writing-original draft, writing-reviewing and editing. S.M. conceptualization and methodology. J.R. data collection. M.J.R data collection. R.K. resources, funding, writing-reviewing and editing.

TABLES

Genetic Line	Markers on Genetic Map	Corresponding position in base pairs	Number of known genes
NIL 1-3	0-15.1 cM	0 – 3,400,000 bp	995
NIL 1-2.8	0-10.9 cM	0 – 2,500,000 bp	747
NIL 1-2.5	0-7.8 cM	0 – 1,800,000 bp	525
NIL 1-2.1	0- 3.9 cM	0 – 1,200,000 bp	355

Table 3.1 Genomic markers and number of genes for NILs on chromosome 1. The genomic markers and the number of gene sequences for NILs in chromosome 1.

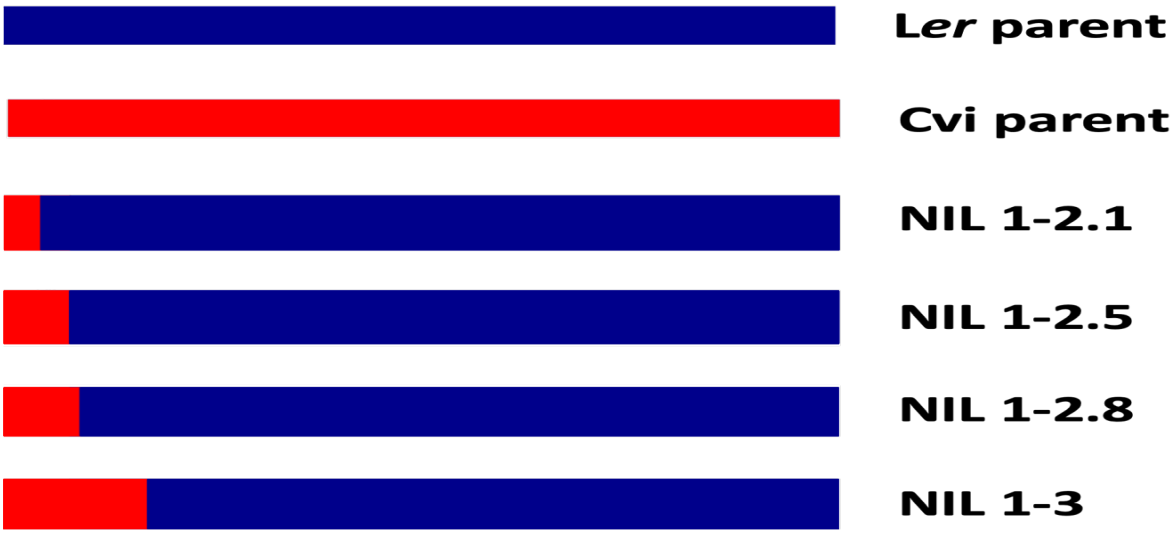


Figure 3.1. Cvi, L*er* and Near isogenic lines on chromosome 1. Near isogenic lines on chromosome 1 that were obtained from the cross between L*er* (blue) and Cvi (red).

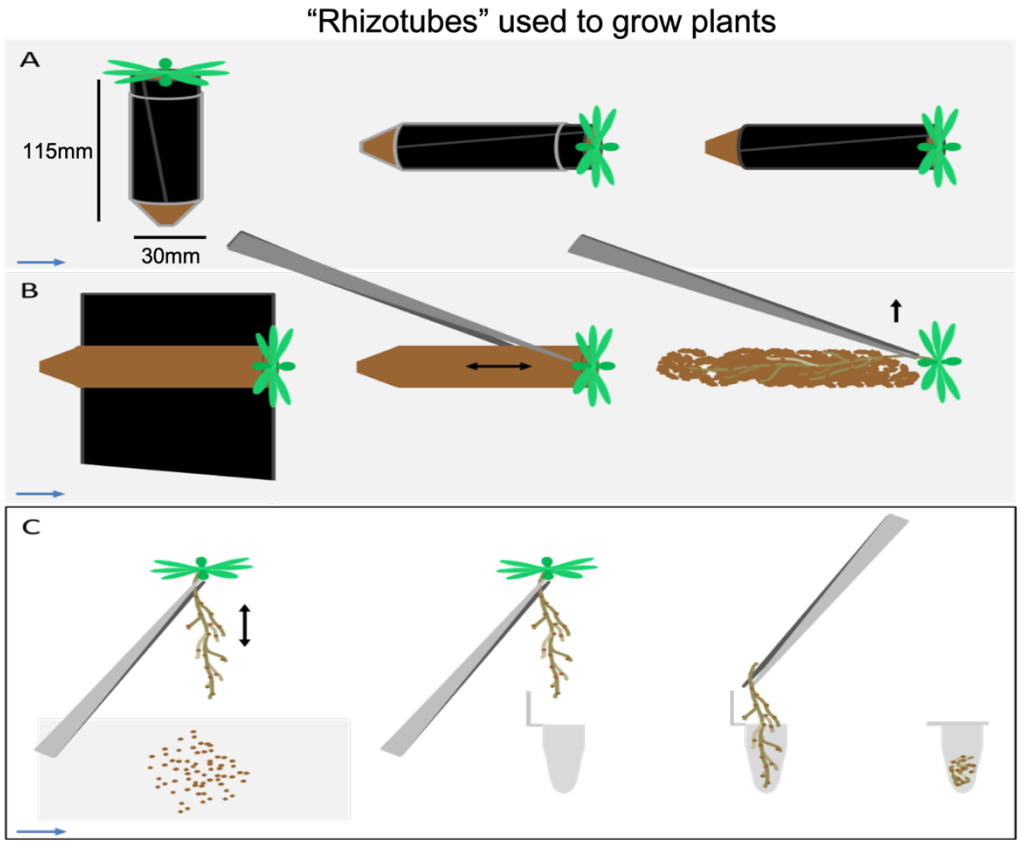
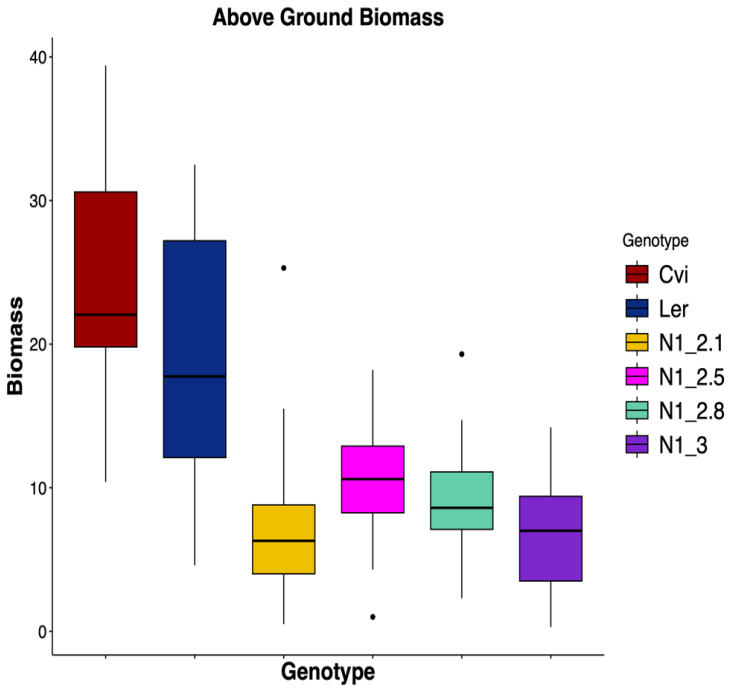


Figure 3.2. Rhizotube apparatus. The purpose of this illustration is to demonstrate how the custom-made rhizotubes facilitate the collection of plant rhizosphere. The black insert within the tube is unwrapped to extract the rhizosphere, which is then separated from the bulk soil by shaking the roots. After removing the above-ground portion of the plant, the root-soil complex (comprising the rhizosphere and the endosphere) was obtained by shaking off the excess soil and then placed in a sterile 5ml tube and bulk soil samples were moved to a sterile Ziplock bag. Both were immediately transferred to dry ice and then stored at -80^{o}C for DNA analysis, Stuart Morey (Boston, USA).



Group1	Group2	p adj
Ler	Cvi	0.044
N1_2.1	Cvi	0.000
N1_2.5	Cvi	0.000
N1_2.8	Cvi	0.000
N1_3	Cvi	0.000
N1_2.1	Ler	0.000
N1_2.5	Ler	0.000
N1_2.8	Ler	0.000
N1_3	Ler	0.000
N1_2.5	N1_2.1	0.412
N1_2.8	N1_2.1	0.849
N1_3	N1_2.1	1.000
N1_2.8	N1_2.5	0.981
N1_3	N1_2.5	0.264
N1_3	N1_2.8	0.705

Figure 3.3. Above ground biomass of Cvi, L*er* and NILs. Above-ground biomass plotted for all samples and colored by genotype and results from Tukey test shown in table on the right for all pairwise comparisons. The biomass of the NILs is in general much lower than that of the parents. (ANOVA (Biomass ~ Genotype + Harvest Date; pairwise significance using Tukey test).

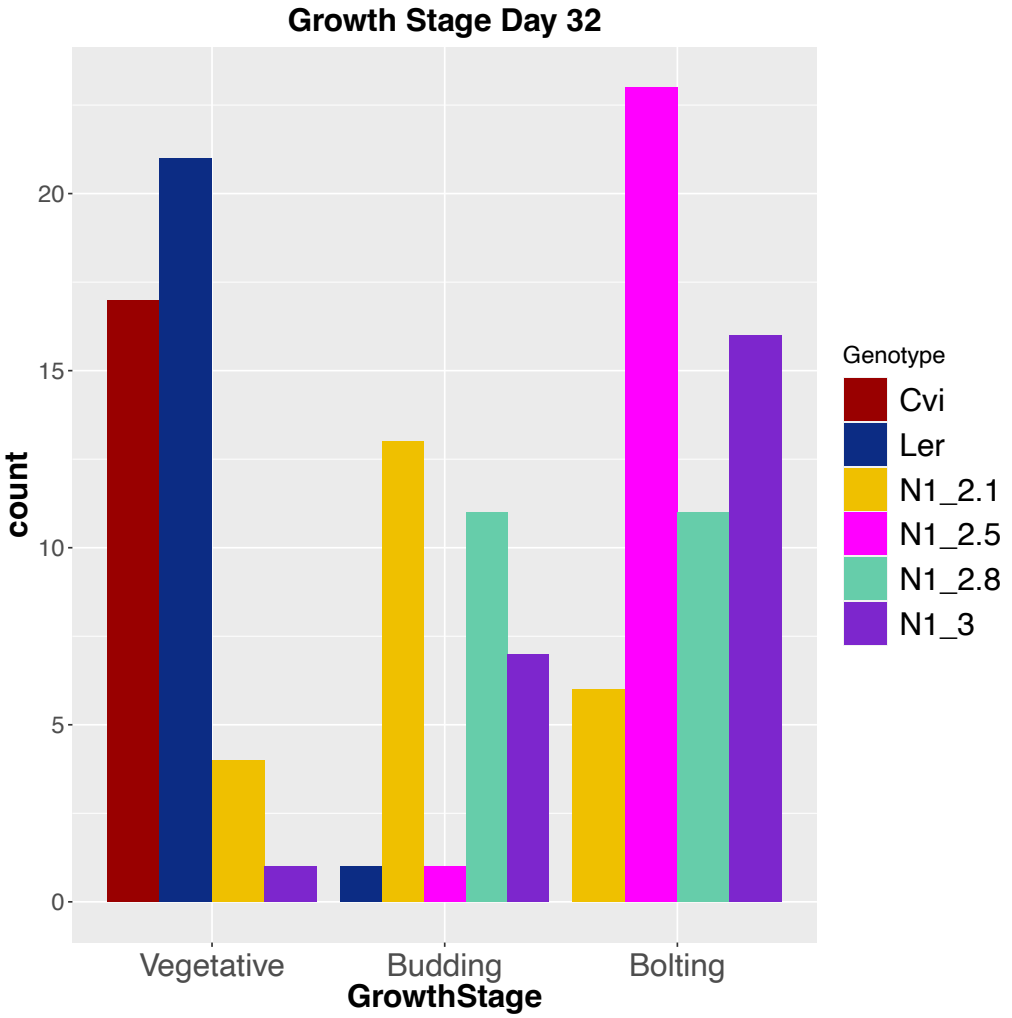


Figure 3.4. Growth stage at day 32 for Cvi, L*er* and NILs. The growth stage (Vegetative, Budding or Bolting) on day 32 is depicted for all samples and differentiated by genotype using colors. The plot illustrates that both Cvi and L*er* exhibit a substantial number of plants in the vegetative stage of development, while a majority of the NILs have entered the reproductive stage, either through budding or bolting.

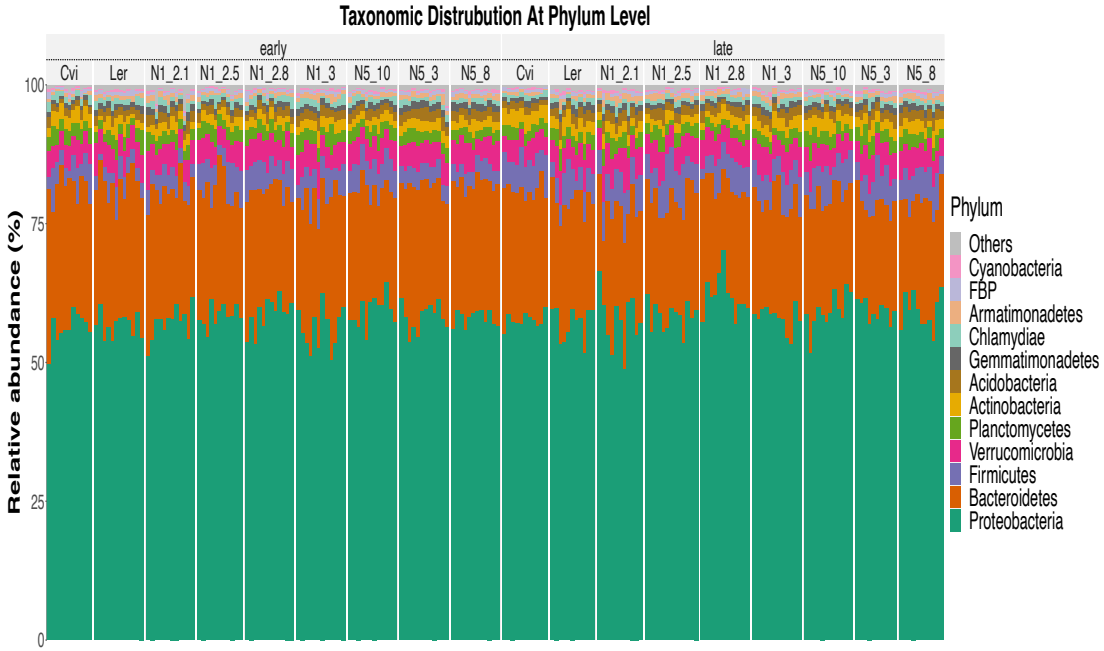


Figure 3.5. Taxonomic distribution at Phylum level. Taxonomic distribution of the microbial community represented in terms of the relative abundance at the Phylum level. The microbial community is dominated by Proteobacteria, Bacteroidetes, Acidobacteria , Verrucomicrobia*, and* Actinobacteria which together comprise nearly 78% of the bacterial community. There do not appear to be large differences at the Phylum level between different genotypes.

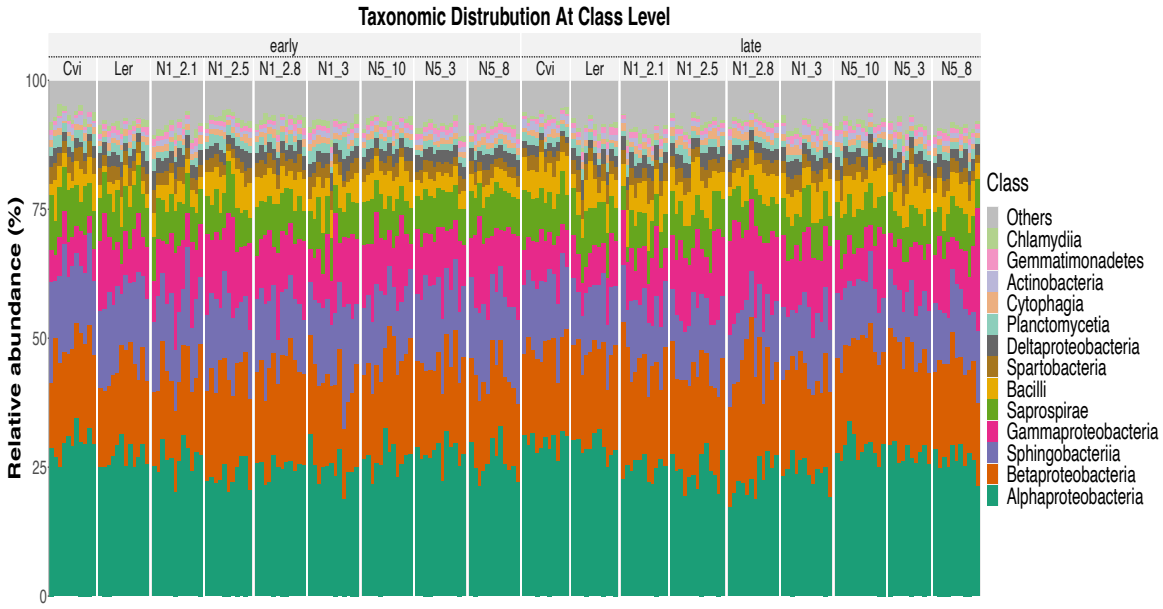
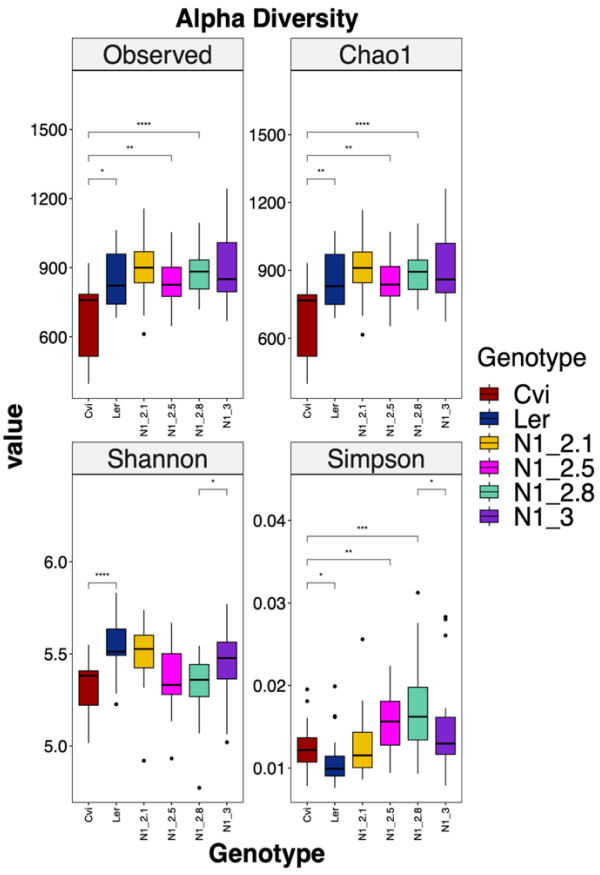


Figure 3.6. Taxonomic distribution at Class level. Taxonomic distribution of the microbial community represented in terms of the relative abundance at the Class level. The microbial community is dominated by Alphaproteobacteria, Saprospirae, Deltaproteobacteria, Gammaproteobacteria, and Betaproteobacteria which together comprise above 50% of the Classes. There do not appear to be large differences at the Class level between different genotypes.



Group1	Group2	p adj value			
		Observed	Chao1	Shannon	Simpson
Ler	Cvi	0.110	0.100	0.000	0.160
N1_2.1	Cvi	0.000	0.000	0.003	0.890
N1_2.5	Cvi	0.025	0.021	1.000	0.064
N1_2.8	Cvi	0.001	0.001	1.000	0.007
N1_3	Cvi	0.008	0.005	0.110	0.650
N1_2.1	Ler	1.000	1.000	1.000	0.260
N1_2.5	Ler	1.000	1.000	0.009	0.000
N1_2.8	Ler	1.000	1.000	0.000	0.000
N1_3	Ler	1.000	1.000	0.440	0.007
N1_2.1	N1_2.5	0.610	0.690	0.031	0.044
N1_2.1	N1_2.8	1.000	1.000	0.002	0.006
N1_2.5	N1_2.8	1.000	1.000	1.000	0.890
N1_2.1	N1_3	1.000	1.000	1.000	0.360
N1_2.5	N1_3	1.000	1.000	0.390	0.650
N1_2.8	N1_3	1.000	1.000	0.110	0.260

Figure 3.7. Alpha diversity of Cvi, L*er* and NILs. Four Alpha diversity indices were plotted, Observed, Chao1, Shannon, and Simpson indices and associated adjusted p-values of pairwise comparisons are shown in the table on the right. The Shannon index, which was more responsive to species richness, shows a significant difference between the L*er* and Cvi genotypes. Additionally, Shannon diversity of Cvi , N1-2.5, N1-2.8 and N1-3 were more similar than Cvi Vs N1-2.1. (Wilcox test *<0.05, **<0.01, *** <0.001).

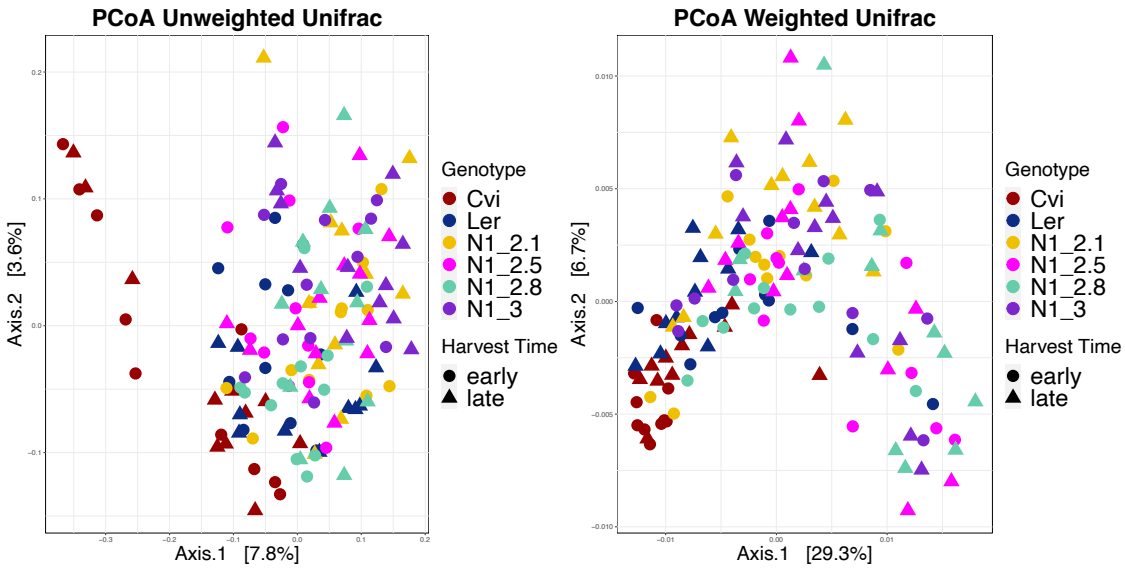
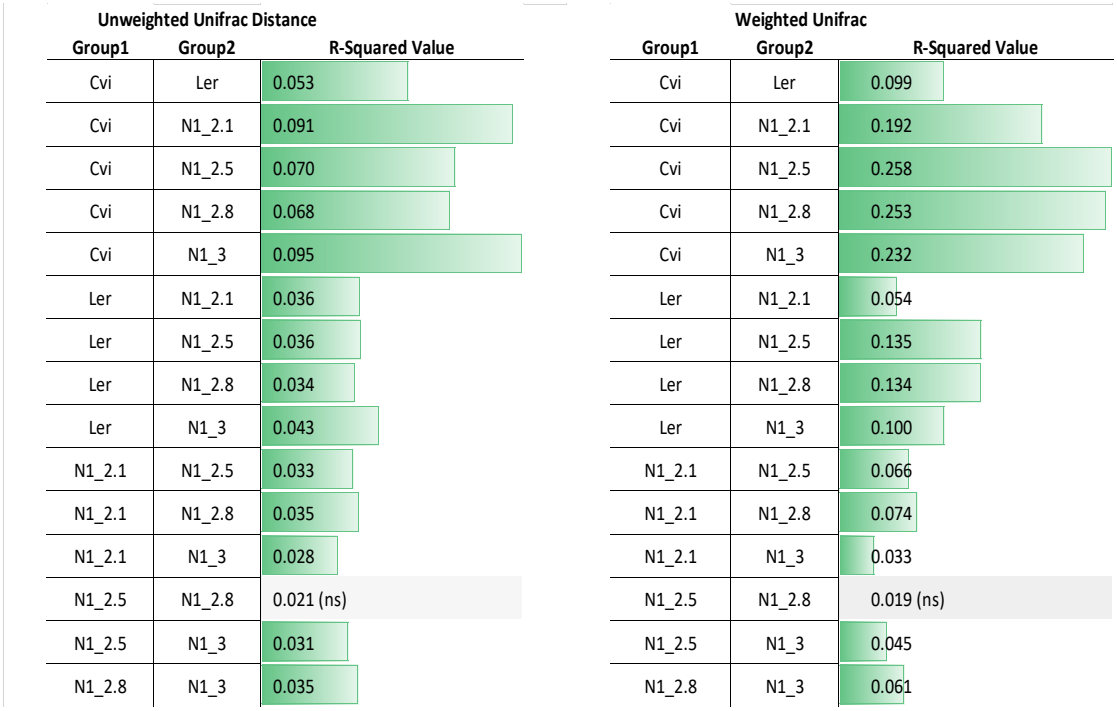


Figure 3.8. Principle Coordinate Analysis (PCoA) using weighted and unweighted UniFrac. Each point represents the microbial community. The colors represent the different genotypes and the shape represents whether it is in the late or early harvest group. The first two coordinate axes plotted accounted highest variation in the data, ranging from 10% for unweighted UniFrac and 35.4% for weighted UniFrac.



Unweighted Unifrac Distance

Group1	Group2	R-Squared Value
Cvi	Ler	0.053
Cvi	N1_2.1	0.091
Cvi	N1_2.5	0.070
Cvi	N1_2.8	0.068
Cvi	N1_3	0.095
Ler	N1_2.1	0.036
Ler	N1_2.5	0.036
Ler	N1_2.8	0.034
Ler	N1_3	0.043
N1_2.1	N1_2.5	0.033
N1_2.1	N1_2.8	0.035
N1_2.1	N1_3	0.028
N1_2.5	N1_2.8	0.021 (ns)
N1_2.5	N1_3	0.031
N1_2.8	N1_3	0.035

Weighted Unifrac

Group1	Group2	R-Squared Value
Cvi	Ler	0.099
Cvi	N1_2.1	0.192
Cvi	N1_2.5	0.258
Cvi	N1_2.8	0.253
Cvi	N1_3	0.232
Ler	N1_2.1	0.054
Ler	N1_2.5	0.135
Ler	N1_2.8	0.134
Ler	N1_3	0.100
N1_2.1	N1_2.5	0.066
N1_2.1	N1_2.8	0.074
N1_2.1	N1_3	0.033
N1_2.5	N1_2.8	0.019 (ns)
N1_2.5	N1_3	0.045
N1_2.8	N1_3	0.061

Figure 3.9. R^2 values obtained from pairwise PERMANOVA. This was done between group 1 and group 2 were obtained using the adonis2 test for both unweighted (left) and weighted UniFrac (right) distances. The magnitude of R-squared values was proportional to the differences in the microbial community explained by a particular comparison. The results reveal larger differences between Cvi Vs NILs compared to L*er* Vs NILs. The differences among the NILs were smaller in magnitude, with no significant differences between N1-2.5 and N1-2.8. (p-value <0.001 (green bars) ; ns : not significant)

Comparison Between Left	Vs Right	Number of Differentially Abundant Microbes	Δ Number of Genes
Ler	NIL 1-2.1	3	355
NIL 1-2.1	NIL 1-2.5	15	170
NIL 1-2.5	NIL 1-2.8	0	222
NIL 1-2.8	NIL 1-3	36	248

Figure 3.10. Differentially abundant OTUs on chromosome 1. These were determined using DeSeq2. Comparison between genotypes is represented in the left column. The number of differentially abundant OTUs, and the number of genes corresponding to a given comparison made are represented on right two columns. There are 15 and 36 differentially abundant OTUs between the NIL 1-2.5 Vs NIL 1-2.1 and NIL 1-3 Vs 1-2.8 respectively. Bonferroni correction for multiple tests.

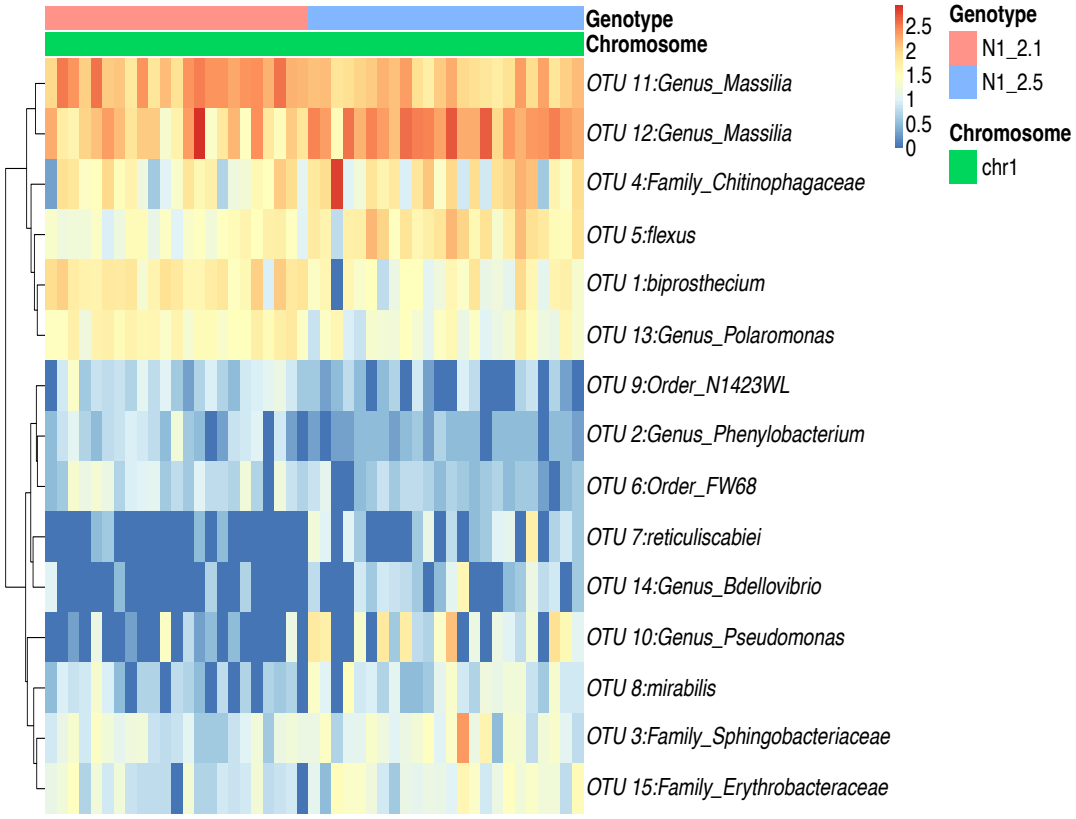


Figure 3.11. Heatmap for differentially abundant OTUs between the N1_2.1 and N1_2.5. The data are clustered by Genotype N1_2.1 in pink and N1_2.5 in blue. The chromosome number is highlighted in green. Taxa from the families *Chitinophagaceae*, *Sphingobacteriaceae*, *Pseudomonadaceae,* and *Bdellovibrio* are present in higher abundance in the

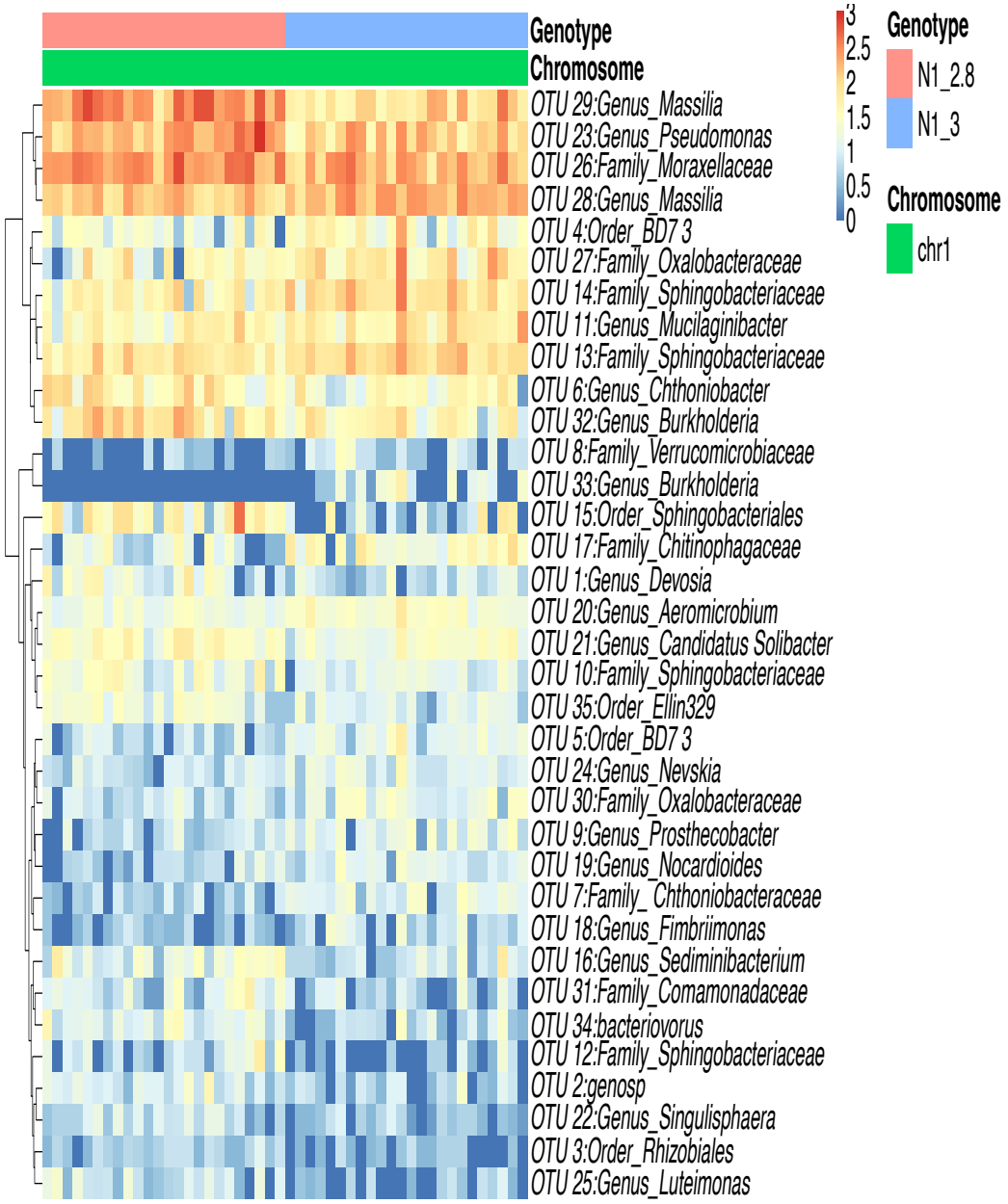
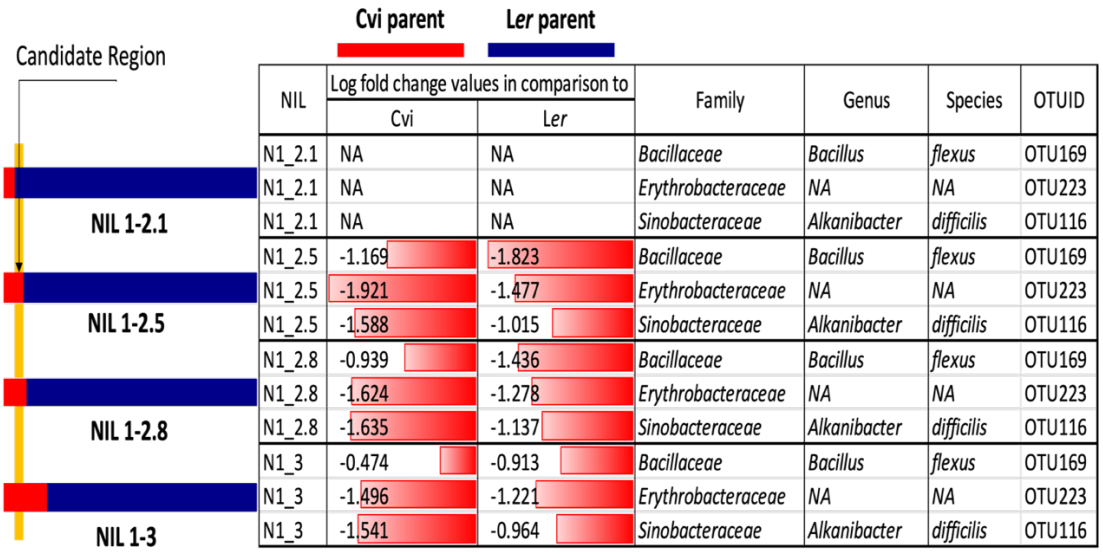


Figure 3.12. Heatmap for differentially abundant OTUs between the N1_2.8 and N1_3. The data are clustered by Genotype N1_2.1 in pink and N1_2.5 in blue. The chromosome number is highlighted in green. *Sphingobacteriales* , *Chthoniobacterales*, and *Burkholderiales* are present in higher abundance in N1_3



NIL	Log fold change values in comparison to		Family	Genus	Species	OTUID
	Cvi	Ler				
N1_2.1	NA	NA	*Bacillaceae*	*Bacillus*	*flexus*	OTU169
N1_2.1	NA	NA	*Erythrobacteraceae*	*NA*	*NA*	OTU223
N1_2.1	NA	NA	*Sinobacteraceae*	*Alkanibacter*	*difficilis*	OTU116
N1_2.5	-1.169	-1.823	*Bacillaceae*	*Bacillus*	*flexus*	OTU169
N1_2.5	-1.921	-1.477	*Erythrobacteraceae*	*NA*	*NA*	OTU223
N1_2.5	-1.588	-1.015	*Sinobacteraceae*	*Alkanibacter*	*difficilis*	OTU116
N1_2.8	-0.939	-1.436	*Bacillaceae*	*Bacillus*	*flexus*	OTU169
N1_2.8	-1.624	-1.278	*Erythrobacteraceae*	*NA*	*NA*	OTU223
N1_2.8	-1.635	-1.137	*Sinobacteraceae*	*Alkanibacter*	*difficilis*	OTU116
N1_3	-0.474	-0.913	*Bacillaceae*	*Bacillus*	*flexus*	OTU169
N1_3	-1.496	-1.221	*Erythrobacteraceae*	*NA*	*NA*	OTU223
N1_3	-1.541	-0.964	*Sinobacteraceae*	*Alkanibacter*	*difficilis*	OTU116

Figure 3.13. Transgressive segregation of microbes between NILs and parents. Differential abundance expressed as $\log_2$ fold changes of three microbial taxa between each NIL and L*er* or Cvi genotypes are shown above. The red bars indicate the magnitude of $\log_2$ fold changes for a given comparison. The taxa *Bacillus flexus*, *Erythrobacteraceae*, and *Alkanibacter difficillis* were found in higher abundance in the Cvi and L*er* parents compared to the NILs N1-2.5, N1-2.8, and N1-3, but no significant difference was observed for N-1.2.1. This suggests that the interaction of the Cvi region between N1-2.1 and N1-2.5 with the L*er* background plays a role in the differential abundance of these taxa. This region is highlighted in orange as a potential the "Candidate Region".

SUPPLEMENTARY MATERIAL

comparison	p adj
Ler-Cvi	0.0435177553039974
N1_2.1-Cvi	4.42978986825437e-13
N1_2.5-Cvi	7.53714091139557e-10
N1_2.8-Cvi	6.09875483448263e-11
N1_3-Cvi	1.77968750847413e-13
N1_2.1-Ler	2.62033202913514e-08
N1_2.5-Ler	5.88694278736979e-05
N1_2.8-Ler	4.71399678647622e-06
N1_3-Ler	7.1347413532763e-09
N1_2.5-N1_2.1	0.412333044478635
N1_2.8-N1_2.1	0.849019822961342
N1_3-N1_2.1	0.999831845367001
N1_2.8-N1_2.5	0.981214807205271
N1_3-N1_2.5	0.26390996370132
N1_3-N1_2.8	0.704775868885817

Supplementary Table 3.1. Statistical comparisons of biomass for NILs, Cvi, and L*er*. (tukey method)

Effect	DFn	DFd	F	p	p<.05
Genotype	5	117	26.829	5.43e-18	*
HarvestDate	1	117	4.119	4.50e-02	*

Supplementary Table 3.2. Summary of ANOVA for biomass. ANOVA shows there is a significant influence of genotype on biomass and a weaker but, significant influence of harvest date. (Dfn: degrees of freedom numerator; Dfd : degrees of freedom denominator ; F: F statistic; p – pvalue ; * $p<0.05$)

.y.	group1	group2	p	p.adj	p.format	p.signif	method
Observed	N1_2.1	N1_2.5	6.107101e-02	0.61000	0.06107	ns	Wilcoxon
Observed	N1_2.1	N1_2.8	3.328085e-01	1.00000	0.33281	ns	Wilcoxon
Observed	N1_2.1	N1_3	5.512096e-01	1.00000	0.55121	ns	Wilcoxon
Observed	N1_2.1	Ler	1.145298e-01	1.00000	0.11453	ns	Wilcoxon
Observed	N1_2.1	Cvi	2.596501e-05	0.00039	2.6e-05	****	Wilcoxon
Observed	N1_2.5	N1_2.8	2.198467e-01	1.00000	0.21985	ns	Wilcoxon
Observed	N1_2.5	N1_3	2.524101e-01	1.00000	0.25241	ns	Wilcoxon
Observed	N1_2.5	Ler	9.211668e-01	1.00000	0.92117	ns	Wilcoxon
Observed	N1_2.5	Cvi	2.050400e-03	0.02500	0.00205	**	Wilcoxon
Observed	N1_2.8	N1_3	9.178783e-01	1.00000	0.91788	ns	Wilcoxon
Observed	N1_2.8	Ler	2.812234e-01	1.00000	0.28122	ns	Wilcoxon
Observed	N1_2.8	Cvi	8.280075e-05	0.00120	8.3e-05	****	Wilcoxon
Observed	N1_3	Ler	3.387129e-01	1.00000	0.33871	ns	Wilcoxon
Observed	N1_3	Cvi	6.168248e-04	0.00800	0.00062	***	Wilcoxon
Observed	Ler	Cvi	1.008789e-02	0.11000	0.01009	*	Wilcoxon

Supplementary Table 3.3. Statistical comparisons of Observed number of OTUs. This was done between NILs, Cvi, and L*er* samples. (Wilcoxon test. p-value * < 0.05; **<0.01;*** <0.001; ****<0.0001).

.y.	group1	group2	p	p.adj	p.format	p.signif	method
Chao1	N1_2.1	N1_2.5	6.853598e-02	6.9e-01	0.06854	ns	Wilcoxon
Chao1	N1_2.1	N1_2.8	3.691674e-01	1.0e+00	0.36917	ns	Wilcoxon
Chao1	N1_2.1	N1_3	5.336122e-01	1.0e+00	0.53361	ns	Wilcoxon
Chao1	N1_2.1	Ler	1.151966e-01	1.0e+00	0.11520	ns	Wilcoxon
Chao1	N1_2.1	Cvi	6.140147e-06	9.2e-05	6.1e-06	****	Wilcoxon
Chao1	N1_2.5	N1_2.8	1.994301e-01	1.0e+00	0.19943	ns	Wilcoxon
Chao1	N1_2.5	N1_3	2.911710e-01	1.0e+00	0.29117	ns	Wilcoxon
Chao1	N1_2.5	Ler	9.219499e-01	1.0e+00	0.92195	ns	Wilcoxon
Chao1	N1_2.5	Cvi	1.717431e-03	2.1e-02	0.00172	**	Wilcoxon
Chao1	N1_2.8	N1_3	9.755592e-01	1.0e+00	0.97556	ns	Wilcoxon
Chao1	N1_2.8	Ler	2.696287e-01	1.0e+00	0.26963	ns	Wilcoxon
Chao1	N1_2.8	Cvi	4.030823e-05	5.6e-04	4.0e-05	****	Wilcoxon
Chao1	N1_3	Ler	3.534376e-01	1.0e+00	0.35344	ns	Wilcoxon
Chao1	N1_3	Cvi	3.942524e-04	5.1e-03	0.00039	***	Wilcoxon
Chao1	Ler	Cvi	9.357290e-03	1.0e-01	0.00936	**	Wilcoxon

Supplementary Table 3.4. Statistical comparisons of Chao1 index. This was done between NILs, Cvi, and L*er* samples. (Wilcoxon test. p-value * < 0.05; **<0.01;*** <0.001; ****<0.0001).

.y.	group1	group2	p	p.adj	p.format	p.signif	method
Shannon	N1_2.1	N1_2.5	3.058262e-03	3.1e-02	0.00306	**	Wilcoxon
Shannon	N1_2.1	N1_2.8	1.791027e-04	2.3e-03	0.00018	***	Wilcoxon
Shannon	N1_2.1	N1_3	3.255769e-01	1.0e+00	0.32558	ns	Wilcoxon
Shannon	N1_2.1	Ler	4.642959e-01	1.0e+00	0.46430	ns	Wilcoxon
Shannon	N1_2.1	Cvi	2.495619e-04	3.0e-03	0.00025	***	Wilcoxon
Shannon	N1_2.5	N1_2.8	6.456199e-01	1.0e+00	0.64562	ns	Wilcoxon
Shannon	N1_2.5	N1_3	5.596157e-02	3.9e-01	0.05596	ns	Wilcoxon
Shannon	N1_2.5	Ler	8.478786e-04	9.3e-03	0.00085	***	Wilcoxon
Shannon	N1_2.5	Cvi	6.603589e-01	1.0e+00	0.66036	ns	Wilcoxon
Shannon	N1_2.8	N1_3	1.278123e-02	1.1e-01	0.01278	*	Wilcoxon
Shannon	N1_2.8	Ler	4.375269e-06	6.6e-05	4.4e-06	****	Wilcoxon
Shannon	N1_2.8	Cvi	9.910162e-01	1.0e+00	0.99102	ns	Wilcoxon
Shannon	N1_3	Ler	7.289375e-02	4.4e-01	0.07289	ns	Wilcoxon
Shannon	N1_3	Cvi	1.254117e-02	1.1e-01	0.01254	*	Wilcoxon
Shannon	Ler	Cvi	6.417869e-06	9.0e-05	6.4e-06	****	Wilcoxon

Supplementary Table 3.5. Statistical comparisons of Shannon index. This was done between NILs, Cvi, and L*er* samples. (Wilcoxon test. p-value * < 0.05; **<0.01;*** <0.001; ****<0.0001).

.y.	group1	group2	p	p.adj	p.format	p.signif	method
Simpson	N1_2.1	N1_2.5	4.416250e-03	0.04400	0.00442	**	Wilcoxon
Simpson	N1_2.1	N1_2.8	4.591134e-04	0.00600	0.00046	***	Wilcoxon
Simpson	N1_2.1	N1_3	7.190862e-02	0.36000	0.07191	ns	Wilcoxon
Simpson	N1_2.1	Ler	3.685759e-02	0.26000	0.03686	*	Wilcoxon
Simpson	N1_2.1	Cvi	7.097201e-01	0.89000	0.70972	ns	Wilcoxon
Simpson	N1_2.5	N1_2.8	4.428819e-01	0.89000	0.44288	ns	Wilcoxon
Simpson	N1_2.5	N1_3	1.784738e-01	0.65000	0.17847	ns	Wilcoxon
Simpson	N1_2.5	Ler	1.591497e-05	0.00022	1.6e-05	****	Wilcoxon
Simpson	N1_2.5	Cvi	7.140272e-03	0.06400	0.00714	**	Wilcoxon
Simpson	N1_2.8	N1_3	4.148767e-02	0.26000	0.04149	*	Wilcoxon
Simpson	N1_2.8	Ler	3.328315e-06	0.00005	3.3e-06	****	Wilcoxon
Simpson	N1_2.8	Cvi	6.386752e-04	0.00700	0.00064	***	Wilcoxon
Simpson	N1_3	Ler	5.395759e-04	0.00650	0.00054	***	Wilcoxon
Simpson	N1_3	Cvi	1.632836e-01	0.65000	0.16328	ns	Wilcoxon
Simpson	Ler	Cvi	2.056764e-02	0.16000	0.02057	*	Wilcoxon

Supplementary Table 3.6. Statistical comparisons of Simpson index. This was done between NILs, Cvi, and L*er* samples. (Wilcoxon test. p-value * < 0.05; **<0.01;*** <0.001; ****<0.0001).

	Df	SumOfSqs	R2	F	Pr(>F)
Genotype	5	1.3439221	0.07273744	1.904517	0.001
Biomass	1	0.2227830	0.01205774	1.578566	0.001
HarvestDate	1	0.2563053	0.01387208	1.816094	0.001
Residual	118	16.6533351	0.90133274	NA	NA
Total	125	18.4763455	1.00000000	NA	NA

Supplementary Table 3.7 PERMANOVA for unweighted unifrac distances. PERMANOVA done with adonis2 with unweighted UniFrac distances. Design adonis2(weighted UniFrac ~Genotype+Biomass+Harvest Date, data = meta, permutations=999).

	Df	SumOfSqs	R2	F	Pr(>F)
Genotype	5	0.0054925630	0.1740780	5.322322	0.001
Biomass	1	0.0009948445	0.0315300	4.820048	0.001
HarvestDate	1	0.0007100376	0.0225035	3.440151	0.003
Residual	118	0.0243548732	0.7718885	NA	NA
Total	125	0.0315523184	1.0000000	NA	NA

Supplementary Table 3.8 PERMANOVA for weighted unifrac distances. PERMANOVA done with adonis2 with weighted UniFrac distances. Design adonis2(unweighted UniFrac ~Genotype+Biomass+Harvest Date, data = meta, permutations=999).

baseMean	log2FoldCha	Order	Family	Genus	Species	Numerator	Denom
246.848362	-0.6445895	Pseudomonadales	Moraxellaceae	NA	NA	N1_3	N1_2.8
214.40171	-1.7312545	Burkholderiales	Oxalobacteraceae	Massilia	NA	N1_3	N1_2.8
199.475356	-1.0940705	Streptophyta	NA	NA	NA	N1_3	N1_2.8
190.366742	0.99649209	Burkholderiales	Oxalobacteraceae	Massilia	NA	N1_2.5	N1_2.1
180.315392	-0.8994015	Pseudomonadales	Pseudomonadaceae	Pseudomonas	NA	N1_3	N1_2.8
163.326502	0.62605179	Burkholderiales	Oxalobacteraceae	Massilia	NA	N1_3	N1_2.8
149.890284	-0.6482187	Burkholderiales	Oxalobacteraceae	Massilia	NA	N1_2.5	N1_2.1
86.3398442	0.5744668	Sphingobacteriales	Sphingobacteriaceae	NA	NA	N1_3	N1_2.8
76.4156992	1.12561839	Sphingobacteriales	Sphingobacteriaceae	NA	NA	N1_3	N1_2.8
65.9671499	-1.0501977	Burkholderiales	Burkholderiaceae	Burkholderia	NA	N1_3	N1_2.8
62.234786	1.31584394	[Saprospirales]	Chitinophagaceae	NA	NA	N1_2.5	N1_2.1
59.3911068	1.29244083	Burkholderiales	Oxalobacteraceae	NA	NA	N1_3	N1_2.8
58.2528681	0.79330699	Sphingobacteriales	Sphingobacteriaceae	Mucilaginibacter	NA	N1_3	N1_2.8
46.1738006	-0.9207537	[Chthoniobacterales]	[Chthoniobacteraceae]	Chthoniobacter	NA	N1_3	N1_2.8
45.002533	1.50169182	Bacillales	Bacillaceae	Bacillus	flexus	N1_2.5	N1_2.1
41.459178	-0.9613305	Caulobacterales	Caulobacteraceae	Asticcacaulis	biprosthecium	N1_2.5	N1_2.1
35.396954	-2.1203345	Sphingobacteriales	NA	NA	NA	N1_3	N1_2.8
31.2440672	1.15524957	BD7-3	NA	NA	NA	N1_3	N1_2.8
27.0302688	-0.6513389	Burkholderiales	Comamonadaceae	Polaromonas	NA	N1_2.5	N1_2.1
26.9502744	-0.5941849	Solibacterales	Solibacteraceae	Candidatus Solibacter	v	N1_3	N1_2.8
22.2975599	1.4656586	[Saprospirales]	Chitinophagaceae	NA	NA	N1_3	N1_2.8
21.5406071	0.58478543	Actinomycetales	Nocardioidaceae	Aeromicrobium	NA	N1_3	N1_2.8
16.9207436	-0.9477718	Sphingobacteriales	Sphingobacteriaceae	NA	NA	N1_3	N1_2.8
16.8424703	1.2813233	Sphingobacteriales	Sphingobacteriaceae	NA	NA	N1_2.5	N1_2.1
14.3333779	-0.7678285	Ellin329	NA	NA	NA	N1_3	N1_2.8
13.9979452	3.26160481	Pseudomonadales	Pseudomonadaceae	Pseudomonas	NA	N1_2.5	N1_2.1
13.4392741	-1.0571916	[Saprospirales]	Chitinophagaceae	Sediminibacterium	NA	N1_3	N1_2.8
13.2078078	-1.2913903	Rhizobiales	Hyphomicrobiaceae	Devosia	NA	N1_3	N1_2.8
12.482244	1.03912613	Sphingomonadales	Erythrobacteraceae	NA	NA	N1_2.5	N1_2.1
11.6365224	1.06776329	Burkholderiales	Oxalobacteraceae	NA	NA	N1_3	N1_2.8
11.5107897	2.88179513	Burkholderiales	Comamonadaceae	Aquabacterium	NA	N1_2.1	Ler
10.8317118	-3.2026318	Sphingobacteriales	Sphingobacteriaceae	NA	NA	N1_2.1	Ler
10.5782444	1.2654115	BD7-3	NA	NA	NA	N1_3	N1_2.8
9.86612871	-1.2186489	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrio	bacteriovorus	N1_3	N1_2.8

Supplementary Table 3.9. Comprehensive results from differential abundance analyses between NILs. Results from differential abundance analyses conducting in DeSeq2. The baseMean is the average number of reads for a specific OTU, the log2FoldChange – the log (base 2) fold change, The Denominator is compared to the Numerator in all comparisons.

Literature Cited Chapter 3

1. Grayston SJ, Vaughan D, Jones D. Rhizosphere carbon flow in trees, in comparison with annual plants: The importance of root exudation and its impact on microbial activity and nutrient availability. Applied Soil Ecology [Internet]. 1997;5(1):29–56. Available from: <Go to ISI>://WOS:A1997WB60400003\nhttp://ac.els-cdn.com/S0929139396001266/1-s2.0-S0929139396001266-main.pdf?_tid=59e83e82-5c1b-11e3-9c6d-00000aab0f6b&acdnat=1386076009_ec91f3130df82d6b03d875a0ad36575a\nhttp://ac.els-cdn.com/S0929139396001266/1-s2.0-S0929
2. Grant MR, Jones JDG. Hormone (dis)harmony moulds plant health and disease. Science. 2009;324(5928):750–2.
3. Bakker P a HM, Berendsen RL, Doornbos RF, Wintermans PC a, Pieterse CMJ. The rhizosphere revisited: root microbiomics. Front Plant Sci [Internet]. 2013;4(May):165. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3667247&tool=pmcentrez&rendertype=abstract
4. Grayston SJ, Wang S, Campbell CD, Edwards AC. Selective Influence of Plant Species on Microbial Diversity in the Rhizosphere. 1998;30(3):369–78.
5. Baetz U, Martinoia E. Root exudates: The hidden part of plant defense. Trends Plant Sci [Internet]. 2014;19(2):90–8. Available from: http://dx.doi.org/10.1016/j.tplants.2013.11.006
6. Bowen GD, Rovira AD. The rhizosphere and its management to improve plant growth. Advances in agronomy. 1999;66(1):102.
7. Smalla K, Wieland G, Buchner a, Zock a, Parzy J, Roskot N, et al. Bulk and Rhizosphere Soil Bacterial Communities Studied by Denaturing Gradient Gel Electrophoresis : Plant-Dependent Enrichment and Seasonal Shifts Revealed Bulk and Rhizosphere Soil Bacterial Communities Studied by Denaturing Gradient Gel Electrophoresis. Appl Environ Microbiol. 2001;67(10):4742–51.
8. Hartmann A, Schmid M, van Tuinen D, Berg G. Plant-driven selection of microbes. Vol. 321, Plant and Soil. 2009. p. 235–57.
9. Haichar F, Fochesato S, Achouak W. Host Plant Specific Control of 2,4-Diacetylphloroglucinol Production in the Rhizosphere. Agronomy [Internet]. 2013;3(4):621–31. Available from: http://www.mdpi.com/2073-4395/3/4/621/
10. Berg G, Smalla K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. FEMS Microbiol Ecol. 2009;68(1):1–13.
11. Chaparro JM, Badri D V, Vivanco JM. Rhizosphere microbiome assemblage is affected by plant development. ISME J [Internet]. 2014;8(4):790–803. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24196324
12. Micallef SA, Channer S, Shiaris MP, Colón-Carmona A. Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. Plant Signal Behav. 2009 Aug;4(8):777–80.

13. De Bruijn FJ, Szabados L, Schell J. Chimeric genes and transgenic plants are used to study the regulation of genes involved in symbiotic plant-microbe interactions (nodulin genes). Dev Genet [Internet]. 1990;11(3):182–96. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2279354
14. Read D. plant-microbe mutualisms and community structure. Biodiversity and ecosystem function [Internet]. 1993;181–209. Available from: http://link.springer.com/chapter/10.1007/978-3-642-58001-7_9
15. Siqueira JO, Nair MG, Hammerschmidt R, Safir GR, Putnam AR. Significance of phenolic compounds in plant-soil-microbial systems. Vol. 10, Critical Reviews in Plant Sciences. 1991. p. 63–121.
16. Quigley NB, Roberts DM, Ownley BH. Molecular plant—microbe interactions. Vol. 1, Trends in Plant Science. 1996. p. 366–8.
17. Veronese P, Ruiz MT, Coca MA, Hernandez-Lopez A, Lee H, Ibeas JI, et al. In defense against pathogens. Both plant sentinels and foot soldiers need to know the enemy. Plant Physiol. 2003;131(4):1580–90.
18. Mark GL, Dow JM, Kiely PD, Higgins H, Haynes J, Baysse C, et al. Transcriptome profiling of bacterial responses to root exudates identifies genes involved in microbe-plant interactions. Proc Natl Acad Sci U S A. 2005;102(48):17454–9.
19. Panke-buisse K, Poole AC, Goodrich JK, Ley RE, Kao-kniffin J. Selection on soil microbiomes reveals reproducible impacts on plant function. ISME J [Internet]. 2014;1–10. Available from: http://dx.doi.org/10.1038/ismej.2014.196
20. Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M, et al. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. Science (1979) [Internet]. 2015;349(6250):860–4. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.aaa8764
21. Millet YA, Danna CH, Clay NK, Songnuan W, Simon MD, Werck-Reichhart D, et al. Innate immune responses activated in Arabidopsis roots by microbe-associated molecular patterns. Plant Cell. 2010;22(3):973–90.
22. Adams-Phillips L, Wan J, Tan X, Dunning FM, Meyers BC, Michelmore RW, et al. Discovery of ADP-ribosylation and other plant defense pathway elements through expression profiling of four different Arabidopsis-Pseudomonas R-avr interactions. Mol Plant Microbe Interact. 2008;21(5):646–57.
23. Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, et al. Plant ABC Transporters. The Arabidopsis book / American Society of Plant Biologists [Internet]. 2011;9:e0153. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3268509&tool=pmcentrez&rendertype=abstract
24. Yazaki K, Sugiyama A, Morita M, Shitan N. Secondary transport as an efficient membrane transport mechanism for plant secondary metabolites. Phytochemistry Reviews. 2008;7(3):513–24.
25. Smith KP, Handelsman J, Goodman RM. Genetic basis in plants for interactions with disease-suppressive bacteria. Proc Natl Acad Sci U S A [Internet]. 1999 Apr 27 [cited 2015 Dec 14];96(9):4786–90. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=21769&tool=pmcentrez&rendertype=abstract

26. Balint-Kurti P, Simmons SJ, Blum JE, Ballaré CL, Stapleton AE. Maize leaf epiphytic bacteria diversity patterns are genetically correlated with resistance to fungal pathogen infection. Mol Plant Microbe Interact [Internet]. 2010 Apr 1 [cited 2015 Dec 14];23(4):473–84. Available from: http://apsjournals.apsnet.org/doi/abs/10.1094/MPMI-23-4-0473
27. Li P, Wind JJ, Shi X, Zhang H, Hanson J, Smeekens SC, et al. Fructose sensitivity is suppressed in Arabidopsis by the transcription factor ANAC089 lacking the membrane-bound domain. Proc Natl Acad Sci U S A [Internet]. 2011;108(8):3436–41. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3044370&tool=pmcentrez&rendertype=abstract
28. Bergelson J, Mittelstrass J, Horton MW. Characterizing both bacteria and fungi improves understanding of the Arabidopsis root microbiome. Sci Rep. 2019 Dec 1;9(1).
29. Wallace JG, Kremling KA, Kovar LL, Buckler ES. Quantitative genetics of the maize leaf microbiome. Phytobiomes J. 2018;2(4):208–24.
30. Khasanova A, Edwards J, Bonnette J, Singer E, Haque T, Juenger TE. Quantitative genetic-by-soil microbiome interactions in a perennial grass affect functional traits. Proceedings of the Royal Society B: Biological Sciences. 2023 Jan 25;290(1991).
31. Oyserman BO, Flores SS, Griffioen T, Pan X, van der Wijk E, Pronk L, et al. Disentangling the genetic basis of rhizosphere microbiome assembly in tomato. Nat Commun. 2022 Dec 1;13(1).
32. Keurentjes JJB, Bentsink L, Alonso-Blanco C, Hanhart CJ, Vries HB De, Effgen S, et al. Development of a near-isogenic line population of Arabidopsis thaliana and comparison of mapping power with a recombinant inbred line population. Genetics. 2007;175(2):891–905.
33. Lisec J, Meyer RC, Steinfath M, Redestig H, Becher M, Witucka-Wall H, et al. Identification of metabolic and biomass QTL in Arabidopsis thaliana in a parallel analysis of RIL and IL populations. Plant Journal. 2008;53(6).
34. Franklin KA, Quail PH. Phytochrome functions in Arabidopsis development. Vol. 61, Journal of Experimental Botany. 2010.
35. Silva H, Yoshioka K, Dooner HK, Klessig DF. Characterization of a new Arabidopsis mutant exhibiting enhanced disease resistance. Molecular Plant-Microbe Interactions. 1999;12(12).
36. Vílchez JI, Yang Y, He D, Zi H, Peng L, Lv S, et al. DNA demethylases are required for myo-inositol-mediated mutualism between plants and beneficial rhizobacteria. Nat Plants. 2020;6(8).
37. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature [Internet]. 2013 Aug 1 [cited 2016 Oct 6];488(7409):86–90. Available from: http://www.nature.com/doifinder/10.1038/nature11237
38. Wagner MR, Lundberg DS, Del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat Commun. 2016 Jul 12;7.
39. Wang TT, Ding P, Chen P, Xing K, Bai JL, Wan W, et al. Complete genome sequence of endophyte Bacillus flexus KLBMP 4941 reveals its plant growth promotion mechanism and genetic basis for salt tolerance. J Biotechnol. 2017;260.
40. LIU H, ZHONG X, HUANG Y, QIAO C, SHAO C, LI R, et al. Production of Free Amino Acid and Short Peptide Fertilizer from Rapeseed Meal Fermentation Using Bacillus flexus NJNPD41 for Promoting Plant Growth. Pedosphere. 2018;28(2).

41. Xiong YW, Li XW, Wang TT, Gong Y, Zhang CM, Xing K, et al. Root exudates-driven rhizosphere recruitment of the plant growth-promoting rhizobacterium Bacillus flexus KLBMP 4941 and its growth-promoting effect on the coastal halophyte Limonium sinense under salt stress. Ecotoxicol Environ Saf. 2020;194.
42. Xu H, Gao J, Portieles R, Du L, Gao X, Borras- Hidalgo O. Endophytic bacterium Bacillus aryabhattai induces novel transcriptomic changes to stimulate plant growth. PLoS One. 2022 Aug 1;17(8 August).
43. Parada AE, Needham DM, Fuhrman JA. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol. 2016 May 1;18(5):1403–14.
44. Apprill A, Mcnally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquatic Microbial Ecology. 2015 Jun 4;75(2):129–37.
45. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME Journal. 2012 Aug;6(8):1621–4.
46. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth's multiscale microbial diversity. Nature [Internet]. 2017 Nov 1 [cited 2018 Jun 12];551(7681):457. Available from: http://www.nature.com/doifinder/10.1038/nature24621
47. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Vol. 37, Nature Biotechnology. Nature Publishing Group; 2019. p. 852–7.
48. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jun 29;13(7):581–3.
49. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinformatics. 2022 Nov 30;38(23):5315–6.
50. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020 May 1;37(5):1530–4.
51. Bunn AG. A dendrochronology program library in R (dplR). Dendrochronologia (Verona). 2008 Oct 1;26(2):115–24.
52. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014 Dec 5;15(12).
53. Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, et al. The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis. 2015 Aug 1;53(8):474–85.
54. Weigel D, Mott R. The 1001 genomes project for Arabidopsis thaliana. Genome Biol. 2009;10(5):107.
55. Sushanta Kumar B. Involvement of Disease Resistant Quantitative trait loci (QTLs) and resistant (R) Genes in key Signaling pathways of Arabidopsis thaliana during pathogen infection: An Overview [Internet]. Vol. 3, International Research Journal of Biological Sciences. 2014. Available from: www.isca.me

56. Cordovez V, Rotoni C, Dini-Andreote F, Oyserman B, Carrión VJ, Raaijmakers JM. Successive plant growth amplifies genotype-specific assembly of the tomato rhizosphere microbiome. Science of the Total Environment. 2021 Jun 10;772.
57. Badri D V., Vivanco JM. Regulation and function of root exudates. Plant Cell Environ [Internet]. 2009;32(6):666–81. Available from: http://doi.wiley.com/10.1111/j.1365-3040.2009.01926.x
58. Bulgarelli D, Garrido-Oter R, Münch PC, Weiman A, Dröge J, Pan Y, et al. Structure and function of the bacterial root microbiota in wild and domesticated barley. Cell Host Microbe. 2015;17(3):392–403.
59. Zachow C, Müller H, Tilcher R, Berg G. Differences between the rhizosphere microbiome of Beta vulgaris ssp. maritima-ancestor of all beet crops-and modern sugar beets. Front Microbiol [Internet]. 2014;5(August):415. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4144093&tool=pmcentrez&rendertype=abstract
60. Mitchum MG, Yamaguchi S, Hanada A, Kuwahara A, Yoshioka Y, Kato T, et al. Distinct and overlapping roles of two gibberellin 3-oxidases in Arabidopsis development. Plant Journal. 2006;45(5).
61. Bandaranayake CK, Koumproglou R, Wang XY, Wilkes T, Kearsey MJ. QTL analysis of morphological and developmental traits in the Ler x Cvi population of Arabidopsis thaliana. In: Euphytica. 2004.
62. Alonso-Blanco C, El-Assal SED, Coupland G, Koornneef M. Analysis of natural allelic variation at flowering time loci in the Landsberg erecta and Cape Verde Islands ecotypes of Arabidopsis thaliana. Genetics. 1998;149(2).
63. Juenger TE, Sen S, Stowe KA, Simms EL. Epistasis and genotype-environment interaction for quantitative trait loci affecting flowering time in Arabidopsis thaliana. In: Genetica. 2005.
64. Ungerer MC, Halldorsdottir SS, Modliszewski JL, Mackay TFC, Purugganan MD. Quantitative Trait Loci for Inflorescence Development in Arabidopsis thaliana.
65. El-Din El-Assal S, Alonso-Blanco C, Peeters AJM, Raz V, Koornneef M. A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. Nat Genet. 2001;29(4):435–40.
66. Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T. Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. Nat Genet. 2008;40(12).
67. Shalit A, Rozman A, Goldshmidt A, Alvarez JP, Bowman JL, Eshed Y, et al. The flowering hormone florigen functions as a general systemic regulator of growth and termination. Proc Natl Acad Sci U S A. 2009;106(20).
68. Micallef SA, Shiaris MP, Colón-Carmona A. Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. J Exp Bot. 2009 Apr;60(6):1729–42.
69. Miller SB, Heuberger AL, Broeckling CD, Jahn CE. Non-targeted metabolomics reveals sorghum rhizosphere-associated exudates are influenced by the belowground interaction of substrate and sorghum genotype. Int J Mol Sci. 2019 Jan 2;20(2).
70. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature [Internet]. 2012;488(7409):86–90. Available from: http://dx.doi.org/10.1038/nature11237

71. Thiergart T, Zgadzaj R, Bozsóki Z, Garrido-Oter R, Radutoiu S, Schulze-Lefert P. Lotus japonicus Symbiosis Genes Impact Microbial Interactions between Symbionts and Multikingdom Commensal Communities. 2019; Available from: https://doi.org/10.1128/mBio
72. Huang XF, Chaparro JM, Reardon KF, Zhang R, Shen Q, Vivanco JM. Rhizosphere interactions: Root exudates, microbes, and microbial communities1. Botany [Internet]. 2014;92(4):267–75. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84898743467&partnerID=40&md5=40968131de51f105acde45bb457c2205
73. Sun H, Jiang S, Jiang C, Wu C, Gao M, Wang Q. A review of root exudates and rhizosphere microbiome for crop production. Available from: https://doi.org/10.1007/s11356-021-15838-7

CHAPTER 4

CONCLUSIONS

In wrapping up this thesis, all three chapters took us into the intricate world of plant-microbe interactions within *Arabidopsis thaliana*. Chapter 1 set the stage, painting a comprehensive picture of these interactions, and highlighting their significance in the broader context. Plant-microbe interactions took place in two main arenas: the rhizosphere (root zone) and the phyllosphere (leaf surface). These interactions could be both positive and negative, emerging from the interplay between plants and microbes. Plant pathogens could lead to devastating diseases such as late blight in potatoes and tomatoes, banana Panama disease, and Dutch elm disease, resulting in reduced crop yields and economic setbacks. Lately, there had been a focused effort in research aimed at utilizing the advantageous elements of plant-microbiome interactions.

Strategies for harnessing beneficial plant-microbial relationships have far-reaching implications for improving agricultural yields and ecosystem health. For instance, certain microbes acted as plant bodyguards, fending off harmful pathogens. Nitrogen-fixing bacteria were like nature's nitrogen factories, converting atmospheric nitrogen into a plant-usable form, while other microbes aided in nutrient uptake and stress tolerance. Aiding in phosphorous and iron absorption, as well as drought and salt resistance, these microbial allies enhanced plant resilience.

The review highlighted two exciting research endeavors. Firstly, it delved into a noteworthy study using artificial ecosystem selection, where plant traits were modified over generations through microbial interactions. By examining the shifting microbial communities

and their impact on plant biomass, researchers gained insights into the intricate web of microbe-mediated plant traits. Secondly, the study explored plant genetic influence on the rhizosphere bacteria community using traditional genetic mapping techniques, striving to pinpoint the genetic regions responsible for shaping these microbial communities.

Chapter 2 centered around examining how the plant host's selection influences microbial responses. To investigate this, I employed the system of Arabidopsis thaliana Cvi and L*er*. I delved into the intricate relationship between plants and their rhizosphere microbial communities, aiming to understand how factors such as plant genotype, environment, and biomass treatment (high or low biomass selection pressures) influence the assembly of these communities. I conducted an artificial ecosystem selection experiment over eight generations using *Arabidopsis thaliana* plants. My main objective was to select soil microbiomes associated with either higher or lower plant biomass. The findings shed light on the interplay between genetics, environmental factors, and microbial communities, providing insights that could be applied to enhance plant growth and soil restoration in agricultural settings.

In my artificial ecosystem selection experiment, I closely monitored the shifts in plant biomass and the dynamics of microbial communities across multiple generations. Despite inherent variability stemming from random environmental variations (REV), I successfully achieved the goal, selecting for either higher or lower plant biomass. Particularly intriguing was the observation of a gradual amplification in the impact of both plant genotype and biomass treatment on the composition of microbial communities over the course of the experiment. This escalating influence gained prominence, eventually accounting for approximately 40% of the total observed variation.

My analysis of alpha and beta diversity uncovered fascinating trends. At the outset of the experiment, we observed a decline in microbial species richness, which later gave way to stabilization in the number of species. This intriguing shift in the pattern highlighted the intricate interplay within microbial communities subjected to prolonged selection pressure over multiple generations. To thoroughly understand these dynamics, I utilized sophisticated statistical methods, including Principal Coordinate Analysis (PCoA). This approach visually depicted the evolving structure of microbial communities across generations, providing valuable insights into their complex responses. Furthermore, I utilized a neutral model in my study to evaluate the influence of random processes in comparison to deterministic selective forces on microbial community assembly. As the experiment progressed through generations, I observed a diminishing fit of the neutral model, indicating a growing impact of selective forces on the microbiome.

When examining specific microbial taxa, I observed a notable increase in plant growth-promoting rhizobacteria (PGPR) within the high biomass treatment, accompanied by potential genotype-specific interactions. This finding suggests the feasibility of cultivating microbial communities that promote plant growth in a relatively short time through artificial ecosystem selection growth cycles. To emphasize the significance of this Chapter, my study deepens our comprehension of the intricate interplay among plant genetics, environment, and rhizosphere microbial communities. Through the multi-generational ecosystem selection experiment, I unveiled the evolving dynamics of microbial community assembly and its influence on plant biomass. These findings hold implications for enhancing agricultural practices and promoting plant health through the strategic utilization of specific microbial communities.

As opposed to microbe-mediated plant-microbe interactions, in Chapter 3, I focused on understanding the influence of the plant genetic component on the microbial community. Plant genetic regulation of the microbiome pertains to how the genetic traits of a plant impact the structure and behavior of the microbial population linked to both its root system and aboveground portions. Through the secretion of root exudates, plants shape the composition of microbial communities in the rhizosphere. Notably, plant genes produce products that can engage in interactions with distinct microorganisms, resulting in either cooperative or antagonistic associations. The role of plant genetic control is pivotal in shaping the presence and variety of advantageous microorganisms that enhance processes such as nutrient uptake, defense against diseases, and the ability to withstand stress in plants.

In the quest to unravel the impact of plant genetics on specific microbes, traditional methods have involved techniques like quantitative trait loci (QTL) mapping through the use of recombinant inbred lines (RILs) or near-isogenic lines (NILs). These strategies have proven effective in pinpointing smaller chromosomal regions in a variety of plant model systems. Nevertheless, a notable gap existed in applying these techniques to study entire microbial communities.

NILs, a collection of closely related plant lines or varieties that diverged only at a precise targeted gene or genomic region, offered a powerful tool for investigating the effects of specific genetic elements on various traits, while minimizing the influence of genetic background noise. The NILs we employed in our study were originally created by Keurentjes and colleagues, using the parental strains L*er* and Cvi. These NILs carried a limited portion of Cvi genetic material on the upper section of chromosome 1, while maintaining a genomic

background primarily resembling L*er*. Consequently, all the NILs we utilized shared a closer genetic resemblance to L*er* than to Cvi.

My research aimed to narrow down quantitative trait loci (QTL) on chromosome 1 of *Arabidopsis thaliana* Cvi and L*er*. This study was built upon a previous lab member's (Stuart Morey) finding, which suggested that this region had a role to play in manipulating the microbial community associated with the rhizosphere of *Arabidopsis thaliana*. The outcomes of my research validated the conclusions drawn in prior investigations, underscoring the modest yet noteworthy influence of plant genotype on the structure of the microbial community. Notably, I identified distinct microbial taxa that exhibited varying associations with the NILs.

In summary, my study provided deeper insights into the precise genetic impacts of NILs situated at the upper portion of chromosome 1 on plant-microbe interactions. Additionally, it underscored the effectiveness of traditional mapping techniques for pinpointing the potential regions accountable for plant-microbe interactions. Further research endeavors aimed at understanding the mechanisms underlying plant genetic control over the microbiome would involve intricate mapping, complemented by a comprehensive analysis of plant gene expression and metagenomic sequencing.

This thesis addresses both plant-mediated as well as microbe-mediated interactions. Looking ahead, I foresee that advancements in computational and experimental methodologies will surmount certain methodological and technological obstacles. The utilization of a multi-omics approach is increasing to comprehend the complexity of plant-microbe interactions at the genomic, proteomic, and metabolomic levels. There is an increasing interest in the functional understanding of the microbiome and the development of

synthetic microbial communities. Studies focusing on responses of multi-generation selection on microbial community assembly will enhance the potential use case of incorporating engineered microbial communities into large-scale field settings for improved and sustainable plant production.

BIOGRAPHICAL SKETCH OF THE AUTHOR

Nachiket Shankar was born in Karjat, Maharashtra, India. He relocated to Bangalore, India at the age of four and resided there until the age of 23. He obtained a bachelor's degree in engineering with a focus on biotechnology from M.S. Ramaiah Institute of Technology. During this period, he undertook summer internships at one of India's premier research institutions, the National Centre for Biological Sciences (NCBS), where he engaged in projects spanning from molecular cloning to the regeneration of Planaria. Toward the conclusion of his bachelor's degree, he developed an interest in the evolution and ecology of microbes while working in Dr. Deepa Agashe's lab. After completing his bachelor's degree in 2012, he enrolled as a research fellow at NCBS, collaborating with Dr. Deepa Agashe on DNA evolution in bacteria. This collaboration resulted in a second-author publication in the Journal of Experimental Zoology, fostering his fascination with host-microbe interactions in the context of evolution and ecology.

In 2013, he applied to the University of Massachusetts, Boston (UMB) within the molecular, cellular, and organismal biology program, commencing his tenure at Rick Kesseli's lab in the autumn of 2014. Here, he undertook projects exploring various facets of plant-microbe interactions. During his time at UMB, he also contributed as an instructor for several semesters of Genetics laboratory sessions and delivered guest lectures for the Genetics lecture component. Upon successfully completing his Ph.D. from UMB in December 2023, he aspires to embark on a career in research centered around host-microbe interactions.