

Structure, Variation, and Dynamics of the Root-Associated Microbiota of
the Crop Plant Rice

By

JOSEPH ANDREW ARTHUR EDWARDS
B.S. (Virginia Polytechnic Institute and State University) 2011

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Plant Biolgy

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Chair Venkatesan Sundaresan, Chair

Siobhan Brady

Kate Scow

Committee in Charge

2016

Copyright © 2016 by
Joseph Andrew Arthur Edwards
All rights reserved.

*To my family
For your inexhaustible love and support*

CONTENTS

List of Figures	vi
List of Tables	vii
Abstract	viii
Acknowledgments	x
1 Introduction	1
1.1 The beginnings of microbiology and microbial ecology	1
1.2 The foundation of modern microbial ecology experiments	2
1.3 The rise in modern microbial ecology	3
1.4 Plant Associated Microbiomes	6
2 Structure, variation, and assembly of the root-associated microbiomes of rice	9
2.1 Abstract	9
3 The impact of host interspecies diversity on the root associated microbiome and potential implications in agriculture.	65
3.1 Abstract	65
3.2 Introduction	66
3.3 Results	67
3.3.1 Root-associated microbiomes are influenced by environmental conditions.	67
3.3.2 Different plant species host distinct microbiotas when growing in the same environment.	70
3.3.3 Rice acquires an outlier microbiome compared to native plant species.	72
3.3.4 Rice roots are enriched for methanogenic archaea but not methanotrophic bacteria compared to native plant species.	73
3.3.5 Soil cultivation history affects plant root microbial assemblages.	76
3.4 Discussion	78
3.5 Materials and Methods	81

3.5.1	Meta-Analysis	81
3.5.2	Native Plant Study	82
3.5.3	Soil Domestication Study	82
3.5.4	Seed Sterilization and Germination	83
3.5.5	16S rRNA gene amplification and sequencing	83
3.5.6	Sequence processing	83
4	Successional shifts in the root-associated microbiota throughout the life cycle of the crop plant <i>Oryza sativa</i>	88
4.1	Abstract	88
4.2	Introduction	89
4.3	Results	90
4.3.1	Experimental Design and Sequencing	90
4.3.2	Microbiomes are shaped by distance from the root and plant age	90
4.3.3	Microbiota shifts over the season are marked by increasing and decreasing relative abundance of specific phyla	92
4.3.4	The composition of the root-associated microbiota is an accurate predictor of plant age	95
4.3.5	Developmental stage is a driver of microbiota succession	99
4.3.6	Developmental stage is a driver of microbiota succession	101
4.4	Discussion	103
4.5	Materials and Methods	106
4.5.1	Plant growth and sampling during the 2014 and 2015 seasons	106
4.5.2	Plant growth during the 2016 Season	106
4.5.3	Separation of the root-associated compartments and DNA extraction	107
4.5.4	PCR amplification and sequencing	107
4.5.5	Sequence Processing, OTU clustering, and OTU filtering	108
4.5.6	Statistical Analysis	108

5 Conclusion	113
5.1 Chapter 1	113
5.2 Chapter 2	115
5.3 Chapter 3	116

LIST OF FIGURES

1.1	Figure 1.1	4
1.2	Figure 1.2	5
3.1	Figure 3.1	69
3.2	Figure 3.2	71
3.3	Figure 3.3	73
3.4	Figure 3.4	74
3.5	Figure 3.5	77
3.6	Figure 3S1	85
3.7	Figure 3S2	86
3.8	Figure 3S3	87
4.1	Figure 4.1	91
4.2	Figure 4.2	93
4.3	Figure 4.3	95
4.4	Figure 4.4	98
4.5	Figure 4.5	99
4.6	Figure 4.6	102
4.7	Figure 4S1	110
4.8	Figure 4S2	111
4.9	Figure 4S3	112
5.1	Figure 5.1	115
5.2	Figure 5.2	116
5.3	Figure 5.3	118

LIST OF TABLES

3.1 Table 3.1	68
3.2 Table 3.2	70

ABSTRACT

Structure, Variation, and Dynamics of the Root-Associated Microbiota of the Crop Plant Rice

Soil microbes are important mediators of plant nutrition, growth, and resistance to biotic and abiotic stresses, thus understanding the composition and dynamics of plant root-associated microbial communities is of importance to scientists and farmers. Plants acquire root-associated microbiota that are distinct from that of the surrounding soil. In this thesis I detail general characteristics of the rice root-associated microbiota.

In the first chapter I introduce the importance of both animal and plant microbiota and I discuss the technological breakthroughs that have allowed scientists to study complex microbial ecosystems.

In the second chapter, I present the results of a study characterizing the rice root-associated microbiota across three root compartments: the soil adjacent to the root (the rhizosphere), the root surface (the rhizoplane), and the root interior (the endosphere). I present how the composition of the microbiota in these compartments varies across soil type, cultivation practice, and host plant genotype.

In the third chapter, I present the results of a study characterizing the rice root-associated microbiota compared to diverse species of both crop plants and native plants growing in the same environment as rice. The results suggest that rice assimilates a unique root-associated microbiota and that plant phylogenetic variation cannot fully explain microbiota variation between the tested host plants. A comparison of the root-associated microbiomes of rice with microbiota of soil from cultivated fields and non-cultivated soils, suggests that rice cultivation changes the overall soil microbiome in rice fields not only through agricultural inputs and water submergence, but also through specific enrichment of microbial taxa by rice plants.

In the fourth chapter, I demonstrate how the rice root-associated microbiota shifts across the lifecycle of the host plants under field conditions. We find that the composition of the microbiome shifts throughout vegetative growth until the initiation of reproductive growth,

whereupon the microbiome stabilizes. These results suggest that plants select for different microbiota as they undergo developmental changes from germination to reproduction.

I present concluding remarks on the various models of microbiota acquisition and discuss future implications of this research in the fifth and final chapter of this dissertation.

ACKNOWLEDGMENTS

Acknowledgements to those who helped you get to this point. They should be listed by chapter when appropriate.

Chapter 1

Introduction

1.1 The beginnings of microbiology and microbial ecology

For centuries, humans have been curious to understand the unseen organisms that surround them: this is the basic objective of microbial ecology. Up until the 17th century, however, prominent scientists were naive to the lifeforms that could not be observed with the naked eye. The field of microscopy began to expand when the Englishman Robert Hooke developed one of the first compound microscopes and published the book *Micrographia* [23] in 1665 with detailed drawings of never before observed images of magnified organisms and structures. It was in this book that he coined the term cell after observing plant cells from cork tissue. He based this term on the resemblance to the walled structure in honeycomb. Unfortunately, very little is known of Hooke's career and accomplishments (there does not even exist a portrait of him) due to his falling out with contemporary scientists of prominence, namely Sir Isaac Newton [20].

Hooke's microscopy work inspired others, however. Antonie van Leeuwenhoek took an interest in Hooke's *Micrographia* in order to apply microscopy to his business of owning and operating a drapery [16]. Leeuwenhoek, a Dutch businessman, was intent to use the finest quality thread and fabric for his work, and believed he could have an advantage over his competitors if he were to use an instrument with higher magnification than modern magnifying glasses to inspect the material for imperfections. Leeuwenhoek was skilled at grinding lenses from glass and was able to craft microscopes with 200X magnification, thus

surpassing the capability of the best microscopes at the time, the most powerful of which had 30X magnification. He began to discover many microscopic structures which had yet to be observed. Most notable among these observations were blood cells, sperm cells, and protists and algae from pond water. After he would record his observations, he would write letters in Dutch to the Royal Society of London, which would be translated to English or Latin and then published by the society. In 1683, Leeuwenhoek wrote a letter to the Royal Society detailing the life forms which he had observed within plaque sampled from between his teeth. He wrote of the motility and the dense occurrence of at least two types of “animalcules”, what we now know as bacteria. These observations led to the emergence of the microbiology and microbial ecology fields.

1.2 The foundation of modern microbial ecology experiments

Microbial ecology is the broad study of microbial organisms in their natural environments [2]. One of the primary objectives of microbial ecology is to understand the complete diversity of microbial organisms cohabitating a single environment. This objective includes understanding the taxonomic classification of the detectable microbes. Culturing individual microbial isolates from environmental samples, which was established by Robert Koch in the late 19th century [41], was a promising technique for surveying the microorganisms present in each sample. However, it was found apparent that the diversity levels capable of being cultured on media were underwhelming compared to the diversity of microorganisms observed under a microscope. Specifically, in soil, <1% of microorganisms are capable of being cultured under laboratory conditions [25]. Thus culturing isolated microbes is currently a fruitless task for surveying overall diversity, while microscopy will give little to no answers about the taxonomic information of the observed microbes.

In order to get a better grasp of the microorganisms present in environments and to what taxonomic classifications they belong, scientists turned to culture-independent surveys [25]. One groundbreaking and seminal technique was the use of 16S ribosomal RNA to classify microorganisms phylogenetically by Carl Woese and colleagues. Woese, in the mid 1970’s, was skeptical of the dogma that all organisms separate into 2 domains solely based

on the presence or absence of a nucleus, i.e. eukaryotes and prokaryotes. In his seminal study [64], Woese complained that “Dividing the living world into Prokaryotae and Eukaryotae has served, if anything, to obscure the problem of what extant groupings represent the various primeval branches from the common line of descent.” He made this complaint because using presence or absence of nuclei as a classifier of domain is not a phylogenetic measure, and thus scientists cannot resolve evolutionary relationships in any meaningful way using this metric. Woese and his colleague, John Fox, proved their point by comparing 16S rRNA sequences found in prokaryotic microorganisms to 18S rRNA sequences from eukaryotes (the homologous eukaryotic rRNA to 16S rRNA in prokaryotes) [64]. They used 16S rRNA as the phylogenetic marker because they needed a biological molecule that was conserved across the tree of life, slowly evolved over time, and was easily extractable from cells. No known protein fit these descriptions at the time. They found that, through analysis of 16S rRNA, prokaryotic organisms had 2 distinct lineages which were similarly distant from each other as each lineage was to eukaryotes. They termed these lineages as eubacteria and archaeabacteria. Not only did Woese and Fox discover a new domain of life, but they laid the groundwork for using the phylogenetically conserved 16S rRNA subunit as a marker for topologically describing the tree of life.

Woese conducted his fundamental studies on cultured organisms. Norman Pace, a colleague of Woese, began to apply 16S rRNA sequencing to unculturable organisms to identify evolutionary lineages of uncharacterized bacteria and archaea [43]. The scientific community soon adopted his approach and the vast topology of the tree of life began to fill in. In the 1980’s, when these studies were taking place, DNA sequencing was a tedious, non-trivial process. Two decades later, DNA sequencing technology reached major breakthroughs causing the rapid expansion of genomic studies and importantly microbial ecology studies.

1.3 The rise in modern microbial ecology

Relatively recently, the use of culture-independent techniques to study microbial community structure has been rebranded under new field names. The Google search terms “mi-

crobiome” and “microbiota” are trending over the last decade and are still rising. The search term for “microbial ecology”, on the other hand, is decreasing. I will define the terms microbiome and microbiota and give a few potential reasons as to why these fields have expanded in popularity.

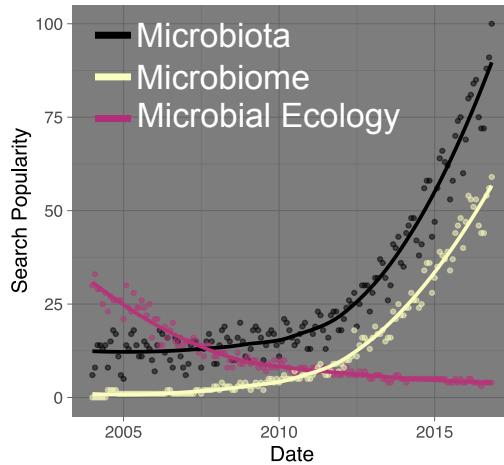


Figure 1.1: Popularity of Google searches for terms related to microbial ecology over the last decade. Data was collected from Google Trends (<https://www.google.com/trends/>).

A microbiota is the complete catalog of microorganisms inhabiting an environment. The definition of the environment can range depending on the study system, but typically the term microbiota refers to microbial communities that are associated with a multicellular host. The study of the human gut microbiota by Jeffrey Gordon's lab has led to growing interests in host-associated microbiomes. Peter Turnbaugh, Ruth Ley and Fredrik Backhed, all members of Gordon's Laboratory, drew a direct connection between the gut microbiota and the host species' physiology and phenotype [57, 4, 31]. They showed that lean and obese mice hosted significantly different communities of microbes in their guts and when an obese mouse gut microbiota was inoculated into lean mice or germ free mice (i.e. mice that have no microbes in their gut) that the recipients rapidly gained body fat. These studies definitively showed that the gene content of the microbiota residing in a host animal's gut, i.e. the microbiome, has the ability to significantly influence the host's physiology and health. These studies were instrumental for initiating the Human Microbiome Project [56] as well as providing a significant motivation for studying all other host-associated microbiomes, including plant microbiomes.

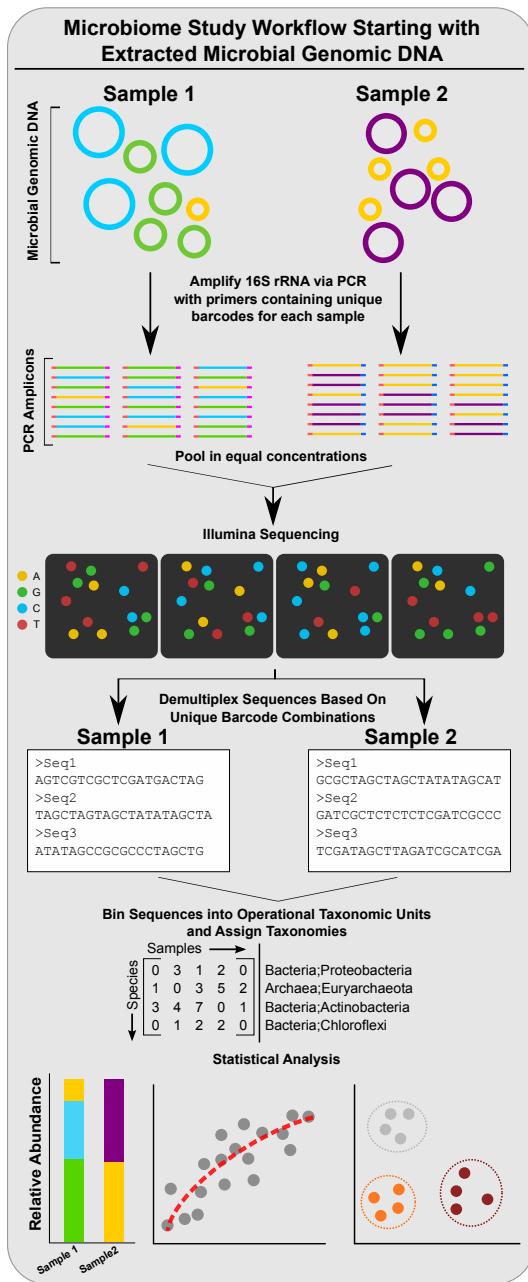


Figure 1.2: Standard workflow for profiling microbial communities from environmental DNA.

While biological advances drew a considerable amount of interest to the field of microbiomes, sequencing capability still posed technical barriers for understanding host-associated microbial communities in depth. 16S rRNA sequencing was the method of choice for profiling microbial communities but researchers were burdened by the process of cloning 16S rRNA PCR amplicons into plasmids, cloning into *Escherichia coli*, and using Sanger sequencing to survey microbial diversity within samples. This method is relatively tedious,

very low throughput, and expensive. At the time, a standard study would profile 300-400 different clones for each sample. Next generation sequencing again revolutionized the technical aspect of microbial ecology field. The first application of pyrosequencing to 16S rRNA sequencing was by Sogin et al in 2006 [54]. They were able to sequence over 119,000 16S rRNA sequence tags and importantly they were able to barcode separate samples into a single sequencing run. Pyrosequencing became the preferred method for researchers investigating microbiomes, but this technology would soon be surpassed. Illumina sequencing, which had previously not been applicable to profiling 16S rRNA fragments due to read length limitations, began to emerge as the standard method for sequencing once the MiSeq platform was introduced in 2011 [12]. Compared to the previous iteration of Illumina chemistry, the MiSeq platform allowed for greater read lengths, while maintaining the high throughput standard with Illumina. The general workflow for community profiling using 16S rRNA gene sequencing is displayed in Figure 2.

1.4 Plant Associated Microbiomes

The impact of host-associated microbial communities on animal health is a robust area of research. Investigators are similarly interested in plant-associated microbiota. As sessile organisms, plants rely on their environment for the uptake of critical nutrients [19]. Co-habitating the soil with plants are thousands of different bacterial and archaeal strains, some of which are known to aid host plants in tolerating otherwise inhospitable conditions. For example, the genus *Lupinus* in the Fabaceae family, thrives in nitrogen poor soil where other plants struggle to survive. The lupine received its name because it was thought to act as a wolf, plundering nutrients in the soil and thus making the environment unfavorable for the growth of other plants [3]. This hypothesis was false. In fact, lupines and other members of the Fabaceae family, form symbiotic relationships with nitrogen fixing soil bacteria thereby allowing them to survive otherwise nitrogen limited conditions [58]. This is an important and widespread example of how plants form beneficial relationships with soil microbes.

Plants also rely on soil microbes to survive attacks from pathogens [7, 9]. In some crop-

ping systems, such as sugar beet and wheat, continuous monoculture cultivation of soil leads to reduced loss of yield from pathogen infection [24]. It was found that this phenomenon is transmissible to other, non-suppressive soils and that the effect is lost upon pasteurization at 80°C, thus suggesting that the mechanism is biological in nature [24]. Mendes et al. tested the root-associated bacterial communities between sugar beet seedlings grown in soils suppressive to *Rhizoctonia solani* infection and soils conducive for infection (sampled from the margin of the fields where suppressive soils were located) [37]. They found that no single bacterial or archaeal strain was responsible for the difference in microbial community structure between suppressive and conducive soils. Many of the microbial strains they detected differed in relative abundance between the soils, mainly within the Gamma and Beta Proteobacterial classes. These results were instrumental in demonstrating how broader microbial communities, i.e. the root microbiota, are important for soil and plant health.

Because soil microbial communities were found to impact plant health, deep characterizations of the plant root microbiota began to take place. In 2012, two groups published studies characterizing the root-associated microbiota of *Arabidopsis thaliana* [9, 34, ?]. These studies found that the consortia of microbes inside of the root (the endosphere) is significantly different and of lower diversity than the microbial communities in the soil adhering to the root (the rhizosphere). Each study found that variability in microbial community structure between different soils is also reflected in the root-associated microbial assemblages of the plants grown in the respective soils. The studies found plant genotype to have a small but significant effect on the microbial communities associated with the roots. It was later found in rice that there is a third discernible niche on the rhizoplane (the root surface) and that these microbial communities are intermediate in composition between the rhizosphere and endosphere [18]. A longitudinal study of early microbiome acquisition in rice detailed the dynamic changes in microbiota composition within the rhizosphere, rhizoplane, and endosphere compartments [18]. Together this data suggested a three-step model for microbiota assembly in the root: microbes are first attracted to the roots in the rhizosphere, perhaps by root exudates, after which a select set of microbes can

attach to the rhizoplane and a subset of rhizoplane attached microbes can enter the root endosphere.

Open questions still remain for the plant root-associated microbiota. Functionally, there are only limited examples of how root associated bacterial and archaeal communities may impact plant fitness. Further studies are needed to assess how the root-associated microbiota may aid plants in tolerating abiotic and biotic stresses. Once a better understanding of how microbiomes may help to protect plants from various stresses is attained, researchers are intent to leverage crop genetics to breed for beneficial root microbiotas. Because intraspecies genetic variation has a minimal impact on the root-associated microbiota, it is of importance to better understand how host interspecies variation impacts the root-associated microbiome and what physiological traits are causal of microbiome divergence between plant species. It is also critical to understand how root associated microbial communities change over the life cycle of the host plant. Plants face different nutritional requirements at various stages of their lives, but it is yet to be characterized how root-associated microbial communities shift over the growing season and how these potential changes are tied to plant developmental shifts.

This thesis is separated into three research chapters. In the first chapter, I characterize the root-associated microbiota of rice plants and quantify the factors affecting root assemblages. In the second chapter, I characterize how the rice root microbiota differs from other crop species as well as how the rice root microbiota differs from native plants growing in the same flooded environment. In this chapter I also show how continuous and intensive cultivation of rice enriches a soil microbiota with rice specific microbes. In the final chapter, I demonstrate the dynamics of the root-associated microbiota over the life cycle of the rice plant and how the abundances of specific microbial taxa are tuned by the developmental stage of the host plant.

Chapter 2

Structure, variation, and assembly of the root-associated microbiomes of rice

Joseph Edwards¹, Cameron Johnson¹, Christian Santos¹, Eugene Lurie¹, Natraj Kumar Podishetty¹, Srijak Bhatnagar¹ Jonathan A. Eisen² and Venkatesan Sundaresan¹

2.1 Abstract

Plants depend upon beneficial interactions between roots and microbes for nutrient availability, growth promotion, and disease suppression. High-throughput sequencing approaches have provided recent insights into root microbiomes, but our current understanding is still limited relative to animal microbiomes. Here we present a detailed characterization of the root-associated microbiomes of the crop plant rice by deep sequencing, using plants grown under controlled conditions as well as field cultivation at multiple sites. The spatial resolution of the study distinguished three root-associated compartments, the endosphere (root interior), rhizoplane (root surface), and rhizosphere (soil close to the root surface), each of which was found to harbor a distinct microbiome. Under controlled greenhouse conditions, microbiome composition varied with soil source and genotype. In field conditions, geographical location and cultivation practice, namely organic vs. conventional, were factors contributing to microbiome variation. Rice cultivation is a major source of global methane emissions, and methanogenic archaea could be detected in all spatial com-

¹Department of Plant Biology, University of California, Davis

¹Department Medical Microbiology and Immunology, University of California, Davis

parts of field-grown rice. The depth and scale of this study were used to build coabundance networks that revealed potential microbial consortia, some of which were involved in methane cycling. Dynamic changes observed during microbiome acquisition, as well as steady-state compositions of spatial compartments, support a multistep model for root microbiome assembly from soil wherein the rhizoplane plays a selective gating role. Similarities in the distribution of phyla in the root microbiomes of rice and other plants suggest that conclusions derived from this study might be generally applicable to land plants.

Structure, variation, and assembly of the root-associated microbiomes of rice

Joseph Edwards^a, Cameron Johnson^{a,1}, Christian Santos-Medellín^{a,1}, Eugene Lurie^{a,2}, Natraj Kumar Podishetty^b, Srijak Bhatnagar^c, Jonathan A. Eisen^c, and Venkatesan Sundaresan^{a,b,3}

Departments of ^aPlant Biology, ^bPlant Sciences, and ^cMedical Microbiology and Immunology, University of California, Davis, CA 95616

Edited by Jeffery L. Dangl, Howard Hughes Medical Institute and The University of North Carolina at Chapel Hill, Chapel Hill, NC, and approved December 22, 2014 (received for review July 31, 2014)

Plants depend upon beneficial interactions between roots and microbes for nutrient availability, growth promotion, and disease suppression. High-throughput sequencing approaches have provided recent insights into root microbiomes, but our current understanding is still limited relative to animal microbiomes. Here we present a detailed characterization of the root-associated microbiomes of the crop plant rice by deep sequencing, using plants grown under controlled conditions as well as field cultivation at multiple sites. The spatial resolution of the study distinguished three root-associated compartments, the endosphere (root interior), rhizoplane (root surface), and rhizosphere (soil close to the root surface), each of which was found to harbor a distinct microbiome. Under controlled greenhouse conditions, microbiome composition varied with soil source and genotype. In field conditions, geographical location and cultivation practice, namely organic vs. conventional, were factors contributing to microbiome variation. Rice cultivation is a major source of global methane emissions, and methanogenic archaea could be detected in all spatial compartments of field-grown rice. The depth and scale of this study were used to build coabundance networks that revealed potential microbial consortia, some of which were involved in methane cycling. Dynamic changes observed during microbiome acquisition, as well as steady-state compositions of spatial compartments, support a multistep model for root microbiome assembly from soil wherein the rhizoplane plays a selective gating role. Similarities in the distribution of phyla in the root microbiomes of rice and other plants suggest that conclusions derived from this study might be generally applicable to land plants.

microbiomes | rice | soil microbial communities | methane cycling | microbiome assembly

Land plants grow in soil, placing them in direct proximity to a high abundance of microbial diversity (1). Plants and microbes have both adapted to use their close association for their mutual benefit. Critical nutrients are converted to more usable forms by microbes before assimilation by plants (2–4). In turn, bacteria in the rhizosphere receive carbon metabolites from the plant through root exudates (5). Beneficial soil microbes also contribute to pathogen resistance, water retention, and synthesis of growth-promoting hormones (6–8).

Recent studies have used high-throughput sequencing to provide new insights into the bacterial composition and organization of different plant microbiomes, including *Arabidopsis*, *Populus*, and maize (9–14). Detailed characterization of the core root microbiome of *Arabidopsis* (9–11) showed that the dominant phyla inside the root (the endosphere) are much less diverse than the phyla in the soil around the root (the rhizosphere), and a potential core root microbiome could be identified. In *Arabidopsis*, the endophytic microbiome exhibits some genotype-dependent variation within the species and an increased variation when other related species are examined (9–11). A recent study in maize examined microbiome variation across many different inbred lines at different sites and found a large variation arising from geographical location between three different states in the United States and a relatively smaller dependence on the genotype (12). Although the

microbiomes examined in the maize study consisted of combined rhizospheric and endospheric microbes (12), a study in poplar found that the variation between locations in two different states affected both rhizospheric and endospheric microbes (14).

These studies have opened the way toward a new understanding of the composition and structure of plant microbiomes and the factors that affect them. However, this understanding is still at the initial stages, and several key questions are as yet unanswered. One such question regards the mechanism of microbiome acquisition and assembly in plants. Unlike animals, where the gut microbiome is assembled internally and is transmissible through birth (15, 16), the root microbiome is predominantly assembled from the external microbes in the soil. Based on the composition of the endospheric and rhizospheric microbiomes, it has been proposed that plants might assemble their microbiomes in two steps, with the first step involving a general recruitment to the vicinity of the root and a second step for entry inside the root that involves species-specific genetic factors (7). Although this is a plausible hypothesis, direct support for this model through detailed dynamic studies has not yet been provided. Additionally, the role of the root surface or rhizoplane, which forms the critical interface between plants and soil, remains poorly understood, and the microbial composition of the rhizoplane in relation to those of the rhizosphere and endosphere is unknown.

Significance

Land plants continuously contact beneficial, commensal, and pathogenic microbes in soil via their roots. There is limited knowledge as to how the totality of root-associated microbes (i.e., the microbiome) is shaped by various factors or its pattern of acquisition in the root. Using rice as a model, we show that there exist three different root niches hosting different microbial communities of eubacteria and methanogenic archaea. These microbial communities are affected by geographical location, soil source, host genotype, and cultivation practice. Dynamics of the colonization pattern for the root-associated microbiome across the three root niches provide evidence for rapid acquisition of root-associated microbiomes from soil, and support a multistep model wherein each root niche plays a selective role in microbiome assembly.

Author contributions: J.E., C.J., C.S.-M., and V.S. designed research; J.E., C.S.-M., E.L., and N.K.P. performed research; S.B. and J.A.E. contributed new reagents/analytic tools; J.E. analyzed data; and J.E. and V.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the National Center for Biotechnology Information Short Read Archive (accession no. [SRP044745](#)).

See Commentary on page 2299.

¹C.J. and C.S.-M. contributed equally to this work.

²Present address: Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030.

³To whom correspondence should be addressed. Email: sundar@ucdavis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1414592112/-DCSupplemental.

To address some of these questions, we have undertaken an exhaustive characterization of the root-associated microbiome of rice. Rice is a major crop plant and a staple food for half of the world's population. Metagenomic and proteomic approaches have been used to identify different microbial genes present in the rice microbiome (17, 18), but an extensive characterization of microbiome composition and variation has not been performed. Rice cultivation also contributes to global methane, accounting for an estimated 10–20% of anthropogenic emissions, due to the growth of methanogenic archaea in the vicinity of rice roots (19). Here we have used deep sequencing of microbial 16S rRNA genes to detect over 250,000 operational taxonomic units (OTUs), with a structural resolution of three distinct compartments (rhizosphere, rhizoplane, and endosphere) and extending over multiple factors contributing to variation, both under controlled greenhouse conditions as well as different field environments. The large datasets from the different conditions sampled in this study were used for identification of putative microbial consortia involved in processes such as methane cycling. Through dynamic studies of the microbiome composition, we provide insights into the process of root microbiome assembly.

Results

Root-Associated Microbiomes Form Three Spatially Separable Compartments Exhibiting Distinct and Overlapping Microbial Communities. Sterilized rice seeds were germinated and grown under controlled greenhouse conditions in soil collected from three rice fields across the Central Valley of California (*SI Appendix*, Fig. S1). We analyzed the bacterial and archaeal microbiomes from three separate rhizocompartments: the rhizosphere, rhizoplane, and endosphere (Fig. 1A). Because the root microbiome has been shown to correlate with the developmental stage of the plant (10), the root-associated microbial communities were sampled at 42 d (6 wk), when rice plants from all genotypes were well-established in the soil but still in their vegetative phase of growth. For our study, the rhizosphere compartment was com-

posed of ~1 mm of soil tightly adhering to the root surface that is not easily shaken from the root (*SI Appendix*, Fig. S2). The rhizoplane compartment microbiome was derived from the suite of microbes on the root surface that cannot be removed by washing in buffer but is removed by sonication (*SI Appendix*, Materials and Methods). The endosphere compartment microbiome, composed of the microbes inhabiting the interior of the root, was isolated from the same roots left after sonication. Unplanted soil pots were used as a control to differentiate plant effects from general edaphic factors.

The V4–V5 region of the 16S rRNA gene was amplified using PCR and sequenced using the Illumina MiSeq platform. A total of 10,554,651 high-quality sequences was obtained with a median read count per sample of 51,970 (range: 2,958–203,371; **Dataset S2**). The high-quality reads were clustered using >97% sequence identity into 101,112 microbial OTUs. Low-abundance OTUs (<5 total counts) were discarded, resulting in 27,147 OTUs. The resulting OTU counts in each library were normalized using the trimmed mean of M values method. This method was chosen due to its sensitivity for detecting differentially abundant taxa compared with traditional microbiome normalization techniques such as rarefaction and relative abundance (20). Measures of within-sample diversity (α -diversity) revealed a diversity gradient from the endosphere to the rhizosphere (Fig. 1B and **Dataset S4**). Endosphere communities had the lowest α -diversity and the rhizosphere had the highest α -diversity. The mean α -diversity was higher in the rhizosphere than bulk soil; however, the difference in α -diversity between these two compartments cannot be considered as statistically significant (Wilcoxon test; **Dataset S4**).

Unconstrained principal coordinate analyses (PCoAs) of weighted and unweighted UniFrac distances were performed to investigate patterns of separation between microbial communities (*SI Appendix*, Materials and Methods). The UniFrac distance is based on taxonomic relatedness, where the weighted UniFrac (WUF) metric takes abundance of taxa into consideration whereas the unweighted UniFrac (UFF) does not and is thus more sensitive to rare taxa. In both the WUF and UUF PCoAs, the rhizocompartments separate across the first principal coordinate, indicating that the largest source of variation in root-associated microbial communities is proximity to the root (Fig. 1C, WUF and *SI Appendix*, Fig. S4, UUF). Moreover, the pattern of separation is consistent with a gradient of microbial populations from the exterior of the root, across the rhizoplane, and into the interior of the root. Permutational multivariate analysis of variance (PERMANOVA) corroborates that rhizospheric compartmentalization comprises the largest source of variation within the microbiome data when using a WUF distance metric (46.62%, $P < 0.001$; **Dataset S5A**). PERMANOVA using a UUF distance, however, describes rhizospheric compartmentalization as having the second largest source of variation behind soil type (18.07%, $P < 0.001$; **Dataset S5H**). In addition to PERMANOVA, we also performed partial canonical analysis of principal coordinates (CAP) on both the WUF and UUF metrics to quantify the variance attributable to each experimental variable (*SI Appendix*, Materials and Methods). This technique differs from unconstrained PCoA in that technical factors can be controlled for in the analysis and the analysis can be constrained to any factor of interest to better understand the quantitative impact of the factor on the microbial composition. Using this technique to control for soil type, cultivar, and technical factors (biological replicate, sequencing batch, and planting container), we found that in agreement with the PERMANOVA results, microbial communities vary significantly between rhizocompartments (34.2% of variance, $P = 0.005$, WUF, *SI Appendix*, Fig. S5A and 22.6% of variance, $P = 0.005$, UUF, *SI Appendix*, Fig. S5C).

There are notable differences in the proportions of various phyla across the compartments that are consistent across every tested soil (Fig. 1D). The endosphere has a significantly greater proportion of Proteobacteria and Spirochaetes than the rhizosphere or bulk soil, whereas Acidobacteria, Planctomycetes, and Gemmatimonadetes are mostly depleted in the endosphere

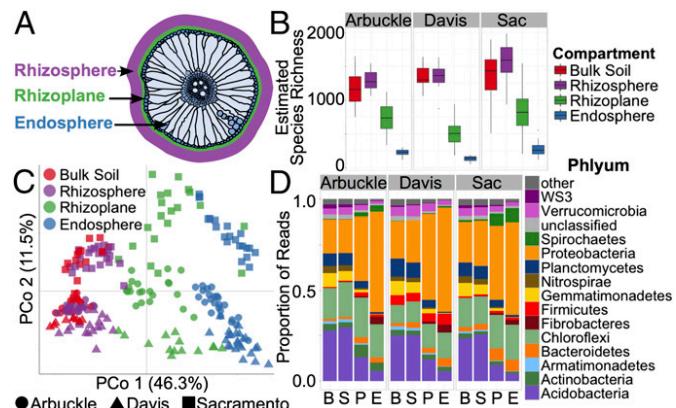


Fig. 1. Root-associated microbial communities are separable by rhizo-compartment and soil type. (A) A representation of a rice root cross-section depicting the locations of the microbial communities sampled. (B) Within-sample diversity (α -diversity) measurements between rhizospheric compartments indicate a decreasing gradient in microbial diversity from the rhizosphere to the endosphere independent of soil type. Estimated species richness was calculated as $e^{\text{Shannon_entropy}}$. The horizontal bars within boxes represent median. The tops and bottoms of boxes represent 75th and 25th quartiles, respectively. The upper and lower whiskers extend 1.5x the interquartile range from the upper edge and lower edge of the box, respectively. All outliers are plotted as individual points. (C) PCoA using the WUF metric indicates that the largest separation between microbial communities is spatial proximity to the root (PCo 1) and the second largest source of variation is soil type (PCo 2). (D) Histograms of phyla abundances in each compartment and soil. B, bulk soil; E, endosphere; P, rhizoplane; S, rhizosphere; Sac, Sacramento.

compared with the bulk soil or rhizosphere (Wilcoxon test; [Dataset S6](#)). The reduction in relative abundance of these phyla across the compartments is consistent with the observation that microbial diversity decreases from the rhizosphere to the endosphere.

Association of Significantly Enriched OTUs with Different Rhizocompartments. To identify OTUs that are correlated with community separation between the compartments, we conducted differential OTU abundance analysis by fitting a generalized linear model with a negative binomial distribution to normalized values for each of the 27,147 OTUs in the greenhouse experiment and testing for differential abundance using a likelihood ratio test ([Dataset S7](#)). Using OTU counts from unplanted soil as a control and an adjusted P value cutoff of 0.01, there were 578 OTUs that were significantly enriched in at least one compartment. The rhizosphere was the most similar to bulk soil, as indicated by the “tail” in the MA plot (Fig. 2A); however, an enrichment effect in the rhizosphere is implied by the high ratio of statistically significant enriched OTUs compared with depleted OTUs (152 vs. 17). In comparison, the rhizoplane enriches for many OTUs while simultaneously depleting a larger proportion of OTUs (422 vs. 730). The endosphere is the most exclusive compartment, enriching for 394 OTUs while depleting 1,961 OTUs (Fig. 2A).

There were noteworthy overlaps in differentially abundant OTUs between the compartments (Fig. 2B and C). The OTUs enriched in the rhizosphere are very successful at colonizing the root, as 119 out of the 152 OTUs enriched in the rhizosphere are also enriched in either the rhizoplane or endosphere communities or both (Fig. 2B). There was a set of 96 OTUs mainly consisting of Bacteroidetes, Firmicutes, Chloroflexi, and Betaproteobacteria that were differentially enriched in all rhizocompartments compared with bulk soil ([SI Appendix](#), Fig. S6A). The Betaproteobacterial OTUs that are enriched in every rhizocompartment correspond mainly to *Rhodocyclaceae* and *Comamonadaceae* ([SI](#)

[Appendix](#), Fig. S6B). OTUs belonging to the genus *Pleomorphomonas* are also enriched in all of the rhizocompartments ([Dataset S7](#)). Some members within the *Pleomorphomonas* genus are capable of nitrogen fixation (21–23). The rhizoplane and endosphere were the most similar rhizocompartments, sharing 271 enriched OTUs. Most of the OTUs enriched between the rhizoplane and endosphere compartments belonged to Alpha-, Beta-, and Deltaproteobacterial classes, Chloroflexi, and Bacteroidetes. Not surprisingly, a subset of the OTUs enriched in the endosphere and rhizoplane belong to Fibrobacteres and Spirochaetes, phyla that are associated with cellulose degradation (24, 25).

It was possible to quantify exclusionary effects of each compartment by analyzing OTU abundance relative to bulk soil (Fig. 2A). The rhizosphere had a small effect on excluding microbes, as only 17 OTUs were significantly depleted compared with bulk soil. These OTUs were mainly in the Proteobacteria and Acidobacteria phyla ([SI Appendix](#), Fig. S7). Many more OTUs were reduced in relative abundance in the rhizoplane (730 OTUs, mainly Acidobacteria and Planctomycetes; [SI Appendix](#), Fig. S7), and even more OTUs (1,961 OTUs, mainly Acidobacteria, Planctomycetes, Chloroflexi, and Verrucomicrobia; [SI Appendix](#), Fig. S7) were reduced in the endosphere. There are considerable overlaps in the OTUs that are excluded from each compartment (Fig. 2C). Nearly all of the OTUs depleted from the rhizosphere are also depleted in the rhizoplane and endosphere communities. The rhizoplane shares 713 of the 1,961 OTUs that are significantly depleted from the endosphere. These results indicate that plant-controlled changes in the rhizospheric soil are the first level of exclusion of microbial colonization of the root and that selectivity at the rhizoplane might act effectively as a gate for controlling entry of the microbes into the root endosphere. Similar patterns were present when each soil type was analyzed separately ([SI Appendix](#), Fig. S8 B–G).

Microbial Communities Colonizing Rhizocompartments Vary by Soil Type. To investigate how soil variation might affect the root microbiome, plants were grown in soil collected from rice fields at three locations from across the California Sacramento Valley: Davis, Sacramento, and Arbuckle ([SI Appendix](#), Fig. S1). Both the Arbuckle and Sacramento fields primarily grow rice and were drained 3 wk before soil collection. The Davis site had previously grown rice for several years; however, it was not cultivated the year before soil collection. By growing the plants in controlled greenhouse conditions, we aimed to control for climatic variations between the sites and identify only changes attributable to the different soils.

α -Diversity measurements comparing microbial communities between each soil type revealed a small but significant difference in diversity (1.54% variance explained, $P = 2.54\text{E-}5$, ANOVA; [Dataset S3](#)). The two cultivated fields (Arbuckle and Sacramento) had significantly higher diversities in the rhizoplane and endosphere microbial communities than the uncultivated Davis field (Fig. 1B and [Dataset S4](#)). PCoA shows that rhizosphere samples from plants grown in the distinct soils separate along the second axis and that the separation pattern manifests in every compartment (Fig. 1C, WUF and [SI Appendix](#), Fig. S4, UUF). The Arbuckle and Davis rhizosphere microbiomes were most similar to each other despite being the most geographically separated. PERMANOVA using the WUF distance supports the PCoA results that the soil effect describes the second largest source of variation in the tested factors of the experiment ([Dataset S5A](#)). PERMANOVA using the UUF distance measure indicates that the soil effect has the largest source of variation in the factors tested (20.90%, $P < 0.001$; [Dataset S5H](#)). CAP analysis constrained to soil type and controlling for rhizocompartment, cultivar, and technical factors agreed with the PERMANOVA result in that soil type explained the second largest source of variation in the microbial communities behind compartment when using the WUF metric (20.2%, $P = 0.005$; [SI Appendix](#), Fig. S5B) and the largest source of variation when

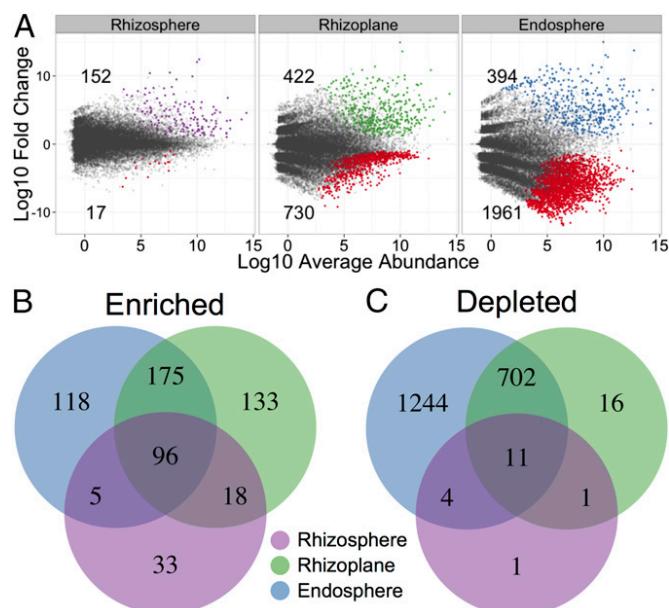


Fig. 2. Rhizocompartments are enriched and depleted for certain OTUs. (A) Enrichment and depletion of the 27,147 OTUs included in the greenhouse experiment for each rhizospheric compartment compared with bulk soil controls as determined by differential abundance analysis. Each point represents an individual OTU, and the position along the y axis represents the abundance fold change compared with bulk soil. (B) Numbers of differentially enriched OTUs between each compartment compared with bulk soil. (C) Numbers of differentially depleted OTUs between each compartment.

using the UUF metric (26.7%, $P = 0.005$; *SI Appendix*, Fig. S5D). This discrepancy is likely due to differences between the WUF and UUF distance metric: Soil type might have more of an effect on frequency of rare taxa than abundant taxa, and thus the UUF metric has a larger effect size for soil type. Compartments of plants grown in distinct soils have commonalities in differentially abundant OTUs (**Dataset S9**), sharing 92 endosphere-enriched OTUs, 71 rhizoplane-enriched OTUs, and 10 rhizosphere OTUs (*SI Appendix*, Fig. S8 J, I, and H, respectively, and *SI Appendix*, Fig. S9). In agreement with the PCoA analysis, Davis and Arbuckle shared a significant overlap in OTUs enriched in the endosphere and rhizoplane ($P = 2.22 \times 10^{-16}$ and 7.86×10^{-7} , respectively, hypergeometric test; *SI Appendix*, Fig. S8 I and J) but not the rhizosphere ($P = 0.52$, hypergeometric test; *SI Appendix*, Fig. S8H). The Sacramento soil did not share significant overlaps in compartment-enriched OTUs with the other sites.

The enrichment/depletion effects within each rhizosphere compartment vary by soil. Rhizosphere compartments of plants in Davis and Arbuckle soils exhibited higher enrichment/depletion ratios (72/3 and 53/17, respectively) than plants in Sacramento soil (78/116) (*SI Appendix*, Fig. S8A). The level of enrichment is similar between each soil in the rhizosphere; however, the depletion level is higher in Sacramento soil than in Arbuckle or Davis. Chemical analysis of the soils showed that the nutrient compositions of the soils did not show any exceptional trends (**Dataset S7**). The Davis and Arbuckle fields were similar in pH and nitrate, magnesium, and phosphorus content, whereas the Arbuckle and Sacramento fields were similar in potassium, calcium, and iron content. Taken together, these results indicate that each soil contains a different pool of microbes and that the plant is not restricted to specific OTUs but instead draws from available OTUs in the pool to organize its microbiome. Nevertheless, the distribution of phyla across the different compartments was similar for all three soil types (Fig. 1D), suggesting that the overall recruitment of OTUs is governed by a set of factors that result in a consistent representation of phyla independent of soil type.

Microbial Communities in the Rhizocompartments Are Influenced by Rice Genotype. To investigate the relationship between rice genotype and the root microbiome, domesticated rice varieties cultivated in widely separated growing regions were tested. Six cultivated rice varieties spanning two species within the *Oryza* genus were grown for 42 d in the greenhouse before sampling. Asian rice (*Oryza sativa*) cultivars M104, Nipponbare (both temperate *japonica* varieties), IR50, and 93-11 (both *indica* varieties) were grown alongside two cultivars of African cultivated rice *Oryza glaberrima*, TOG7102 (Glab B) and TOG7267 (Glab E). PERMANOVA indicated that rice genotype accounted for a significant amount of variation between microbial communities when using WUF (2.41% of the variance, $P < 0.001$; **Dataset S5A**) and UUF (1.54% of the variance, $P < 0.066$; **Dataset S5H**); however, visual representations for clustering patterns of the genotypes were not evident on the first two axes of unconstrained PCoA ordinations (*SI Appendix*, Fig. S10). We then used CAP analysis to quantify the effect of rice genotype on the microbial communities. By focusing on rice cultivar and controlling for compartment, soil type, and technical factors, we found that genotypic differences in rice have a significant effect on root-associated microbial communities (5.1%, $P = 0.005$, WUF, Fig. 3A and 3.1%, $P = 0.005$, UUF, *SI Appendix*, Fig. S11A). Ordination of the resulting CAP analysis revealed clustering patterns of the cultivars that are only partially consistent with genetic lineage for both the WUF and UUF metrics. The two *japonica* cultivars clustered together and the two *O. glaberrima* cultivars clustered together; however, the *indica* cultivars were split, with 93-11 clustering with the *O. glaberrima* cultivars and IR50 clustering with the *japonica* cultivars.

To analyze how the genotypic effect manifests in individual rhizocompartments, we separated the whole dataset to focus on each compartment individually and conducted CAP analysis controlling for soil type and technical factors. The rhizosphere

had the greatest genotypic effect on the microbiome (30.3%, $P = 0.005$, WUF, *SI Appendix*, Fig. S11B and 10.5%, $P = 0.005$, UUF, *SI Appendix*, Fig. S11E). The clustering patterns of the cultivars in the rhizosphere were similar to the clustering patterns exhibited when conducting CAP analysis on the whole data using all rhizocompartments. Again, the *japonica* and *O. glaberrima* cultivars clustered separately, whereas the *indica* cultivars were split between the *japonica* and *O. glaberrima* clusters. This clustering pattern is maintained in the rhizoplane communities (*SI Appendix*, Fig. S11 C and F); however, it breaks down in the endosphere compartment communities, which coincidentally are the least affected by rice genotype (12.8%, $P = 0.005$, WUF, *SI Appendix*, Fig. S11D and 8.5%, $P = 0.028$, UUF, *SI Appendix*, Fig. S11G). α -Diversity measurements within the rhizosphere show a notable difference between the cultivars ($P = 3.12E-06$, ANOVA), with the *O. glaberrima* cultivars exhibiting high diversity relative to the *japonica* cultivars, especially in Arbuckle soil (Fig. 3B and **Dataset S11**). Again, the two *japonica* cultivars were more similar to the *indica* cultivar IR50, and the two *O. glaberrima* cultivars were more similar to the *indica* cultivar 93-11. These patterns in α -diversity were not evident when examining other compartments (*SI Appendix*, Fig. S12). To explain which OTUs accounted for the genotypic effects in each rhizocompartment, we performed differential OTU abundance analyses between the cultivars (**Dataset S12**). In total, we found 125 OTUs that were affected by the plant genotype in at least one rhizocompartment. The rhizosphere had the most OTUs that were significantly impacted by genotype (*SI Appendix*, Fig. S13). This is consistent with the results from PERMANOVA and the CAP analyses.

Geographical Effects on the Microbiomes of Field-Grown Plants. We sought to determine whether the results from greenhouse plants were generalizable to cultivated rice and to investigate other factors that might affect the microbiome under field conditions. We therefore characterized the root-associated microbiomes of field rice plants distributed across eight geographically separate sites across California's Sacramento Valley (Fig. 4A). These eight sites were operated under two cultivation practices: organic cultivation and a more conventional cultivation practice termed “ecofarming” (see below). Because genotype explained the least variance in the greenhouse data, we limited the analysis to one cultivar, S102, a California temperate *japonica* variety that is widely cultivated by commercial growers and is closely related to M104 (26). Field samples were collected from vegetatively growing rice plants in flooded fields and the previously defined rhizocompartments were analyzed as before. Unfortunately, collection of bulk soil controls for the field experiment was not

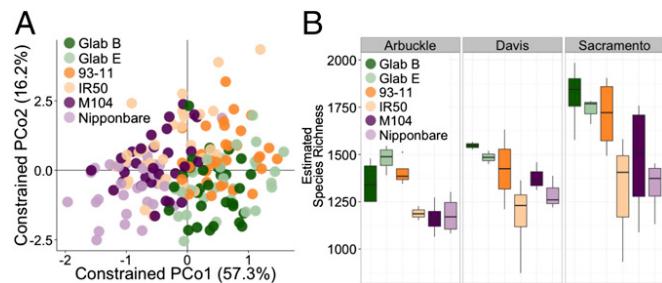


Fig. 3. Host plant genotype significantly affects microbial communities in the rhizospheric compartments. (A) Ordination of CAP analysis using the WUF metric constrained to rice genotype. (B) Within-sample diversity measurements of rhizosphere samples of each cultivar grown in each soil. Estimated species richness was calculated as $e^{Shannon_entropy}$. The horizontal bars within boxes represent median. The tops and bottoms of boxes represent 75th and 25th quartiles, respectively. The upper and lower whiskers extend $1.5 \times$ the interquartile range from the upper edge and lower edge of the box, respectively. All outliers are plotted as individual points.

possible, because planting densities in California commercial rice fields are too high to find representative soil that is unlikely to be affected by nearby plants. Amplification and sequencing of the field microbiome samples yielded 13,349,538 high-quality sequences (median: 54,069 reads per sample; range: 12,535–148,233 reads per sample; **Dataset S13**). The sequences were clustered into OTUs using the same criteria as the greenhouse experiment, yielding 222,691 microbial OTUs and 47,983 OTUs with counts >5 across the field dataset.

We found that the microbial diversity of field rice plants is significantly influenced by the field site. α -Diversity measurements of the field rhizospheres indicated that the cultivation site significantly impacts microbial diversity (*SI Appendix*, Fig. S14A, $P = 2.00E-16$, ANOVA and **Dataset S14**). Unconstrained PCoA using both the WUF and UUF metrics showed that microbial communities separated by field site across the first axis (Fig. 4B, WUF and *SI Appendix*, Fig. S14B, UUF). PERMANOVA agreed with the unconstrained PCoA in that field site explained the largest proportion of variance between the microbial communities for field plants (30.4% of variance, $P < 0.001$, WUF, **Dataset S5O** and 26.6% of variance, $P < 0.001$, UUF, **Dataset S5P**). CAP analysis constrained to field site and controlled for rhizocompartment, cultivation practice, and technical factors (sequencing batch and biological replicate) agreed with the PERMANOVA results in that the field site explains the largest proportion of variance between the root-associated microbial communities in field plants (27.3%, $P = 0.005$, WUF, *SI Appendix*, Fig. S15A and 28.9%, $P = 0.005$, UUF, *SI Appendix*, Fig. S15E), suggesting that geographical factors may shape root-associated microbial communities.

Rhizospheric Compartmentalization Is Retained in Field Plants. Similar to the greenhouse plants, the rhizospheric microbiomes of field plants are distinguishable by compartment. α -Diversity of the field plants again showed that the rhizosphere had the highest microbial diversity, whereas the endosphere had the least

diversity for all fields tested (*SI Appendix*, Fig. S14A and **Dataset S15**). PCoA of the microbial communities from field plants using the WUF and UUF distance metrics showed that the rhizocompartments separate across PCo 2 (Fig. 4C, WUF and *SI Appendix*, Fig. S14C, UUF). PERMANOVA indicated that the separation in the rhizospheric compartments explained the second largest source of variation of the factors that were tested (20.76%, $P < 0.001$, WUF, **Dataset S5O** and 7.30%, $P < 0.001$, UUF, **Dataset S5P**). CAP analysis of the field plants' microbiomes constrained to the rhizocompartment factor and controlled for field site, cultivation practice, and technical factors agreed with PERMANOVA that a significant proportion of the variance between microbial communities is explained by rhizocompartment (20.9%, $P = 0.005$, WUF, *SI Appendix*, Fig. S15C and 10.9%, $P = 0.005$, UUF, *SI Appendix*, Fig. S15G).

Taxonomic distributions of phyla for the field plants were overall similar to the greenhouse plants: Proteobacteria, Chloroflexi, and Acidobacteria make up the majority of the rice microbiota. The taxonomic gradients from the rhizosphere to the endosphere are maintained in the field plants for Acidobacteria, Proteobacteria, Spirochaetes, Gemmatimonadetes, Armatimonadetes, and Planctomycetes. However, unlike for greenhouse plants, the distribution of Actinobacteria generally showed an increasing trend from the rhizosphere to the endosphere of field plants (*SI Appendix*, Fig. S14E and **Dataset S16**).

We again performed differential abundance tests between the OTUs in the compartments of field-grown plants (*SI Appendix*, Fig. S16). We found a set of 32 OTUs that were enriched in the endosphere compartment between every cultivation site, potentially representing a core field rice endospheric microbiome (*SI Appendix*, Fig. S17). The set of 32 OTUs consisted of Deltaproteobacteria in the genus *Anaeromyxobacter* and Spirochaetes, Actinobacteria, and Alphaproteobacteria in the family *Methylocystaceae*. Interestingly, 11 of the 32 core field endosphere OTUs were also found to be enriched in the endosphere compartment of greenhouse plants (*SI Appendix*, Fig. S18). Three of these OTUs were classifiable at the family level. These OTUs consisted of taxa in the families *Kineosporiaceae*, *Rhodocyclaceae*, and *Myxococcaceae*, all of which are also enriched in the *Arabidopsis* root endosphere microbiome (10).

Cultivation Practice Results in Discernible Differences in the Microbiomes. The rice fields that we sampled from were cultivated under two practices, organic farming and a variation of conventional cultivation called ecofarming (27). Ecofarming differs from organic farming in that chemical fertilizers, fungicide use, and herbicide use are all permitted but growth of transgenic rice and use of post-harvest fumigants are not permitted. Although cultivation practice itself does significantly affect α -diversity of the rhizospheric compartments overall ($P = 0.008$, ANOVA; **Dataset S14**), there is also a significant interaction between the cultivation practice used and the rhizocompartments ($P = 3.52E-07$, ANOVA; **Dataset S14**), indicating that the α -diversities of some rhizocompartments are affected differentially by cultivation practice. The α -diversity within the rhizosphere compartment varied significantly by cultivation practice, with the mean α -diversity being higher in ecofarmed rhizospheres than organic rhizospheres ($P = 0.001$, Wilcoxon test; **Dataset S14**), whereas not in the endosphere and rhizoplane microbial communities ($P = 0.51$ and 0.75 , respectively, Wilcoxon tests; **Dataset S14**). Under nonconstrained PCoA, the cultivation practices are separable across principal coordinates 2 and 3 for both the WUF metric (Fig. 4D) and UUF metric (*SI Appendix*, Fig. S14D). PERMANOVA of the microbial communities was in agreement with the PCoAs in that cultivation practice has a significant impact on the rhizospheric microbial communities of field rice plants (8.47%, $P < 0.001$, WUF, **Dataset S5O** and 6.52%, $P < 0.001$, UUF, **Dataset S5P**). CAP analysis of the field plants constrained to cultivation practice agreed with the PERMANOVA results that there are significant differences between microbial communities from organic and ecofarmed rice plants (6.9% of the variance,

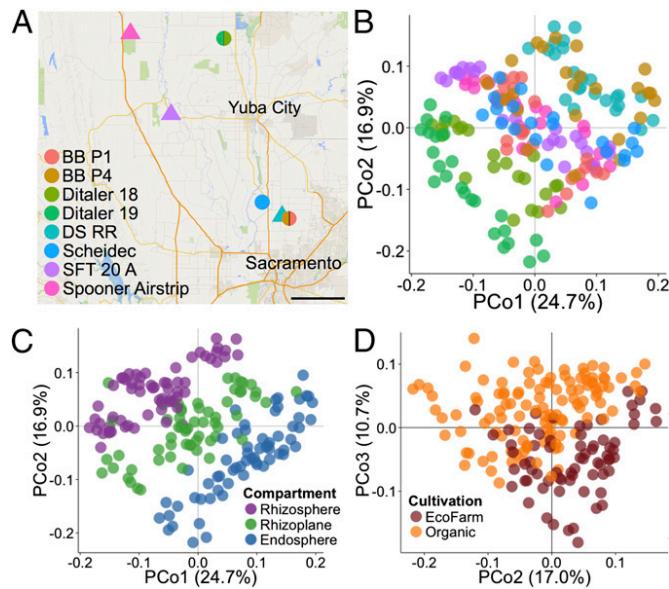


Fig. 4. Root-associated microbiomes from field-grown plants are separable by cultivation site, rhizospheric compartment, and cultivation practice. (A) Map depicting the locations of the field experiment collection sites across California's Central Valley. Circles represent organic-cultivated sites whereas triangles represent ecofarm-cultivated sites. (Scale bar, 10 mi.) (B) PCoA using the WUF method colored to depict the various sample collection sites. (C) Same PCoA in B colored by rhizospheric compartment. (D) Same PCoA in B and C depicting second and third axes and colored by cultivation practice.

$P = 0.005$, WUF, *SI Appendix*, Fig. S15D and 7.0% of the variance, UUF, $P = 0.005$, *SI Appendix*, Fig. S15H).

To better understand which taxa account for the separation between organic and ecofarmed cultivated plants, we used a generalized linear model with a negative binomial distribution to identify OTUs that had significantly different abundance between the two cultivation practices in each rhizocompartment (*SI Appendix*, Fig. S19 and Dataset S18). Notably, organic cultivation farms were enriched for Alphaproteobacteria, Actinobacteria, and Gemmatimonadetes whereas ecofarmed samples were enriched in Deltaproteobacteria, Chloroflexi, and Spirochaetes (*SI Appendix*, Fig. S20). We examined OTUs from genera that have potential implications in promoting plant growth and appear to depend upon cultivation practice. OTUs belonging to known nitrogen-fixing genera were enriched in organically cultivated samples including the cyanobacterium *Anabaena* and the Alphaproteobacterial genera *Azospirillum* and *Rhodobacter* (*SI Appendix*, Fig. S21A). Organically cultivated samples were also enriched for Actinobacterial OTUs belonging to the genus *Streptomyces* (*SI Appendix*, Fig. S21 B and C). Species within *Streptomyces* are known for their wide diversity of secondary metabolite production, many of which include antibiotic compounds (28). Microbial communities from ecofarmed rhizocompartments were enriched for OTUs that play a part in methane cycling belonging to the families *Syntrophorhabdaceae* (*SI Appendix*, Fig. S21 D and E) and *Syntrophaceae*. Members from both of these families are known for syntrophic growth with methanogenic archaea (see below), providing these archaea with H₂ and formate essential for methane production (29–31).

Microbes Involved in Methane Emissions. Global rice cultivation accounts for a significant proportion of annual greenhouse gas emissions, primarily methane (32). Methane (CH₄) emitted from rice paddies is mostly synthesized by methanogenic archaea. Although the majority of the CH₄ produced in rice paddy soil is oxidized by CH₄-using eubacteria (33), much of the remaining CH₄ diffuses through the soil into the plant roots and is expelled via the aerenchyma tissue of the rice plant itself (19). We found a large difference in abundance of methanogenic archaea between the field plants and the greenhouse plants. Methanogenic archaea were in much higher abundance in the field plants than in the greenhouse plants, indicating that either the greenhouse conditions are not a suitable growth habitat for methanogenic archaea or that the soils used for the greenhouse study had a low initial abundance of methanogens (*SI Appendix*, Fig. S22 A–D).

Methanogenic archaea require anaerobic conditions to function, and thus we did not expect methanogens to inhabit the endosphere or the rhizoplane. Surprisingly, the field plants exhibited high relative abundances of OTUs assigned to the genus *Methanobacterium* in the endosphere and rhizoplane that were comparable to or greater than the relative abundance of *Methanobacterium* in the rhizosphere (*SI Appendix*, Fig. S22A). This result is unexpected, given that the interior of rice roots is expected to have the highest levels of O₂ relative to other spatial compartments when grown in flooded soil. This pattern is not maintained with OTUs in other methanogenic genera including *Methanosarcina*, *Methanocella*, and *Methanosaeta*, however (*SI Appendix*, Fig. S22 B–D). Relative abundance of these genera in field plants was always highest in the rhizosphere, where the O₂ concentration is expected to be relatively lower. To further confirm the presence of methanogenic archaea in the endosphere, we used an alternative approach by PCR amplification of the methyl coenzyme-M reductase (*mcrA*) gene, a gene essential for CH₄ production (34), from endosphere samples. By cloning and sequencing the *mcrA* amplicons, we again found that sequences derived from *Methanobacteriaceae* were in higher relative abundance in the endosphere than in the rhizosphere of field plants, whereas *Methanosarcina* sequences exhibited the opposite distribution (Dataset S19).

The relatively oxygenated environment within the rice roots might be expected to be suitable for CH₄ oxidation by methanotrophs. We found that in general the methanotrophic bacteria followed the expected trend of distribution across the rhizocompartments, with the relative abundance lowest in the rhizosphere and highest in the endosphere (*SI Appendix*, Fig. S22 E and F). The field plants had higher abundances of methanotrophic bacteria compared with greenhouse plants, consistent with the low levels of methanogenic archaea and presumptive low levels of CH₄ production under greenhouse growth.

Identification of Coabundance Networks of OTUs. The sequence-based characterization of plant-associated microbiomes has primarily focused on how individual taxa within the microbiome associate with the host plant; however, the complex interactions that occur between taxa in the context of microbial communities are not revealed through this approach. Identification of microbial consortia is important for understanding their biological impact on the plant, but has been hindered by the inability to culture the vast majority of the microbes. It is likely that an important reason for the inability to grow many microbial species in pure culture can be attributed to consortium dynamics (35). Microbes cooperate in networks, supplying each other with critical nutrients for growth and survival. For example, methanogenic archaea cooperate with syntrophic partners to obtain H₂ and formate for CH₄ synthesis; hydrogen consumption by the methanogen allows the secondary fermentation process by the syntroph to become energetically favorable (29, 35).

We hypothesized that it might be possible to identify consortia that are involved in CH₄ cycling by generating OTU coabundance modules from the list of differentially abundant OTUs from the field experiment. We used an approach similar to gene coexpression network construction to generate the OTU coabundance modules. Briefly, pairwise Pearson correlations between OTUs determined to be differentially abundant between experimental factors included in the field experiment (10,848 OTUs; Datasets S16 and S17) were calculated and used as a distance metric for hierarchical clustering into a dendrogram, which was then dynamically pruned to form 284 coabundance modules (Dataset S20). The resulting modules were queried for clusters containing taxa that belonged to methanogenic archaea including *Methanobacterium*, *Methanosarcina*, *Methanocella*, and *Methanosaeta*. We identified 15 modules containing methanogenic OTUs with a correlation of 0.6 or greater to other OTUs within the same module (11 shown in Fig. 5A; Dataset S21). Modules containing methanogenic taxa were enriched for OTUs with methanotrophic, syntrophic, and CH₄ cycling potential (hypergeometric test, Dataset S22; Fig. 5B and *SI Appendix*, Fig. S23). Many of the remaining OTUs in the methanogenic modules have limited taxonomic information or functional information available, thus making functional inference difficult. Taken together, these results show that an OTU coabundance network approach is successful in generating associations that can recapitulate empirical data, and therefore might have predictive value for identifying novel microbial associations.

Acquisition of Microbiomes by the Root. We have shown that rice seedlings grown in field soil in the greenhouse acquire microbiomes that exhibit characteristics similar to those of field rice plants, in terms of the general distribution of phyla in the rhizocompartments. To understand the dynamics of microbiome acquisition from soil, we performed a time-series experiment. We transplanted steriley germinated seedlings of the *japonica* M104 cultivar into soil collected from the rice field in Davis in the greenhouse and collected samples at increasing intervals (0, 1, 2, 3, 5, 8, and 13 d). To monitor general changes in the soil microbial communities, we sampled from pots containing unplanted soil in the same container at each time point.

After collection, 16S rRNA gene sequencing of the samples from different compartments was performed as before. We used the proportion of microbial reads to organellar (plastidial

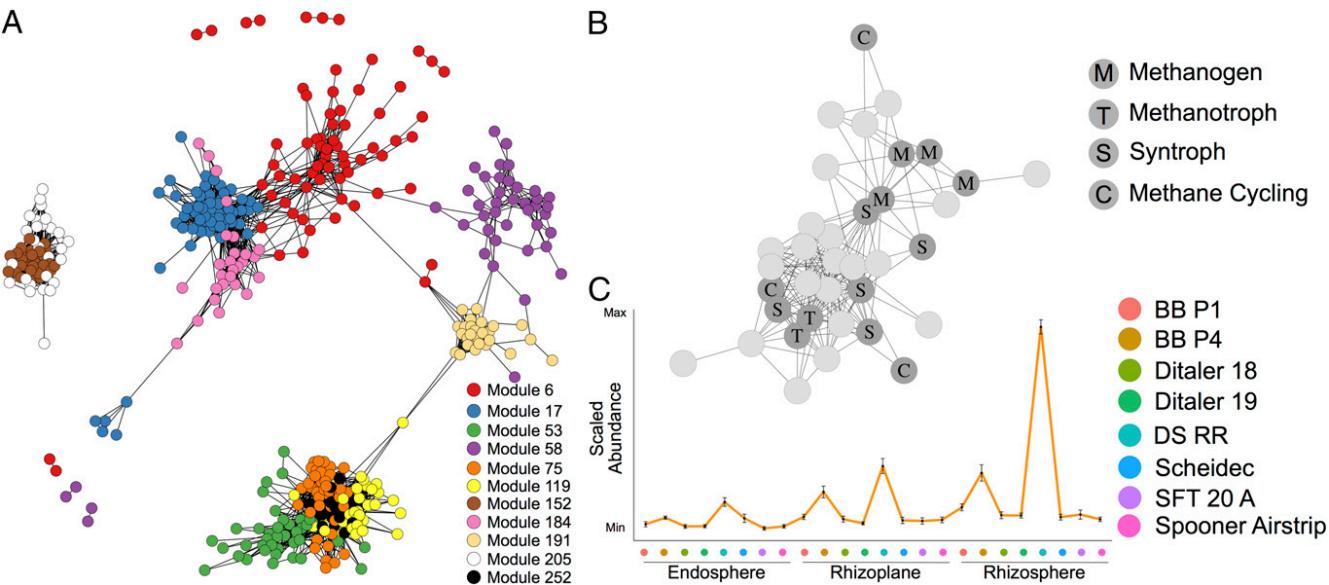


Fig. 5. OTU coabundance network reveals modules of OTUs associated with methane cycling. (A) Subset of the entire network corresponding to 11 modules with methane cycling potential. Each node represents one OTU and an edge is drawn between OTUs if they share a Pearson correlation of greater than or equal to 0.6. (B) Depiction of module 119 showing the relationship between methanogens, syntrophs, methanotrophs, and other methane cycling taxonomies. Each node represents one OTU and is labeled by the presumed function of that OTU's taxonomy in methane cycling. An edge is drawn between two OTUs if they have a Pearson correlation of greater than or equal to 0.6. (C) Mean abundance profile for OTUs in module 119 across all rhizocompartments and field sites. The position along the x axis corresponds to a different field site. Error bars represent SE. The x and y axes represent no particular scale.

and mitochondrial) reads to analyze microbial abundance in the endosphere over time (Fig. 6*A*). Using this technique, we confirmed the sterility of seedling roots before transplantation. We found that microbial penetration into the endosphere occurred at or before 24 h after transplantation and that the proportion of microbial reads to organellar reads increased over the first 2 wk after transplantation (Fig. 6*A*). To further support the evidence for microbiome acquisition within the first 24 h, we sampled root endospheric microbiomes from steriley germinated seedlings before transplanting into Davis field soil as well as immediately after transplantation and 24 h after transplantation (*SI Appendix*, Fig. S24). The root endospheres of steriley germinated seedlings, as well as seedlings transplanted into Davis field soil for 1 min, both had a very low percentage of microbial reads compared with organellar reads (0.22% and 0.71%), with the differences not statistically significant ($P = 0.1$, Wilcoxon test). As before, endospheric microbial abundance increased significantly, by >10-fold after 24 h in field soil (3.95%, $P = 0.05$, Wilcoxon test). We conclude that brief soil contact does not strongly increase the proportion of microbial reads, and therefore the increase in microbial reads at 24 h is indicative of endophyte acquisition within 1 d after transplantation.

α -Diversity significantly varied by rhizocompartment ($P < 2E-16$; Dataset S23) and there was a significant interaction between rhizocompartment and collection time ($P = 0.042$; Dataset S23); however, when each rhizocompartment was analyzed individually, the bulk soil was the only compartment that showed a significant amount of variation in α -diversity over time (*SI Appendix*, Fig. S25 and Dataset S23). The above results suggest that a diverse microbiota can begin to colonize the rhizoplane and endosphere as early as 24 h after transplanting into soil. We next evaluated how β -diversities shift over time in each rhizocompartment. We compared the time-series microbial communities with the previous greenhouse experiment microbial communities of M104 in Davis soil (Fig. 6 *B* and *C*). β -Diversity measurements of the time-series data indicated that microbiome samples from each compartment are separable by time. Furthermore, the rhizoplane and endosphere microbiomes from the later time point in the time-series data

(13 d) approach the endosphere and rhizoplane microbiome compositions for plants that have been grown in the greenhouse for 42 d.

There are slight shifts in the distribution of phyla over time; however, there are significant distinctions between the compartments starting as early as 24 h after transplantation into soil (Fig. 6*D*, *SI Appendix*, Figs. S24*B* and S26, and Dataset S24). Because each phylum consists of diverse OTUs that could exhibit very different behaviors during acquisition, we next examined the dynamics and colonization patterns of specific OTUs within the time-course experiment. The core set of 92 endosphere-enriched OTUs obtained from the previous greenhouse experiment (*SI Appendix*, Fig. S9*C*) was analyzed for relative abundances at different time points (Fig. 6*E*). Of the 92 core endosphere-enriched microbes present in the greenhouse experiment, 53 OTUs were detectable in the endosphere in the time-course experiment. The average abundance profile over time revealed a colonization pattern for the core endospheric microbiome. Relative abundance of the core endosphere-enriched microbiome peaks early (3 d) in the rhizosphere and then decreases back to a steady, low level for the remainder of the time points. Similarly, the rhizoplane profile shows an increase after 3 d with a peak at 8 d with a decline at 13 d. The endosphere generally follows the rhizoplane profile, except that relative abundance is still increasing at 13 d. These results suggest that the core endospheric microbes are first attracted to the rhizosphere and then locate to the rhizoplane, where they attach before migration to the root interior. To summarize, microbiome acquisition from soil appears to occur relatively rapidly, initiating within 24 h and approaching steady state within 14 d. The dynamics of accumulation suggest a multistep process, in which the rhizosphere and rhizoplane are likely to play key roles in determining the compositions of the interior and exterior components of the root-associated microbiome (*Discussion*).

Discussion

Factors Affecting the Composition of Root-Associated Microbiomes. The data presented here provide a characterization of the microbiome of rice, involving the combination of finer structural

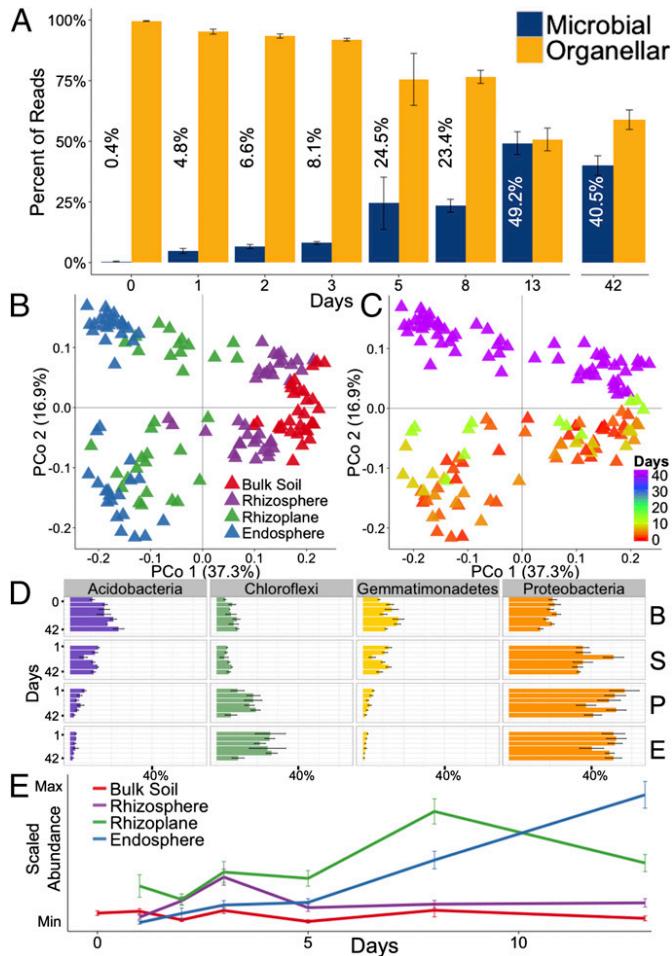


Fig. 6. Time-series analysis of root-associated microbial communities reveals distinct microbiome colonization patterns. (A) Ratios of microbial to organellar (plastidial and mitochondrial) 16S rRNA gene reads in the endosphere after transplantation into Davis soil. The 42-d time point corresponds to the earlier greenhouse experiment data (Fig. 1) subsetted to M104 in Davis soil. Mean percentages of the ratios are depicted with each bar. (B) PCoA of the time-series experiment and the greenhouse experiment subsetted to plants growing in Davis soil and colored by rhizospheric compartment. (C) The same PCoA as in B colored by collection day after transplantation into soil. (D) Average relative abundance for select phyla over the course of microbiome acquisition. (E) Average abundance profile of 53 out of the 92 core endosphere-enriched OTUs in each rhizospheric compartment. Error bars represent SE.

resolution and deeper sequencing than previous plant microbiome studies and using both controlled greenhouse and field studies covering a geographical range of cultivation. Specifically, we have been able to characterize in-depth the compositions of three distinct rhizocompartments—the rhizosphere, rhizoplane, and endosphere—and gain insights into the effects of external factors on each of these compartments. We note that a detailed characterization of plant rhizoplane microbiota in relation to the rhizosphere and the endosphere has not been previously attempted. To achieve this, we successfully adapted protocols for removal of rhizoplane microbes from the endosphere of *Arabidopsis* roots (9, 10). Because the fractional abundance of organellar reads in the rhizosphere, rhizoplane, and endosphere exhibits a clear increasing gradient (*SI Appendix*, Fig. S27), we hypothesize that we are isolating the rhizoplane fraction via disruption of the rhizodermis, consistent with direct EM observations on *Arabidopsis* roots following sonication (9, 10). The fine structure approach we have used combined with depth of sequencing allowed us to analyze over 250,000 OTUs, an order

of magnitude greater than in any single plant species to date. Under controlled greenhouse conditions, the rhizocompartments described the largest source of variation in the microbial communities sampled (*Dataset S5A*). The pattern of separation between the microbial communities in each compartment is consistent with a spatial gradient from the bulk soil across the rhizosphere and rhizoplane into the endosphere (Fig. 1C). Similarly, microbial diversity patterns within samples hold the same pattern where there is a gradient in α -diversity from the rhizosphere to the endosphere (Fig. 1B). Enrichment and depletion of certain microbes across the rhizocompartments indicates that microbial colonization of rice roots is not a passive process and that plants have the ability to select for certain microbial consortia or that some microbes are better at filling the root colonizing niche. Similar to studies in *Arabidopsis*, we found that the relative abundance of Proteobacteria is increased in the endosphere compared with soil, and that the relative abundances of Acidobacteria and Gemmatimonadetes decrease from the soil to the endosphere (9–11), suggesting that the distribution of different bacterial phyla inside the roots might be similar for all land plants (Fig. 1D and *Dataset S6*). Under controlled greenhouse conditions, soil type described the second largest source of variation within the microbial communities of each sample. However, the soil source did not affect the pattern of separation between the rhizospheric compartments, suggesting that the rhizocompartments exert a recruitment effect on microbial consortia independent of the microbiome source.

By using differential OTU abundance analysis in the compartments, we observed that the rhizosphere serves an enrichment role for a subset of microbial OTUs relative to bulk soil (Fig. 2). Further, the majority of the OTUs enriched in the rhizosphere are simultaneously enriched in the rhizoplane and/or endosphere of rice roots (Fig. 2B and *SI Appendix*, Fig. S16B), consistent with a recruitment model in which factors produced by the root attract taxa that can colonize the endosphere. We found that the rhizoplane, although enriched for OTUs that are also enriched in the endosphere, is also uniquely enriched for a subset of OTUs, suggesting that the rhizoplane serves as a specialized niche for some taxa. Conversely, the vast majority of microbes depleted in the rhizoplane are also depleted in the endosphere (Fig. 2C and *SI Appendix*, Fig. S16C), suggesting that the selectivity for colonization of the interior occurs at the rhizoplane and that the rhizoplane may serve an important gating role for limiting microbial penetration into the endosphere. It is important to note that the community structure we observe in each compartment is likely not simply caused by the plant alone. Microbial community structural differences between the compartments may be attributable also to microbial interactions involving both competition and cooperation.

In the case of field plants, we observed that the largest source of microbiome separation was due to cultivation site, rather than the spatial compartments (*Dataset S5 O* and *P*). These results are in contrast to the controlled greenhouse experiment where the soil effect was the second largest source of variation, suggesting the geography may be more important for determining the composition of the root microbiome than soil structure alone (*Dataset S5A*). These results differ from the results in the maize microbiome study, where microbial communities showed clear separation by state but not very much by geographic location within the same state (12). However, we note that in our study the locations within California were separated by distances of up to ~125 km, vs. a maximum separation of ~40 km in the intrastate locations of the maize study. Other factors that might account for the different results in our study include the number of field sites examined (eight, vs. three intrastate fields examined in the maize study), increased sequencing depth, different resolution because spatial compartments in maize roots were not separately analyzed, or possibly intrinsic differences between cultivated rice and maize.

Our design of the field experiment allowed us to test for cultivation practice effects on the rice root-associated microbiome,

specifically between organic cultivation and ecofarming, a more conventional cultivation practice. We found that cultivation practice described a significant amount of variation between the microbiomes, and that this effect was exhibited across every rhizospheric compartment (Fig. 4D and *SI Appendix*, Fig. S14D). These results are in contrast to the maize microbiome study, where an organically cultivated field had no significant separation from a conventionally cultivated field in the same state (12). Again, the difference in results between the two studies might be due to differences in sequencing depth and structural resolution or to intrinsic differences between the crops.

Microbial Consortia Associated with Methane Cycling in Rice Fields. Rice cultivation is a large contributor to agricultural CH₄ emissions due to the anaerobic conditions resulting from flooded fields. Currently, the model for CH₄ emissions under rice cultivation is that CH₄ is formed in the soil away from the roots and diffuses through the soil to the rice root, where it is expelled via the aerenchyma in the roots (19, 32, 36). We used our dataset to examine the abundances of taxa related to CH₄ cycling. As predicted, we observed that the relative abundance of some methanogens (*Methanocella*, *Methanosarcina*, and *Methanosaeta*) is higher in the rhizosphere than in the endosphere (*SI Appendix*, Fig. S22). However, the relative abundance of the methanogen *Methanobacterium* was found to be typically equal or higher in the endosphere than in the rhizosphere. This is an unexpected finding, because the root interior is relatively O₂-rich and therefore unfavorable for anaerobic archaea. How these archaea might survive within the root and whether they are contributing to CH₄ production at this interior location are questions that should be investigated. These results indicate that our understanding of rice-associated CH₄ production is still incomplete, and emphasize the need for more detailed studies to elucidate the different players and their roles in the process, perhaps using metagenomic and metatranscriptomic sequencing to better understand the metabolic potential of endosphere microbial communities.

An approach toward unraveling methanogen interactions with rice roots that we have explored is the construction of OTU coabundance networks. We hypothesized that microbes producing or using CH₄ might form consortia that would be revealed through this analysis. The sequencing depth and variation provided by the field experiment allowed us to test this approach by generating clusters of OTUs and identification of clusters containing methanogenic archaea. Inspection of individual OTUs within such clusters revealed additional taxonomies known to be involved in CH₄ cycling, such as methanotrophic, syntrophic, and CH₄-cycling eubacteria (Fig. 5 and *SI Appendix*, Fig. S23). The clusters also contained OTUs with no known role in CH₄ cycling. Possibly these OTUs represent novel players that have been revealed through this approach, although larger-scale studies with additional field sites will be needed to validate the significance of their associations. Even with this limited scale, the results of this analysis are very encouraging. This approach might be generalizable for culture-independent identification of other microbial consortia involved in different types of geochemical cycles that interact with plant roots. However, this approach might be more applicable to those microbial lineages that are tightly linked to certain traits, such as CH₄ cycling, and might not be successful when a trait is much more widespread independent of lineage, such as nitrogen fixation.

A Multistep Model for Microbiome Acquisition. Studies on mammals have established that the fetus at birth already has a microbiome derived from the mother, and that diet and other environmental factors further shape the gut microbiome (37–39). For plants, the process of microbiome acquisition is not well-understood (7). Microbiome transmission through the embryo seems unlikely, because surface sterilization of the seeds resulted in axenic plants (Fig. 6A and *SI Appendix*, Fig. S23). Our success in obtaining in-depth sequence coverage for all three different spatial compartments in the root provided us with an opportunity to address

this question by carrying out a time-staged profiling of the microbial compositions of the compartments during the early stages of microbiome acquisition. We find that microbiome acquisition from soil is rapid, namely that rice plants begin to assemble an endospheric microbiome within a day after transplantation from sterile media to soil, and the relative level approaches steady state within 2 wk. The 13-d endosphere and rhizoplane microbiomes were most similar to the older, 42-d, microbiomes from the previous greenhouse experiment that we consider to represent steady state, as the latter microbiomes are from well-established vegetatively growing plants (Fig. 6B and C). Because separate batches of soil were used between the greenhouse and the acquisition experiment, it is important to note that the microbiomes may not be directly comparable across different experiments, due to seasonal variation between soil batches (9–11). The patterns of relative abundance in phyla between the compartments (e.g., higher proportions of Acidobacteria in the rhizosphere than endosphere and lower proportions of Proteobacteria and Chloroflexi in the rhizosphere than endosphere) are evident within 24 h after the rice plants were transplanted into soil (Fig. 6C and *SI Appendix*, Figs. S24B and S26). Based on studies of the steady-state microbiomes of the rhizosphere and endosphere, it has been proposed that plants might assemble their microbiomes in two steps, with the first step involving a general recruitment to the vicinity of the root and a second step for entry inside the root that involves species-specific genetic factors (7). The dynamics of microbiome acquisition in our study provide experimental support for this model, in that the general concept of recruitment and assembly as separate steps is consistent with our data. We observed an initial enrichment in the rhizosphere, consistent with the attraction of a diverse set of microbes to the vicinity of the plant, followed by slower rates of accumulation of OTUs in the endosphere (Fig. 6E). The selectivity of the latter process is also implied by the extensive depletion of rhizospheric OTUs in the endosphere. However, when rhizoplane microbiome data are taken into consideration, microbiome acquisition in the root appears to be a more complex multistep process, in which the rhizoplane plays a key role. After initial recruitment to the rhizosphere, only a subset of these microbes initially recruited to the rhizosphere are bound to the root surface at the rhizoplane, suggesting selectivity for direct physical association with the root. This selection may occur by the plant, or may occur through the ability to form biofilms, as certain microbes are known to form biofilms along root surfaces (40, 41). A further selective step must operate to account for the additional depletion of rhizoplane OTUs from the endosphere at steady state, implying that binding is not sufficient for entry. Binding at the rhizoplane might, however, act as a necessary prerequisite for endospheric OTUs, as the time course shows that accumulation of OTUs at the rhizoplane precedes that in the endosphere. We suggest that the rhizoplane serves a critical gating role; of the microbes that are attracted to the rhizosphere, only a subset can bind the rhizoplane, and a fraction of these are permitted to enter and proliferate in the endosphere. Each of these steps likely involves molecular signals from the plant, presumptively components of root exudates and possibly cell-wall components or membrane proteins. The signals could consist of general plant metabolites as well as species- and genotype-specific molecules. In the model proposed by Bulgarelli et al. (7), genotype signals are proposed to regulate entry into the endosphere. However, in our study, the rhizosphere composition showed the largest variation in the comparison of different cultivars, suggesting that the genotypic factors appear to also act at the level of general microbial recruitment. In summary, the dynamic shifts in microbiomes across each compartment indicate that the actions of three or more selective steps at distinct spatial locations from the root interior, and responding to multiple signals from the plant, coordinate the assembly of the root microbiome.

Materials and Methods

Microbial communities from the rhizosphere, rhizoplane, and root endosphere of various rice cultivars grown in the greenhouse and field were profiled by amplifying fragments of the 16S rRNA gene and sequenced using the Illumina MiSeq platform. The sequence analysis was carried out using the QIIME pipeline (42) and all statistical analyses were performed using R v3.1.0 (43). Further details of materials and methodology are explained in *SI Appendix, Materials and Methods*. The raw sequencing reads for this project were submitted to the National Center for Biotechnology Information Short Read Archive under accession no. SRP044745. The custom scripts used for analyzing the data presented here can be found at github.com/ricemicromicrobiome/edwards-et-al.-2014.

1. Tringe SG, et al. (2005) Comparative metagenomics of microbial communities. *Science* 308(5721):554–557.
2. Zhang H, et al. (2009) A soil bacterium regulates plant acquisition of iron via deficiency-inducible mechanisms. *Plant J* 58(4):568–577.
3. Long SR (1989) Rhizobium-legume nodulation: Life together in the underground. *Cell* 56(2):203–214.
4. Bolan N (1991) A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants. *Plant Soil* 134(2):189–207.
5. Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* 57: 233–266.
6. Berendsen RL, Pieterse CM, Bakker PA (2012) The rhizosphere microbiome and plant health. *Trends Plant Sci* 17(8):478–486.
7. Bulgarelli D, Schlaepff K, Spaepen S, Ver Loren van Themaat E, Schulze-Lefert P (2013) Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* 64: 807–838.
8. Mendes R, et al. (2011) Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* 332(6033):1097–1100.
9. Bulgarelli D, et al. (2012) Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488(7409):91–95.
10. Lundberg DS, et al. (2012) Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488(7409):86–90.
11. Schlaepff K, Dombrowski N, Oter RG, Ver Loren van Themaat E, Schulze-Lefert P (2014) Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proc Natl Acad Sci USA* 111(2):585–592.
12. Peiffer JA, et al. (2013) Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci USA* 110(16):6548–6553.
13. Gottel NR, et al. (2011) Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. *Appl Environ Microbiol* 77(17):5934–5944.
14. Shakya M, et al. (2013) A multifactor analysis of fungal and bacterial community structure in the root microbiome of mature *Populus deltoides* trees. *PLoS ONE* 8(10): e76382.
15. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO (2007) Development of the human infant intestinal microbiota. *PLoS Biol* 5(7):e177.
16. Dominguez-Bello MG, et al. (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci USA* 107(26):11971–11975.
17. Sessitsch A, et al. (2012) Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Mol Plant Microbe Interact* 25(1):28–36.
18. Knief C, et al. (2012) Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J* 6(7):1378–1390.
19. Neue H-U (1993) Methane emission from rice fields. *Bioscience* 43(7):466–474.
20. McMurdie PJ, Holmes S (2014) Waste not, want not: Why rarefying microbiome data is inadmissible. *PLOS Comput Biol* 10(4):e1003531.
21. Im W-T, Kim S-H, Kim MK, Ten LN, Lee S-T (2006) *Pleomorphomonas koreensis* sp. nov., a nitrogen-fixing species in the order *Rhizobiales*. *Int J Syst Evol Microbiol* 56(Pt 7):1663–1666.
22. Madhaiyan M, et al. (2013) *Pleomorphomonas diazotrophica* sp. nov., an endophytic N-fixing bacterium isolated from root tissue of *Jatropha curcas* L. *Int J Syst Evol Microbiol* 63(Pt 7):2477–2483.
23. Xie C-H, Yokota A (2005) *Pleomorphomonas oryzae* gen. nov., sp. nov., a nitrogen-fixing bacterium isolated from paddy soil of *Oryza sativa*. *Int J Syst Evol Microbiol* 55(Pt 3):1233–1237.
24. Ransom-Jones E, Jones DL, McCarthy AJ, McDonald JE (2012) The *Fibrobacteres*: An important phylum of cellulose-degrading bacteria. *Microb Ecol* 63(2):267–281.
25. Leschine SB (1995) Cellulose degradation in anaerobic environments. *Annu Rev Microbiol* 49:399–426.
26. Kim S-I, Tai TH (2013) Identification of SNPs in closely related *temperate japonica* rice cultivars using restriction enzyme-phased sequencing. *PLoS ONE* 8(3):e60176.
27. Lundberg Family Farms (2014) Organic vs. eco-farmed. Available at www.lundberg.com/commitment/ecofarmed.aspx. Accessed July 14, 2014.
28. Wave MG, Tickoo R, Jog MM, Bhole BD (2001) How many antibiotics are produced by the genus *Streptomyces*? *Arch Microbiol* 176(5):386–390.
29. Stams AJ, Plugge CM (2009) Electron transfer in syntrophic communities of anaerobic bacteria and archaea. *Nat Rev Microbiol* 7(8):568–577.
30. Gray ND, et al. (2011) The quantitative significance of *Syntrophaceae* and syntrophic partnerships in methanogenic degradation of crude oil alkanes. *Environ Microbiol* 13(11):2957–2975.
31. Schink B (1997) Energetics of syntrophic cooperation in methanogenic degradation. *Microbiol Mol Biol Rev* 61(2):262–280.
32. Conrad R (2009) The global methane cycle: Recent advances in understanding the microbial processes involved. *Environ Microbiol Rep* 1(5):285–292.
33. Holzapfel-Pschorn A, Conrad R, Seiler W (1985) Production, oxidation and emission of methane in rice paddies. *FEMS Microbiol Lett* 31(6):343–351.
34. Thauer RK (1998) Biochemistry of methanogenesis: A tribute to Marjory Stephenson. *Microbiology* 144(Pt 9):2377–2406.
35. McInerney MJ, Sieber JR, Gunsalus RP (2009) Syntropy in anaerobic global carbon cycles. *Curr Opin Biotechnol* 20(6):623–632.
36. Lee HJ, Kim SY, Kim PJ, Madsen EL, Jeon CO (2014) Methane emission and dynamics of methanotrophic and methanogenic communities in a flooded rice field ecosystem. *FEMS Microbiol Ecol* 88(1):195–212.
37. Gueimonde M, et al. (2006) Effect of maternal consumption of *Lactobacillus GG* on transfer and establishment of fecal bifidobacterial microbiota in neonates. *J Pediatr Gastroenterol Nutr* 42(2):166–170.
38. Vaishampayan PA, et al. (2010) Comparative metagenomics and population dynamics of the gut microbiota in mother and infant. *Genome Biol Evol* 2:53–66.
39. Koenig JE, et al. (2011) Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci USA* 108(Suppl 1):4578–4585.
40. Bais HP, Fall R, Vivanco JM (2004) Biocontrol of *Bacillus subtilis* against infection of *Arabidopsis* roots by *Pseudomonas syringae* is facilitated by biofilm formation and surfactin production. *Plant Physiol* 134(1):307–319.
41. Walker TS, et al. (2004) *Pseudomonas aeruginosa*-plant root interactions. Pathogenicity, biofilm formation, and root exudation. *Plant Physiol* 134(1):320–331.
42. Caporaso JG, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5):335–336.
43. R Core Team (2012) R: A Language and Environment for Statistical Computing (R Found Stat Comput, Vienna).

ACKNOWLEDGMENTS. We thank Lance Benson and Jessica Lundberg of Lundberg Family Farms; George Tibbets and Michael Lear for generously providing samples and access to their field facilities; Derek Lundberg and Jeffery Dangl (The University of North Carolina) for sharing their protocol for root sonication prior to publication; Kelsey Galimba, John Jaeger, Cassandra Ramos, and Paul Fisher for technical assistance at the early stages of the project; Gurdev Khush and Kate Scow for valuable advice; and members of the J.A.E. and V.S. laboratories for helpful discussions. J.A.E. and V.S. acknowledge the support of National Science Foundation Awards DBI-0923806 and IOS-1444974; J.E. acknowledges partial support from the Henry A. Jastro Graduate Research Award; and C.S.-M. acknowledges support from the University of California Institute for Mexico (UCMEXUS)/Consejo Nacional de Ciencia y Tecnología (CONACYT) and Secretaría de Educación Pública (Mexico).

1 **Supplementary Information**

2 Edwards et al. PNAS

3

4 **Soil collection for greenhouse experiment and microbiome acquisition experiment**

5 Soil from the rice field in Sacramento (38.58575 degrees north and -121.596911 West) was
6 collected on 3/15/2013 using shovels to gather down to a depth of approximately 8 inches. Soil from the
7 rice field in Arbuckle, CA (39.011732 degrees North and -121.92212 degrees West) was collected on
8 3/18/2013 using a front-end loader to gather down to a depth of approximately 8 inches. Soil from a rice
9 field in Davis, CA (38.543864 degrees North and -121.81223 West) was collected on 3/19/2013 using
10 shovels to gather down to a depth of approximately 8 inches. All soils were transported back to the
11 greenhouse and stored until planting on 3/28/2013. All soils were mixed individually in clean tubs to
12 homogenize the soil. The soil was placed into new 5 x 5 inch pots that were then placed into tubs (24
13 pots each). Each tub contained only one soil type in order to avoid microbial mixing between the soils.
14 Each tub was watered in order to submerge the soils as suited to rice cultivation. Soil from the Davis field
15 was collected for the microbiome acquisition experiment on 11/26/2013 using the same method as
16 described above. Soil samples for the Arbuckle, Davis, and Sacramento fields were analyzed at the UC
17 Davis Analytical Lab for chemical content (Dataset S8).

18 **Plant germination, transplantation, and cultivation in the greenhouse and microbiome acquisition
19 experiment**

20 Seeds from 6 cultivated varieties (M104, Nipponebare, IR50, 93-11, TOg 7102, and TOg 7267)
21 were dehulled, surface sterilized in 70% bleach for 5 minutes and steriley germinated on MS agar media
22 in the dark. After germination, the rice seedlings were transplanted into the various soils in the
23 greenhouse. The tubs were watered every other day and nutrients were supplied to each tub on 2-week
24 basis on 4/12/2013 and 4/26/2013. All weeds were manually removed from the pots when identified.

25 For the microbiome acquisition experiment, M104 seeds were dehulled and surface sterilized in
26 bleach for 5 minutes and subsequently germinated on MS agar media in the dark. The seedlings were
27 transplanted into Davis soil in the greenhouse and sampled according to the time series using the same
28 protocol for sample collection detailed above.

29 **Experimental design for greenhouse and microbiome acquisition experiments**

30 The greenhouse experiment was designed as a split-split plot experiment. Briefly, there were 12
31 tubs total so that each soil had 4 tubs. We collected only one rhizocompartment from each pot such that
32 each every rhizocompartment was taken from every cultivar once per tub, giving a total of 18
33 rhizocompartment samples and 6 bulk soil samples per tub (Dataset S1). Because the selected cultivars
34 flower at various times, to avoid confounding issues between developmental stages and cultivar effects
35 we collected all samples at 42 days while all cultivars were still vegetatively growing.

36 The plants for the acquisition experiment were all contained in one large tub along with unplanted
37 pots for bulk soil controls. Each plant collected had all three rhizocompartments sampled.

38 Experimental design of the field experiment

39 All fields sampled in the field experiment are managed by Lundberg Family Farms (Richvale,
40 CA, USA). All of these fields are subject to typical California rice cultivation practices (presoaked seeds,
41 aerial seeding, dense planting, etc), with the cultivation differences being between “eco farming” and
42 organic farming of the fields. 8 individual plants were sampled per field site (Dataset S1).

43 **Sample Collection of Rhizosphere, Rhizoplane, and Endosphere Fractions**

44 Samples were collected over a 4-day period from 5/6/2013 to 5/9/2013. The soil and plant were
45 removed from each pot and the roots were removed from the soil. We avoided collecting any roots that
46 were at the interface of the pot and the soil in order to avoid false environments. The excess soil was
47 manually shaken from the roots, leaving approximately 1mm of soil still attached to the roots (Fig. S2).
48 We separated the 1mm of soil from the roots directly in the greenhouse by placing the roots with soil still
49 attached in a sterile flask with 50 ml of sterile Phosphate Buffered Saline (PBS) solution. The roots were
50 then stirred vigorously with sterile forceps in order to clean all the soil from the root surfaces. The soil

51 that was cleaned from the roots was poured into a 50ml Falcon tube and stored as the rhizosphere
52 compartment at 4°C until DNA extraction the same day.

53 The roots designated for rhizoplane collection were cleaned in the greenhouse and placed in a
54 Falcon tube with 15 ml PBS, and tightly adhering microbes at the root surface were removed using a
55 sonication protocol originally developed for Arabidopsis roots (1-3). The roots in the Falcon tube were
56 sonicated for 30 s at 50-60 Hz (output frequency 42 kHz, power 90 W, Branson Unltrasonics). The
57 sonication procedure strips the rhizoplane microbes from the root surface as well as portions of the
58 rhizodermis as evidenced by the gradient of organellar reads from the rhizoplane to the endosphere (Fig.
59 S27). The roots were then removed and discarded and the liquid PBS fraction was kept as the rhizoplane
60 compartment.

61 The roots designated for the endosphere collection were cleaned and sonicated as described
62 before. Two more sonication procedures using clean PBS solution were used to ensure that all microbes
63 were removed from the root surface. CARD-FISH on whole non-sonicated roots and thrice sonicated
64 roots was used to analyze the efficacy of this procedure for removing microbes from the rhizoplane (Fig.
65 S3). The sonicated roots were then stored at -80°C until DNA extraction the same day.

66 Bulk soil samples were collected from unplanted pots approximately 2 inches below the soil
67 surface. The samples were placed in 15 ml tubes and stored at 4°C until DNA extraction the same day.

68 Samples for the field experiment were collected over a two-day period. The roots of plants in the
69 field were collected with a bulb planter (Fiskars). The soil was shaken off the roots to leave ~1mm of soil
70 still attached. These roots were placed in sterile PBS solution and brought back to the laboratory for
71 isolation of the rhizocompartments as described above. Each rhizocompartment was isolated from each
72 plant sampled and had total DNA extracted.

73 **DNA Extraction from Rhizocompartments**

74 The rhizosphere soil was concentrated by pipetting 1mL of the PBS / rhizosphere soil into a 2mL
75 tube and centrifuging for 30 seconds at 10,000 g. The supernatant was discarded leaving only the soil
76 fraction behind. The rhizoplane compartment was concentrated in the same manner, except all 15mL of

77 the sample was concentrated in the same 2 mL tube using multiple centrifugations. The endosphere
78 fraction was pre-homogenized before the DNA extraction by bead beating for 1 minute (Mini Beadbeater,
79 Biospec Products). The DNA for each sample was then extracted using the MoBio PowerSoil DNA
80 isolation kit and eluted in 50 µL of elution buffer. The rhizoplane samples typically had low DNA yield
81 and were subsequently concentrated in a speedvac down to 10 µL.

82 **16S rRNA gene V4 amplification, quantitation, and sequencing**

83 Targeted metagenomic profiling of the samples was carried out by sequencing the V4 region of
84 the 16S rRNA gene. V4 amplification was carried out using primers modified from Caporaso et al, 2010
85 (4). Briefly, these primers are designed to amplify from 515 to 806 of the 16S rRNA gene and they
86 include a barcode an adaptor for annealing to the Illumina flow cell. Our primers differed in that both the
87 primers contained a 12bp barcode instead of only the reverse primer (Dataset S25). This allowed us to
88 pool many samples together using unique barcode combinations instead of relying on a multitude of
89 reverse primers with unique barcodes. PCR reaction mixes were made using Qiagen HotStar HiFidelity
90 polymerase. Each mix was done in a volume of 25 µL using 14 µL H₂O, 5 µL HotStar PCR Buffer, 2.5
91 µL forward primer (10 µM), 2.5 µL reverse primer (10 µM), 1 µL sample DNA, and 0.5 µL HotStar
92 polymerase. We used a touchdown PCR program on a Biometra TProfessional Basic Gradient
93 thermocycler: 95°C for 5 min, then 7 cycles of 95°C for 45 sec, 65°C for 1 min (decreasing at 2°C /
94 cycle), and 72°C for 90 sec, followed by 30 cycles of 95°C for 45 sec, 50°C for 30 sec, and 72°C for 90
95 sec. A final extension at 72°C was used for 10 min and the reactions were held at 4°C. The reactions
96 were run on a 1% agarose gel in order to ensure the amplification was successful. Unsuccessful reactions
97 were attempted once more, but removed from the experiment if unsuccessful a second time.

98 The amplicons libraries were diluted 40x and quantified using an Agilent Bioanalyzer for the
99 greenhouse libraries, or a Caliper LabChip GX for the field experiment libraries at the DNA Technologies
100 Core at the Genome Center, UC Davis. The libraries were then pooled at equimolar concentrations into 4
101 pooled libraries (2 libraries for the greenhouse experiment and 2 libraries for the field experiment). To
102 remove any primer dimer from the pooled amplicon libraries we ran the 4 pooled libraries on 1.8%

103 agarose gels and extracted a 400 bp band. The bands were purified (Macherey-Nagel Nucleospic Gel and
104 PCR Cleanup kit) and bioanalyzed as a final quality control check. Each library was submitted to the UC
105 Davis DNA Technologies core for 250 x 250 paired end, dual index sequencing on an Illumina MiSeq
106 instrument.

107 **Sequence Analysis**

108 The sequences obtained from the MiSeq runs were demultiplexed based on the barcode sequences
109 using a custom Perl script based upon exact matching. The sequences were overlapped to form
110 contiguous reads using MOTHUR's command make.contigs (5). Reads containing any ambiguous bases
111 were then discarded along with any reads that were over 275 bp. The sequences were then clustered into
112 operational taxonomic units (OTUs) by UCLUST (6) based on 97% pairwise identity using QIIME's (7)
113 open reference OTU picking strategy which used the Greengenes 16S rRNA database (13_5 release) as a
114 reference (8). Taxonomic classification of the representative sequence for each OTU was done using
115 QIIME's version of the Ribosomal Database Project's classifier (9) against the Greengenes 16S rRNA
116 database (13_5 release) using default parameters. All OTUs identified as belonging to chloroplast and
117 mitochondria were removed from the data set. The representative sequences for each OTU were aligned
118 using PyNAST (10) in QIIME. Chimeric OTUs were identified using QIIME's implementation of
119 ChimeraSlayer (11) and removed from the OTU table and OTU representative sequences file. A
120 phylogenetic tree was generated from the alignment file by FastTree (12).

121 **Statistical Analysis**

122 The resulting OTU table was divided by experiment and analyzed separately except when
123 comparing methanogenic and methanotrophic OTUs. Low abundance OTUs were eliminated from the
124 OTU table if they did not have a total of at least 5 counts across all the samples in the experiment.. OTU
125 tables for each experiment were normalized by the trimmed mean of M values (TMM) method using the
126 BioConductor package EdgeR in R (13). Weighted and Unweighted UniFrac (14) distances were
127 calculated from the normalized OTU tables for each experiment. α -diversity measurements were
128 calculated by the function diversity() using the "Shannon" method in the R package Vegan (15).

129 Rarefaction curves were calculated using custom R scripts. Principal coordinate analyses utilizing the
130 weighted and unweighted UniFrac distances were calculated using the pcoa() function from the R
131 package Ape (16). CAP analysis was performed using the function capscale() from the R Package Vegan.
132 When specifying CAP models, we constrained the analysis to the factor of interest while controlling for
133 all other experimental factors and technical factors (MiSeq runs). Variance partitioning and significances
134 for experimental factors was performed by running Vegan's permutest() function over the CAP model
135 using a maximum of 500 permutations. Bulk soil samples were omitted from the CAP analysis when
136 analyzing the greenhouse data. This was done because the bulk soil samples provided a confounding
137 level within the *Cultivar* factor. Additionally, permutational MANOVA was carried out to using Vegan's
138 function adonis() to measure effect size and significances on β -diversity. Differentially abundant OTUs
139 were detected using EdgeR's generalized linear model (GLM) approach. This approach allows the user to
140 test for differential OTU abundance between different levels of factors by employing a design matrix to
141 account for complex experimental designs.

142 **Co-abundance network analysis**

143 Only OTUs that were determined to be differentially abundant in experimental factors
144 encompassed in the field experiment were used for network analysis, thus subsetting the data to OTUs
145 with high variance (10,848 OTUs). Pairwise Pearson correlations were calculated between the remaining
146 OTUs. The Pearson correlations were used as a distance metric to build a hierarchically clustered
147 dendrogram using average linkage. The dendrogram was dynamically pruned using the R package
148 'dynamicTreeCut' (17). This tree cutting technique was employed due to its ability to detect nested
149 clusters within larger clusters. A hypergeometric test was used to detect taxonomies that were
150 significantly enriched in specific clusters. Taxonomies were queried for their involvement in methane
151 metabolism and cycling using the BioCyc (18), MetaCyc (18), or KEGG Pathway (19) databases unless
152 otherwise noted.

153 **CARD-FISH**

154 Roots designated for CARD-FISH were fixed using 4% formaldehyde in PBS for 4 hours,

155 washed twice with PBS, and stored in 1:1 ethanol:PBS at 4° C. CARD-FISH treatments were done in
156 accordance with previous studies in *Arabidopsis* (2, 3, 20) using the eubacterial probe Eub338 (5'-
157 GCTGCCTCCCGTAGGAGT-3', 35% formamide, Biomers Ulm, Germany) and its nonsense sequence
158 as a negative control, NON338 (5' - ACTCCTACGGGAGGCAGC-3', 30% formamide) labeled with
159 horseradish peroxidase at the 5' end (Biomers Ulm, Germany). Signal amplification was carried out
160 using fluorescently labeled tyramide (Fluorescent solutions). All microscopy images were taken using a
161 confocal laser scanning microscope in the Department of Plant Biology at UC Davis (Zeiss LSM 710).

162 **Amplification, cloning, and sequencing of *mcrA*.**

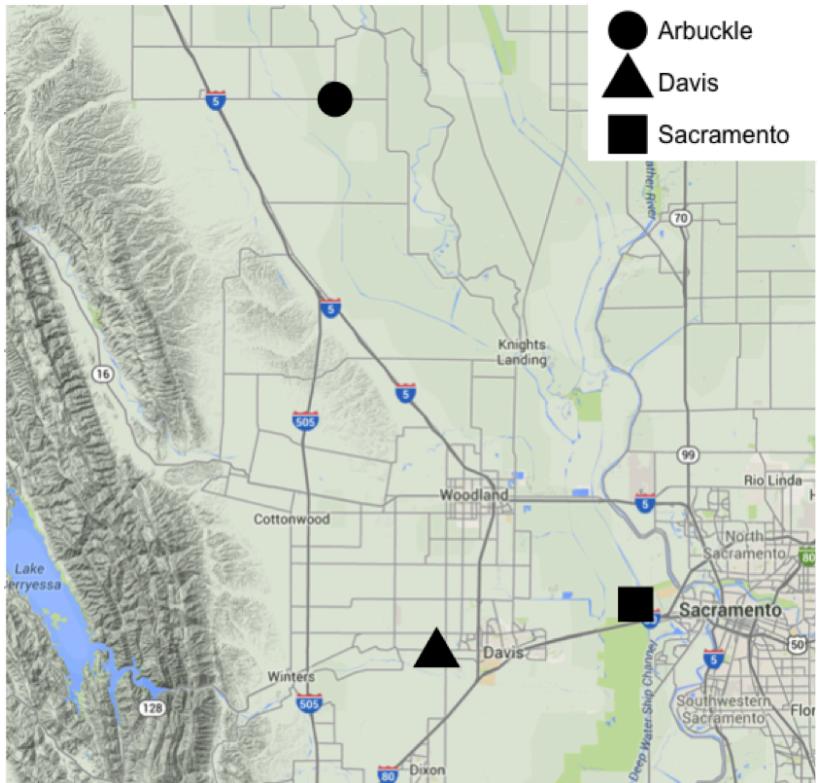
163 Total community DNA extracted from rhizosphere and endosphere samples from the DS RR field
164 was used as a template to amplify fragments of the *mcrA* gene. PCR was performed following the
165 protocol described in Juottonen et al. (21) using the primers designed in Springer et al. (22) The amplified
166 products were cloned with the TOPO TA cloning kit (Invitrogen), and plasmid DNA was recovered from
167 47 clones (29 from endosphere samples and 18 from rhizosphere samples) using GeneJET plasmid
168 miniprep kit (Thermo Scientific). The cloned fragments were sequenced by the UC Davis Sequencing
169 Facility using the M13 primers. A BLAST search was performed using the NCBI nucleotide database,
170 and the top alignment was reported for each sequence (alignments without a defined taxonomy were
171 excluded).

172 **References**

- 173 1. Bulgarelli D, Schlaepf K, Spaepen S, van Themaat EVL, & Schulze-Lefert P (2013) Structure
174 and functions of the bacterial microbiota of plants. *Annual review of plant biology*
175 64:807-838.
- 176 2. Bulgarelli D, et al. (2012) Revealing structure and assembly cues for *Arabidopsis* root-
177 inhabiting bacterial microbiota. *Nature* 488(7409):91-95.
- 178 3. Lundberg DS, et al. (2012) Defining the core *Arabidopsis thaliana* root microbiome. *Nature*
179 488(7409):86-90.
- 180 4. Caporaso JG, et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of
181 sequences per sample. *Proceedings of the National Academy of Sciences* 108(Supplement
182 1):4516-4522.
- 183 5. Schloss PD, et al. (2009) Introducing mothur: open-source, platform-independent, community-
184 supported software for describing and comparing microbial communities. *Applied and
185 environmental microbiology* 75(23):7537-7541.

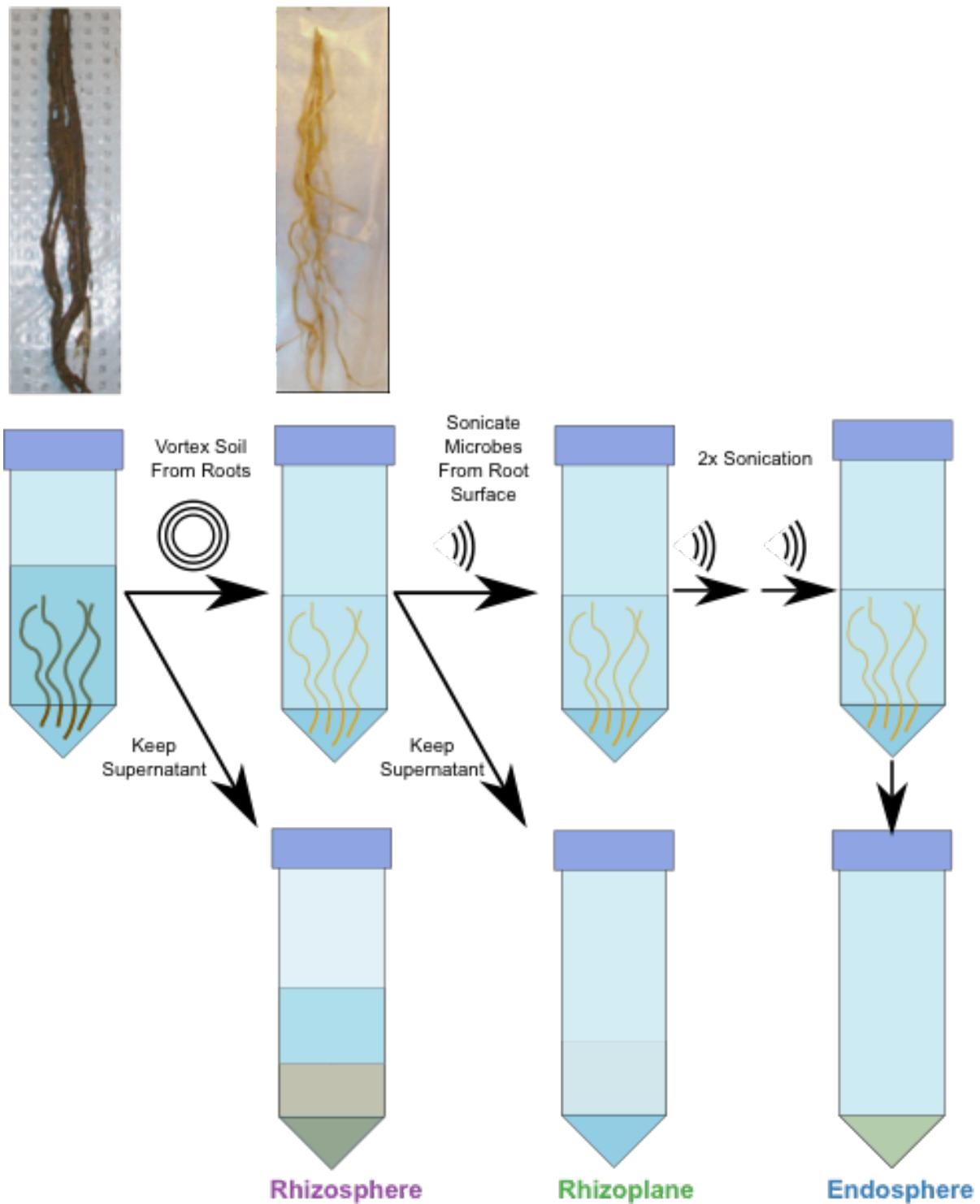
- 186 6. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST.
187 *Bioinformatics* 26(19):2460-2461.
- 188 7. Caporaso JG, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing
189 data. *Nature methods* 7(5):335-336.
- 190 8. DeSantis TZ, *et al.* (2006) Greengenes, a chimera-checked 16S rRNA gene database and
191 workbench compatible with ARB. *Applied and environmental microbiology* 72(7):5069-
192 5072.
- 193 9. Cole JR, *et al.* (2009) The Ribosomal Database Project: improved alignments and new tools
194 for rRNA analysis. *Nucleic acids research* 37(suppl 1):D141-D145.
- 195 10. Caporaso JG, *et al.* (2010) PyNAST: a flexible tool for aligning sequences to a template
196 alignment. *Bioinformatics* 26(2):266-267.
- 197 11. Haas BJ, *et al.* (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and
198 454-pyrosequenced PCR amplicons. *Genome research* 21(3):494-504.
- 199 12. Price MN, Dehal PS, & Arkin AP (2009) FastTree: computing large minimum evolution trees
200 with profiles instead of a distance matrix. *Molecular biology and evolution* 26(7):1641-
201 1650.
- 202 13. Robinson MD, McCarthy DJ, & Smyth GK (2010) edgeR: a Bioconductor package for
203 differential expression analysis of digital gene expression data. *Bioinformatics* 26(1):139-
204 11.
- 205 14. Lozupone C & Knight R (2005) UniFrac: a new phylogenetic method for comparing
206 microbial communities. *Applied and environmental microbiology* 71(12):8228-8235.
- 207 15. Oksanen J, *et al.* (2007) The vegan package. *Community ecology package*.
- 208 16. Paradis E, Claude J, & Strimmer K (2004) APE: analyses of phylogenetics and evolution in R
209 language. *Bioinformatics* 20(2):289-290.
- 210 17. Langfelder P, Zhang B, & Horvath S (2008) Defining clusters from a hierarchical cluster tree:
211 the Dynamic Tree Cut package for R. *Bioinformatics* 24(5):719-720.
- 212 18. Caspi R, *et al.* (2008) The MetaCyc Database of metabolic pathways and enzymes and the
213 BioCyc collection of Pathway/Genome Databases. *Nucleic acids research* 36(suppl
214 1):D623-D631.
- 215 19. Kanehisa M & Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic
216 acids research* 28(1):27-30.
- 217 20. Eickhorst T & Tippkötter R (2008) Improved detection of soil microorganisms using
218 fluorescence *in situ* hybridization (FISH) and catalyzed reporter deposition (CARD-
219 FISH). *Soil Biology and Biochemistry* 40(7):1883-1891.
- 220 21. Juottonen H, Galand PE, & Yrjälä K (2006) Detection of methanogenic Archaea in peat:
221 comparison of PCR primers targeting the *mcrA* gene. *Research in microbiology*
222 157(10):914-921.
- 223 22. Springer E, Sachs MS, Woese CR, & Boone DR (1995) Partial gene sequences for the A
224 subunit of methyl-coenzyme M reductase (*mcrI*) as a phylogenetic tool for the family
225 Methanosarcinaceae. *International Journal of Systematic Bacteriology* 45(3):554-559.
- 226
- 227
- 228
- 229
- 230

231 **Supplementary Figures**



232

233 **Fig. S1 Map depicting soil collection locations for greenhouse experiment.**



234

235 **Fig. S2. Sampling and collection of the rhizocompartments.** Roots are collected from rice
236 plants and soil is shaken off the roots to leave ~1mm of soil around the roots. The ~1 mm of soil

237 is washed off in PBS and kept as the rhizosphere compartment. The clean roots are then washed
238 twice more to remove remaining soil and placed into clean PBS in a 50 mL Falcon tube. The
239 rhizoplane microbes are extracted by sonicating the roots with the rhizosphere compartment
240 removed. The sonicated roots are then placed in a new, clean Falcon tube and sonicated twice
241 more, decanting the PBS in the tube between sonications and refilling with clean PBS. These
242 roots are then kept for extracting the endospheric microbes.

243

244

245

246

247

248

249

250

251

252

253

254

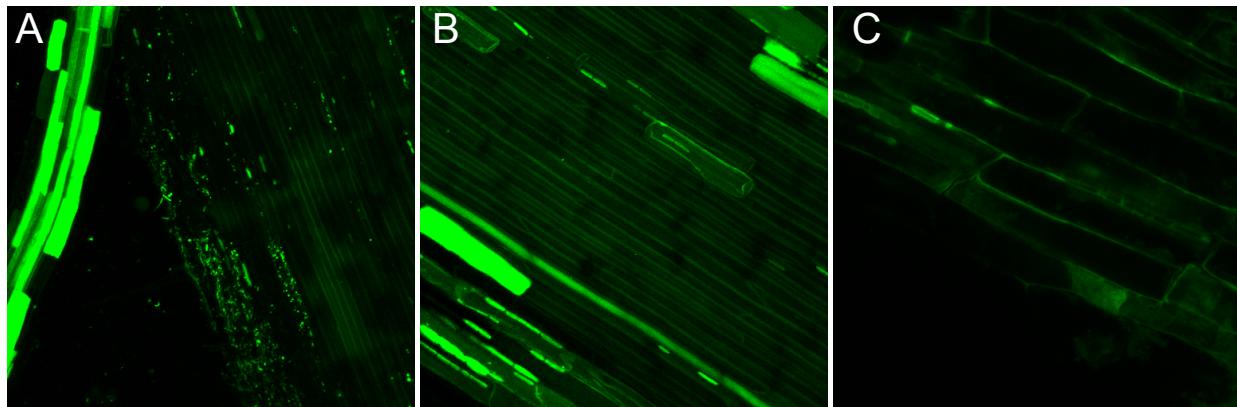
255

256

257

258

259



260

261 **Fig. S3. CARD-FISH reveals that rhizoplane microbes are removed after sonication of rice**
262 **roots.** (A) Pre-sonicated root incubated with the Eub338 eubacterial probe. (B) Thrice sonicated
263 root incubated with the Eub338 eubacterial probe. (C) Pre-sonicated root probed with the
264 antisense Eub338 probe as a negative control. Files of root cells showing bright signals are
265 presumed to be dead cells damaged during the removal of rhizosphere soil.

266

267

268

269

270

271

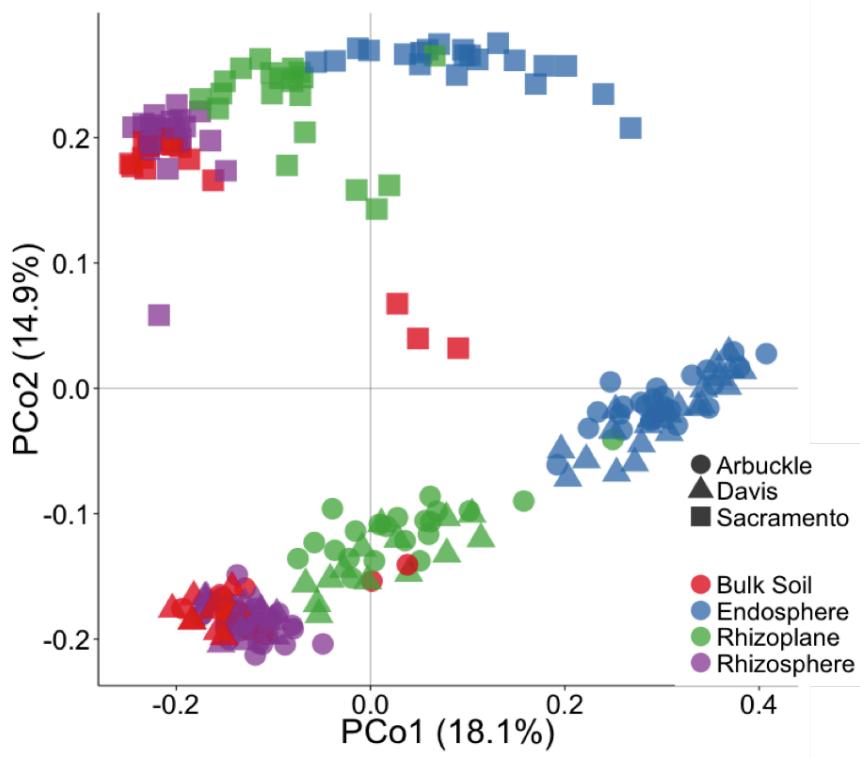
272

273

274

275

276



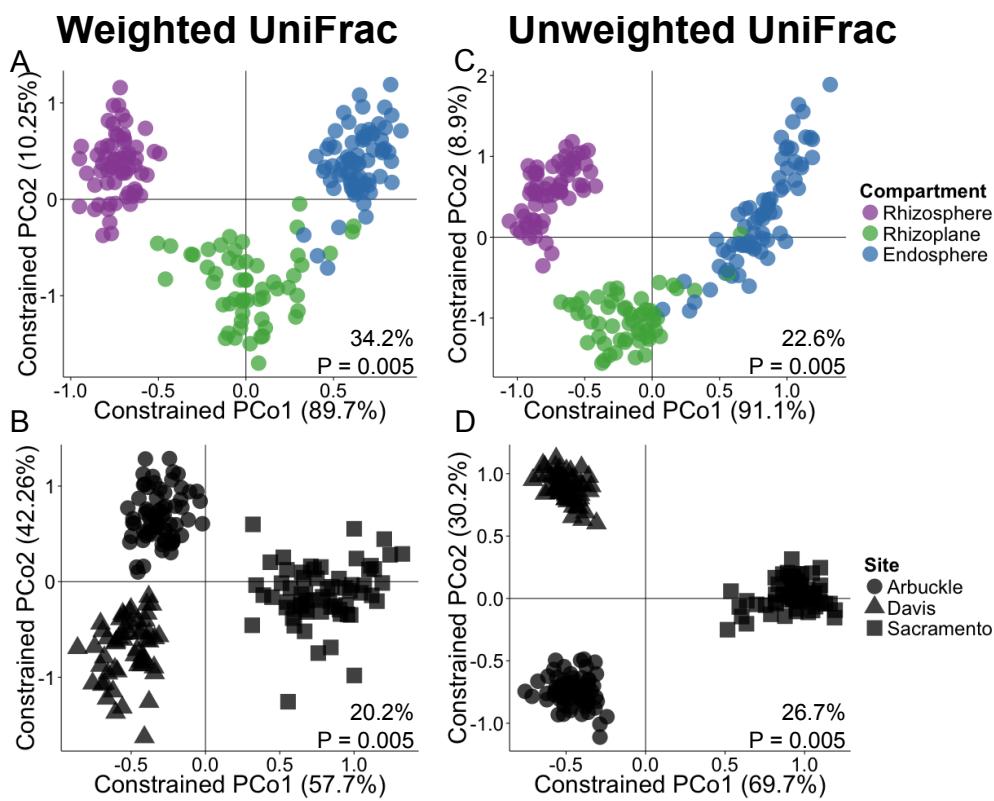
277

278 **Fig. S4. Rice root-associated microbiomes vary by rhizocompartment and site in the**
 279 **greenhouse experiment.** PCoA using the unweighted UniFrac distance metric indicates that
 280 microbiomes separate by rhizocompartment and soil source.

281

282

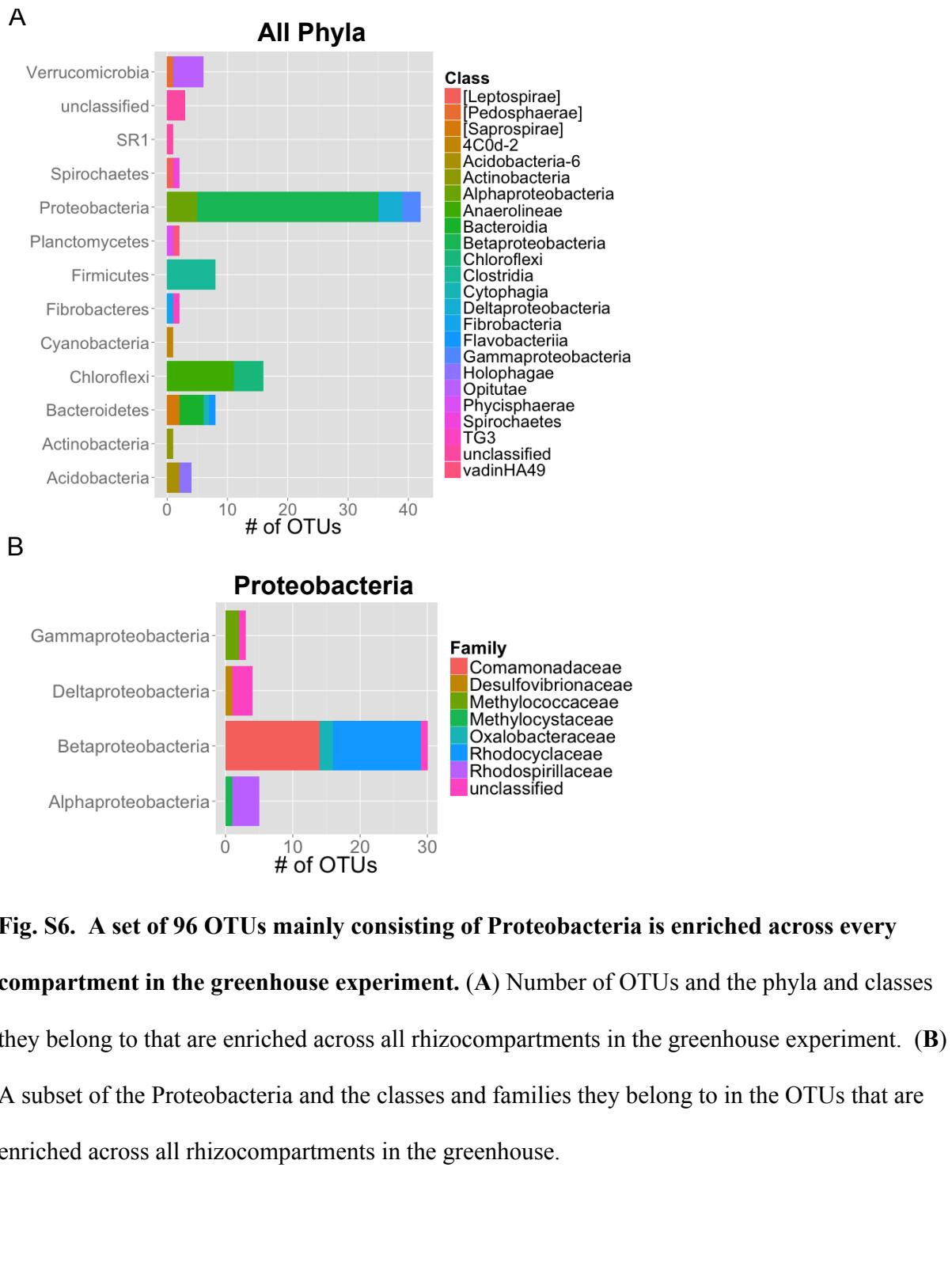
283

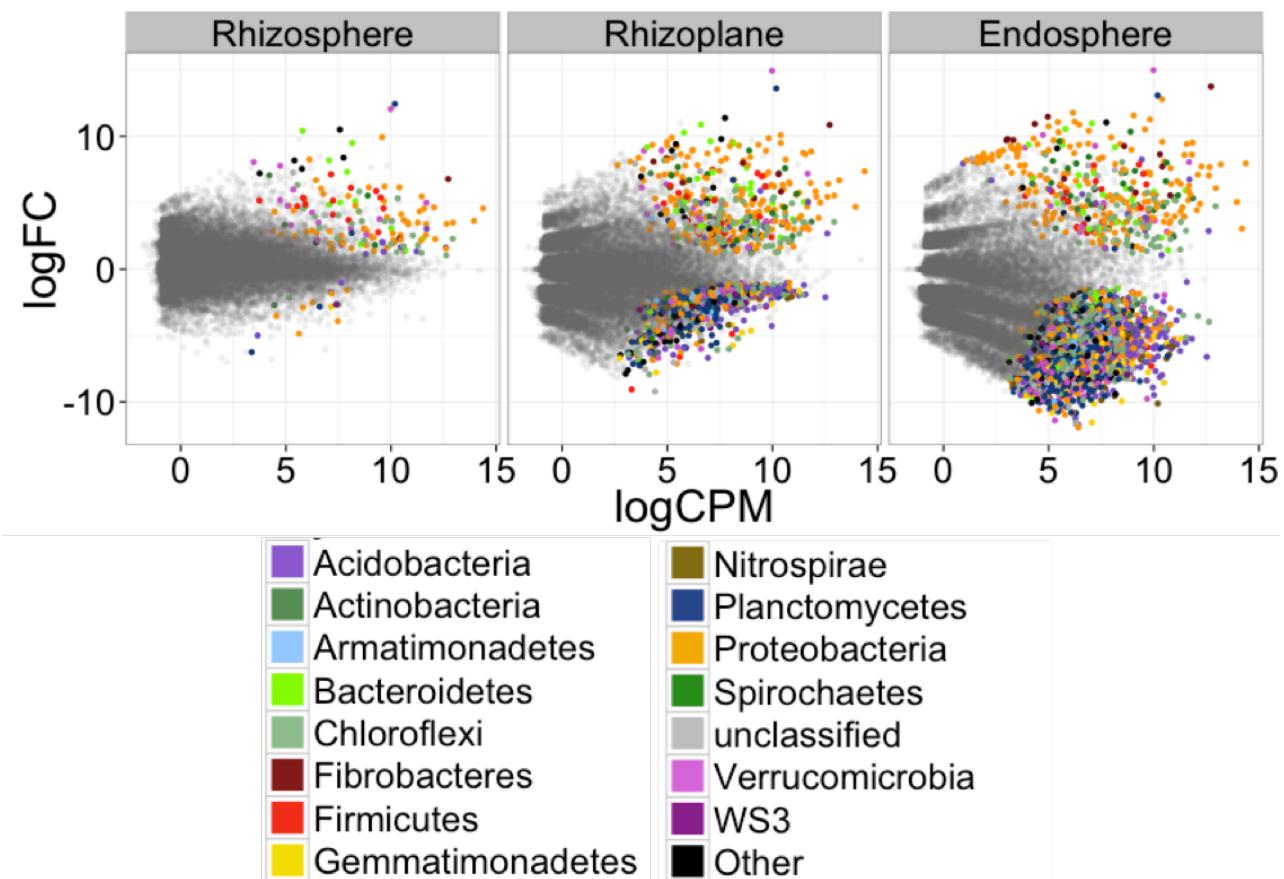


284

285 **Fig. S5. CAP analysis confirms that rice root microbiomes vary by compartment and soil**
 286 **source. (A)** CAP analysis ordination constrained to rhizocompartment and conditioned on soil
 287 source, cultivar, and technical factors using the weighted UniFrac distance metric. **(B)** CAP
 288 analysis ordination constrained to soil source and conditioned on rhizocompartment, cultivar,
 289 and technical factors using the weighted UniFrac distance metric. **(C)** CAP analysis ordination
 290 constrained to rhizocompartment and conditioned on soil source, cultivar, and technical factors
 291 using the unweighted UniFrac distance metric. **(D)** CAP analysis ordination constrained to soil
 292 source and conditioned on rhizocompartment, cultivar, and technical factors using the weighted
 293 UniFrac distance metric. All variances attributable to the constrained factor and the significance
 294 of the factor are portrayed in each plot.

295





304

305 **Fig. S7 Microbes enriched and depleted in the rhizocompartments compared to bulk soil**

306 **have taxonomic patterns.** Each point represents one OTU and the color of the point represents
307 the OTU's assigned Phyla.

308

309

310

311

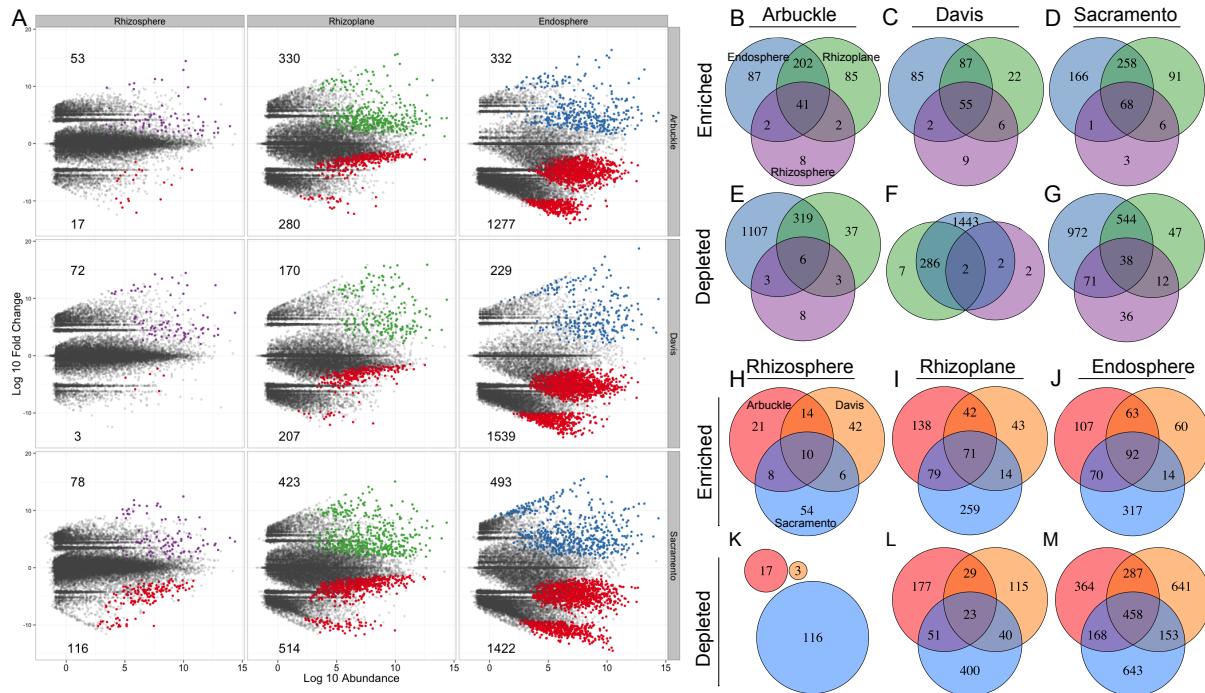
312

313

314

315

316 **Fig. S8. The different soil sources have commonalities and differences in differentially
317 abundant OTUs.** (a) MVA plots representing enrichment and depletion of OTUs in each
318 compartment compared to bulk soil across each soil source in the greenhouse experiment. (b) A
319 venn diagram comparing differentially enriched OTUs in each compartment in Arbuckle soil. (c)
320 A venn diagram comparing differentially enriched OTUs in each compartment in Davis soil. (d)
321 A venn diagram comparing differentially enriched OTUs in each compartment in Sacramento
322 soil. (e) A venn diagram comparing differentially depleted OTUs in each compartment in
323 Arbuckle soil. (f) A venn diagram comparing differentially depleted OTUs in each compartment
324 in Davis soil. (g) A venn diagram comparing differentially depleted OTUs in each compartment
325 in Sacramento soil. (h) A venn diagram comparing differentially enriched OTUs in each soil for
326 the rhizosphere compartment. (i) A venn diagram comparing differentially enriched OTUs in
327 each soil for the rhizoplane compartment. (j) A venn diagram comparing differentially enriched
328 OTUs in each soil for the endosphere compartment. (k) A venn diagram comparing differentially



329 depleted OTUs in each soil for the rhizosphere compartment. (**I**) A venn diagram comparing
330 differentially depleted OTUs in each soil for the rhizoplane compartment. (**m**) A venn diagram
331 comparing differentially depleted OTUs in each soil for the endosphere compartment. Coloration
332 is consistent for rhizocompartments across (b) to (g) and consistent for soil sources across (h) to
333 (m).

334

335

336

337

338

339

340

341

342

343

344

345

346

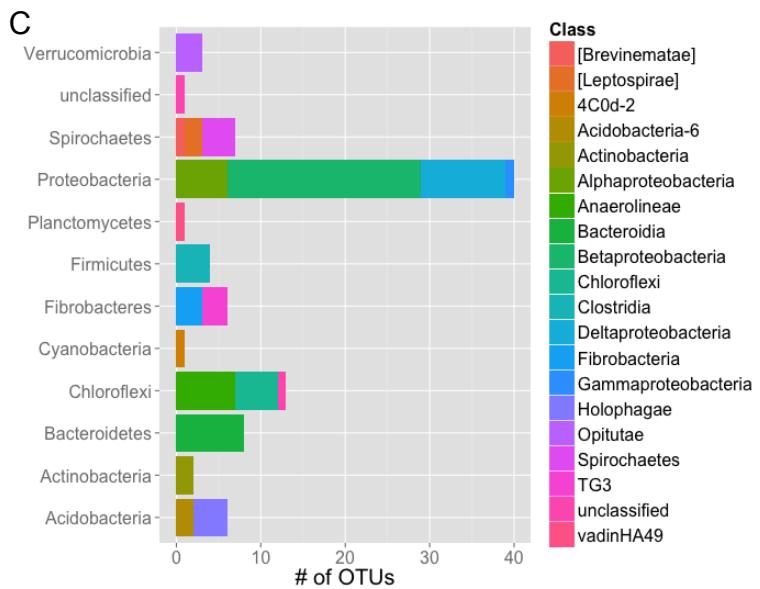
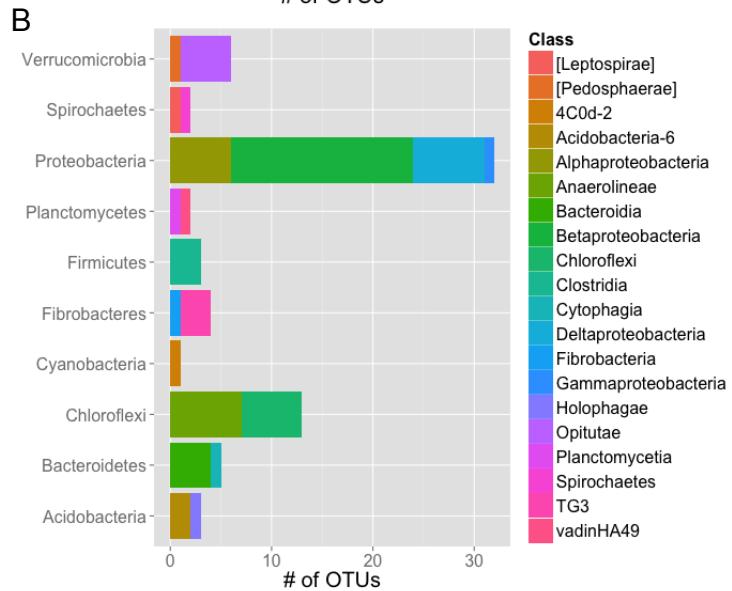
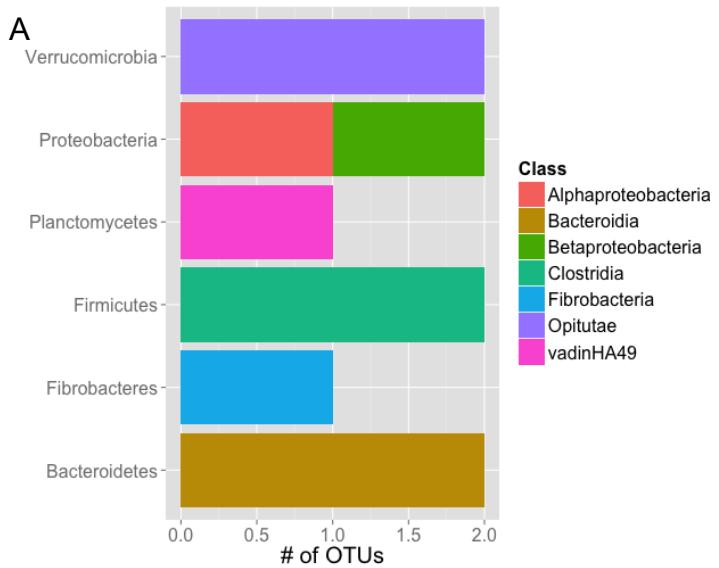
347

348

349

350

351



353 **Fig. S9. Rice plants grown in diverse soil sources have commonalities in enriched OTUs in**
354 **each rhizocompartment.** Plants grown in Davis, Arbuckle, and Sacramento soil share enriched
355 OTUs in the (A) rhizosphere, (B) rhizoplane, and (C) endosphere.

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

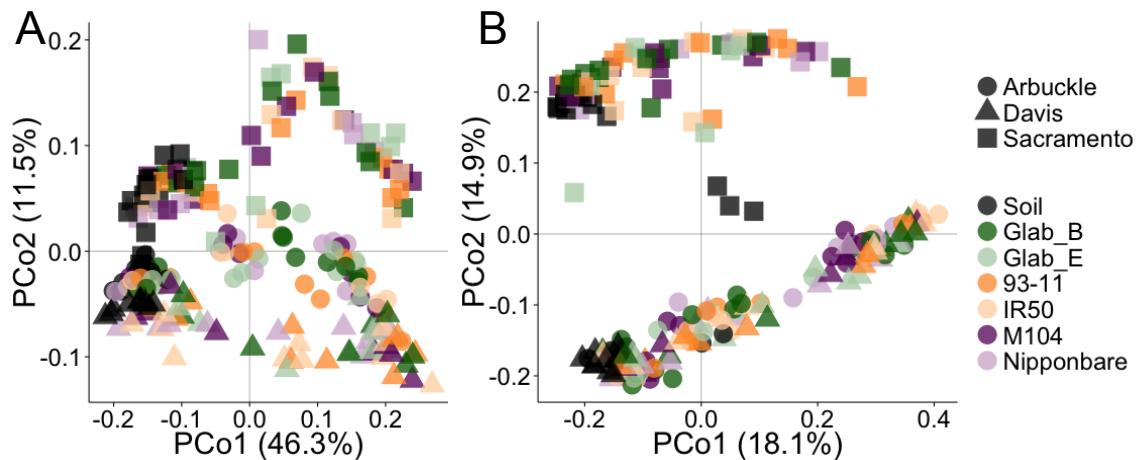
371

372

373

374

375



376

377 **Fig S10. Unconstrained PCoA reveals no distinct clustering of microbiomes of different rice
378 cultivars. (A)** Unconstrained PCoA using the weighted UniFrac distance metric. **(B)**
379 Unconstrained PCoA using the unweighted UniFrac distance metric.

380

381

382

383

384

385

386

387

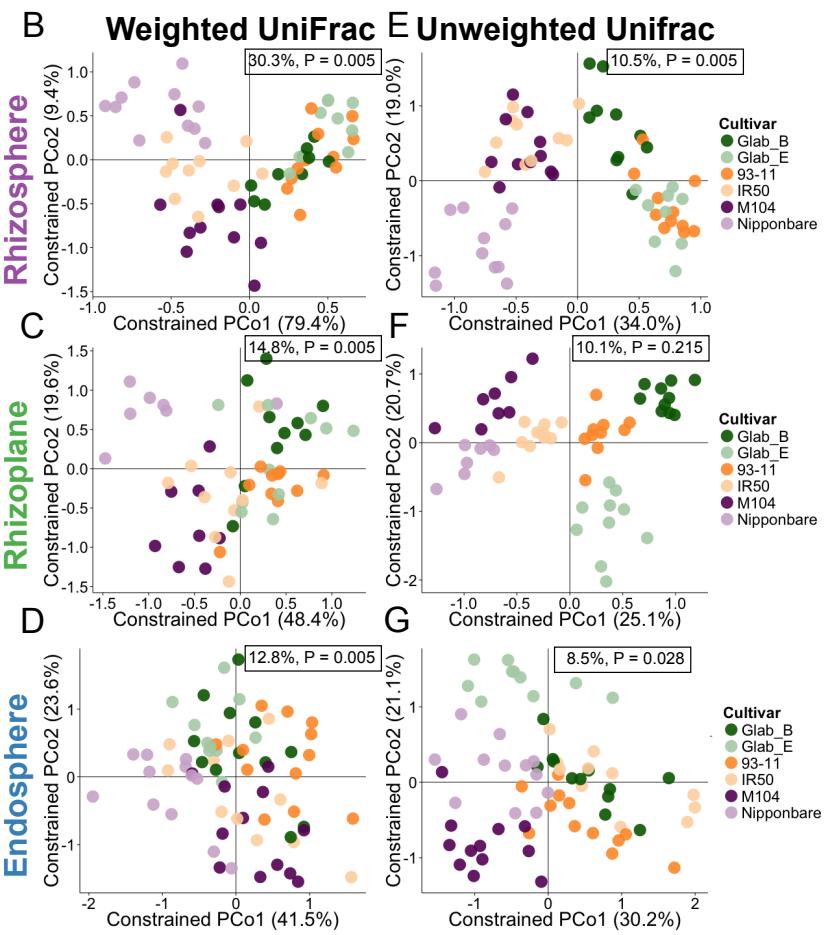
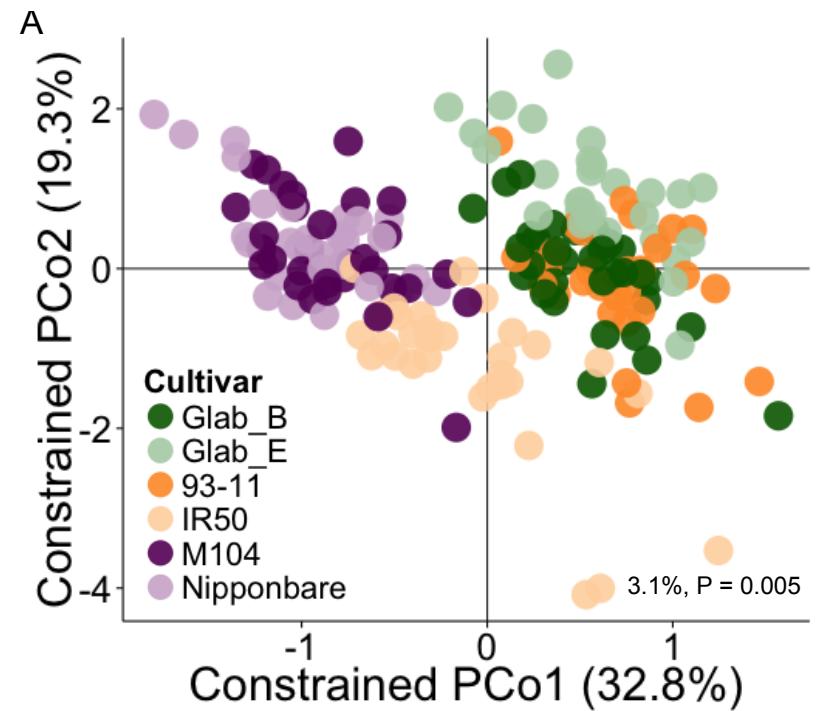
388

389

390

391

392



394 **Fig S11. CAP analysis constrained to rice cultivar while conditioning on**
395 **rhizocompartment, soil source, and technical factors reveals distinct clustering patterns of**
396 **microbiomes between rice genotypes.** (A) CAP analysis of the whole data using the
397 unweighted UniFrac distance metric. (B - D) CAP analysis constrained to rice cultivar using the
398 weighted UniFrac distance metric for (B) the rhizosphere samples, (C) the rhizoplane samples,
399 (D) the endosphere samples. (E – G) CAP analysis constrained to rice cultivar using the
400 unweighted distance metric for (E) the rhizosphere samples, (F) the rhizoplane samples, (G) the
401 endosphere samples. All variances attributable to the constrained factor and the significance of
402 the factor are portrayed in each plot.

403

404

405

406

407

408

409

410

411

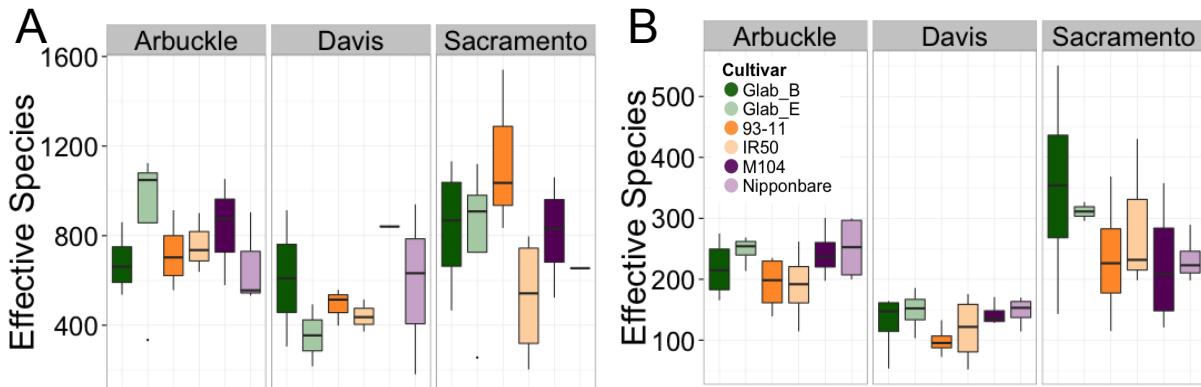
412

413

414

415

416



417

418 **Fig S12. Alpha diversities microbes on the rhizoplane and endosphere of rice cultivars**
419 **grown in different show no genotypic patterns. (A) Rhizoplane alpha diversities. (B)**
420 Endosphere alpha diversities.

421

422

423

424

425

426

427

428

429

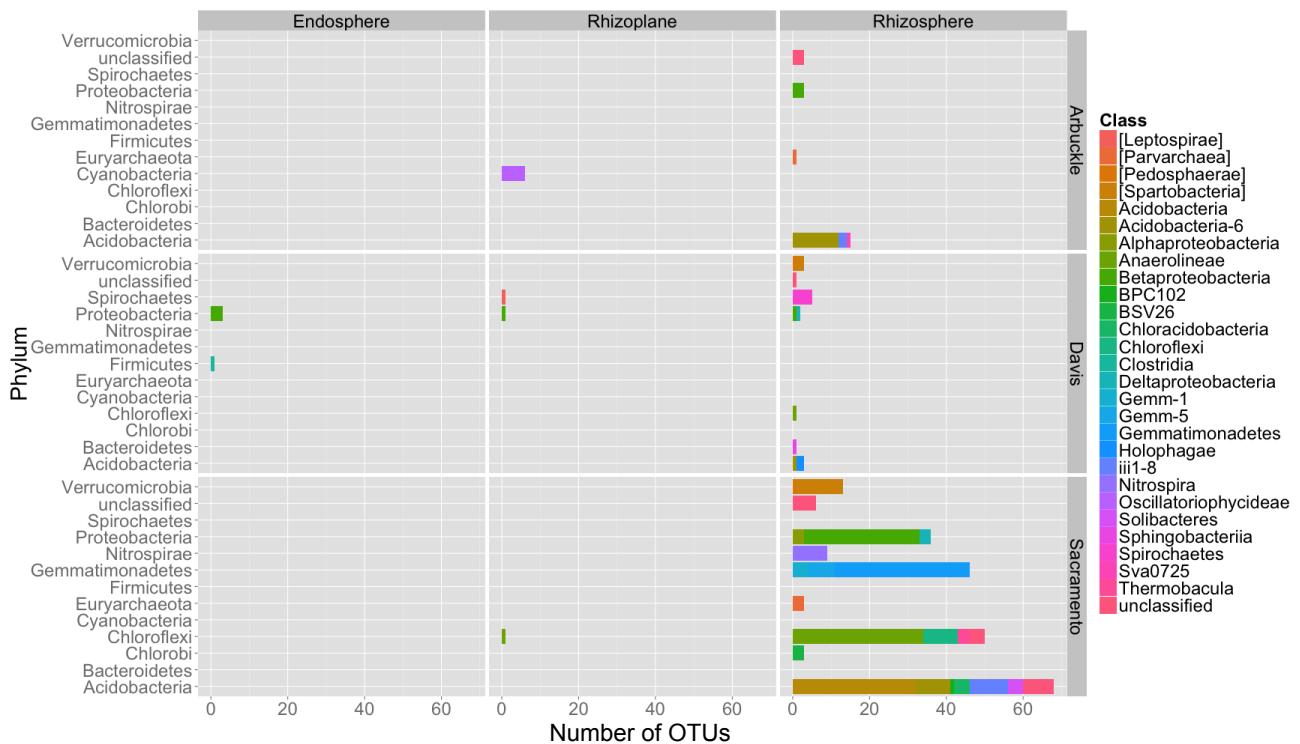
430

431

432

433

434



435

436 **Fig. S13. Differentially abundant OTUs between rice cultivars in each rhizocompartment
437 and soil source.**

438

439

440

441

442

443

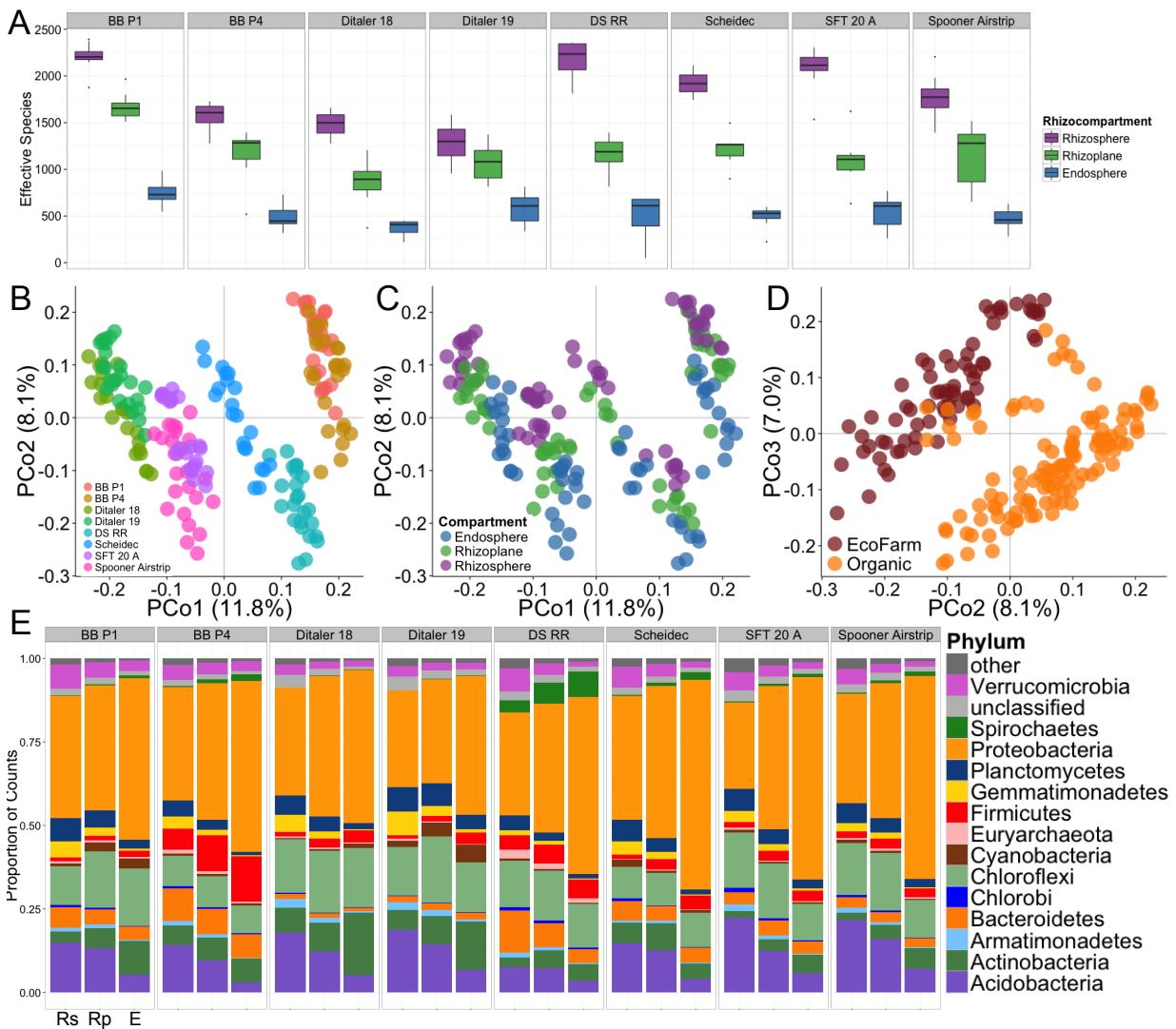
444

445

446

447

448

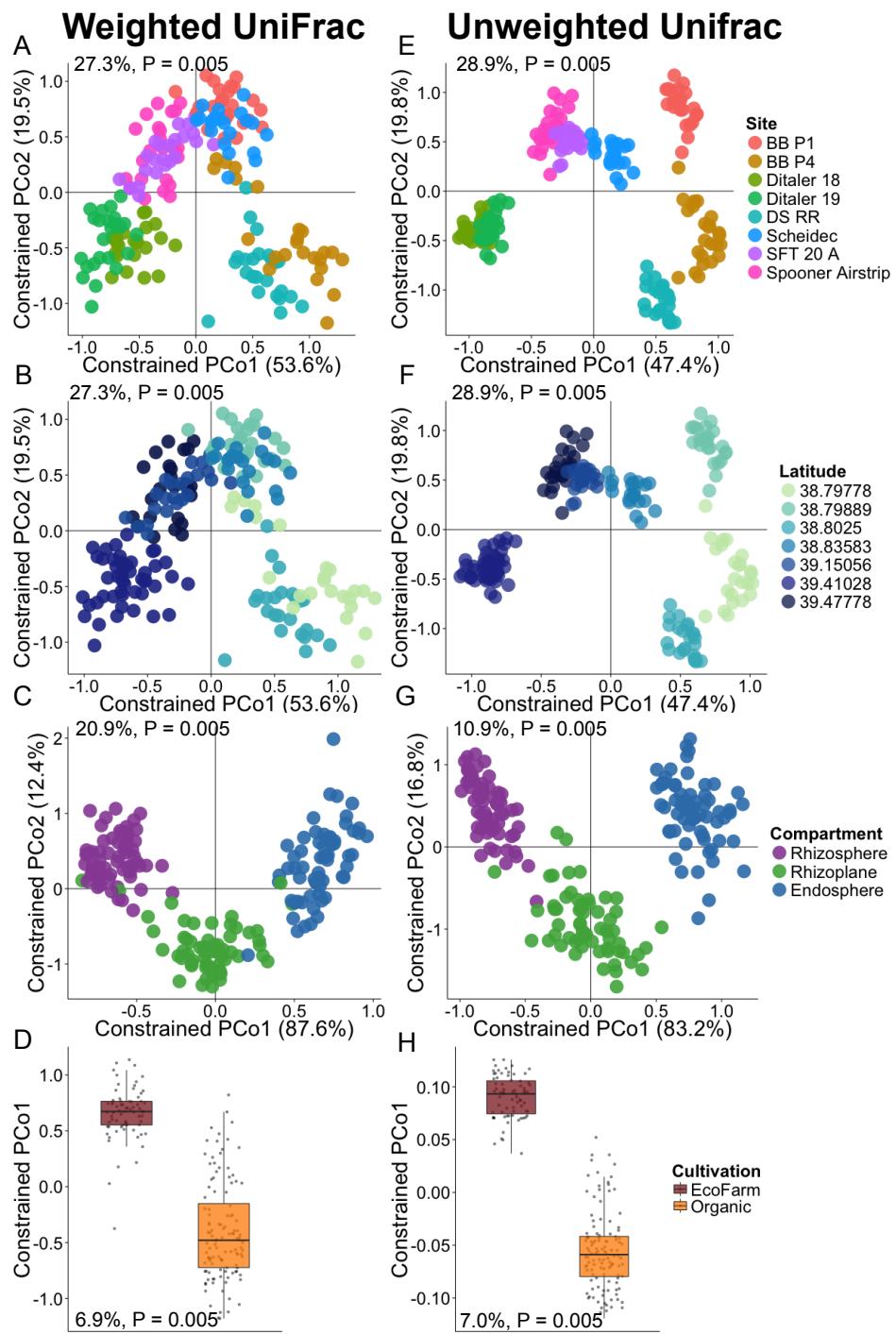


449 **Fig. S14. A gradient of diversity exists in the rhizocompartments of field grown rice. (a)** α -
450 diversity (Shannon) of all rhizocompartments sampled from all fields. **(b)** PCoA using the
451 weighted UniFrac metric colored by field site. **(c)** Same PCoA and axes as represented in (b),
452 colored by rhizocompartment. **(d)** Same PCoA represented in (b) and (c), axes 2 and 3 are shown
453 and colored by cultivation practice. **(e)** Distribution of phyla across each rhizocompartment and
454 every field. Rs, Rhizosphere; Rp Rhizoplane; E, Endosphere.

455

456

457



458

459 **Fig. S15. CAP analysis of the field data reveals that microbiomes vary by**
 460 **rhizocompartment, field site, and cultivation practice. (A – D) CAP analysis using the**
 461 **weighted UniFrac metric. (A) CAP analysis constrained to field site and conditioning on**

462 rhizocompartment, cultivation practice, and technical factors. **(B)** Same analysis as **(A)** but
463 colored by latitude. **(C)** CAP analysis constrained to rhizocompartment conditioned on field site,
464 cultivation practice, and technical factors. **(D)** CAP analysis on constrained to cultivation
465 practice. **(E – H)** CAP analysis using the unweighted UniFrac metric. **(E)** CAP analysis
466 constrained to field site and conditioning on rhizocompartment, cultivation practice, and
467 technical factors. **(F)** Same analysis as **(E)** but colored by latitude. **(G)** CAP analysis constrained
468 to rhizocompartment conditioned on field site, cultivation practice, and technical factors. **(H)**
469 CAP analysis on constrained to cultivation practice. All variances attributable to the constrained
470 factor and the significance of the factor are portrayed in each plot.

471

472

473

474

475

476

477

478

479

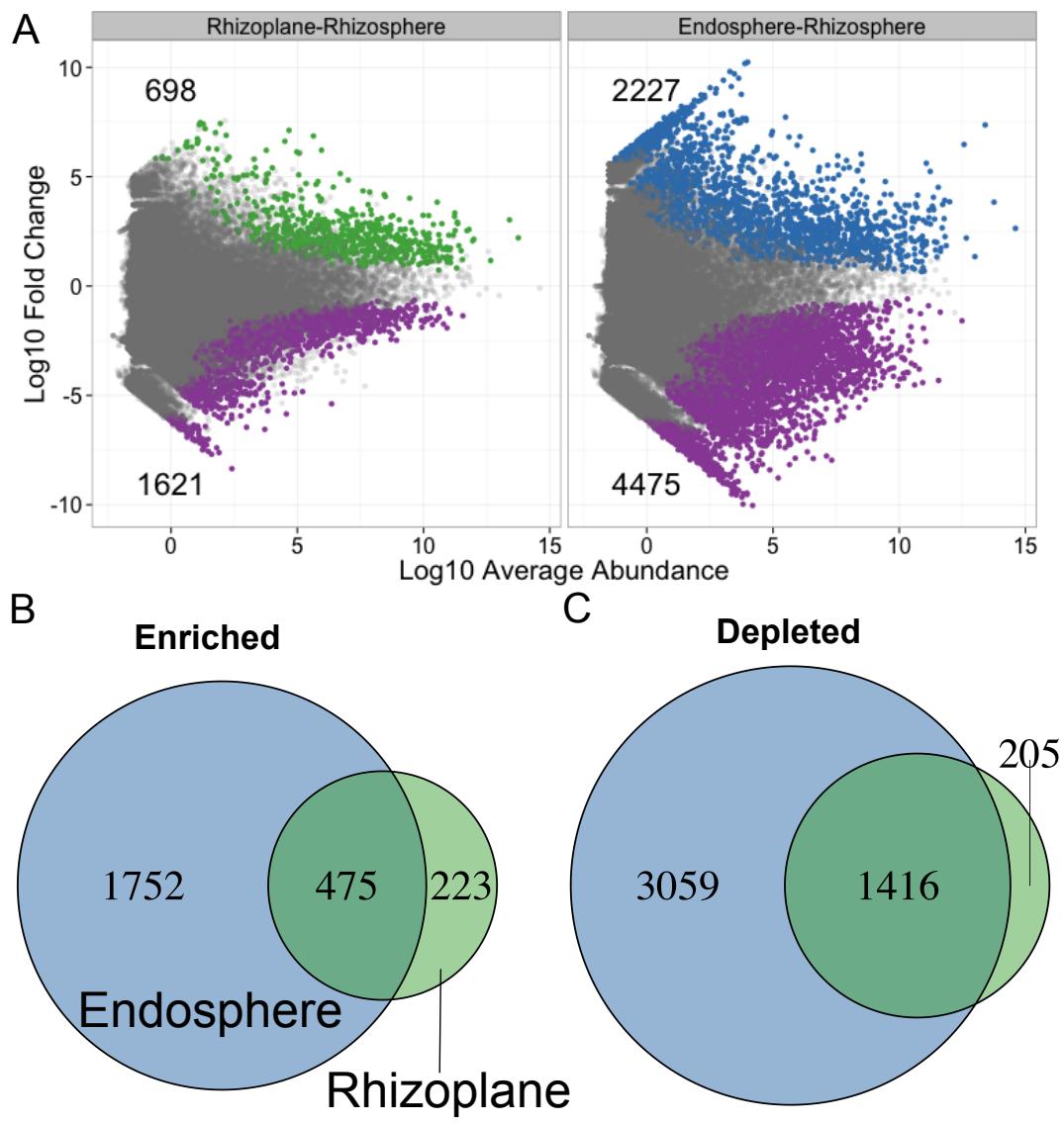
480

481

482

483

484



485

486 **Fig. S16. The rhizocompartments of field grown plants are enriched and depleted for**

487 **OTUs.** (a) MVA plot displaying enriched OTUs in the endosphere and the rhizoplane compared

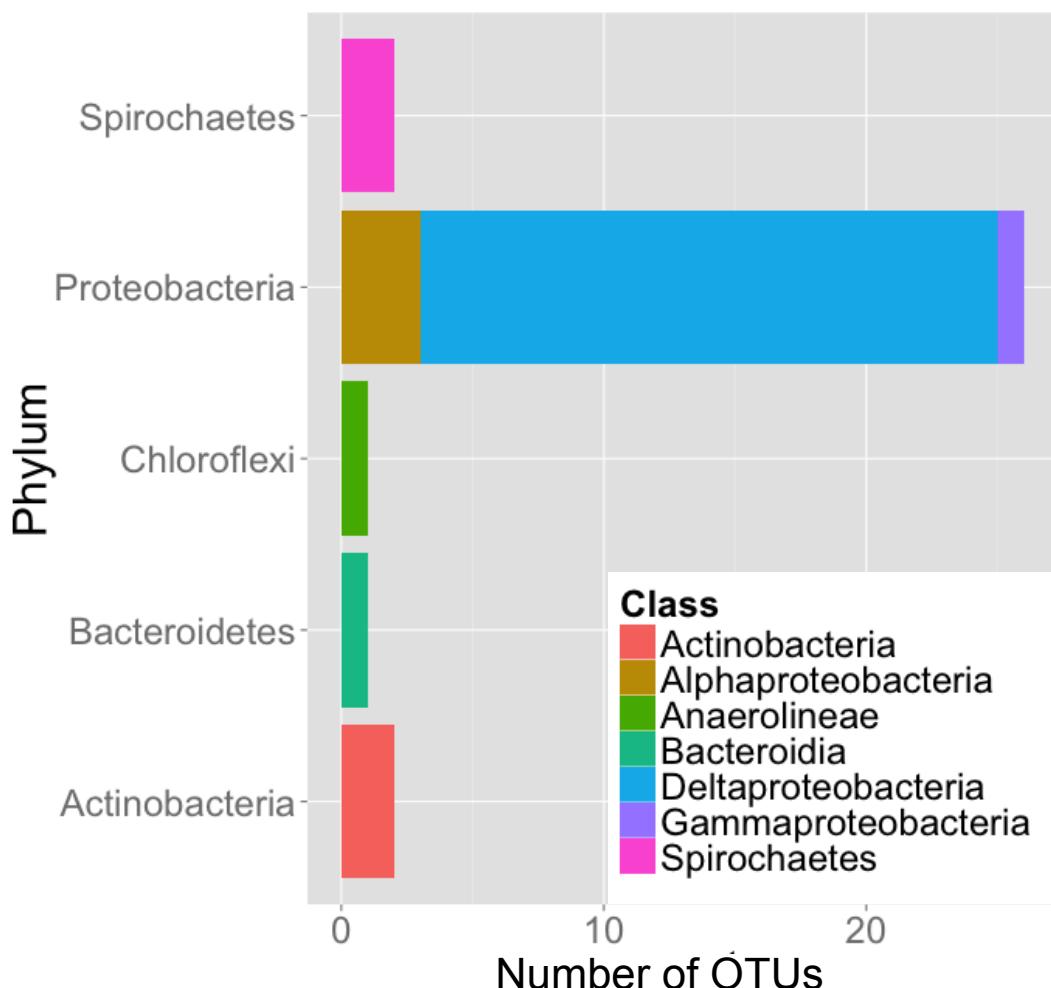
488 to the rhizosphere. (b) Venn diagram displaying similarities and differences among significantly

489 enriched OTUs in the rhizoplane and endosphere. (c) Venn diagram displaying the similarities

490 and differences among significantly depleted OTUs in the rhizoplane and endosphere. The color

491 scheme is consistent of the rhizocompartments in the venn diagrams.

492



493

494 **Fig. S17. A core endospheric microbiome consisting of 32 OTUs enriched across all field
sites displayed by phylum and class.**

496

497

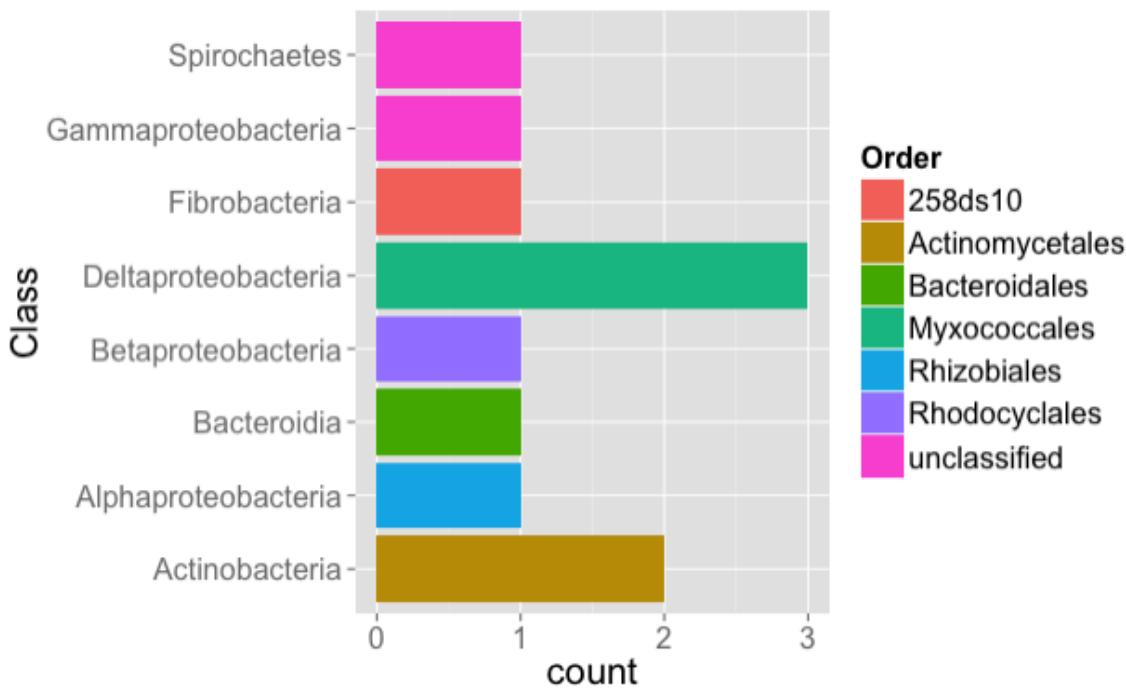
498

499

500

501

502



503

504 **Fig. S18 The greenhouse core endosphere enriched microbiome shares 11 OTUs with the**
 505 **field enriched endosphere microbiome.**

506

507

508

509

510

511

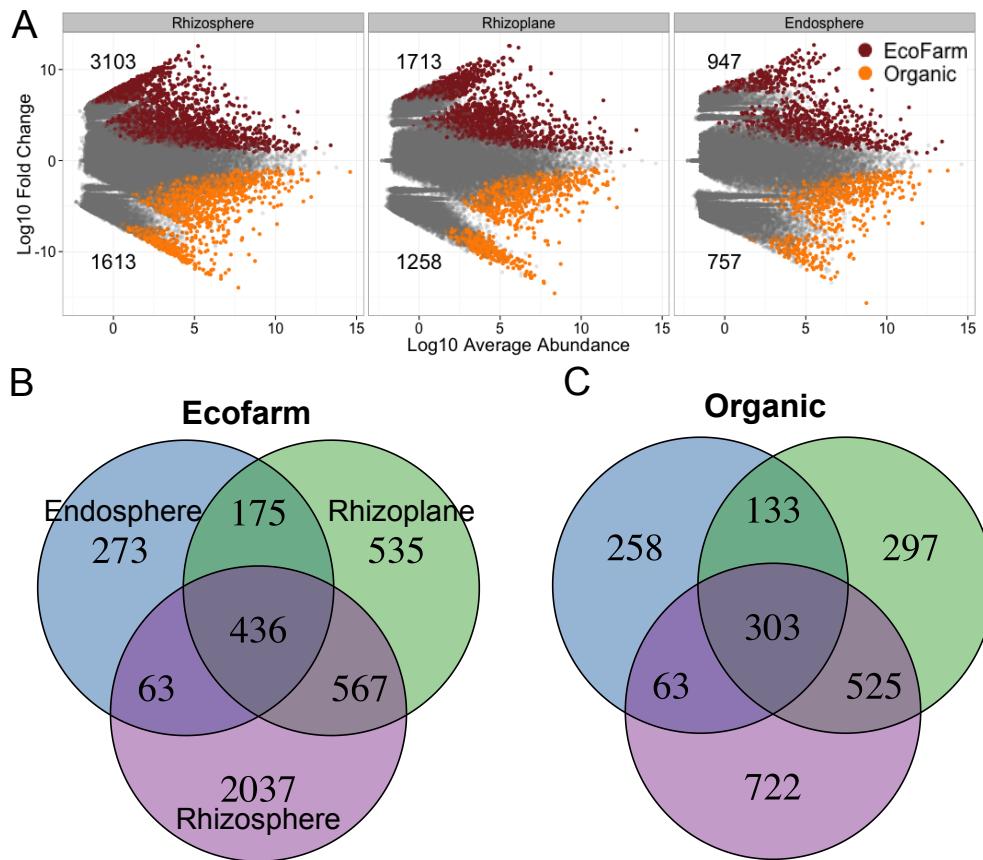
512

513

514

515

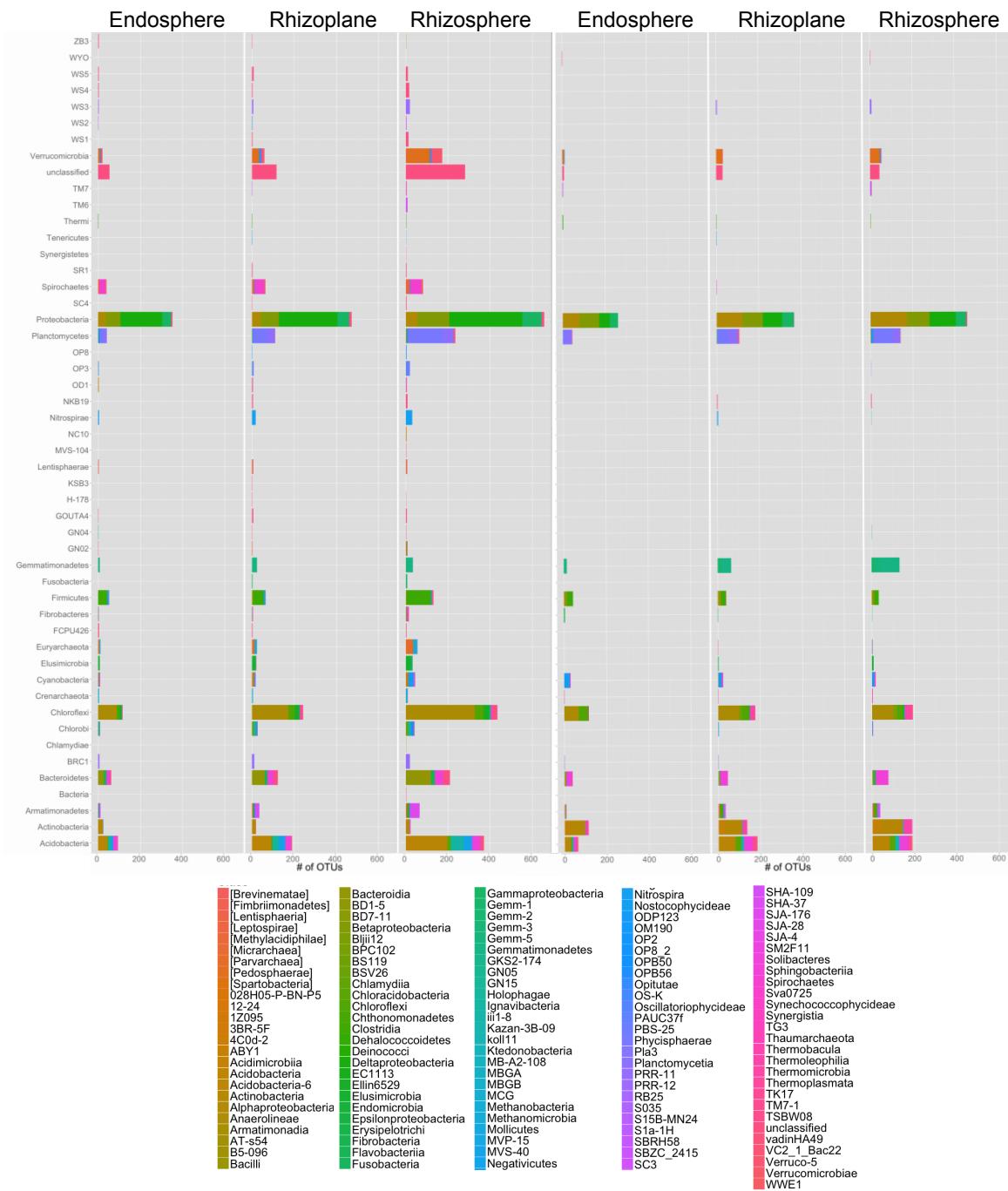
516



518 **Fig. S19. Differential OTU abundance between cultivation practices.** (a) MVA plot
 519 displaying OTUs enriched in either organic or ecofarming practices across each
 520 rhizocompartment. (b) Venn diagram indicating similarities of enriched OTUs between
 521 rhizocompartments under ecofarm cultivation. (c) Venn diagram indicating similarities of
 522 enriched OTUs between rhizocompartments under organic cultivation. The color scheme is
 523 consistent of the rhizocompartments in the venn diagrams.
 524
 525
 526
 527
 528

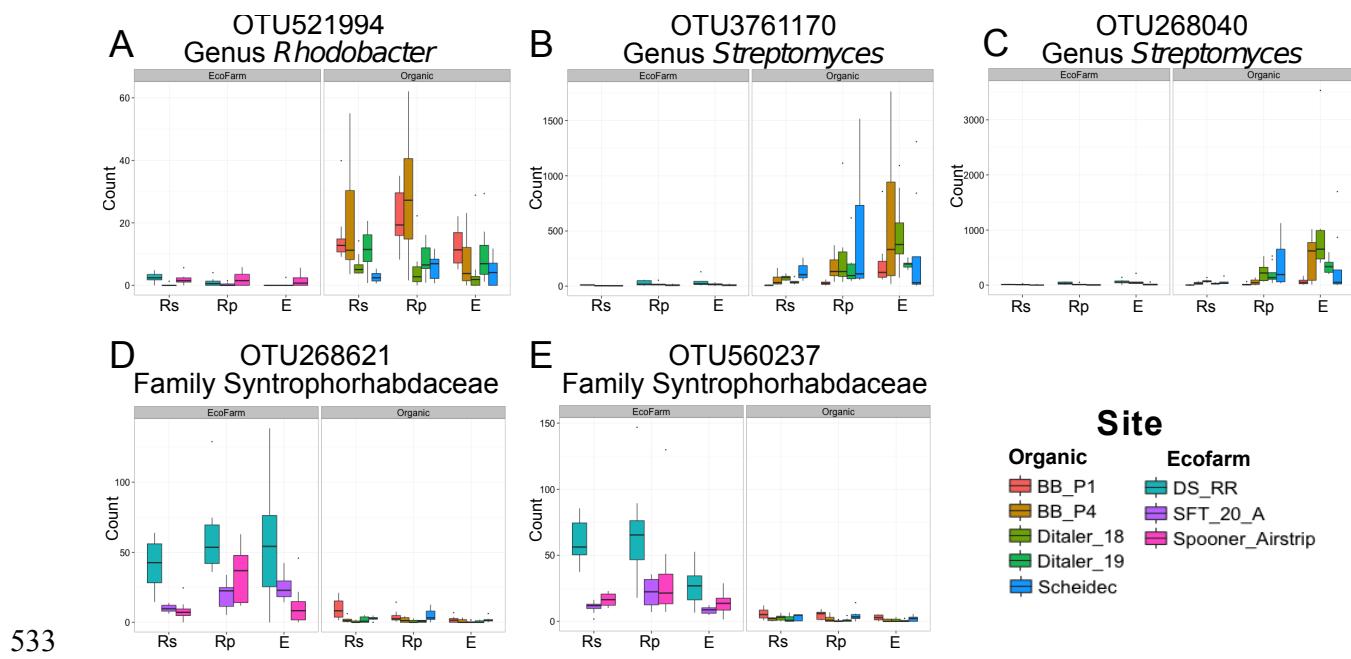
Ecofarm

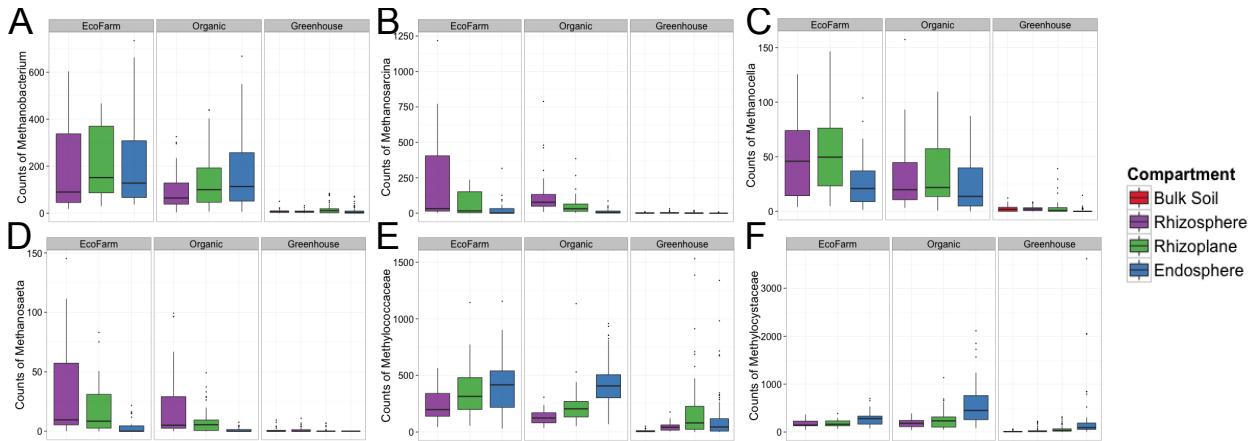
Organic



529

530 **Fig. S20. OTUs that are significantly differentially abundant between cultivation practices**
 531 mainly vary within the phyla of Proteobacteria, Acidobacteria, Actinobacteria, and
 532 Bacteroidetes.





548

549 **Fig. S22. OTUs involved in methane formation and oxidation have various patterns of**
 550 **abundance across the different rhizocompartments and growth conditions. (a)** The sum of
 551 the abundance of all OTUs within the methanogenic genus *Methanobacterium*. **(b)** The sum of
 552 the abundance of all OTUs within them methanogenic genus *Methanosarcina*. **(c)** The sum of the
 553 abundance of all OTUs within the methanogenic genus *Methanocella*. **(d)** The sum of the
 554 abundance of all OTUs within the methanogenic genus *Methanosaeta*. **(e)** The sum of the
 555 abundance of all OTUs within the methanotrophic family Methylococcaceae. **(f)** The sum of the
 556 abundance of all OTUs within the methanotrophic family Methylocystaceae.

557

558

559

560

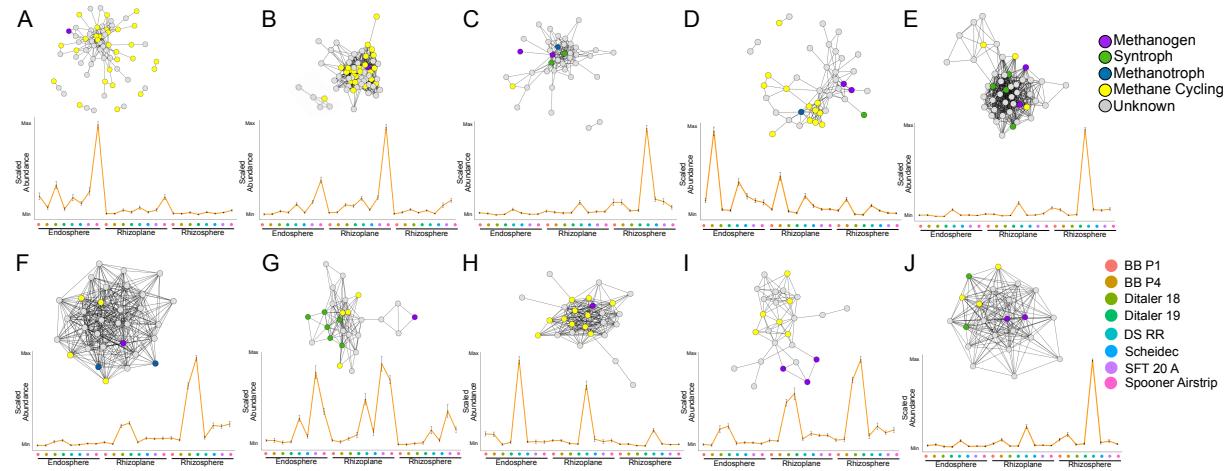
561

562

563

564

565



566

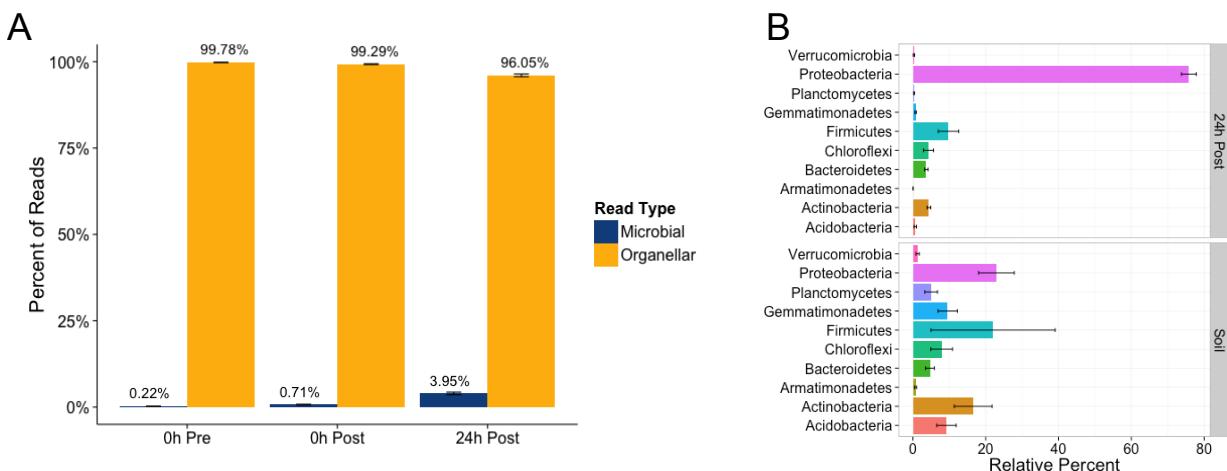
567 **Fig. S23. Modules of the co-abundance network associated with methane cycling.** Each
 568 node represents an OTU and is colored by that OTUs presumed function in methane cycling. An
 569 edge is drawn between OTUs if they have a Spearman correlation value of 0.6 or greater. **(a)**
 570 Module 6 and the average abundance profile for OTUs within the module. **(b)** Module 17 and
 571 the average abundance profile for OTUs within the module. **(c)** Module 53 and the average
 572 abundance profile for OTUs within the module. **(d)** Module 58 and the average abundance
 573 profile for OTUs within the module. **(e)** Module 75 and the average abundance profile for OTUs
 574 within the module. **(f)** Module 152 and the average abundance profile for OTUs within the
 575 module. **(g)** Module 184 and the average abundance profile for OTUs within the module. **(h)**
 576 Module 191. **(i)** Module 205 and the average abundance profile for OTUs within the module. **(j)**
 577 Module 252 and the average abundance profile for OTUs within the module. All error bars
 578 represent standard error.

579

580

581

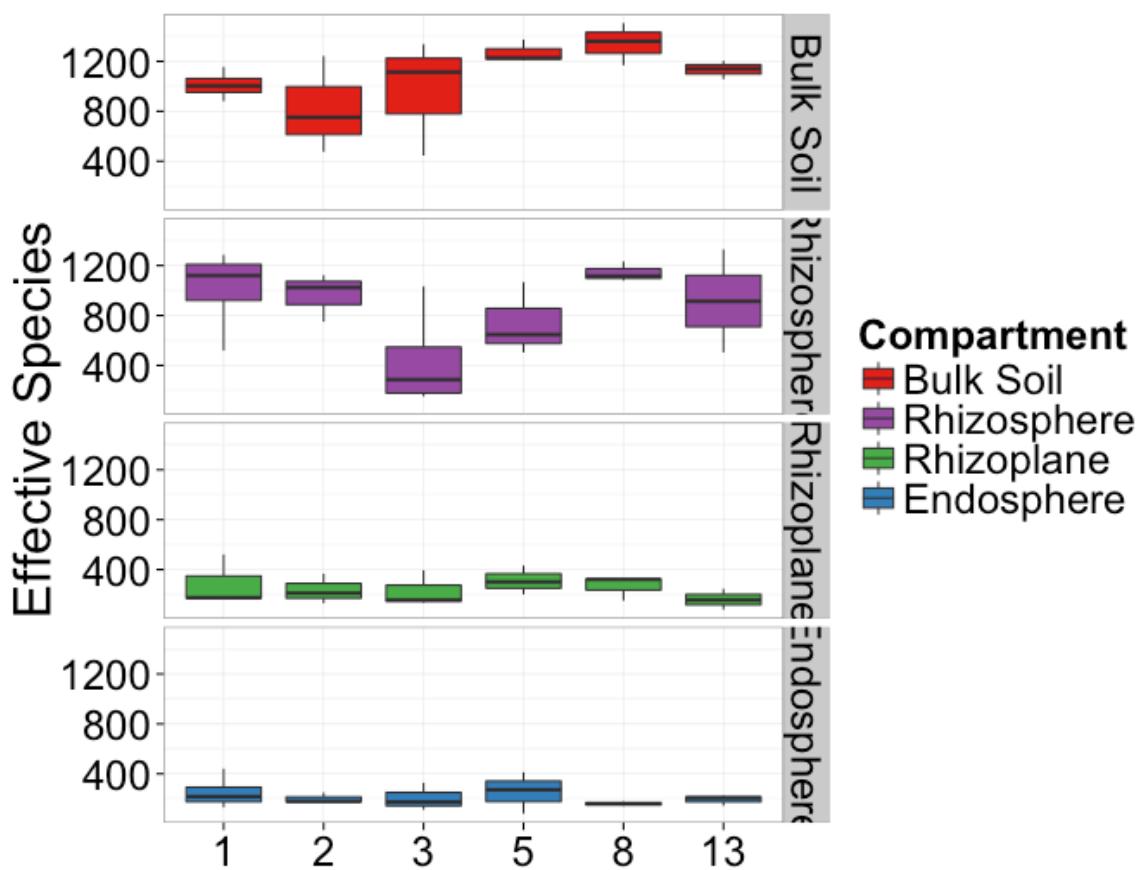
582



583

584 **Fig. S24 Microbe assembly into the endosphere at or before 24 hours is not a consequence**
 585 **of carryover from soil contact.** (A) Microbe ratios in the interior of roots before transplantation
 586 into soil, just after transplantation into soil, and after 24 hours in the soil. Mean percentages of
 587 each read type are displayed above each bar. (B) Relative abundance of phyla between bulk soil
 588 and 24 hours post transplantation into soil.

589



590

591 **Fig S25 Alpha diversity measurements of microbial communities in all compartments over**

592 **time.** Effective species = $e^{\text{Shannon_diversity}}$

593

594

595

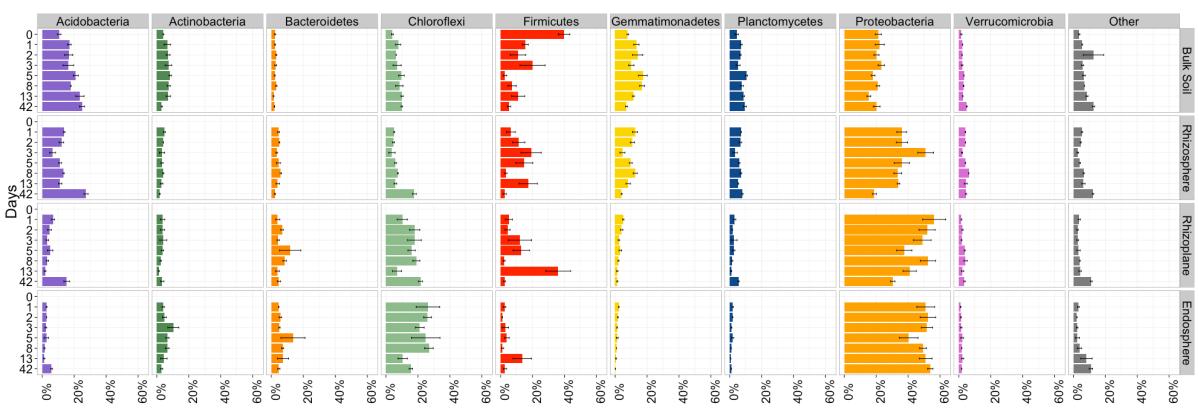
596

597

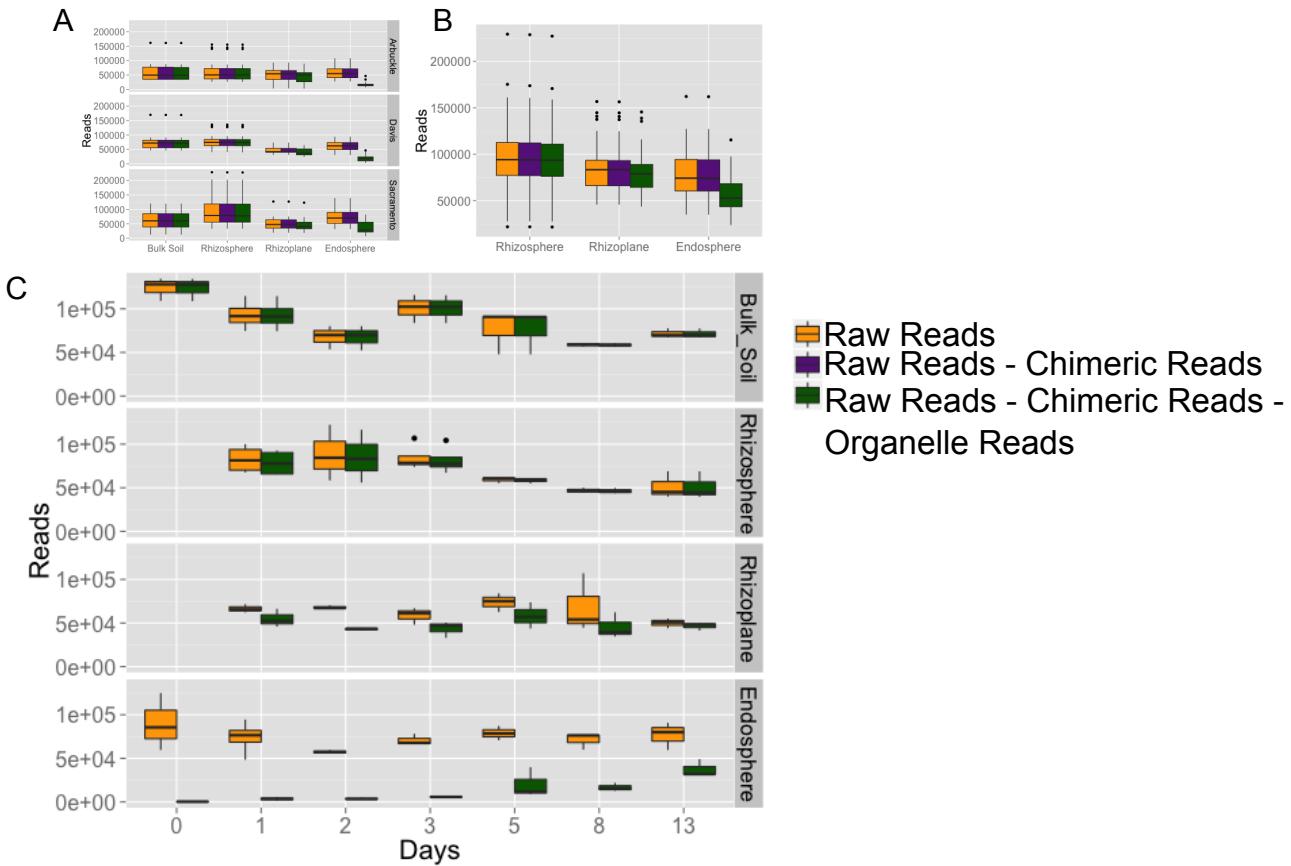
598

599

600



601
602 **Fig. S26.** There are slight shifts in the relative abundance of different phyla during the
603 acquisition of root-associated microbiomes in each rhizocompartment.
604
605
606
607
608
609
610
611
612
613
614
615



629 **Supplementary Dataset Legends**

630

631 **Dataset S1. Table showing number of replicates per factor in the greenhouse and field**
632 **experiment.**

633

634 **Dataset S2. Table displaying sequencing effort for each sample in the greenhouse**
635 **experiment.**

636

637 **Dataset S3. ANOVA results for how various factors affect alpha diversity in the**
638 **greenhouse experiment.**

639

640 **Dataset S4. Pairwise comparisons of alpha diversities between each compartment in each**
641 **soil of the greenhouse experiment.** Hypothesis testing was carried out using Wilcoxon rank
642 sum tests and corrected for multiple testing using the Benjamini-Hochberg method.

643

644 **Dataset S5. Permutational MANOVA results using weighted and unweighted UniFrac as a**
645 **distance metric for the greenhouse and field experiments.** (A) Weighted UniFrac on whole
646 greenhouse data. (B) Weighted UniFrac on Greenhouse data subsetted to bulk soil and
647 rhizosphere samples. (C) Weighted UniFrac on Greenhouse data subsetted to bulk soil and
648 rhizoplane samples. (D) Weighted UniFrac on Greenhouse data subsetted to bulk soil and
649 endosphere samples. (E) Weighted UniFrac on Greenhouse data subsetted Arbuckle samples. (F)
650 Weighted UniFrac on Greenhouse data subsetted Sacramento samples. (G) Weighted UniFrac on
651 Greenhouse data subsetted to Davis samples. (H) Unweighted UniFrac on whole greenhouse
652 data. (I) Unweighted UniFrac on Greenhouse data subsetted to bulk soil and rhizosphere
653 samples. (J) Unweighted UniFrac on Greenhouse data subsetted to bulk soil and rhizoplane
654 samples. (K) Unweighted UniFrac on Greenhouse data subsetted to bulk soil and endosphere
655 samples. (L) Unweighted UniFrac on Greenhouse data subsetted Arbuckle samples. (M)
656 Unweighted UniFrac on Greenhouse data subsetted Sacramento samples. (N) Unweighted
657 UniFrac on Greenhouse data subsetted to Davis samples. (O) Weighted UniFrac on whole Field
658 Experiment data. (P) Unweighted UniFrac on whole Field Experiment data.

659
660 **Dataset S6. Comparisons of phyla differential abundance between compartments in the**
661 **greenhouse experiment.** Hypothesis testing was carried out using Wilcoxon rank sum tests and
662 corrected for multiple testing using the Benjamini-Hochberg method.
663
664 **Dataset S7. OTUs that are significantly differentially abundant between**
665 **rhizocompartments in the greenhouse experiment.**
666
667 **Dataset S8. Results of soil chemical analysis from the greenhouse experiment.**
668
669 **Dataset S9. OTUs that are significantly differentially abundant between**
670 **rhizocompartments for each soil tested in the greenhouse experiment.**
671
672 **Dataset S10. GPS coordinate locations for all the rice fields where soil or plant material**
673 **was collected.**
674
675 **Dataset S11. Pairwise comparisons of alpha diversities between each cultivar in each**
676 **compartment in each soil.**
677
678 **Dataset S12. OTUs that are significantly differentially abundant in each cultivar of each**
679 **rhizocompartment of each soil in the greenhouse experiment.**
680
681 **Dataset S13. Table displaying sequencing effort in the field experiment.**
682
683 **Dataset S14. Impacts of tested factors on alpha diversities in the field experiment.** ANOVA
684 results are shown along with Wilcoxon rank sum tests between cultivation practices in each
685 compartment.
686
687 **Dataset S15. Pairwise comparisons of alpha diversities between each compartment of each**
688 **field site for the field experiment.** Hypothesis testing was carried out using Wilcoxon rank sum
689 tests and corrected for multiple testing using the Benjamini-Hochberg method.

690
691 **Dataset S16. Comparisons of phyla differential abundance between compartments in the**
692 **greenhouse experiment.** Hypothesis testing was carried out using Wilcoxon rank sum tests and
693 corrected for multiple testing using the Benjamini-Hochberg method.
694
695 **Dataset S17 OTUs that are significantly differentially abundant between the**
696 **rhizocompartments in each field site tested of the field experiment.**
697
698 **Dataset S18 OTUs that are significantly differentially abundant between cultivation**
699 **practices in each rhizocompartment of the field experiment.**
700
701 **Dataset S19. Taxonomies that belong to clones of *mcrA* sequenced from the rhizosphere**
702 **and endosphere of plants grown in the DS RR field.**
703
704 **Dataset S20. OTUs in the co-abundance network and the modules they are assigned to.**
705
706 **Dataset S21. OTUs modules containing methanogenic archaea.** OTUs are labeled for their
707 known relationships to methane cycling.
708
709 **Dataset S22. Taxonomies significantly enriched (FDR <= 0.05) in OTU network modules**
710 **containing the methanogenic archaea genera *Methanobacterium*, *Methanosarcina*,**
711 ***Methanocella*, and *Methanosaeta*.**
712
713 **Dataset S23. Taxonomies significantly enriched (FDR <= 0.05) in OTU network modules**
714 **containing the methanogenic archaea genera *Methanobacterium*, *Methanosarcina*,**
715 ***Methanocella*, and *Methanosaeta*.**
716
717 **Dataset S24. Comparisons of phyla differential abundance between compartments in the**
718 **timecourse experiment.** Hypothesis testing was carried out using Wilcoxon rank sum tests and
719 corrected for multiple testing using the Benjamini-Hochberg method.
720

721 **Dataset S25 Sequencing primers used to amplify the 16S rRNA gene.**

Chapter 3

The impact of host interspecies diversity on the root associated microbiome and potential implications in agriculture.

Joseph Edwards¹, John Kilmer², Christian Santos¹, Bao Nguyen¹, Gregory Phillips², and Venkatesan Sundaresan¹

3.1 Abstract

Plant roots assemble a distinct root-associated microbiota from the soil which can contribute beneficial properties to plant growth and performance. Previous studies have indicated that root-associated microbiomes are mainly constrained by environmental factors that act upon the broader soil microbial communities and that various root compartments pose as different niches for microbial colonization with the rhizosphere hosting the highest diversity of microbes and the root endosphere hosting the least. It was consistently found across various plant microbiome studies that genetic diversity within a host plant species contributes a minimal proportion of variance to root-associated microbiome differences. Thus, we were interested in quantifying the effect of host species specificity on root-associated microbiomes and whether this effect outweighs microbiome diversity within host species. We compiled published datasets from various plant microbiome stud-

¹Department of Plant Biology, University of California, Davis

²College of Agriculture and Technology, Arkansas State University

ies and found that the environment that various host species are growing in is the largest determinant of root-associated microbiome structure and that rice has a relatively unique microbiome, presumably because it is cultivated under flooded (i.e. anaerobic conditions). To better understand how root-associated microbiomes vary across host plant species within a single environment, we sampled the roots of native plant species from a rice field along with roots from rice plants. We found host specificity to be a large factor contributing to microbiome variation and that rice plants host a larger proportion of methanogenic archaea than native plants growing in the same environment. The rhizosphere microbiome of rice was significantly more similar to unplanted soil microbial communities than were the rhizosphere microbiomes of native plants. By comparing rice plants grown in rice field soil vs rice plants grown in soil which has never cultivated rice, we established that this observation was at least in part due to continuous rice cultivation establishing a microbiome more similar to the rice rhizosphere. Together, these results indicate that different host plants acquire largely dissimilar root-associated microbiomes and that when rice is grown in dense monoculture, it shifts the overall soil microbiome to be more resemblant of the rice rhizosphere microbiome.

3.2 Introduction

Land plants growing in soil interact with a vast array of microbial diversity via their roots. Individual soil microbes have been shown to provide plants with beneficial effects including general growth promotion, pathogen exclusion, and protection against abiotic stresses [7]. Plant-associated microbes, however, do not act in isolation and are part of a broader interconnected community, collectively known as the rhizosphere or root-associated microbiome [10]. Harnessing the beneficial properties of the root-associated microbiome is of agricultural and biotechnological interest, yet the parameters governing their assembly are not fully understood [10].

Plant roots host microbiomes that are distinct from the surrounding soil. Furthermore, root-associated microbiomes differ in composition between the rhizosphere and the root endosphere, with the latter having a significantly reduced microbial diversity than the for-

mer [34, 9, 18, 44, 60, 66]. Genotypic diversity within a host plant species has a significant role in shaping the root-associated microbiome [34, 60, 44, 18]. It was recently reported that root-endosphere enriched microbiomes of *Arabidopsis* relatives are compositionally distinct in only a few taxonomic members, and the patterns of these differences are contradictory to host species phylogeny [48]. Consistent with other studies, the microbiome variance partitioned to plant genotype was largely overshadowed by the variance partitioned to environmental factors (i.e. soil type, batch, field location and treatments), suggesting that plants from different, but related, species assemble a largely taxonomically conserved microbiome derived from a subset of the soil microbiome reservoir, which is affected by various environmental factors.

Asian rice (*Oryza sativa*) is primarily cultivated under flooded conditions, a growth habit unique for cereal crop plants. We have previously characterized the rice root-associated microbiome, finding that root-associated microbiomes can be separated into 3 spatially and compositionally distinct compartments: the rhizosphere, the rhizoplane, and the endosphere. Rice genotype, despite including another cultivated rice species in this study (*Oryza glaberrima*), was a very small explanatory variable for root-associated microbiome differences. These results suggest that interspecies variation between domesticated members of the *Oryza* genus minimally affects the root-associated microbiome; however, it is unknown how root-associated microbiomes vary outside of the *Oryza* genus when grown under the same flooded conditions. In this study, we sought to characterize the rice root-associated microbiome compared to other plant species.

3.3 Results

3.3.1 Root-associated microbiomes are influenced by environmental conditions.

To examine interspecies variation among root-associated microbiomes, we compiled microbiome sequencing data from 5 published datasets (Table 3.1). Primer choice during 16S rRNA gene sequencing experiments can cause significant variation in detectability of microbial taxa [14], therefore we selected studies that sequenced the V4 region of the 16S

rRNA gene. The sequences were clustered into operational taxonomic units (OTUs) based upon 97% sequencing identity using the NINJA-OPS pipeline [1]. The authors of each study used consistent definitions of the root-associated compartments with the maize study being the only exception. Bulk soil was defined as soil with no associations with plant roots, the rhizosphere was defined as the soil tightly adhering to the root, and the endosphere was defined as the root interior. The authors of the maize study defined the rhizosphere as the combination of the soil adhering to the root, the root surface, and the endosphere.

Table 3.1: Plant hosts included in the meta-analysis study.

Organism	Common Name	Study	Sequencing Technology
<i>Oryza sativa</i>	Asian Rice	Edwards et al. [18]	Illumina
<i>Zea mays</i>	Maize	Peiffer et al. [44]	454
<i>Boechera stricta</i>	Drummond's Rockcress	Wagner et al. [60]	Illumina
<i>Cannabis spp.</i>	Cannabis	Winston et al. [63]	454
<i>Vitis vinifera</i>	Grape vine	Zarraonaindia et al. [66]	Illumina

We first examined how alpha diversity varies between the host species in each study (Fig. 3.1A). We found that alpha diversity significantly varied by host species and compartment ($P < 2e-16$ and $P < 2e-16$, respectively); however, this effect is confounded by sequencing depth ($P < 2e-16$). For example, the studies using 454 sequencing rather than Illumina sequencing had overall lower alpha diversity measurements and sequencing depth within samples, thus it was impossible to attribute significant differences in alpha diversity between the various studies to interspecies variation rather than technical or environmental variation.

We next used canonical analysis of principal coordinates (CAP) of unweighted UniFrac distances between samples to investigate how interspecies variation affects root associated microbial communities. After constraining the analysis to only focus on the effect of host species on the root-associated microbiomes, we found that root microbial composition varies significantly by host species ($R^2 = 0.253$, $P = 0.001$), and that their microbiomes cluster distinctly (Fig. 3.1B). We next divided the data to analyze the rhizosphere and endosphere compartment separately. Only three of the included studies sampled the rhizo-

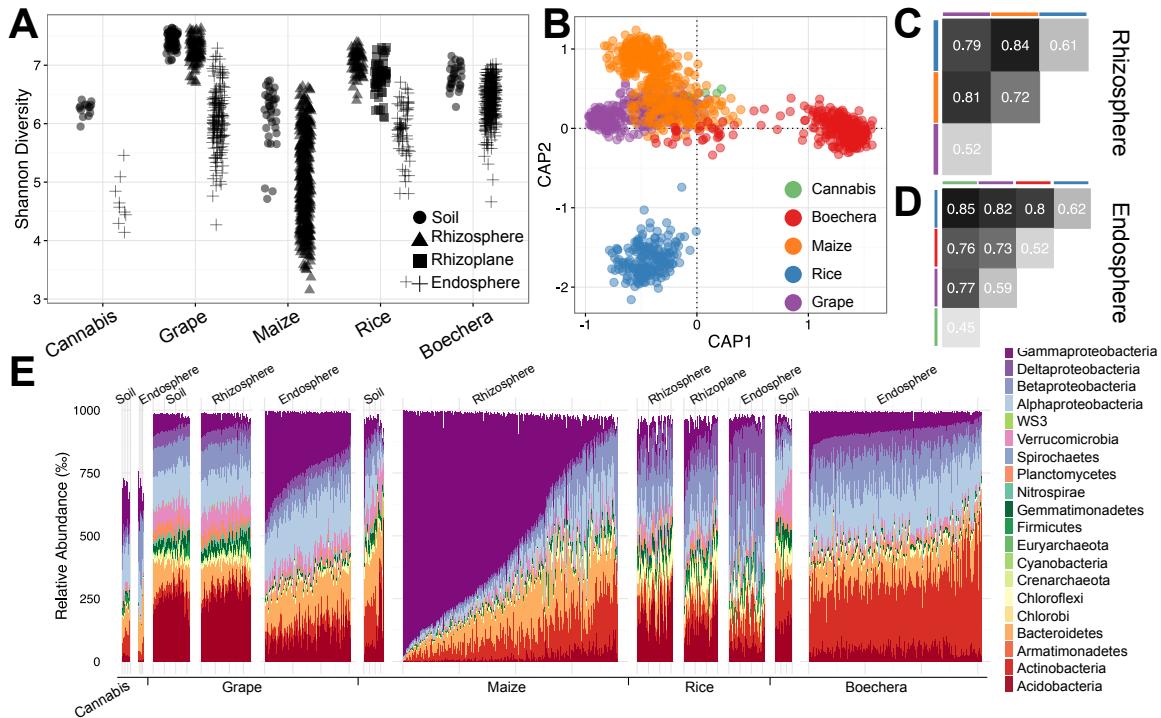


Figure 3.1: Figure 1. Meta-analysis reveals significant differences among plant rhizospheric microbiomes from different studies. (A) Shannon diversity estimates of different compartments of plants. (B) CAP analysis using unweighted UniFrac distances. (C) Mean pairwise unweighted UniFrac distances between rhizosphere samples. Colors around boxes represent host species identity and correspond to the color legend in panel B. (D) Mean pairwise unweighted UniFrac distances between endosphere samples. (E) Phylum level abundance between the different compartments of the different plant species.

sphere compartment. We found that the average unweighted UniFrac distance between microbiomes of the rice rhizosphere and grape rhizosphere was lower than the distance of either the rice rhizosphere or the grape rhizosphere to the maize rhizosphere (Fig. 3.1C). Additionally the maize rhizosphere microbiome had higher variability than the rice and grape rhizosphere microbiomes. For the endosphere fractions, we found that the microbiomes from rice were the most divergent from the other plant species (Fig. 3.1D).

Next, we sought to find which microbial phyla are specifically enriched or depleted in rice compared to the other species (Fig. 3.1E). Because Proteobacteria is the predominant microbial phylum found in the plant root-associated microbiomes, we broke this specific phylum down into its various classes for this analysis. We found that Deltaproteobacteria, Spirochaetes, Firmicutes, Euryarchaeota, Cyanobacteria, Chloroflexi, and Chlorobi were enriched in the rice rhizosphere and endosphere compared to the other plant species, while Armatimonadetes was specifically enriched in the rice rhizosphere. We found that Bacteroidetes was depleted in the rice rhizosphere and endosphere fractions compared to

other plant species, while Alphaproteobacteria and Gammaproteobacteria were specifically depleted in the endosphere of rice compared to the other species. □

Taken together, these data suggest that different plant species acquire significantly distinct root-associated microbiomes and that the rice root endosphere is the most divergent from the other species. While interspecies genetic variation likely has a causal role in the observed variation in the plant root microbiomes, the environment in which plants are grown and technical differences in the studies are also likely to be major drivers of the observed microbiome divergence. For instance, perhaps rice harbors the most divergent microbiome of the included species because it is uniquely growing in submerged soil.

3.3.2 Different plant species host distinct microbiotas when growing in the same environment.

To examine whether rice acquires significantly different root microbial communities than coexisting plant species, we collected roots from rice and various native plant species (Table 3.2) growing in a submerged rice field near Jonesboro, Arkansas (USA). Because plant developmental stage has been implicated in affecting the root microbiota, we collected roots only from plants in the reproductive stage [34]. The bacterial and archaeal communities were analyzed from the rhizosphere (soil adhering to the root after shaking, 1mm from the root surface) and root endosphere compartments. We performed Illumina MiSeq amplicon sequencing of 16S rRNA genes derived from total DNA extracted from the samples and clustered the amplicon sequences into operational taxonomic units (OTUs) based upon 97% sequence identity using the NINJA-OPS pipeline [1].

Table 3.2: Plant hosts sampled from a rice field.

Organism	Common Name	Order
<i>Oryza sativa</i>	Asian Rice	Poales
<i>Heteranthera limosa</i>	Blue Mudplantain	Commelinaceae
<i>Cyperus flavicomus</i>	Whiteedge Flatsedge	Poales
<i>Bacopa rotundifolia</i>	Disk Waterhyssop	Lamiales

To inspect the factors governing the patterns of microbial assembly in our dataset, we

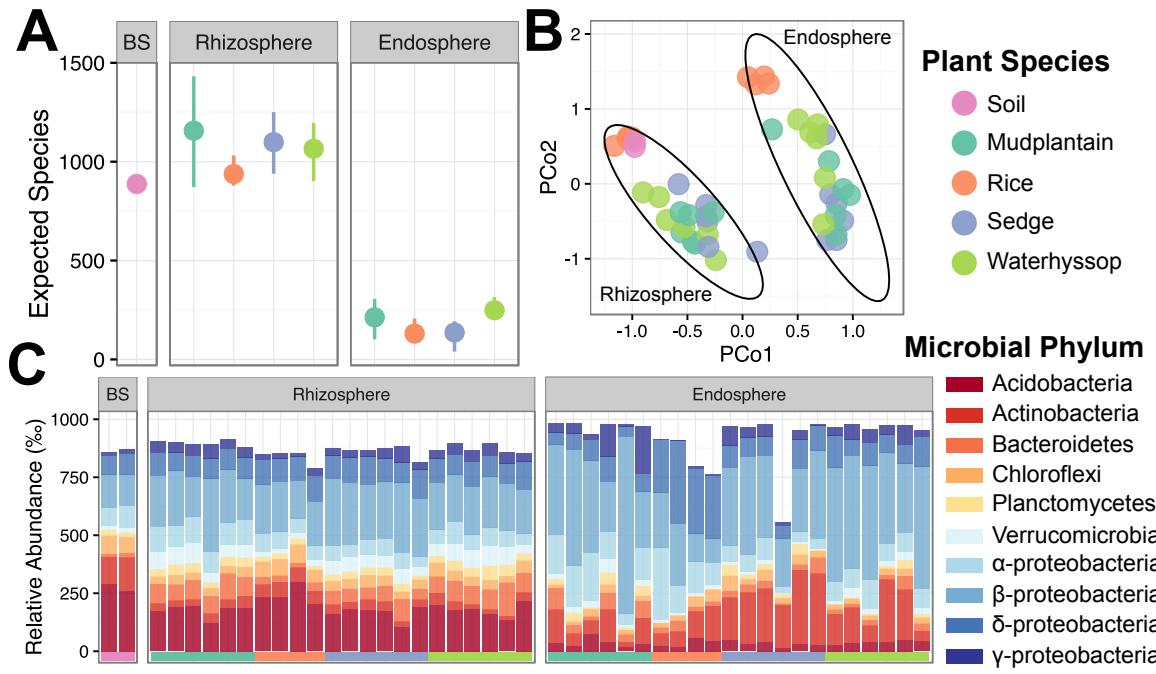


Figure 3.2: Figure 2. Rice plants and native plants growing in a rice field harbor significantly different root-associated microbiomes. (A) Shannon diversity estimates between compartments and plant species. Points show the mean diversity and tips of bars represent maximum and minimum. The colors of the points match the legend in panel B. BS stands for bulk soil. (B) Principal coordinates analysis using unweighted UniFrac distances. Ellipses mark clustering patterns at 95% confidence for the rhizosphere and endosphere compartments. (C) Phylum abundances for the different rhizospheric compartments and plant species. Colored bars beneath the x-axis represent the plant species the sample was gathered from and the colors match the legend in panel B.

calculated unweighted UniFrac distances between all samples. After using principal coordinates analysis (PCoA) to visualize clustering patterns of the microbial communities in our samples, we observed that our samples are distinguishable by rhizospheric compartment and by plant species (Fig. 3.2B). We confirmed the statistical significance of these patterns using perMANOVA (Compartment: $R^2 = 0.23$, $P < 0.001$; Plant Species: $R^2 = 0.22$, $P < 0.001$). There was a significant interaction term between rhizospheric compartment and plant species ($R^2 = 0.10$, $P < 0.001$), suggesting that the magnitude of divergence between microbiomes of the different plant species is dependent upon the root compartment. We compared the effect sizes for host species on microbiome composition between each compartment finding that the rhizosphere microbiome was more affected by host species ($R^2 = 0.45$, $P < 0.001$) than the endosphere microbiome ($R^2 = 0.40$, $P < 0.001$). The differences in microbial community structure were accompanied by divergence in microbial species diversity within samples between compartments ($P < 2e-16$) and also between plant species ($P = 0.026$, Fig. 3.2A). We noticed that there were significant differences in species richness

between the plant species within the endosphere but not the rhizosphere ($P = 0.005$ and 0.15 , respectively, ANOVA). When comparing the samples collected in this study to the publicly available datasets, we found that native plants collected from a rice field cluster more closely with rice, suggesting that the growth environment for the host plant, i.e. submerged soil, is a more important factor than host plant species in determining the root-associated microbiome (Fig. 3S1). Note that the clustering pattern in Fig. 3S1 cannot be explained by geography. The rice samples in the meta-analysis were taken from California, while the rice plants sampled for the native plant study were taken from Arkansas.

We noticed significant differences between the root microbiota of different plant species at the phylum level (Fig. 3.2C). Because root microbiomes are typically dominated by Proteobacterial taxa, we broke the Proteobacteria phylum into its separate classes. Compared to rice, all of the native plants were enriched for Gammaproteobacteria, Bacteroidetes, and Verrucomicrobia in the rhizosphere and enriched for Betaproteobacteria and Gammaproteobacteria in the endosphere. Compared to rice, the native plants were depleted for Acidobacteria, Actinobacteria, and Chloroflexi in the rhizosphere, and depleted for Deltaproteobacteria in the endosphere.

3.3.3 Rice acquires an outlier microbiome compared to native plant species.

We compared unweighted UniFrac distances of microbiotas from native plants compared to other native plants and distances of microbiotas from rice plants to native plants. We found that rice root microbiotas are significantly more dissimilar to those of native plants than the native plant microbiotas are to each other (Rhizosphere: $P = 4.16e-149$, Endosphere: $P = 3.58e-148$ Fig. 3.3A). This observation could be explained by two scenarios: 1) the roots of native plant species enrich for microbes that are not similarly enriched in rice roots, or 2) rice enriches for a root microbiota that is depleted commonly between the native plant species.

To test these two scenarios, we conducted differential abundance analysis to identify members of the microbiotas that are enriched or depleted in rice compared to the different native plant species. We fit a linear model for each OTU in each rhizospheric compartment to test for differential relative abundance between each native plant species compared to

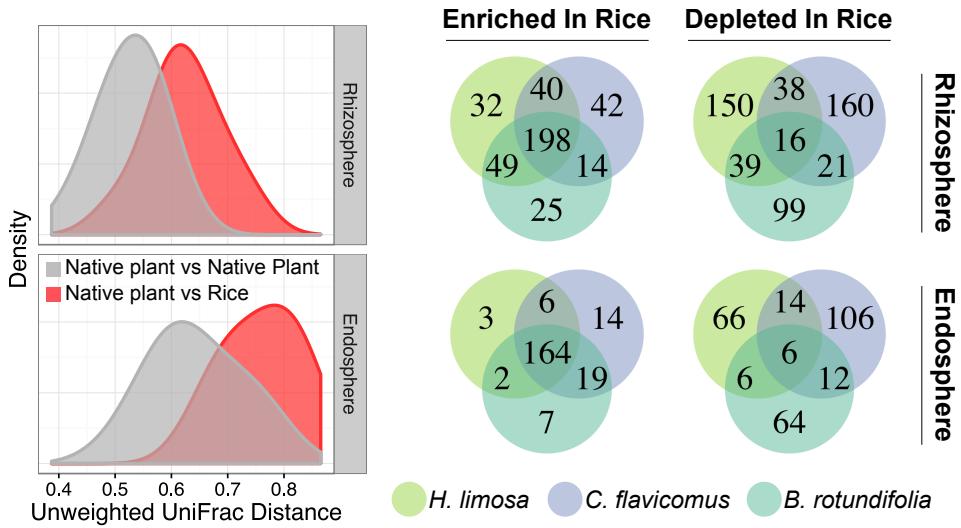


Figure 3.3: Figure 3. **Rice plants acquire an outlier microbiome compared to native plants.** (A) Distribution of unweighted UniFrac distances in each compartment comparing microbiomes from native plants to either rice or other native plants. (B) Results of differential abundance to discover OTUs enriched or depleted in the rice rhizospheric microbiomes compared to native plants.

rice. After correcting for multiple testing, we found 1117 unique OTUs that were either significantly enriched or depleted from the rice rhizosphere or endosphere compared to the native plants. When observing the overlap in microbial taxonomies enriched in each native plant species compared to rice (Fig. 3.3B), we found that 2.4% of OTUs were in common in the rhizosphere and 1.9% in the endosphere. Alternatively, of the OTUs depleted in the native plants compared to rice, we found that 22% of OTUs were in common in the rhizosphere and 28.7% in the endosphere. Together, these results support the second scenario, where rice enriches for microbes that are consistently depleted from the native plant species.

3.3.4 Rice roots are enriched for methanogenic archaea but not methanotrophic bacteria compared to native plant species.

Rice cultivation accounts for a large proportion of anthropogenic methane (CH₄) emissions. CH₄, a potent greenhouse gas with nearly 4 times greater capacity of trapping heat than carbon dioxide [30], is produced under anaerobic conditions by certain archaea clades [38]. In a submerged rice field, methanogens derive their energy from degrading rice root exudates and other organic matter derived from decaying straw or roots [38]. While methanogens have been well-characterized in rice field environments, it remains unclear to how specific

rice field methanogens are to the rice rhizosphere or whether they may also colonize roots of native plants growing in the same environment.

In the set of differentially abundant microbes we found 8 OTUs belonging to methanogenic taxonomies specifically enriched in the rice rhizosphere and 8 OTUs in the endosphere (with 5 overlapping OTUs). These OTUs belong to the *Methanocella*, *Methanobacterium*, and *Methanosarcina* genera in the rhizosphere and the *Methanobacterium* and *Methanocella* genera in the endosphere. The 8 rice-specific enriched rhizosphere methanogenic OTUs were among the top 10 of 41 most abundant methanogens represented in our rhizosphere dataset, while the 8 of 35 rice root enriched endosphere methanogenic OTUs were among the top 9 most abundant methanogens in our endosphere dataset (Fig. 3.4A). We were unable to identify any methanogenic OTUs enriched in the native plants compared to rice. Indeed, the total relative abundance of methanogenic microbes were enriched in the rice rhizosphere and endosphere compared to native plant species.

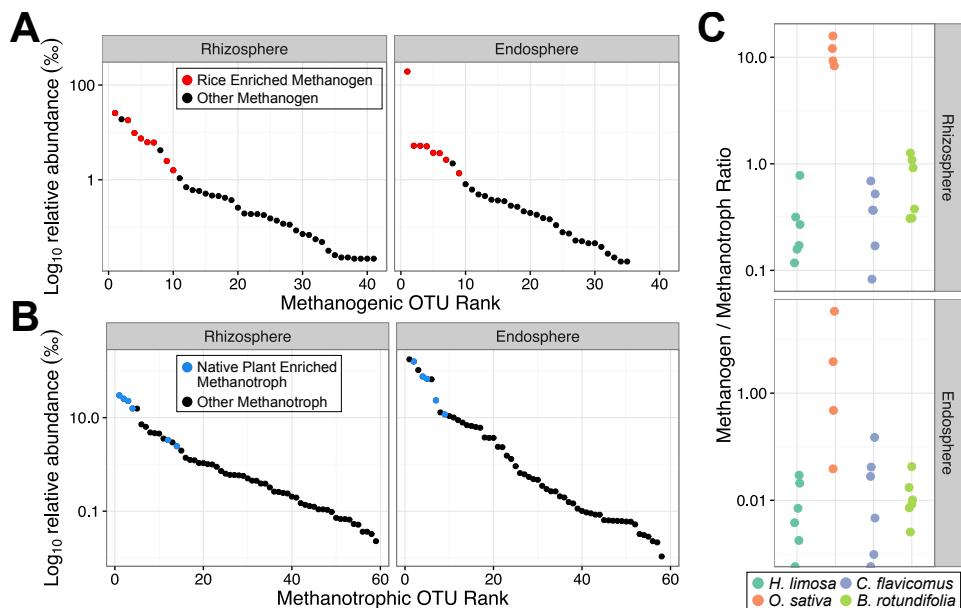


Figure 3.4: Figure 4. **The rice rhizospheric microbiome is enriched for methanogenic microbes.** (A) Rank-abundance plots indicating methanogens enriched in the rice rhizosphere and endosphere microbiomes. (B) Rank-abundance plots indicating methanotrophic eubacteria enriched in the native plant microbiome. (C) Methanogen to methanotroph ratios in each compartment of each plant species.

Methane can be oxidized by methanotrophic eubacteria under aerobic conditions. Methanotrophs have previously been reported to inhabit the rice rhizosphere, rhizoplane, and endosphere and thus may partially mitigate CH₄ release into the atmosphere. We were inter-

ested in how methanotrophic populations differ between host species' rhizospheres and whether different plant species enrich for certain methanotrophic taxa. In rice, we found two OTUs to be significantly enriched in the endosphere compared to at least one native plant species (with one OTU 4752 being enriched compared to *C. flavicomus* and OTU 4736 being enriched compared to *B. rotundifolia*). Both of these OTUs belong to the Methylosinus genus. These two OTUs, however, were not significantly enriched in the rice endosphere compared to the other native plants, suggesting that OTU 4752 and OTU 4736 are specifically depleted in the *C. flavicomus* and *B. rotundifolia* endospheres, respectively.

We next looked for methanotrophs that were significantly enriched in the rhizosphere or the endosphere of at least one native plant compared to rice. We found 6 unique methanotrophic OTUs enriched in the rhizosphere and 5 unique methanotrophic OTUs enriched in the endosphere of at least one native plant compared to rice. The rhizosphere enriched methanotrophs were among the top 14 of 59 most abundant methanotrophs in our rhizosphere dataset and the endosphere enriched methanotrophs were among the top 9 of 58 most abundant methanotrophs in our endosphere dataset (Fig 3.4B). Overall relative abundance of methanotrophs was significantly higher in the rhizosphere of native plants than rice ($P = 4.789e-09$, Welch two sample t-test), but not the endosphere ($P = 0.749$, Welch two sample t-test).

We next compared the relative abundance ratios of methanogenic archaea to methanotrophic bacteria in each plant species. The rhizosphere generally had higher ratios of methanogens to methanotrophs compared to the endosphere (Fig. 3.4C). This is expected as roots contain the highest levels of oxygen in an otherwise anoxic environment and methanotrophs flourish under aerobic conditions (while the opposite is true for methanogens). We found that rice has a significantly higher ratio of methanogenic microbes than methanotrophic bacteria in both the rhizosphere and endosphere compared to native plants growing in the same environment. The native plants had mean ratios < 1 in both the rhizosphere and endosphere, while rice had mean ratios > 1 in both the rhizosphere and endosphere. Without knowing the relative activity of methanogens and methanotrophs in our datasets it is not possible to conclude that the rice root microbiome is a net producer of CH₄ and the

microbiomes of the native plants are net CH₄ sinks, however it is reasonable to hypothesize that the rice root microbiome produces more CH₄ than that of the native plants.

3.3.5 Soil cultivation history affects plant root microbial assemblages.

When comparing the rhizosphere microbial community structure to that of unplanted (bulk) soil from the same field site, we noticed that rice rhizosphere communities are significantly more similar to bulk soil microbial communities than are native plant rhizosphere communities (Fig. 3.2B and Fig. 3.5A). Endosphere microbial community structure of the tested plant species, however, differed equidistantly from bulk soil communities. Two non-mutually exclusive hypotheses arise from these results. The first hypothesis is that dense monoculture of rice over the course of multiple years can change the soil's microbiome to be more similar to that of the rice rhizosphere. Many rice fields in the USA are planted densely such that nearly all areas within a field are impacted by rice roots, thus most of the surface soil is effectively rhizosphere. The second hypothesis is that, compared to the native plants growing in the field, rice has a weaker ability of differentiating its rhizosphere microbiome from that of the surrounding soil microbial communities.

To test the first hypothesis, we grew rice plants (California cultivar M206) in soil from two rice fields (from here on referred to as “domesticated soil”) as well as soil from two non-agricultural sites (i.e. soils in which no crop plants have been grown, from here on referred to as “wild soil”) under controlled conditions. This experiment was carried out in two batches. In the first batch, plants were grown in a domesticated soil from Arbuckle, CA and a wild soil from Sacramento, CA (Fig. 3S2). In the second batch, rice plants were grown in a domesticated soil from Biggs, CA and a wild soil from Arbuckle, CA (Fig. 3S2). Along with unplanted soil controls, we collected DNA from rhizosphere and endosphere samples and sequenced the V4 region of the 16S rRNA gene. The root microbial communities acquired from wild soils were significantly different and clustered distinctly from microbiomes those acquired from domesticated soil (Fig 5C, R² = 11.9, P = 0.001). We noticed a significant interaction term between soil history and root compartment (R² = 0.063, P = 0.001, PerMANOVA). Although each compartment's microbiome was significantly affected by soil history, the rhizosphere communities were more affected by soil history than the

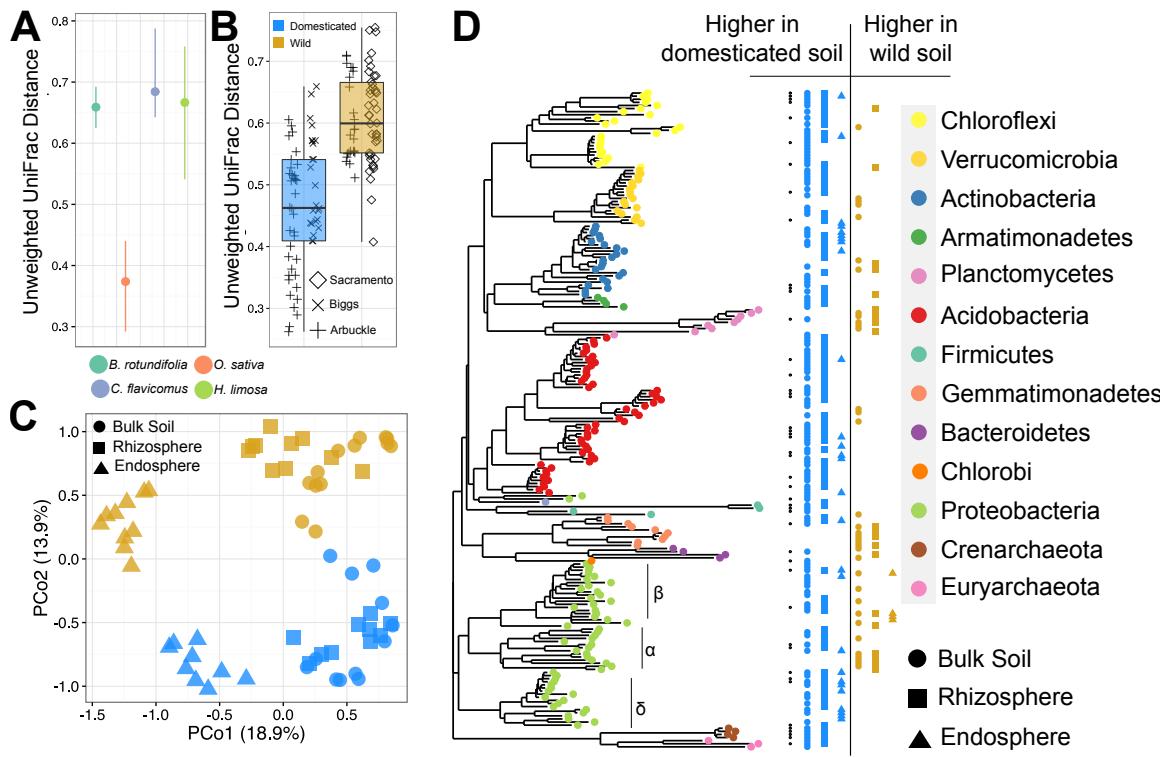


Figure 3.5: Figure 5. Rice cultivation has an impact on the overall soil microbiome. (A) Unweighted UniFrac distances between the rhizosphere microbiomes of rice and native plants compared to unplanted soil from the same field. (B) Unweighted UniFrac distances between the microbiomes of unplanted wild or domesticated soils from different locations and the assembled rhizosphere microbiomes of isogenic rice grown in each soil. (C) Principal coordinates analysis of unweighted UniFrac distances of microbiomes sampled from isogenic rice plants growing in domesticated and wild soil from different sites. (D) Phylogenetic tree showing relationship between OTUs found to be enriched in microbiomes from wild and domesticated soils. Black points represent OTUs found to be specifically enriched in rice compared to native plants (Fig. 3.3B).

endosphere communities ($R^2 = 0.30$ vs 0.22 , respectively, $P = 0.001$, PerMANOVA). The rhizosphere microbiota of plants growing in wild soils were significantly more distant to unplanted soil controls than rhizosphere microbiotas of plants growing in domesticated soils (Fig. 3.5B, $P = 7.29e-16$, ANOVA). Root-associated microbiomes assembled from wild soil had greater variability than those assembled from domesticated soils (Fig. 3S3 Rhizosphere: $P = 0.006$, Endosphere: $P = 0.019$). Taken together, these observations support the hypothesis that rice cultivation \rightarrow domesticates \rightarrow the microbiome of the soil environment to be more similar to the rice rhizosphere microbiome.

We were next interested in finding individual taxa responsible for the acquired microbiome differences between plants grown in domesticated and wild soil. Again we used a linear model approach where the log₂ abundance of each OTU in each rhizospheric com-

partment was modeled as a function of soil history status. Because this experiment was carried out in two batches, we decided to model each experimental batch separately and find the overlap of OTUs that were significantly enriched in each compartment of domesticated and wild soils between the batches. We found a total of 252 OTUs to be enriched in the bulk soil, rhizosphere, and endosphere of domesticated soil (165 unique OTUs) while we found 60 OTUs to be enriched in the bulk soil, rhizosphere, and endosphere of plants grown in wild soil (47 unique OTUs). Soil cultivation history appeared to broadly affect the abundance of individual taxa from several phyla: Acidobacteria, Verrucomicrobia, Chloroflexi, Euryarchaeota, and Actinobacteria members were enriched in the bulk soil, rhizosphere, or endosphere of plants grown in domesticated soils, while Gemmatimonadetes and Planctomycetes members were more present in the microbiomes assembled from wild soils (Fig. 3.5D, hypergeometric test, adjusted $P \leq 0.05$). Similarly, we found that members of the Deltaproteobacteria class are uniquely enriched in the domesticated soil assemblages.

We compared the OTUs that were found to be significantly enriched in the microbiomes assimilated from domesticated soils to those OTUs found to be specifically enriched in the rice rhizosphere and endosphere microbiomes compared to native plants. Of the 165 unique OTUs enriched in microbiomes originating from the domesticated soils, we found an overlap of 46 OTUs between the rice specific taxa (black data points, Fig. 3.5D). The magnitude of this overlap was greater than expected by chance alone ($P = 0.0008$, hypergeometric test), implying that rice plant physiology is a major structuring factor for rice field soil microbiomes.

3.4 Discussion

Our analysis of various plant species cultivated in different environments demonstrates that rice assembles a relatively distinct root-associated microbiome amongst crop species, likely because of the submerged conditions under which the plants are cultivated (Fig. 3.1B). Indeed, after sampling the roots of native plants growing in a rice field, we observed that soil submergence appeared to have an effect on root-assembled microbiomes (Fig. 3S1). However, we found that despite growing under similar submerged conditions, root-

associated microbiomes are significantly dependent upon plant species. The three native plants species, each of which is a common weed found in rice fields across the U.S., assemble microbiomes with profiles clustering distinctly from rice. Through differential abundance analysis, we determined that the rice microbiome clusters separately from the native plant species' microbiomes due to rice roots enriching for a relatively unique microbiome. Conversely, the taxa enriched in the native plant microbiomes (compared to rice) overlap to a much reduced extent. We found that root-associated microbiome differences species were disjunct from host plant phylogeny. Similar results were reported in *Arabidopsis* relatives [48]. In a study comparing domesticated rice species to wild rice species, it was found that microbiome differences only weakly correlated with phylogenetic differences; however, larger microbiome differences were observed based upon domestication status [50]. Together, these results suggest that the culmination of physiological traits governing root-associated microbiome assembly may be arrived at through convergent evolution.

Methane emission from rice paddies is a complex phenomenon governed by many environmental factors [38]. We found that the rice root-associated microbiome is enriched for methanogenic archaea compared to the native plants. Rice plants host a higher methanogen to methanotroph ratio than the native plants, suggesting that methane emission levels from submerged environments may be host species-specific. Early divergent members of the *Oryza* genus were also found to host a higher proportion of methanotrophic eubacteria compared to cultivated members of the genus [50]. Taken together, these results indicated that domestication of rice may have indirectly selected for a microbiota with a higher capacity for hosting methane producing microbes. The root architectural and physiological parameters governing methanogen to methanotroph ratios in root microbiomes remain elusive; however, native plant species growing in similar environments as rice may be a resource for understanding the physiological and root architectural parameters governing associations between plants and methanogens and methanotrophs. Because CH₄ emissions vary significantly across rice genotypes [52], CH₄ emission could be treated as a complex trait and correlated with genomic loci to have a better mechanistic understanding of which plant traits govern the composition of rhizospheric microbiomes in respect to greenhouse

gas emissions.

Intensive agriculture has unknown consequences on the microbiota of agricultural ecosystems. We observed that the rhizosphere microbiome of rice is significantly less dissimilar to unplanted soil communities in agricultural fields than are those of native plants. We proposed two non-mutually exclusive hypotheses for this: 1) intensive cultivation over multiple seasons allows the roots of rice plants to modify the soil microbiome to be more like the rice rhizosphere and 2) that rice has a weaker ability to differentiate its rhizosphere microbiome from the surrounding soil microbiome than native plants. While we were unable to address the latter hypothesis in this study, we tested the first hypothesis by growing rice plants in soils collected from non-agricultural locations (i.e. wild soils, soils from uncultivated land) and soils from rice fields (i.e. domesticated soil). In support of our hypothesis, we found that plants growing in wild soils have a larger rhizosphere effect (i.e. a greater difference between rhizosphere and bulk soil microbiomes) than plants growing in domesticated soils. Furthermore, we found that soil history has a prominent effect on both the rhizosphere and endosphere microbiomes. Community structure variability was significantly higher within the rhizosphere and endosphere microbiomes assembled from wild soils compared to the domesticated soils, suggesting that rice cultivation selects for a consistent microbiome. While our results support the hypothesis that rice plants exert a large effect on the soil microbial communities in rice fields, we cannot exclude the possibility that native plants also have a stronger ability to differentiate their rhizospheric microbiomes from that of the surrounding soil.

Despite variability in wild soils, we found that assembled microbiomes are discernible based upon soil history. These differences are reflected in the taxa that have significantly different relative abundances between the microbiomes assembled from the two soil types. We found many more taxa to be enriched in the microbiomes assembled from domesticated soils than from wild soils, consistent with the observation of higher community variability in wild soils. We observed a significant overlap of 65 OTUs between domesticated soil enriched OTUs in California and the rice specific enriched OTUs from the native plant study in Arkansas. This statistic is impressive given that the studies were conducted in two

disparate, but rice cultivating, regions of the United States. Taken together, these results imply that agricultural soil microbiomes are not only shaped by cultivation practices [18], but also by unique aspects of the crop plants' physiologies. The functional benefits and / or disadvantages of modifying soil microbiomes to better resemble a plant's rhizospheric microbiome are unknown. Amending soil quality and nutrition by rotating cover crops in agricultural settings has been employed by farmers for centuries but, the functional link between soil health and microbial community structure remains largely unresolved. Understanding the relationship between crop physiology, rotation, and productivity and plant-microbe and microbe-microbe interactions should thus be an important focus of future research.

3.5 Materials and Methods

3.5.1 Meta-Analysis

Assembled 16S rRNA gene sequences and metadata were downloaded from the QIITA database for maize (Peiffer et al. 2013, QIITA ID number: 1792), grape (Zarraonaindia, et al., QIITA ID number: 2382), and cannabis (Winston et al., 2013, QIITA ID number: 1001). Sequences for *Boechera* (Wagner et al. 2016) were downloaded from European Nucleotide Archive (study number: PRJEB10570) and metadata was downloaded from the study's associated Dryad repository (<http://datadryad.org/resource/doi:10.5061/dryad.g60r3>). The paired-end sequences were trimmed to remove primer regions and then assembled into contiguous sequences using PANDAseq [35]. For the rice dataset, we only used sequences from field grown plants in Edwards et al. 2015 [18]. Assembled reads were discarded if containing any ambiguous bases or if the read was less than 275 bases in length. The sequences were clustered based upon 97% sequence identity into operational taxonomic units (OTUs) against the GreenGenes [15] 13_8 database using the NINJA-OPS pipeline [1]. NINJA-OPS is a closed-reference, fast, and memory efficient pipeline that uses bowtie2 [29] to map amplicon reads to a simulated chromosome (concatenated reference sequences from the GreenGenes 13_8 database).

3.5.2 Native Plant Study

Roots with attached soil from rice and native plants, along with unplanted soil, were collected from a single rice paddy near Jonesboro, Arkansas (USA) on August 22, 2015. Rhizosphere and endosphere fractions from each host plant was collected according to the methodology used previously [18]. The roots were shaken by hand to remove loosely bound soil and then placed into 50 mL Falcon tubes with 15 mL of sterile PBS solution. The samples were stored on ice and shipped overnight to the University of California, Davis. Upon receiving the samples, the rhizosphere fraction was removed from the roots by vortexing in PBS solution and 500 μ L of the resulting soil slurry was placed into a sample tube for DNA extraction. The roots were vortexed in consecutive washes of fresh PBS solution until all soil was depleted and then sonicated 3 times at (x Hz) for 30 seconds in fresh PBS to remove all rhizoplane microorganisms. The remaining roots were placed into a sample tubes for DNA extraction. All DNA extractions were performed using the MoBio Powersoil DNA isolation kit.

3.5.3 Soil Domestication Study

The soil domestication study was conducted in two batches using 4 different soils. In the first batch, soil was collected on April 10, 2015 from a non-agricultural site near Sacramento, California and from a rice field near Arbuckle, California. The second batch of soils was collected on June 3, 2016 from a non-agricultural site near Arbuckle, California and a rice field at the Rice Experiment Station in Biggs, California. The soils were brought back to the greenhouse, homogenized, placed into pots, and inundated with water. Axenic rice seedlings (see method below) were transplanted into each pot and supplemented with nutrient solution every 14 days. For each batch, roots were harvested from the plants 6 weeks after transplantation. The roots were shaken to remove loosely attached soil, leaving 1mm of soil attached. Compartment separation and DNA isolation was performed as described above.

3.5.4 Seed Sterilization and Germination

Rice seeds were stripped of their hulls by vigorous back and forth rolling in a 5 inch cut portion of a bicycle tire tube. The seeds were shaken in 70% bleach for 5 minutes then washed three times in autoclaved deionized water. Using ethanol flamed forceps, the seeds were then transferred into petri plates containing half strength MS agar (1%). The plates were placed in a dark growth chamber at 30 °C for 7 days. The seeds were then placed upon the bench top in order to de-etiolate. Any plates containing fungal or bacterial growth were discarded.

3.5.5 16S rRNA gene amplification and sequencing

All 16S rRNA gene amplification was performed as noted in Edwards et al., 2015. Briefly, the V4 region of the 16S rRNA gene was amplified using PCR with a dual indexing strategy. For each PCR reaction, a corresponding negative control was also performed. All reactions were checked for amplification by running PCR products out on a 1% agarose gel. If a reaction's negative control succeeded in amplification, then we discarded the particular reaction and reperformed the PCR. The PCR reactions were purified using AMPure beads and measured for concentration using a Qubit. The PCR products were pooled in equimolar concentrations, concentrated using AMPure beads, and then gel extracted from a 2% agarose gel. Sequence libraries were sent to the University of California DNA Technologies Core Laboratory for 250 x 250 bp sequencing on the Illumina Miseq platform.

3.5.6 Sequence processing

The resulting paired end sequences were demultiplexed using custom Python scripts and aligned into contiguous reads using PANDAseq [35]. The contiguous reads were discarded if containing any ambiguous bases or if the length exceeded 275 bases. All reads were then clustered into OTUs based upon 97% sequence identity using NINJA-OPS as described above. OTUs with plastid and mitochondrial taxonomies were removed from all resulting OTU tables. All scripts used for sequence processing can be found at ([github](#)).

Statistical Analysis Weighted and Unweighted UniFrac distances were calculated using QIIME [11]. All other statistical analyses were conducted using R version 3.1 [45]. Unless

otherwise noted, we determined statistical significance at $\alpha = 0.05$ and, where appropriate, corrected for multiple hypothesis testing using the Benjamini and Hochberg method [6]. Shannon diversity was calculated using the `diversity()` function, PCoA and CAP analyses were conducted using the `capscale()` function, perMANOVA was conducted using the `adonis()` function, and community structure variability was measured using the `betadisper()` function all from the Vegan package [42]. Linear models were run using the `lm()` function and ANOVA was run using the `aov()` function. Hypergeometric tests were run using the `phyper()` function. Phylogenetic trees were displayed using the `plot_tree()` command from the PhyloSeq package [36]. All other graphs and plots were generated using the `ggplot2` package [62]. R notebooks for the full analyses can be found ([github](#)).

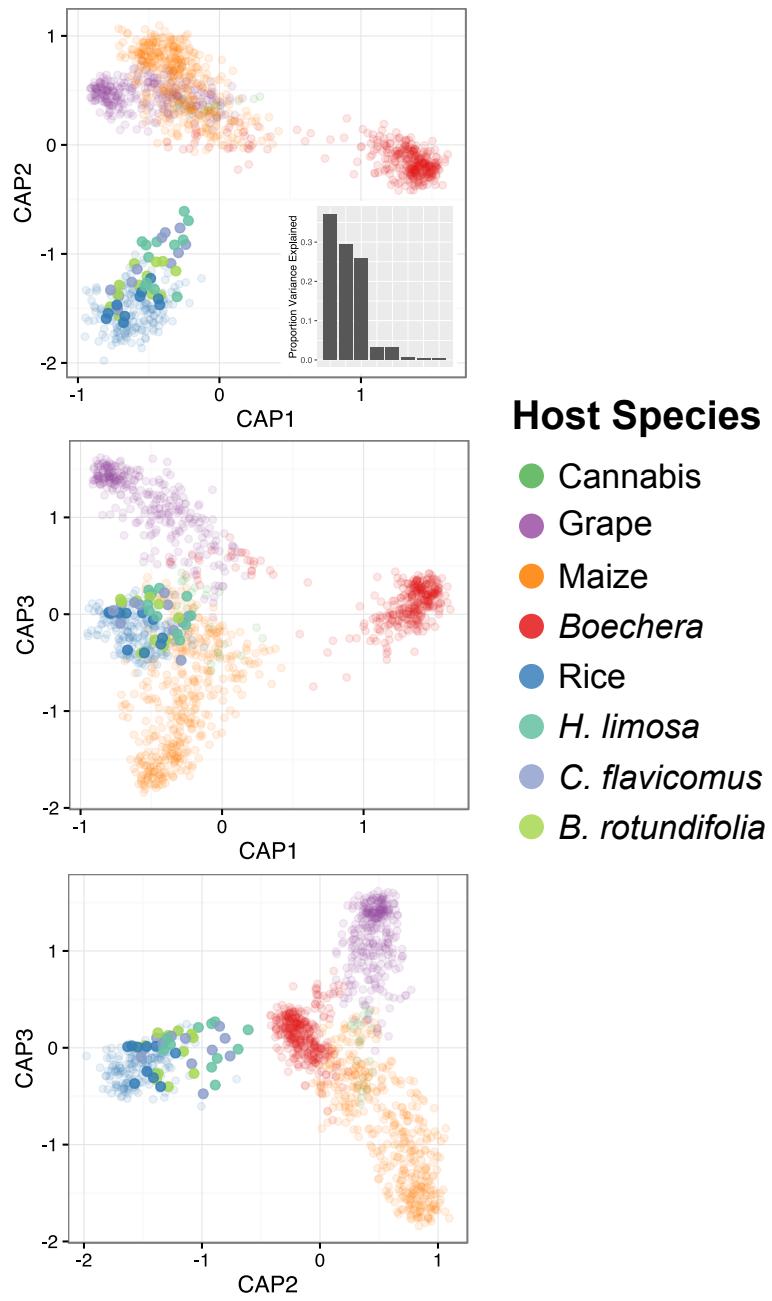


Figure 3S1. CAP analyses of published datasets displaying where the native plants taken from an Arkansas rice field cluster along the first three axes. The inset in the uppermost panel depicts the amount of variance explained by each axis.

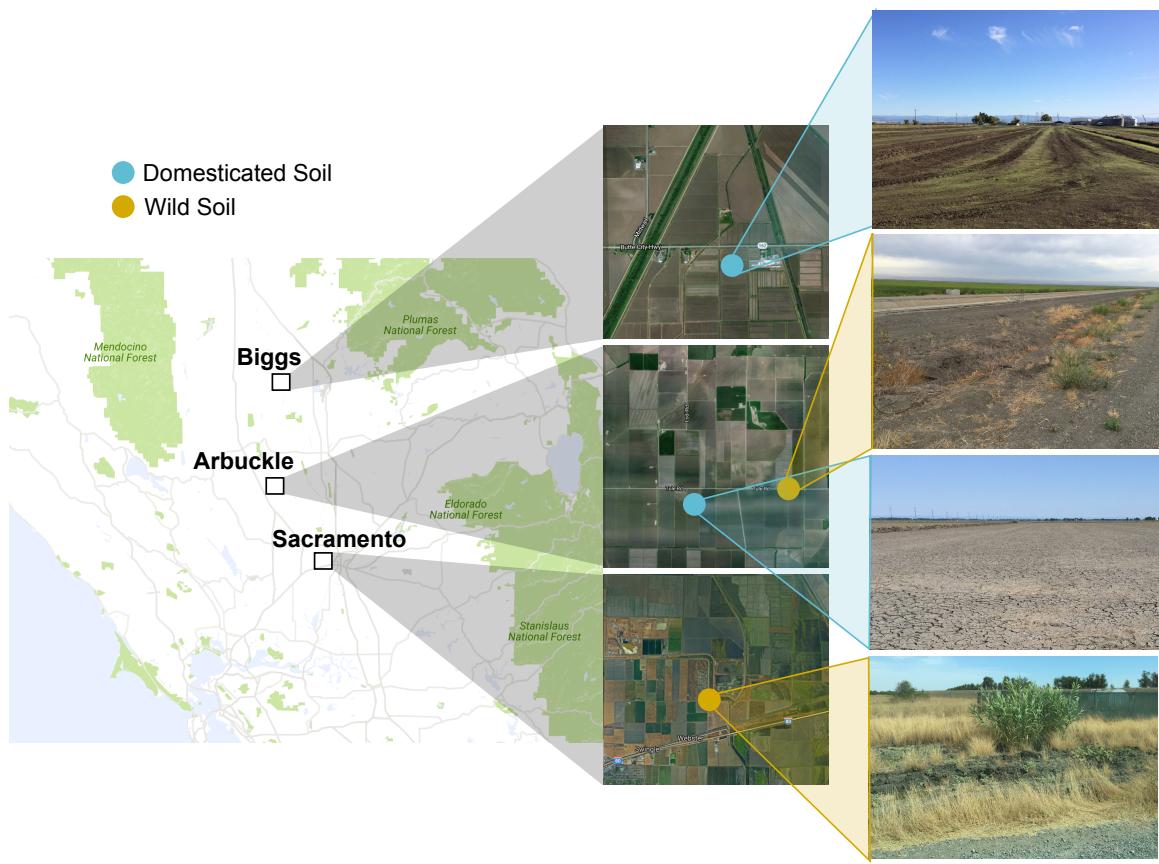


Figure 3S2. Map of where wild and domesticated soils were collected.

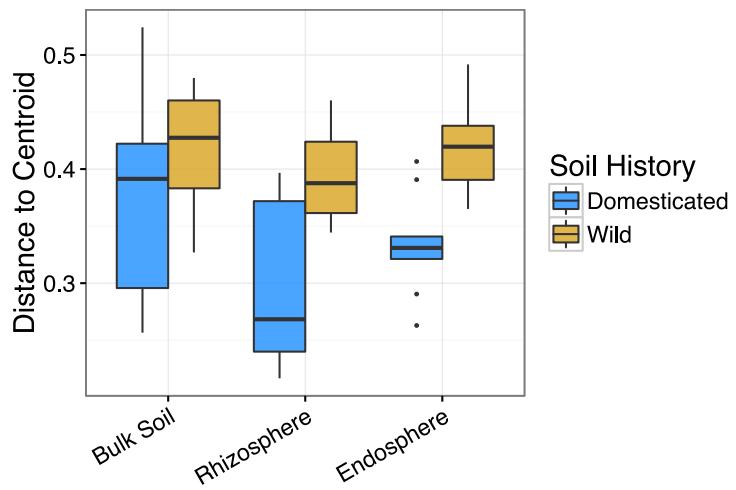


Figure 3S3. Dispersion of beta diversities for each sample type from domesticated and wild soils.

Chapter 4

Successional shifts in the root-associated microbiota throughout the life cycle of the crop plant *Oryza sativa*

Joseph Edwards¹, Christian Santos¹, Zach Liechty¹, Bao Nguyen¹, Nicole Chan¹, Eugene Lurie¹, and Venkatesan Sundaresan¹

4.1 Abstract

Much like how humans host a gut microbiota, plants assemble a root-associated microbiota which can aid the plant in assimilating nutrients, reducing susceptibility to pathogens, and alleviating the severity of abiotic stresses. Plants host microbiota in at least 3 distinct root-associated compartments: the rhizosphere, rhizoplane, and endosphere. It is unclear how the communities in these compartments shift over the course of the plant's life cycle or whether the transition to the reproductive phase has any effect on the root-associated microbiota. Here, we characterize the root-associated microbiota of the cereal crop *Oryza sativa* over the course of 3 seasons. We found that the microbiota in each root-associated compartment shifts significantly over the life cycle of the plant and that these shifts are independent from the shifts in unplanted soil and are consistent from season to season. We found that the root microbiota shifts rapidly until the entry into the reproductive phase, whereupon the microbiota stabilizes. Using a machine learning framework, we identified

¹Department of Plant Biology, University of California, Davis

taxa that were discriminant of plant age. We next investigated whether there exists a developmental stage-specific microbiota by sampling rice varieties with different developmental progression rates. Again using a machine learning approach, we were able to identify bacterial taxa whose abundance patterns are discriminant of plant developmental stage. These results indicate that changes in plant development either directly or indirectly impact the root-associated microbiota.

4.2 Introduction

Plants derive root-associated microbial communities from the soil which are inhabited by thousands of different species and strains of archaea and bacteria [34, 9, 44, 18, 60]. Individual taxa within the root-associated microbiota have been found to be beneficial for plant growth and resistance to biotic and abiotic stresses [10, 7, 37]. In rice, we previously identified three spatially distinct root compartments with significantly different microbiota compositions: the soil adjacent to the root (the rhizosphere), the root surface (the rhizoplane), and the root interior (the endosphere) [18]. It was found that each of these root-associated compartments have significantly different microbiota profiles from the communities in unplanted soil, thus indicating the roots enrich for subsets of the soil microbiota. This enrichment process and how the root-associated microbiota shift throughout the lifecycle of rice plants remain uncharacterized.

In humans, the successional shifts in the gut microbiota have been well studied from infancy into adulthood [27, 40, 5, 65]. Newborns are characterized by a low diversity gut microbiota that increases with age [27]. The composition of the human gut microbiota shifts rapidly during the first three years of life, then stabilizes to become more like an adult microbiota after three years and these shifts correlate well with dietary intake [65]. Whether a similar developmental trend occurs within the root-associated microbiota of plants remains unclear. In this study, we detail the succession of microbial consortia in the three root-associated compartments of rice while being cultivated under field conditions. We also characterize how plant developmental progression is associated with the succession of various taxa.

4.3 Results

4.3.1 Experimental Design and Sequencing

We first were interested in monitoring how root-associated microbiotas shift throughout the life cycle of the rice plant under field conditions and whether changes in community profiles were consistent across multiple seasons. To do this, we collected the rhizosphere, rhizoplane, and endosphere microbiota samples from rice plants grown in a submerged commercial rice field in Arbuckle, California over the 2014 and 2015 seasons. In 2014, the field was sampled weekly until harvest, a total of 19 weeks with each time point consisting of 8 replicates. In 2015, because we were interested in the initial root colonization stage, we sampled every other day for the first 4 weeks of growth, thereafter sampling every other week until harvest. Each time point in 2015 consisted of 4 replicates. For each season, we collected bulk soil (i.e. soil from the submerged field where no roots were growing) until root growth made it impossible to locate soil unaffected by rice roots: around 4 weeks for the 2014 season and 6 weeks for the 2015 season.

PCR amplification and Illumina sequencing of the 16S rRNA gene was conducted to profile microbial communities present in each sample. Operational Taxonomic Unit (OTU) clustering was done at the 97% threshold using the NINJA-OPS pipeline [1]. We detected a total of 24,048 OTUs across 1,290 samples with a total of 48,954,921 sequences. We removed a total of 276 detected OTUs matching mitochondrial or plastidial taxonomies. We then removed samples from our dataset with less than 1000 total reads. One difficulty of microbiome studies is to perform statistics on non-reproducible OTUs. We removed OTUs from our dataset that were not present in at least 5% of the total samples. This step reduced our dataset to 8,554 OTUs for analysis.

4.3.2 Microbiomes are shaped by distance from the root and plant age

To understand the underlying driving forces of microbial community variation in our data, we used Principal Coordinates Analysis (PCoA) of Bray-Curtis distances. We found that samples from 2014 and 2015 display a spatial pattern of divergence along the first principal coordinate, where communities in bulk soil samples cluster on one end of the axis

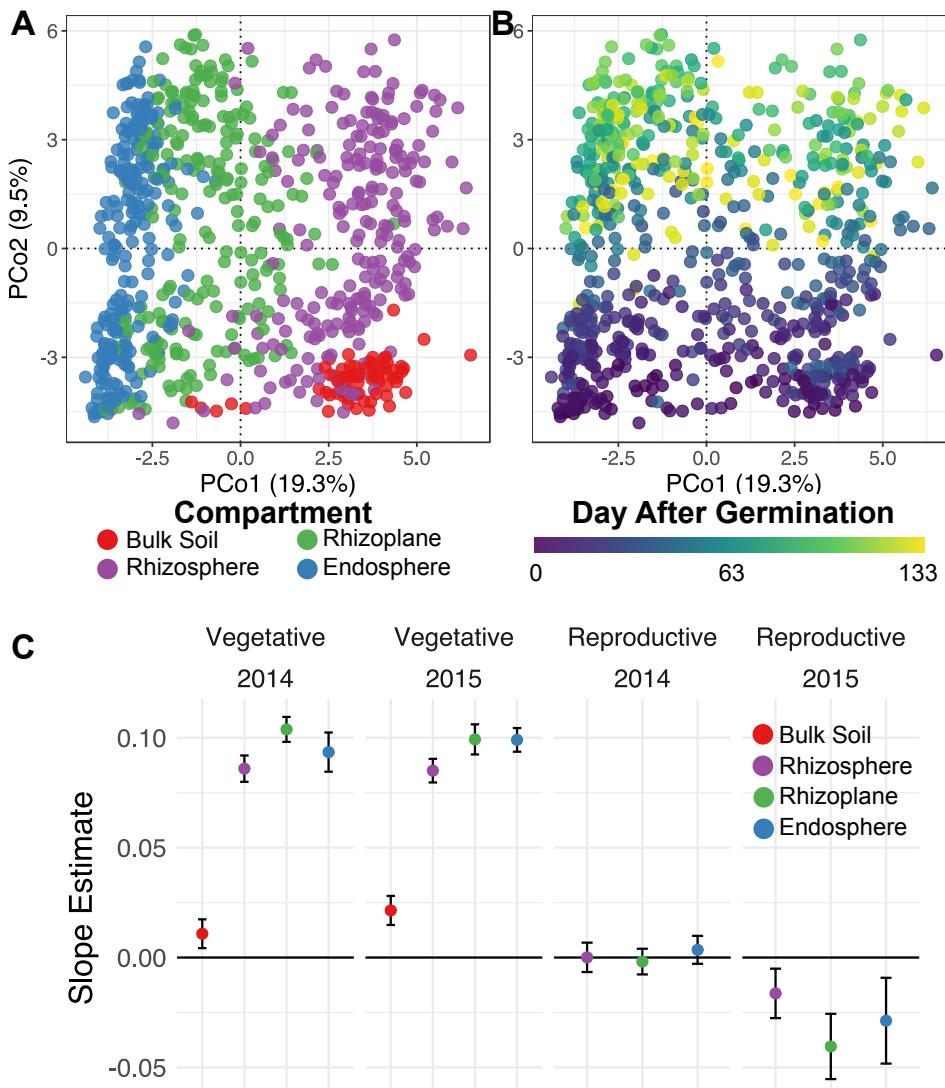


Figure 4.1: The root-associated microbiota is differentiable by root compartment and plant age. (A) Principal coordinate analysis (PCoA) of Bray-Curtis distances between samples colored by root-compartment. (B) The same plot as in A but colored by the age of the plant from which the samples were obtained. (C) The estimated slope of the second principal coordinate of plot A as a function of plant age. We fit separate linear models to estimate that slope during the vegetative phase and the reproductive phase (defined as plants older than 63 days) for each compartment. The 2014 and 2015 seasons' data is included in plots A and B.

and endosphere communities cluster on the other (Fig. 4.1A). In agreement with our previous studies on the rice root-associated microbiomes [18], perMANOVA of pairwise distances between samples indicated that microbial communities differed significantly between root-associated compartments ($R^2 = 0.23$, $P = 0.001$).

We next measured the effect of plant age on the root-associated microbiota. PerMANOVA of pairwise distances between samples revealed that plant age had a significant effect on the root-associated microbiota ($R^2 = 0.21$, $P = 0.001$). We found that this effect was cap-

tured on the second principal coordinate of the PCoA plot (Fig. 4.1B). In the M206 cultivar, panicle initiation (entry into reproduction) occurs 56-63 days after germination [33]. We quantified the slope of the second principal coordinate in response to plant age for the vegetative stage (days < 63) and the reproductive stage (days \geq 63) using a linear model. We found that in each root-associated compartment, the rate of change (i.e. the linear slope estimate) was reduced in the reproductive phase compared to the vegetative phase (Fig. 4.1C). This effect was consistent across the two seasons, suggesting that the root-associated microbiomes shift over the course of the plant's vegetative life cycle, but then stabilize at the onset of reproduction. Although we could not collect bulk samples during the reproductive phase of each season, we noticed a reduced rate of change among bulk soil samples compared to root-associated samples during the vegetative stage of growth. These results suggest that shifts in the root-associated microbiota over the course of the season is mainly driven by plant processes and not directly by changes in edaphic factors.

The multi-season sampling scheme of our experimental design allowed us to quantify the effect of seasonal variation on the root-associated microbiota. Although the root-associated microbiota varied significantly across the two seasons, the effect was small ($R^2 = 0.007$, $P = 0.001$) compared to the other factors analyzed within this experiment. Together, these data suggest that the root-associated microbiota shifts in each root-associated compartment during the vegetative growth stage of the season and stabilizes upon entry into reproduction and that these patterns are consistent across multiple growing seasons.

4.3.3 Microbiota shifts over the season are marked by increasing and decreasing relative abundance of specific phyla

We next sought to characterize the specific phyla responsible for the significant differences between the root-associated compartments and how these various phyla changed in abundance over the lifecycle of the rice plants. Figure 4.2A shows the relative abundance for the top 11 most abundant phyla between the compartments and how their abundance changes over the course of the season. Because the Proteobacteria phylum contains a broad phylogenetic makeup and because Proteobacteria make up the vast proportion of the rice root-associated microbiota, we further dissected the Proteobacteria phylum into its respective

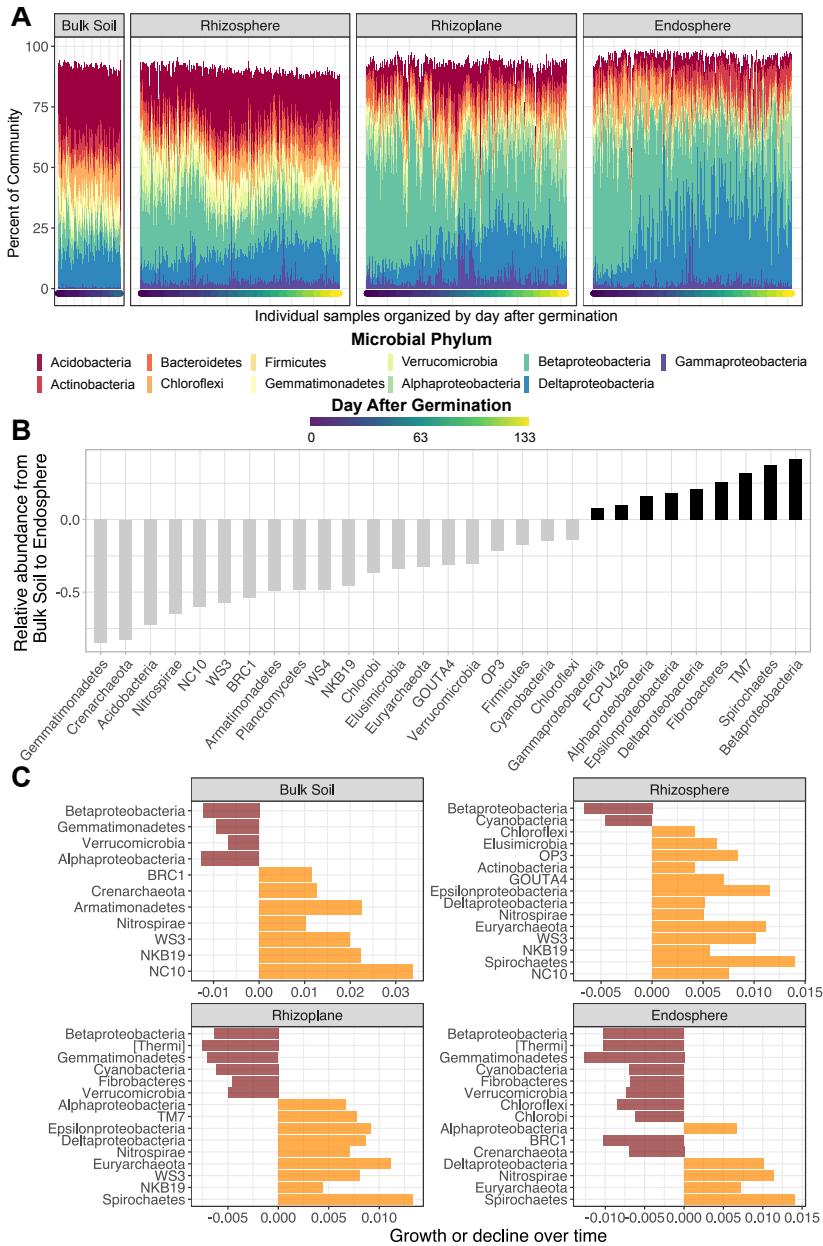


Figure 4.2: Shifts in the microbiota over time are associated with increasing and decreasing phyla. (A) Bar plots of the top 11 phyla abundance over the course of the seasons in each compartment. Each bar represents 1 sample that was taken throughout the course of the growing season. The bars are ordered by the age of the plant as indicated by the colored points beneath each bar. Both the 2014 and 2015 data were used for this graph. (B) Beta regression coefficient estimates for microbial phyla that are either increasing (above 0) or decreasing (below 0) in relative abundance from the outside of the root to the inside of the root. (C) Beta regression coefficient estimates for microbial phyla that are increasing (above 0) or decreasing (below 0) in relative abundance over the course of the seasons in each compartment.

classes for this analysis. To model increasing or decreasing relative abundance of individual phyla between the various root-associated compartments, we assigned each compartment a value relative to its spatial position: bulk soil was position 1, rhizosphere was position 2, rhizoplane was position 3, and endosphere was position 4. We modelled how phyla either

increased or decreased between these positions using beta regression. Beta regression is used for modeling dependent variables which lie in the interval (0, 1) and is thus useful for modelling individual taxa as a relative proportion of the total microbial community.

Using this method, we were capable of identifying phyla which significantly differ in spatial distribution from the exterior to the interior of the root. Figure 4.2B indicates the degree of change for phyla whose relative abundance were significantly different across the root-associated compartments. The phyla are ranked in terms of how pronounced the changes are between the exterior and interior of the root. Those taxa which have values below zero on the y-axis are higher in relative abundance outside of the root, while taxa with a value greater than zero are higher in relative abundance inside of the root. Our analysis revealed more taxa with negative trends from the outside to the inside of the root which is consistent with the hypothesis that the root is a selective niche unsuitable for the growth of a large array of soil microbes. Many of the phyla found to be significantly enriched from the soil to the root interior belong to various classes within the Proteobacteria phylum. We found both Spirochaetes and Fibrobacteres to increase in abundance from the outside to the inside of the root. Both phyla are known to degrade cellulose, thus it is probable that these microbes are using the root tissue as an energy source or as a molecular cue for colonization [28, 46].

We performed the same type of analysis to identify phyla that significantly increased or decreased in relative abundance over the course of the growing season in each compartment. Figure 4.2C shows the result of this analysis where we display the top 15 phyla in each compartment which show the highest degree of change over the course of the season. Below zero on the x-axis represents phyla that were reduced in relative abundance over the season, while values above zero represents taxa that increased in abundance over the season. The plant associated compartments showed similarities in phylum increase or decrease over the season. For instance, in the rhizosphere, rhizoplane, and endosphere Betaproteobacteria consistently decreased over the course of the season, while Deltaproteobacteria, Euryarchaeota, and Spirochaetes all increased. These results indicate that the relatively large variances partitioned to differences in microbial communities across the

root-associated compartments and over the course of the growing season can be explained by significant proportional shifts in various phyla.

4.3.4 The composition of the root-associated microbiota is an accurate predictor of plant age

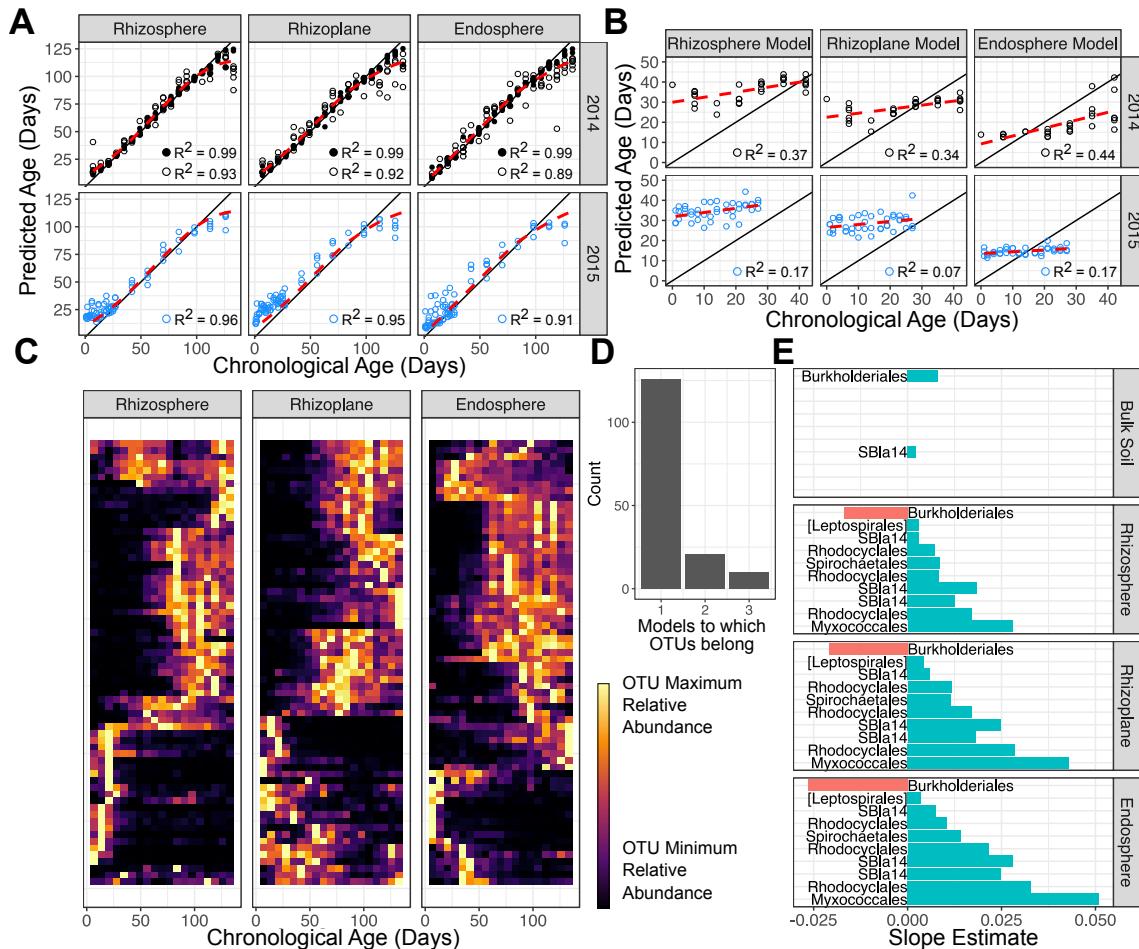


Figure 4.3: Figure 3. Random forests model detect taxa that are accurately predictive of plant age. (A) The result of predicting plant age using the sparse RF models for the 2014 and 2015 season. Each point represents a predicted age value, solid points were the samples on which the models were trained and the hollow points are samples in which the model never encountered during training. The solid black line represents the perfect fit line (i.e. if a sample were to be perfectly predicted, it would fall on this line). The dashed red line is a Loess regression curve fit to the predicted data from the 2014 season to show how the predicted values deviate from the real values. Each Loess curve is specific to the respective compartment from the 2014 data. These same curves are plotted along with the 2015 data to show that the areas of deviation from the perfect fit line are similar in the 2015 data. (B) The values of the bulk soil samples as predicted by the compartment specific RF models. The dashed red line is a linear fit for the bulk soil samples in each compartment's model showing how the predictions deviate from the perfect prediction line. (C) A heatmap representing the abundance profiles of the 66 most age discriminant OTUs in each compartment over the course of the season. (D) The number of models the OTUs belong to. A value of 3 means that the particular OTU was shared in all three compartment specific models. (E) Shows the slope estimates of the 10 shared OTUs between the compartments to indicate whether the OTU is increasing or decreasing as a function of plant age. The Order of the OTUs are shown in text next to the bars.

Because there were notable patterns of microbiota shifts throughout the growing season in each root-associated compartment, we next sought to identify patterns in the abun-

dance of various taxa to discriminate the age of the rice plants. This posed a formidable difficulty due to the underlying structure of our data and the biology of the microorganisms we are studying. In our dataset exist the abundance of many thousands of detectable microbes. The abundance pattern over time is unique for each detected microbe and some of these abundance patterns are useful for predicting the age of the rice plant while others have little to no utility for this purpose. Because of this problem, we settled on using the Random Forests (RF) algorithm, a machine learning technique, to model the age of the rice plants by using the abundance patterns of various microbes [8]. The RF algorithm is particularly suited for this task because it can incorporate multiple variables for consideration (i.e. using multiple OTUs for predictions), it does not assume linear trends in the input data nor does it require the data to have a parametric distribution, and it can return the importance values for the input features (i.e. it outputs how useful an OTU was for creating an accurate model).

To begin, we regressed the abundances of all 8554 detected OTUs in a training set of samples from each root-associated compartment separately as a function of plant age to form the “full” RF models. Because various taxa had different abundance profiles over time between each compartment (Figure 2A), we surmised that each compartment should be modeled separately in order to make the most accurate prediction. The training data for each compartment is a random selection of four of the eight replicates within each time point from the 2014 season. The full models are not optimally accurate because some OTUs which are not useful for predictions can introduce a significant amount of error. We then conducted 10-fold cross validation on each model while subsequently removing non-important OTUs in order to identify how many of the most important OTUs should be included in each model (Fig. 4S1). This analysis revealed a marginal decrease in cross-validation error when using greater than 66 of the most important OTU features for each model. Thus, for each root-associated compartment, we formed a “sparse” RF model using the most important 66 OTUs derived from each compartment’s full RF model. The sparse models were used for downstream analysis and were not subjected to any further parameter optimization.

We used the sparse models to predict the ages of plants from our test set of samples consisting of the remaining samples from the 2014 season which were not included in the training of the model. Each model predicted the new data accurately (Fig. 4.3A), with the lowest adjusted R² value being 0.89 in the test set of the 2014 endosphere data. The least accurate timeframe for each model was after 100 days where each model consistently underpredicted the plants' ages, perhaps due to the overall stability of the root associated microbiome during the reproductive phase. We also predicted the ages of plants from the 2015 season using these sparse model. Again, we found the models to be accurate in their predictions of plant ages based off of the 66 OTUs included in each model. We used the sparse models to predict the ages of bulk soil samples for the 2014 and 2014 seasons (Fig. 4.3B). Each model consistently mis-predicted the age of the bulk soil samples. These results suggest that the 66 age discriminatory OTUs included in each model fluctuate over the growing season, i.e. some OTUs have increasing abundance over the season while other have decreasing patterns (Fig. 4.3C), and these patterns of fluctuations are consistent across different seasons. Furthermore, the 66 OTUs in each model are specifically predictive for the root-associated compartments and not for bulk soil samples, suggesting that the fluctuations in abundance for the sets of 66 OTUs are driven by the host plants and are not directly influenced by edaphic factors.

We next examined the top age discriminatory OTUs in each sparse model. The models consisted of phylogenetically diverse sets of taxa (Fig. 4S2): included in the models were OTUs from 14 Phyla, 33 Classes, 55 Orders, and 58 Families. Interestingly there was minimal overlap in OTUs between the models: we found only 10 OTUs to be conserved between all three models, while 126 OTUs were unique to a model (Fig. 4.3D). Of the ten overlapping OTUs, only one had a significantly decreasing abundance over the season in each root-associated compartment while the remaining nine had significantly increasing abundance patterns over the season in each compartment (Fig. 4.3E). Only two of these OTUs were found to also have either a significant increasing or decreasing trend in the bulk soil samples. Taxonomically, 7 of the 10 OTUs belong to the Betaproteobacteria class, specifically within the orders of SBIa14 and Rhodocyclales. Taken together, these data indicate while

a small set of OTUs are shared between the models, the models use a broad phylogenetic diversity of OTUs to discriminate plant age.

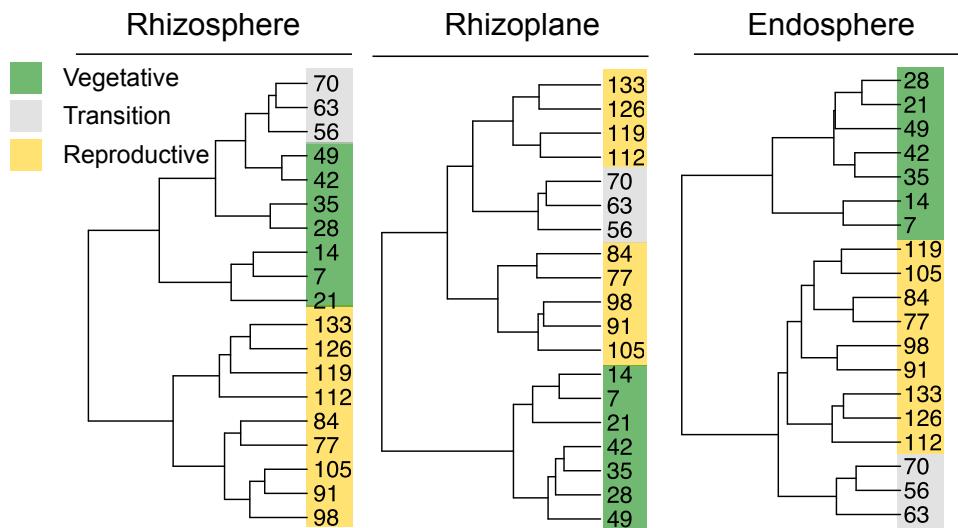


Figure 4.4: Hierarchical clustering of time points based upon abundance of the 66 most age discriminant OTUs in each compartment's sparse RF model. Colors represent the developmental stage that the plants were in when sampled. The numbers represent that host plant's age at the time of sampling.

We were next interested in understanding how the various time points relate to each other in terms of microbial community structure. That is, do time points cluster based upon developmental stages? To address this question, we performed hierarchical clustering using the 66 most age-discriminant taxa in each compartment (Fig. 4.4). The particular variety that was included in this study, M206, starts transitioning to the reproductive phase around 56 to 63 weeks after germination. In each compartment we found clear clustering of the time points during the vegetative phase. Consistently across each compartment, we found that the samples taken on day 56 through 70 cluster together. This is a transition time for the rice plants where they are switching from vegetative to reproductive growth. Interestingly, this transitory cluster falls closer to the reproductive or vegetative time points depending on the observed compartment. For instance, in the rhizosphere the transitory time points cluster with the vegetative samples and in the endosphere and rhizoplane the transitory time points cluster with the reproductive time points. This may suggest that the endosphere and rhizoplane microbiomes are the first to be affected by the transition to reproduction while the rhizosphere follows suit later.

4.3.5 Developmental stage is a driver of microbiota succession

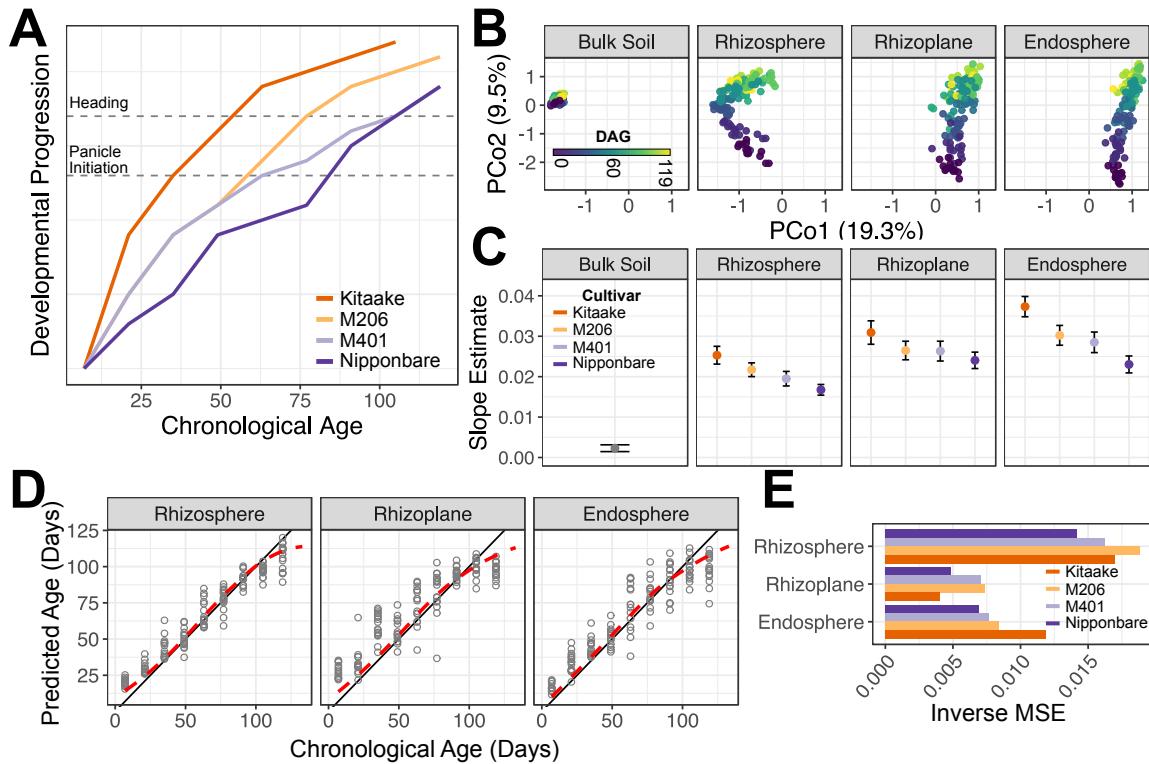


Figure 4.5: Varieties with different developmental rates have different microbiota progressions. (A) The developmental stage of the tested varieties as a function of plant age. We devised a numerical system for developmental progress based up morphological features indicative of developmental stage. (B) PCoA of the 2016 data indicating root-associated compartment and plant age are major determinants of microbiota structure. (C) Linear slope estimates for the principal coordinate 2 in panel B as a function of plant age for each variety and each compartment. A higher slope indicates that the microbiota is progressing faster along PCo2 as a function of plant age. (D) The predicted age of the 2016 data as predicted by the age discriminant sparse RF models. The solid black line is the perfect fit and the red dashed lines are the Loess curves calculated in Fig. 4.3A. (E) The inverse of the mean squared residual errors from panel D. A larger value means the model had more accurate predictions for the respective set of samples.

Our sparse models appear to be detecting shifts in the root-associated microbiota that correlate with developmental shifts in the rice plants. However, shifts in plant development also co-vary with shifting climatic and edaphic factors which may have indirect effects on the microbiota through various plant responses, thus making these factors difficult to uncouple from the effect of plant developmental stage on the shifting root-associated microbiota. Because our study was confined to one commercial field, we were limited to studying the dynamics of one rice variety throughout the 2014 and 2015 seasons which was synced in its developmental progression. To have a better idea of whether plant developmental stage significantly impacts the root-associated microbiota, we needed to un-sync plant age from developmental progression. In order to achieve this goal, we decided to grow rice vari-

eties with different developmental rates in the same field in Arbuckle, CA. Plant genotype, while significant, has a relatively small effect on the root-associated microbiota compared to other environmental effects (such as drought), thus we chose 4 varieties from the tropical *Japonica* sub-species of *Oryza sativa* for this study, all with different developmental progression rates (Figure 5A). Kitaake, a variety often used by research labs because of its relatively fast generation to generation time, reaches the flowering stage faster than the rest of the included varieties. We included the California varieties M206 and M401 into this study. M206 is the variety that we have previously been studying in the Arbuckle field and the variety for which the sparse RF models were generated. Although M401 and M206 reach the panicle initiation stage around the same time, M401 has a much later heading data. We also included the variety Nipponbare, which reaches panicle initiation later than the rest of the varieties, but flowers within the same timeframe as M401. These varieties were water seeded in the Arbuckle field in a complete randomized block design (Figure 5A). We sampled plants within each plot every two weeks throughout the season, collecting rhizosphere, rhizoplane, and endosphere fractions from the plant roots. Our field design allowed us to also collect bulk soil samples throughout the entirety of the season as compared to the 2014 and 2015 seasons, which restricted our soil sampling due to the invasiveness of the rice roots.

The data from the 2016 season had similar trends as exhibited in the 2014 and 2015 seasons. The root-associated compartments hosted distinct microbiotas (Fig. 4.5B, $R^2 = 0.27$, $P < 0.001$) and the microbiota varied significantly due to plant age (Fig. 4.5B, $R^2 = 0.17$, $P < 0.001$). We found that genotype had a very small overall effect on the root-associated microbiota ($R^2 = 0.01$, $P < 0.001$). This level of variance is smaller than what we have previously detected in rice [18]. In our previous study, we grew 6 different cultivars in the greenhouse: 2 temperate *Japonica*, 2 *Indica*, and 2 *Oryza glaberrima*. Our current study only used varieties within the temperate *Japonica* clade, thus we expect that the reduced genetic diversity within this experimental population led to lower levels of microbiota variation compared out the previous study. We noticed that there was a significant statistical interaction between plant age and genotype ($R^2 = 0.05$, $P < 0.001$), suggesting that the trends in mi-

crobiota shifts over the season differ depending on the rice genotype. To further inspect this observation, we calculated the slope of the second principal coordinate of Fig 5B as a function of plant age. We used the second principal coordinate because it was the axis best differentiating between plant age. We hypothesized that if plant developmental rate were to have an effect on the root associated microbiota, then the faster developing varieties would have steeper slopes than the slower developing varieties. Indeed, we found that Kitaake had the steepest slope in each compartment, while M206 and M401 have similar slopes and Nipponbare had the most gradual slope in each compartment. Bulk soil, however, had a much reduced slope over the season and is almost flat. Taken together these results indicate that microbiota shifts throughout the season are directly driven by the host plant and the rate of these shifts correlate with the developmental progression rate of the host plant.

4.3.6 Developmental stage is a driver of microbiota succession

We next used the sparse RF models to predict the age of samples taken from the 2016 season (Fig. 4.5D). The models were remarkably accurate at predicting the ages of plants from the 2016 season despite the differences in the developmental progression rates of the genotypes. We calculated the mean squared error (MSE) of residuals for each genotypes predicted values by the compartment specific sparse RF models (Fig. 4.5E). In the case of the rhizosphere and rhizoplane, the age of the M206 genotype was most accurately predicted by the models (it was the second most accurately predicted variety using the endosphere models). The original models were trained on data collected from M206 during the 2014 season. These results suggest that the accuracies of the models are affected by the differences between the genotypes likely due to variation in developmental progression rate.

To have a better understanding of which microbes are associated with the various developmental stages, we formed new RF models for predicting plant developmental stage rather than plant age. Development in plants is a continuous process marked by numerous stage-specific morphological features. We monitored the developmental stage of the genotypes throughout the season assigning a numerical value from 1-27, with a value of 1 corresponding to a recently germinated seedling and 27 corresponding to a senescent plant.

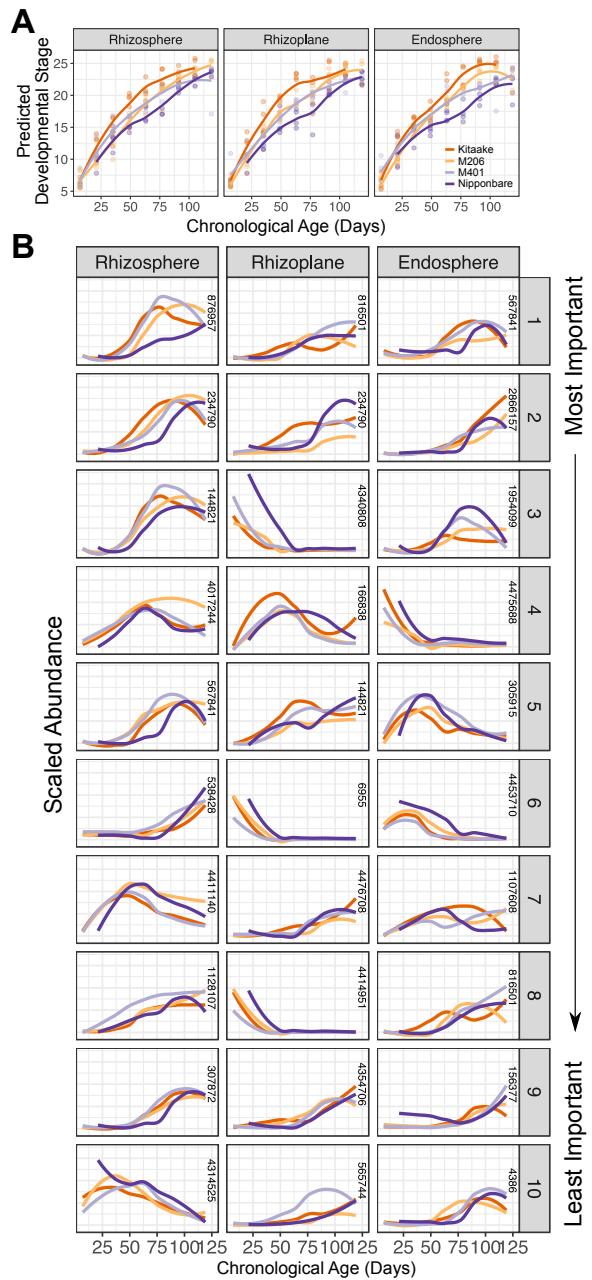


Figure 4.6: Random forests model detect taxa that are accurately predictive of plant developmental stage. (A) Results of the model after predicting developmental stages of the samples taken from the 2016 season in each compartment plotted as a function of plant age. The different varieties separate similar to the pattern exhibited in Figure 5A. (B) The abundance profiles of the top 10 most developmental stage discriminant taxa in each compartment over the course of the 2016 season.

Panicle initiation corresponded to a value of 18, thus a plant that is transitioning to the reproductive phase has a value of 18 or higher and a plant in the vegetative phase has a value of 17 or lower. We followed the same approach as previously mentioned to develop the RF models. Briefly, we trained full RF models in each compartment where we regressed the

full dataset of OTUs against the developmental stage number for a training set of samples. From these full models, we sequentially removed OTUs of lower importance while performing 10-fold cross validation. We found that the models were near peak accuracy when including 29 of the most important OTUs for each compartment (Fig. 4S3). With these 29 most-important OTUs, we developed sparse RF models that modelled the microbiota as a function of plant developmental stage. When the predicted values of plant developmental stage were plotted as a function of the plant's chronological age, we found that the predictions accurately matched the developmental progressions that we witnessed in the field (compare the prediction plot in Fig. 4.6A to the plot of witnessed developmental progressions in Fig. 4.5A). These results indicate that the sparse RF models were able to detect shifts in the microbiota that correlate with developmental shifts in the plant. The abundance profiles for each the top 10 most important OTUs for each compartment's model's accuracy are plotted in Fig. 4.6B. The developmental stage sparse RF models used less OTUs than the plant age models to make accurate predictions. This suggests that the total root-associated microbiome is more affected by the chronological age of the plant rather than the plant's developmental stage.

4.4 Discussion

Annual plants quickly undergo large shifts in morphology and physiology throughout their lifecycle [21, 39]. Developmental shifts are accompanied by different nutritional requirements and source sink relationships [26, 55] which in turn could affect the compositional profile of the root-associated microbiota. Here we have detailed how the root-associated microbiota shift throughout the life cycle of rice plants growing in a single field in Arbuckle, CA. We found that the microbiota shifts significantly over time, with most of these shifts happening in the vegetative phase of developmental and a stabilization of the microbiota occurs after the entry into the reproductive phase. We found that there are large shifts in various phyla throughout the season. Members of the Deltaproteobacteria class predominantly increased over the season and members of the class Betaproteobacteria predominantly decreased over the season. We found that these patterns held over multiple seasons

and that seasonal effects were of minimal importance on the overall root-associated microbiota. The bulk soil community members are relatively stable compared to the root-associated microbiota, suggesting that the plant non-randomly controls the fluctuating abundances of various taxa throughout the season. We found that the microbiota is predictive of plant age by fitting sparse random forests models to each compartment. These models identified many plant age discriminant microbial taxa which reproducibly changed in abundance over the course of the season. The mechanisms by which plants control the structure of their root-associated microbiota is unknown. Using cultured representatives from these age-discriminant microbial taxa in controlled inoculation experiments could help elucidate these processes and may suggest various roles for members of the root-associated microbiota on host health.

From our initial results, it remained unclear whether plant developmental stage was indeed a regulator of the root-associated microbiota structure. A recent study claimed that the entry into flowering likely has no effect on the root-associated microbiota [17]. The authors conducted a clever experiment by growing a very early flowering *Arabis alpina* mutant, *perpetual flowering1 (pep1)*[61], along with wild type *Arabis alpina*. They collected samples from the rhizosphere and endosphere of each genotype after 6, 12, and 28 weeks of growth (the *pep1* mutant flowers after 12 weeks). They found no significant differences in the microbiotas taken from the flowering mutant and the still vegetatively growing wild type plants. They did, however, find significant differences in root-associated microbiotas between the different time points. These results led the authors to conclude that the root-associated microbiota is affected by plant residence time (i.e. the chronological age of the host plant) but not the developmental stage. *Arabis alpina* (a wild *Arabidopsis thaliana* relative) is a perennial plant, and is thus physiologically distinct from rice plants. We tested the hypothesis ourselves of whether fluctuations in the microbiota are correlated with different developmental stages by growing several rice varieties with different flowering times under field conditions and sampling their root microbiota throughout the growing season. We found that the progression of the microbiota throughout the season is correlated with the developmental progression of the plant and that these progressions are not found in the

microbiota of unplanted soil. We developed models to predict the developmental stages of the plants using the abundance of various members within the plants' root-associated microbiota. From these models we were able to extract various microbes whose abundance profiles are predictive of plant developmental stage. Isolating these candidate microbes in pure culture would allow for a deeper investigation of how host plants communicate with their microbiota during various developmental and physiological states.

The random forests models trained to discriminate plant age detected more microbes than the random forests models trained to predict plant development. This discrepancy is indicative that plant age has a larger impact on the overall structure of the root-associated microbiota than plant developmental stage, yet plant developmental stage is still an important factor for differentiating the root-associated microbiota. Our results are clearly different than the aforementioned study that found no microbes to be affected by plant developmental stage. There are several likely causes for this discrepancy. *Arabis alpina* is a wild perennial plant species and rice is a domesticated annual crop species. One hallmark trait of cereal domestication is the selection for varieties with larger sink sizes in the seed [26]. For instance, in wild grass species, source carbohydrates are more evenly distributed to various sink tissues in the plant than domesticated cereals, where the seeds are a predominant sink [51, 26]. The discrepancy between sink-source dynamics in these two host plants could, at least partially, explain why our study detected shifts in the microbiota due to developmental stage while it was undetected in *Arabis*. In rice, accumulation and storage of carbohydrates in the stems occur until the onset of reproduction, after which internal signals program the host plant to begin repartitioning carbohydrates to the developing panicle and to the filling grain [53, 49]. These host signals, along with the shifting nutritional needs at this stage could explain the stabilization in the microbiota that occurs at the onset of reproduction. Here we have provided one of the first links between host plant development and composition of the root-associated microbiota. Understanding the physiological, molecular, and developmental signals from the host plant that affect the root-associated microbiota may unlock the use of microbes to boost plant performance and yield.

4.5 Materials and Methods

4.5.1 Plant growth and sampling during the 2014 and 2015 seasons

We collected samples from a commercially cultivated rice field in Arbuckle, CA. In both seasons the field was water seeded in early May. Water seeding is conducted by first preparing the field by removing winter vegetation, disking the soil to produce baseball sized clods, applying nutrients, and then flooding [22]. The rice seeds are soaked in water overnight and then loaded into an aircraft where they are then applied aerially to the field in an even density. For this particular field, the farmer grew the M206 cultivar, a medium grained California variety that has an average heading date of 86 days after germination [33]. We began sampling plants 7 days after the fields were seeded, which coincided with the emergence of the seminal roots. In the 2015 and 2016 seasons we restricted our area of sampling to a 150 x 150 foot plot. Within this plot we would sample plants from random locations. Our sampling occurred as previously described. Using gloved hands, would scoop under the root mass to separate the plant from the ground. Grabbing the plant by the base of the stem, we would then shake the plant to remove loosely attached soil from the roots. We would then place the roots with tightly adhering soil into 50 mL Falcon tubes with 15 mL autoclaved phosphate buffered saline (PBS) solution. We would then bring the samples back to the laboratory at UC Davis. We collected bulk soil samples as the season allowed. During the later parts of the season it became impossible to find soil samples unaffected by rice roots.

4.5.2 Plant growth during the 2016 Season

We grew various cultivars in the same field in Arbuckle, CA for the 2016 season. Briefly we designed a fully randomized block design to grow these varieties in 1 x 1 m plots in quadruplicate. We left 0.5 m wide walking lanes between each plot which subsequently allowed us to sample bulk soil throughout the course of the growing season. To be consistent with the previous years' methods, we water seeded these varieties. This entailed soaking the seeds in a 2% bleach solution for 4 hours to remove the risk of the field being contaminated *Fusarium moniliforme*, a seed borne fungal pathogen that causes the disease Bakanae. The seeds were then washed 3 times with sterile water and soaked overnight. We then hand seeded

each plot at similar density as what the farmer had applied in previous seasons. At each timepoint the plants were sampled as previously described and transported to the lab for further processing. The plants used for sampling during this season were dissected to score various developmental stages between the cultivars [39].

4.5.3 Separation of the root-associated compartments and DNA extraction

The root associated compartments were separated as previously described [18]. The roots with soil attached were vortexed in PBS solution and 500 uL of the resulting slurry was used for DNA extraction. The roots were cyclically washed in fresh PBS solution until no soil particles were visible in the solution. The roots were then placed into fresh PBS and sonicated for 30 s to remove surface cell layer of the roots. The resulting slurry was centrifuged down to concentrate the biomass and then used as the rhizoplane fraction for DNA extraction. The remaining roots were sonicated twice more, refreshing the PBS solution at each stage, and then ground in the bead beater. The resulting solution was used for DNA extraction as the endosphere fraction. All DNA extraction was performed using the MoBio Powersoil DNA isolation kits.

4.5.4 PCR amplification and sequencing

We amplified the V4 region of the 16S ribosomal RNA gene using the universal 515F and 806R PCR primers. Both our forward and reverse primers contained 12 base pair barcodes, thus allowing us to multiplex our sequencing libraries at well over 150 libraries per sequencing run. Each library was accompanied by a negative PCR control to ensure that the reagents were free of contaminant DNA. We purified the PCR products using AMPure beads to remove unused PCR reagents and resulting primer dimers. After purification, we quantified the concentration of our libraries using a Qubit machine. Our libraries were then pooled into equal concentrations into a single library and concentrated using AMPure beads. The pooled library then went through a final gel purification to remove any remaining unwanted PCR products. Pooled libraries were sequenced using the Illumina MiSeq machine with 250 x 250 paired end chemistry.

4.5.5 Sequence Processing, OTU clustering, and OTU filtering

The resulting sequences were demultiplexed using the barcode sequences by custom Python scripts. The sequences were quality filtered and then assembled into full contigs using the PandaSeq software [35]. Any sequences containing ambiguous bases or having a length of over 275 were discarded from the analysis. The high-quality sequences were clustered into operational taxonomic units (OTUs) using the Ninja-OPS pipeline [1] against the Green Genes database [15] and then assembled into an OTU table. This OTU table was filtered to remove plastidial and mitochondrial OTUs and filtered to remove OTUs that occur in less than 5% of the samples. This process removed the total OTU count from 24,048 to 8,554 taxa. The resulting 8,554 taxa were used for the analysis.

4.5.6 Statistical Analysis

All statistical analyses were conducted using R version 3.1 [45]. Unless otherwise noted, we determined statistical significance at $\alpha = 0.05$ and, where appropriate, corrected for multiple hypothesis testing using the false discovery rate method. Shannon diversity was calculated using the `diversity()` function, PCoA analyses were conducted using the `capscale()` function, perMANOVA was conducted using the `adonis()` function, and community structure variability was measured using the `betadisper()` function all from the Vegan package [42]. Linear models were run using the `lm()` function and ANOVA was run using the `aov()` function. Beta regression was used to model proportion of any given phylum (μ_p) as function of distance from the root (x_2) using the `betareg()` function from the BetaReg package [13]. To do this, we assigned each compartment a value from the outside to the inside of the root: bulk soil had a value of 1 and endosphere had a value of 4. The model is described by the equation

$$g(\mu_p) = \beta_1 + \beta_2 x_{p2} \quad (4.1)$$

We also used beta regression to model trends in proportions for a given phylum (p) in a given compartment (c) as a function of host plant age (x_2) using the equation

$$g(\mu_{pc}) = \beta_1 + \beta_2 x_{pc2} \quad (4.2)$$

For each model, we used a logit link function. Random forests models were generated and analyzed using the randomForest package [32]. All graphs and plots were generated using the ggplot2 package [62]. R notebooks for the full analyses can be found on GitHub (<https://github.com/bulksoil>).

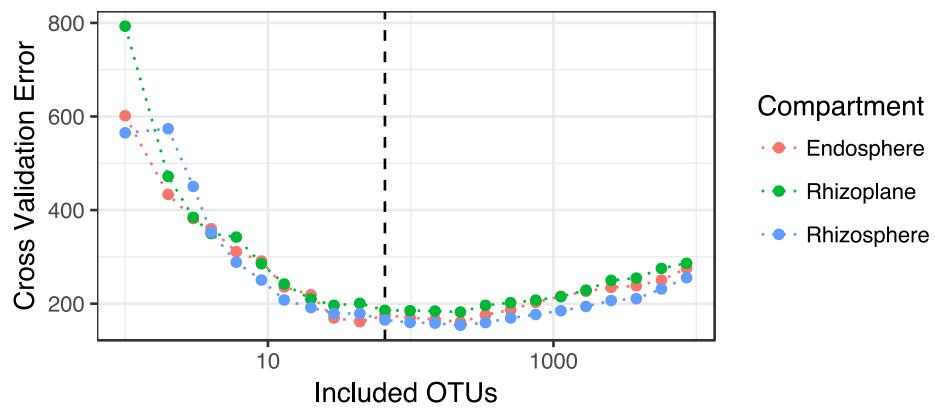


Figure 4S1. Cross validation error of models predicting plant age while sequentially removing less important OTUs for model accuracy. This analysis revealed that 66 of the most important OTUs should be included for the models to be the most accurate for predicting plant age.

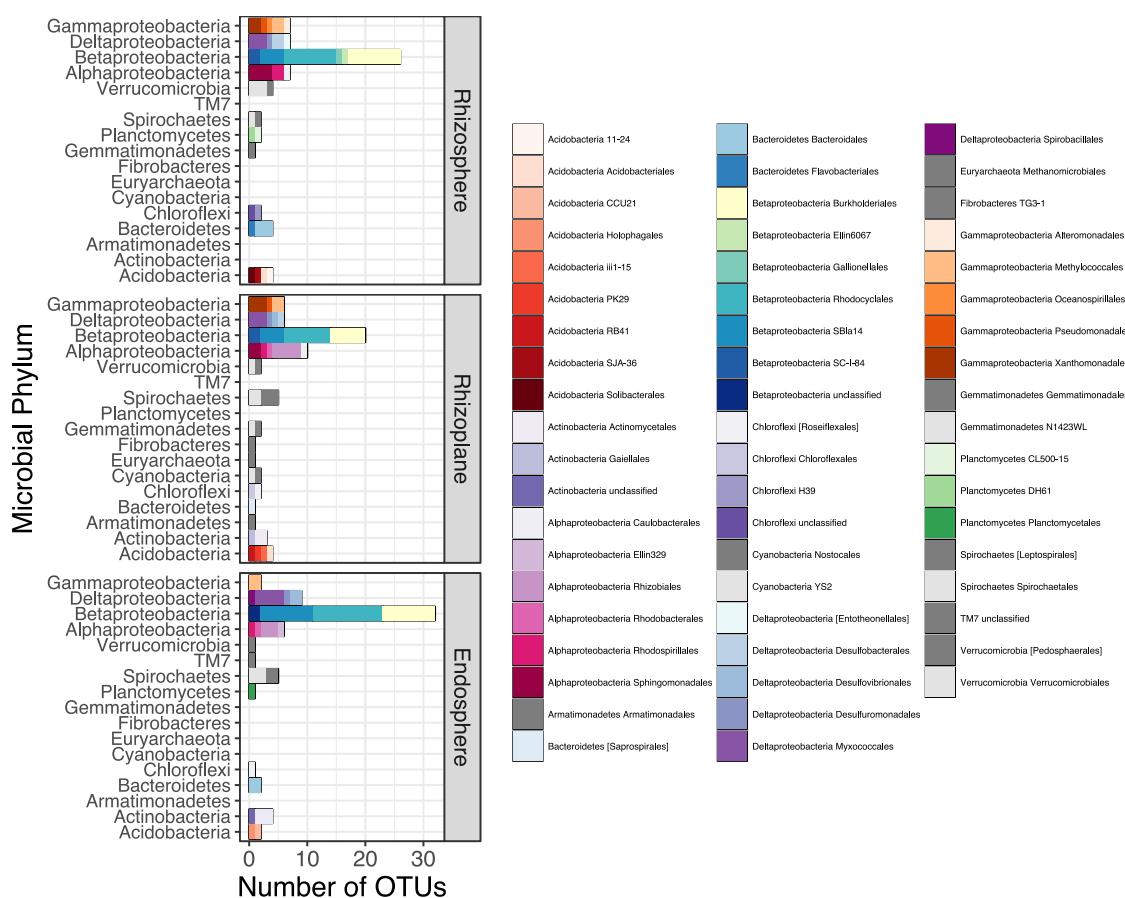


Figure 4S2. The orders of the 66 most important OTUs used in the sparse RF models predicting plant age.

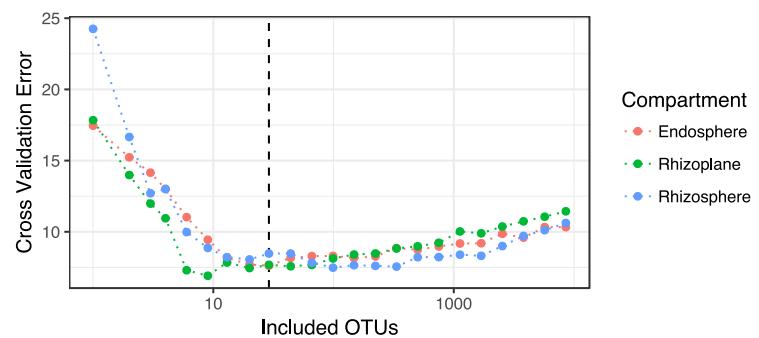


Figure 4S3. Cross validation error of RF models predicting plant developmental stage while sequentially removing less important OTUs for model accuracy. This analysis revealed that 29 of the most important OTUs should be included for the models to be the most accurate for predicting plant age.

Chapter 5

Conclusion

Soil microbes are integral for plant growth, health, and fecundity, thus the understanding the nature of associations between plants and of soil microbes is of agronomic and environmental importance [7, 10, 47, 59]. In this dissertation, I have detailed various ecological characteristics of rice root-associated microbiota which I summarize below. It is important to note that each project within this dissertation was conducted by a team. The experimental ideas and designs, the conducting of the experiments and lab work, the data analysis, and the interpretations and discussions were all contributed to by many diligent and brilliant scientists whom I have noted as authors in each chapter.

5.1 Chapter 1

In the first chapter, I described the general characteristics of rice root-associated microbiota. We began by quantifying how microbial community profiles differ as a function of distance to the root. We detailed the microbiota within three root-associated compartments: the rhizosphere (soil tightly adhering to the root), the rhizoplane (the root surface), and the endosphere (the interior of the root). We found all three of these compartments to host distinct microbial communities and that each community is also distinct from the microbiota of unplanted soil. The rhizoplane and endosphere communities had lower microbial diversity compared to the rhizosphere and bulk soil, suggesting that there is a niche specialization for the microbiota in these root-associated compartments. We next characterized the effect of soil type on the root-associated microbial communities of rice. We

found that soil originating from various rice fields host distinct sets of microbes and this in turn determines the subset of microbes that can colonize the roots of plants growing within these soils. We found similar trends in the trends in the abundance of various phyla across the root-associated compartments of plants grown in the different soils. For instance, in each soil Acidobacteria microbes have a decreasing abundance from the soil to the inside of the root, while Proteobacteria have an increased abundance inside the root and lower abundance in the soil. We also found that soil communities can be affected by agricultural practice. Specifically, root-associated microbiota are differentiable based upon whether the plants were grown under organic or conventional cultivation. We next characterized how the microbiota differs between various domesticated rice varieties. We profiled the microbiota from several cultivated varieties of two domesticated rice species, *Oryza sativa* and *Oryza glaberrima*, and found that rice genotype has a very minimal (although significant) effect on the root-associated microbiota. Many of these results have been corroborated by similar studies in other plant models. The results of the experiments are summarized in the visual model below. The model can be interpreted as such:

- Different soils host different consortia of microbes which can depend on geography or agricultural practice. This is indicated by different colors between the soil microbiota in the model.
- Rice plants assemble a root-associated microbiota from the soil in which the plants are growing and this microbiota resembles the parent soil. The microbes that assemble on the inside of the root are much reduced subset of the soil community. This is indicated by the lighter color within the circle of the root-associated microbiota in the model.
- Rice genotype has only a small effect on the root-associated microbiota even though much visible phenotypic diversity exists between the varieties. In the model, this is portrayed as the different hues in color between the different varieties. Please note that they are still reminiscent of the original hue of the parent soil in which the plants are growing.

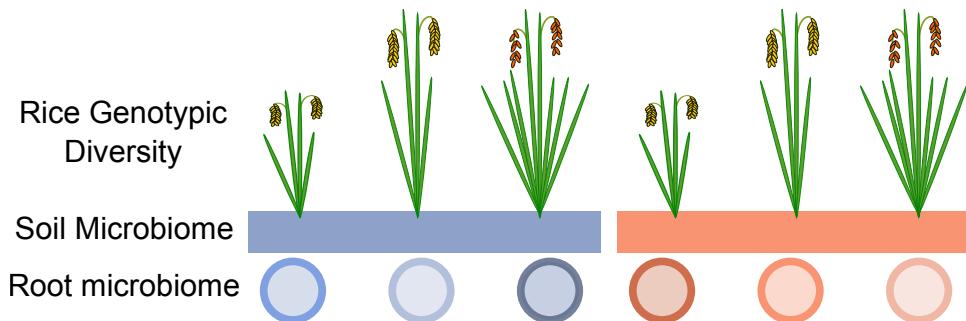


Figure 5.1: Model depicting host and environmental factors affecting the root-associated microbiota

5.2 Chapter 2

In the second chapter, I characterized how the rice root-associated microbiota differs from other plants' microbiotas. We were perplexed that genotypic variation within two species of the *Oryza* genus only had a small impact on the root-associated microbiota. The varieties that we used for our experiment had distinct morphological traits that made each variety easily identifiable, yet their root microbiota was overall remarkably similar. We hypothesized perhaps there was a large conservation of root-associated microbiota across different lineages of flowering plants. We tested this hypothesis by first compiling published data from other plant root-associated microbiota studies that used similar methods to the methods in my first chapter. We found that each plant species assembles a distinct microbiota, but that this effect is confounded by the studies occurring in different geographic regions and environments. The rice microbiota was an outlier in this analysis, potentially due to its semi-submerged growth habit. We therefore collected and analyzed the root-associated microbiotas of three separate host plants growing in the same semi-submerged conditions as rice. We found that the microbiota from these plants, all of which are common rice paddy weeds with broad distributions across the United States, have distinct sets of microbial taxa colonizing their root-associated compartments. Rice, however, hosted an outlier microbiota compared to the other plants. The microbiota of rice plants were enriched for methane-producing microbes while the other plant species hosted microbiotas enriched for methane-utilizing microbes. It was also found from this experiment that the rice rhizosphere microbiota composition is notably similar to the microbiota from unplanted soil. We attributed this effect to continuous and intensive cultivation of rice which

may in turn “domesticate” the soil microbiota to be more like the rhizosphere microbiota from rice. I have summarized the results of this analysis in the visual model below:

- The composition of the soil microbiota in which the different host plants are growing is represented by the orange box.
- Each species assembled a significantly different microbiota as represented by the different colors of the root-associated microbiota circles for each host species. Notice that the different rice genotypes host a microbiota that is more similar to the hue of the bulk soil than the microbiotas of the other included host plant species.
- The differences in microbiota composition between the different host species may have functional and environmental impacts as the rice plants host more methanogenic microbes than the other host species.

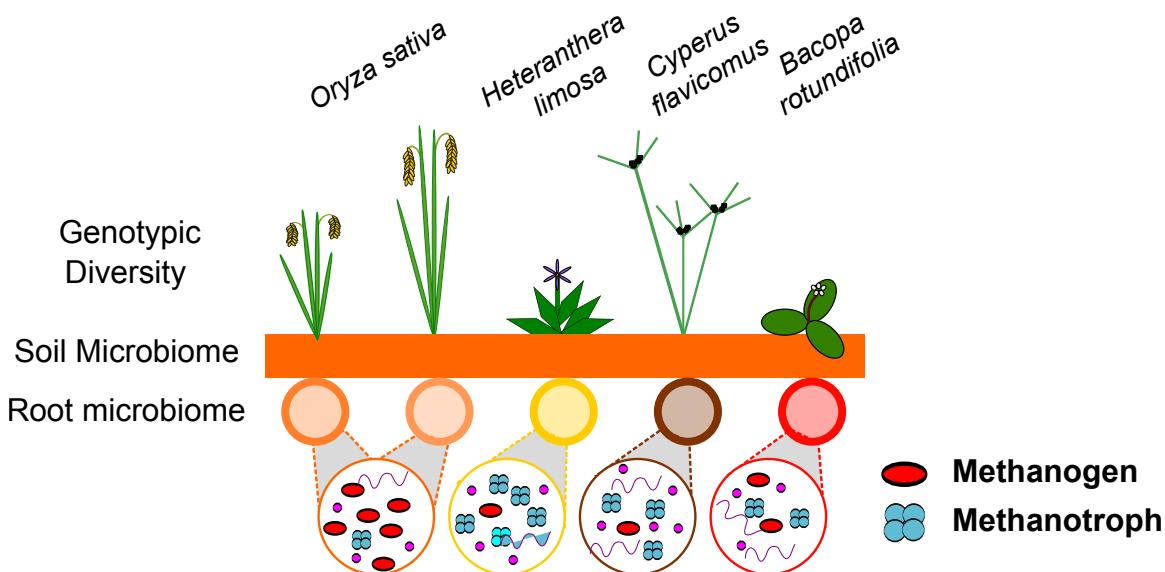


Figure 5.2: Model depicting the impact of host plant genotypic diversity on the root-associated microbiota

5.3 Chapter 3

In the third chapter, I characterized how the root-associated microbiota changes throughout the lifecycle of the rice plant. Plants undergo large morphological and physiological

shifts throughout their life cycle which in turn have stage-specific nutritional and environmental demands [26]. We sampled the microbiota of rice plants growing in single field in California over the course of three seasons. Remarkably, we found there to be only a small amount of variation across the seasons. In each compartment, we found that the microbiota progressively shifts throughout the growing season, especially during the vegetative phase of growth. These shifts are most pronounced in decreasing abundance of Betaproteobacteria and increasing abundance of Deltaproteobacteria over the course of the season. Upon the entry into the reproductive stage of development, we found that the overall composition of root-associated microbiota stabilizes. This shift appeared to be faster in the endosphere and rhizoplane than in the rhizosphere. While this was an interesting correlation, we were unable to definitively associate the shifting microbiota with plant developmental progression in the first two seasons of our trial because all of the sampled plants were synchronized for developmental progression. There are environmental and edaphic factors that fluctuate over the season and there is also a possibility that the microbes require a certain amount of time to stabilize independent of plant developmental stage. To test the effect of plant developmental stage on the root-associated microbiota independent of plant age and fluctuating environmental and edaphic factors, we grew different varieties that progress through development at different rates. Indeed, we found that faster progressing genotypes had faster shifting microbiotas. By using a machine learning approach, we were able to identify individual microbes whose abundance are descriptive of developmental stage. I have summarized our model for root-associated microbiota shifts over the lifecycle of the rice plant in a visual model below.

- The microbiota shifts in a gradient over the course of the season as mimicked by the hue of the horizontal bar shifting from the beginning of the season to the end.
- The microbiota shifts during the vegetative phase of the rice plant's life (i.e. the seedling and tillering stages) until the onset of reproduction where the microbiome remains relatively steady. This pattern is denoted by the shifting hue of the horizontal bar during the vegetative phase which becomes steady at the onset of the panicle initiation stage.

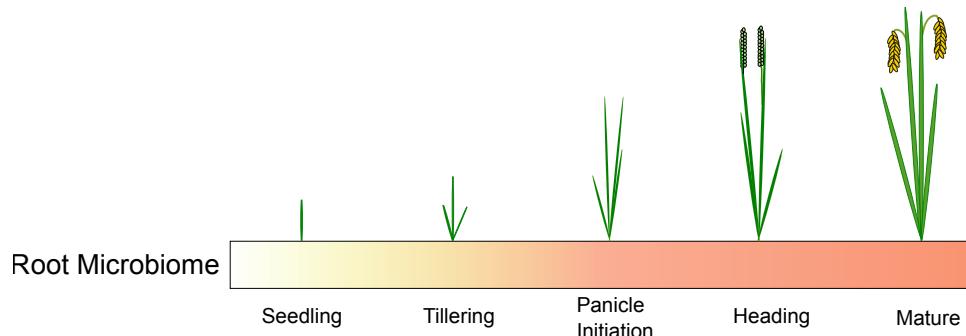


Figure 5.3: Model depicting the dynamic shifts in the microbiota over the course of a growing season as a function of plant developmental stage

This dissertation is a small block in a foundation which further studies can be built upon in order to understand the root-associated microbiota of rice plants. Much work still needs to be conducted to better characterize the functional, agricultural, and environmental implications of the root-associated microbiota. It is my hope that the ecological descriptions detailed in this dissertation can one day provide at least partial inspiration for further studies to untangle the infinitely complex webs and networks of plant-microbe interactions. It is also my hope that some of the approaches adopted within this dissertation, especially the use of machine learning algorithms, can be used to predict the composition of beneficial microbiotas for use in agriculture.

REFERENCES

- [1] Gabriel A. Al-Ghalith, Emmanuel Montassier, Henry N. Ward, and Dan Knights. NINJA-OPS: Fast Accurate Marker Gene Alignment Using Concatenated Ribosomes. *PLoS Computational Biology*, 12(1):e1004658, jan 2016.
- [2] R.M. Atlas and R. Bartha. Microbial ecology: Fundamentals and applications, 1986.
- [3] Daniel Austin. *Florida Ethnobotany*. CRC Press, Boca Raton, 2004.
- [4] Fredrik Bäckhed, Hao Ding, Ting Wang, Lora V Hooper, Gou Young Koh, Andras Nagy, Clay F Semenkovich, and Jeffrey I Gordon. The gut microbiota as an environmental factor that regulates fat storage. *Proceedings of the National Academy of Sciences of the United States of America*, 101(44):15718–23, 2004.
- [5] Fredrik Bäckhed, Josefina Roswall, Yangqing Peng, Qiang Feng, Huijue Jia, Petia Kovatcheva-Datchary, Yin Li, Yan Xia, Hailiang Xie, Huanzi Zhong, Muhammad Tanweer Khan, Jianfeng Zhang, Junhua Li, Liang Xiao, Jumana Al-Aama, Dongya Zhang, Ying Shiuan Lee, Dorota Kotowska, Camilla Colding, Valentina Tremaroli, Ye Yin, Stefan Bergman, Xun Xu, Lise Madsen, Karsten Kristiansen, Jovanna Dahlgren, and Jun Wang. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host & Microbe*, 17(5):690–703, 2015.
- [6] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 57(1):289–300, 1995.
- [7] Roeland L. Berendsen, Corné M.J. Pieterse, and Peter A.H.M. Bakker. The rhizosphere microbiome and plant health. *Trends in Plant Science*, 17(8):478–486, 2012.
- [8] Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001.
- [9] Davide Bulgarelli, Matthias Rott, Klaus Schlaepipi, Emiel Ver Loren van Themaat, Nahnah Ahmadinejad, Federica Assenza, Philipp Rauf, Bruno Huettel, Richard Reinhardt, Elmon Schmelzer, Joerg Peplies, Frank Oliver Gloeckner, Rudolf Amann, Thilo Eickhorst, and Paul Schulze-Lefert. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature*, 488(7409):91–5, aug 2012.
- [10] Davide Bulgarelli, Klaus Schlaepipi, Stijn Spaepen, Emiel Ver Loren van Themaat, and Paul Schulze-Lefert. Structure and functions of the bacterial microbiota of plants. *Annual review of plant biology*, 64(1):807–38, apr 2013.
- [11] J Gregory Caporaso, Justin Kuczynski, Jesse Stombaugh, Kyle Bittinger, Frederic D Bushman, Elizabeth K Costello, Noah Fierer, Antonio Gonzalez Peña, Julia K Goodrich, Jeffrey I Gordon, Gavin A Hutley, Scott T Kelley, Dan Knights, Jeremy E Koenig, Ruth E Ley, Catherine A Lozupone, Daniel McDonald, Brian D Muegge, Meg Pirrung, Jens Reeder, Joel R Sevinsky, Peter J Turnbaugh, William A Walters, Jeremy Widmann,

Tanya Yatsunenko, Jesse Zaneveld, and Rob Knight. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7(5):335–336, may 2010.

- [12] J Gregory Caporaso, Christian L Lauber, Elizabeth K Costello, Donna Berg-Lyons, Antonio Gonzalez, Jesse Stombaugh, Dan Knights, Pawel Gajer, Jacques Ravel, Noah Fierer, Jeffrey I Gordon, and Rob Knight. Moving pictures of the human microbiome. *Genome biology*, 12(5):R50, jan 2011.
- [13] Francisco Cribari-Neto and A Zeileis. Beta regression in R. *Journal of Statistical Software*, 34(2):24, 2010.
- [14] Justine Debelius, Se Jin Song, Yoshiki Vazquez-Baeza, Zhenjiang Zech Xu, Antonio Gonzalez, and Rob Knight. Tiny microbes, enormous impacts: what matters in gut microbiome studies? *Genome Biology*, 17(1):217, dec 2016.
- [15] T Z DeSantis, P Hugenholtz, N Larsen, M Rojas, E L Brodie, K Keller, T Huber, D Dalevi, P Hu, and G L Andersen. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology*, 72(7):5069–72, jul 2006.
- [16] Clifford Dobell. *Antony van Leeuwenhoek and his "Little animals"*. Harcourt, Brace, and Co., New York, 1932.
- [17] Nina Dombrowski, Klaus Schlaeppi, Matthew T Agler, Stéphane Hacquard, Eric Kemen, Ruben Garrido-Oter, Jörg Wunder, George Coupland, and Paul Schulze-Lefert. Root microbiota dynamics of perennial *Arabis alpina* are dependent on soil residence time but independent of flowering time. *Nature Publishing Group*, 2016.
- [18] Joseph Edwards, Cameron Johnson, Christian Santos-Medellín, Eugene Lurie, Natraj Kumar Podishetty, Srijak Bhatnagar, Jonathan A Eisen, and Venkatesan Sundaresan. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences of the United States of America*, 112(8):E911–20, 2015.
- [19] Emanuel Epstein. *Mineral nutrition of plants: principles and perspectives*. John Wiley and Sons, Inc., London, 1971.
- [20] M. 'Espinasse. *Robert Hooke*. University of California Press, Berkeley and Los Angeles, 2nd edition, 1962.
- [21] Michael Freeling. Patterns in plant development (2nd edn). *Trends in Genetics*, 6:97, 1990.
- [22] Jarrod Hardke and Bob Scott. Water-Seeded Rice.
- [23] Robert Hooke. *Micrographia*. Royal Society, London, 1665.
- [24] D Hornby. Suppressive Soils. *Annual Review of Phytopathology*, 21(1):65–85, sep 1983.

- [25] Philip Hugenholtz, Brett M Goebel, and Norman R Pace. MINIREVIEW Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. *JOURNAL OF BACTERIOLOGY*, 180(18):4765–4774, 1998.
- [26] Ken Ishimaru, Naoki Hirotsu, Yuka Madoka, Naomi Murakami, Nao Hara, Haruko Onodera, Takayuki Kashiwagi, Kazuhiro Ujiie, Bun-ichi Shimizu, Atsuko Onishi, Hisashi Miyagawa, and Etsuko Katoh. Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nature Genetics*, 45(6):707–711, apr 2013.
- [27] Jeremy E Koenig, Aymé Spor, Nicholas Scalfone, Ashwana D Fricker, Jesse Stombaugh, Rob Knight, Lergus T Angenent, and Ruth E Ley. Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, 108 Suppl:4578–85, mar 2011.
- [28] H. Kudo, K.-J. Cheng, and J. W. Costerton. Interactions between *Treponema bryantii* and cellulolytic bacteria in the in vitro degradation of straw cellulose. *Canadian Journal of Microbiology*, 33(3):244–248, mar 1987.
- [29] Ben Langmead and Steven L Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4):357–359, mar 2012.
- [30] Daniel A Lashof and Dilip R Ahuja. Relative contributions of greenhouse gas emissions to global warming. 1990.
- [31] Ruth E Ley, F Backhed, Peter Turnbaugh, Catherine A Lozupone, Robin D Knight, and Jeffrey I Gordon. Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A*, 102(31):11070–11075, 2005.
- [32] Andy Liaw, Jingli Yan, Weifeng Li, Lijian Han, F. Schroff, A. Criminisi, Andrew Zisserman, Barrett Lowe, Arun Kulkarni, Anna Bosch, Andrew Zisserman, Xavier Mu, Xavier Munoz, R Devadas, R J Denham, M Pringle, Remote Sensing Centre, Ecosciences Percinct, Tidy Data, Angelos Tzotsos, and X Zhang. Package ‘randomForest’. *R news*, XXXIX(1):54.1–54.10, 2014.
- [33] Bruce Linquist and Robert Hijmans. ANNUAL REPORT COMPREHENSIVE RESEARCH ON RICE. 2012.
- [34] Derek S. Lundberg, Sarah L. Lebeis, Sur Herrera Paredes, Scott Yourstone, Jase Gehring, Stephanie Malfatti, Julien Tremblay, Anna Engelbrektson, Victor Kunin, Tijana Glavina del Rio, Robert C. Edgar, Thilo Eickhorst, Ruth E. Ley, Philip Hugenholtz, Susanah Green Tringe, and Jeffery L. Dangl. Defining the core *Arabidopsis thaliana* root microbiome. *Nature*, 488(7409):86–90, aug 2012.
- [35] Andre P Masella, Andrea K Bartram, Jakub M Truszkowski, Daniel G Brown, and Josh D Neufeld. PANDAseq: paired-end assembler for illumina sequences. *BMC bioinformatics*, 13(1):31, 2012.

- [36] Paul J. McMurdie and Susan Holmes. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4), apr 2013.
- [37] Rodrigo Mendes, Marco Kruijt, Irene de Bruijn, Ester Dekkers, Menno van der Voort, Johannes H. M. Schneider, Yvette M. Piceno, Todd Z. DeSantis, Gary L. Andersen, Peter A. H. M. Bakker, and Jos M. Raaijmakers. Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science*, 332(6033), 2011.
- [38] K. Minami. Methane from rice production. *Fertilizer Research*, 37(3):167–179, 1994.
- [39] Karen Moldenhauer and Nathan Slaton. Rice growth and development. *Rice production handbook*, pages 7–14, 2001.
- [40] Noel T. Mueller, Elizabeth Bakacs, Joan Combellick, Zoya Grigoryan, and Maria G. Dominguez-Bello. The infant microbiome development: mom matters. *Trends in Molecular Medicine*, 21(2):109–117, 2015.
- [41] Nobel Foundation. *Nobel Lectures, Physiology or Medicine 1901-1921*. Elsevier Publishing Company, Amsterdam, 1967.
- [42] Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, and Helene Wagner. Package ‘vegan’. *R package ver. 2.0-8*, page 254, 2013.
- [43] Norman R. Pace, David A. Stahl, David J. Lane, and Gary J. Olsen. The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences. In *Advances in Microbial Ecology*, pages 1–55. Springer US, 1986.
- [44] Jason A Peiffer, Aymé Spor, Omry Koren, Zhao Jin, Susannah Green Tringe, Jeffery L Dangl, Edward S Buckler, and Ruth E Ley. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences of the United States of America*, 110(16):6548–53, apr 2013.
- [45] R Core Team. R: A Language and Environment for Statistical Computing, 2016.
- [46] Emma Ransom-Jones, David L. Jones, Alan J. McCarthy, and James E. McDonald. The Fibrobacteres: An Important Phylum of Cellulose-Degrading Bacteria, feb 2012.
- [47] Klaus Schlaeppi and Davide Bulgarelli. The Plant Microbiome at Work. *Molecular Plant-Microbe Interactions*, 28(3):212–217, mar 2015.
- [48] Klaus Schlaeppi, Nina Dombrowski, Ruben Garrido Oter, Emiel Ver Loren van Themaat, and Paul Schulze-Lefert. Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proceedings of the National Academy of Sciences of the United States of America*, 111(2):585–92, jan 2014.
- [49] Graham N Scofield, Sari A Ruuska, Naohiro Aoki, David C Lewis, Linda M Tabe, and Colin L D Jenkins. Starch storage in the stems of wheat plants: localization and temporal changes. *Annals of botany*, 103(6):859–68, apr 2009.

- [50] Matthew Shenton, Chie Iwamoto, Nori Kurata, and Kazuho Ikeo. Effect of Wild and Cultivated Rice Genotypes on Rhizosphere Bacterial Community Composition. *Rice (New York, N.Y.)*, 9(1):42, dec 2016.
- [51] Ayahiko Shomura, Takeshi Izawa, Kaworu Ebana, Takeshi Ebitani, Hiromi Kanegae, Saeko Konishi, and Masahiro Yano. Deletion in a gene associated with grain size increased yields during rice domestication. *Nature Genetics*, 40(8):1023–1028, aug 2008.
- [52] Maegen B. Simmonds, Merle Anders, Maria Arlene Adviento-Borbe, Chris van Kessel, Anna McClung, and Bruce A. Linquist. Seasonal Methane and Nitrous Oxide Emissions of Several Rice Cultivars in Direct-Seeded Systems. *Journal of Environment Quality*, 44(1):103, 2015.
- [53] Thomas L Slewinski. Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. *Journal of experimental botany*, 63(13):4647–70, aug 2012.
- [54] Mitchell L Sogin, Hilary G Morrison, Julie A Huber, David Mark Welch, Susan M Huse, Phillip R Neal, Jesus M Arrieta, and Gerhard J Herndl. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America*, 103(32):12115–20, aug 2006.
- [55] Mechthild Tegeder and Doris Rentsch. Uptake and partitioning of amino acids and peptides. *Molecular plant*, 3(6):997–1011, nov 2010.
- [56] Peter J Turnbaugh, Ruth E Ley, Micah Hamady, Claire M Fraser-Liggett, Rob Knight, and Jeffrey I Gordon. The human microbiome project. *Nature*, 449(7164):804–10, oct 2007.
- [57] Peter J Turnbaugh, Ruth E Ley, Michael A Mahowald, Vincent Magrini, Elaine R Mardis, and Jeffrey I Gordon. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*, 444(7122):1027–31, 2006.
- [58] C P Vance. Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant physiology*, 127(2):390–7, oct 2001.
- [59] Maggie R. Wagner, Derek S. Lundberg, Devin Coleman-Derr, Susannah G. Tringe, Jeffery L. Dangl, and Thomas Mitchell-Olds. Natural soil microbes alter flowering phenology and the intensity of selection on flowering time in a wild *Arabidopsis* relative. *Ecology Letters*, 17(6):717–726, jun 2014.
- [60] Maggie R Wagner, Derek S Lundberg, Tijana G del Rio, Susannah G Tringe, Jeffery L Dangl, and Thomas Mitchell-Olds. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun*, 7:1–15, 2016.
- [61] Renhou Wang, Sara Farrona, Coral Vincent, Anika Joecker, Heiko Schoof, Franziska Turck, Carlos Alonso-Blanco, George Coupland, and Maria C. Albani. PEP1 regulates perennial flowering in *Arabis alpina*. *Nature*, 459(7245):423–427, may 2009.

- [62] Hadley Wickham. *ggplot2*. 2009.
- [63] Max E. Winston, Jarrad Hampton-Marcell, Iratxe Zarraonaindia, Sarah M. Owens, Corrie S. Moreau, Jack A. Gilbert, Josh Hartsel, Suzanne J. Kennedy, and S. M. Gibbons. Understanding cultivar-specificity and soil determinants of the Cannabis microbiome. *PLoS ONE*, 9(6), jun 2014.
- [64] Carl R Woese and George E Fox. Phylogenetic structure of the prokaryotic domain: The primary kingdoms (archaeabacteria/eubacteria/urkaryote/16S ribosomal RNA/molecular phylogeny). *Proceedings of the National Academy of Sciences of the United States of America*, 74(11):5088–5090, 1977.
- [65] Tanya Yatsunenko, Federico E. Rey, Mark J. Manary, Indi Trehan, Maria Gloria Dominguez-Bello, Monica Contreras, Magda Magris, Glida Hidalgo, Robert N. Baldassano, Andrey P. Anokhin, Andrew C. Heath, Barbara Warner, Jens Reeder, Justin Kuczynski, J. Gregory Caporaso, Catherine A. Lozupone, Christian Lauber, Jose Carlos Clemente, Dan Knights, Rob Knight, and Jeffrey I. Gordon. Human gut microbiome viewed across age and geography. *Nature*, 486(7402):222, may 2012.
- [66] Iratxe Zarraonaindia, Sarah M. Owens, Pamela Weisenhorn, Kristin West, Jarrad Hampton-Marcell, Simon Lax, Nicholas A. Bokulich, David A. Mills, Gilles Martin, Safiyh Taghavi, Daniel van der Lelie, and Jack A. Gilbert. The soil microbiome influences grapevine-associated microbiota. *mBio*, 6(2), mar 2015.

Joseph Andrew Arthur Edwards
December 2016
Plant Biology

Structure, Variation, and Dynamics of the Root-Associated Microbiota of the Crop Plant
Rice

Abstract

Soil microbes are important mediators of plant nutrition, growth, and resistance to biotic and abiotic stresses, thus understanding the composition and dynamics of plant root-associated microbial communities is of importance to scientists and farmers. Plants acquire root-associated microbiota that are distinct from that of the surrounding soil. In this thesis I detail general characteristics of the rice root-associated microbiota.

In the first chapter I introduce the importance of both animal and plant microbiota and I discuss the technological breakthroughs that have allowed scientists to study complex microbial ecosystems.

In the second chapter, I present the results of a study characterizing the rice root-associated microbiota across three root compartments: the soil adjacent to the root (the rhizosphere), the root surface (the rhizoplane), and the root interior (the endosphere). I present how the composition of the microbiota in these compartments varies across soil type, cultivation practice, and host plant genotype.

In the third chapter, I present the results of a study characterizing the rice root-associated microbiota compared to diverse species of both crop plants and native plants growing in the same environment as rice. The results suggest that rice assimilates a unique root-associated microbiota and that plant phylogenetic variation cannot fully explain microbiota variation between the tested host plants. A comparison of the root-associated microbiomes of rice with microbiota of soil from cultivated fields and non-cultivated soils, suggests that rice cultivation changes the overall soil microbiome in rice fields not only through agricultural inputs and water submergence, but also through specific enrichment of microbial taxa by rice plants.

In the fourth chapter, I demonstrate how the rice root-associated microbiota shifts across the lifecycle of the host plants under field conditions. We find that the composition of the microbiome shifts throughout vegetative growth until the initiation of reproductive growth, whereupon the microbiome stabilizes. These results suggest that plants select for different microbiota as they undergo developmental changes from germination to reproduction.

I present concluding remarks on the various models of microbiota acquisition and discuss future implications of this research in the fifth and final chapter of this dissertation.