CHARACTERIZATION OF AXENIC IMMUNE DEFICIENCY IN ARABIDOPSIS THALIANA

By

James Michael Kremer

A DISSERTATION

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

Microbiology and Molecular Genetics – Doctor of Philosophy

2017

ABSTRACT

CHARACTERIZATION OF AXENIC IMMUNE DEFICIENCY
IN *ARABIDOPSIS THALIANA*

By

James Michael Kremer

Evolution of land plants began and has since occurred, in concert with complex communities of microorganisms, giving rise to a vast spectrum of plant-microbe relationships. Over the past decade, plant molecular biologists and microbial ecologists have worked together to identify drivers of microbiome composition that inspire hypotheses about microbiome functional potential, but many fall short of offering empirical evidence of microbiome-mediated influence on host phenotypes. Herein, I introduce a new suite of tools to explore microbiome function and report that many facets of plant immunocompetence are microbiome-dependent.

Chapter One summarizes the current understanding of plant innate immunity and notable progress of plant microbiome research, including: (1) detection and response to microbe-associated molecular patterns, (2) hormone signaling during biotic interactions, (3) technology for exploration of plant microbiome ecology, (4) factors that influence microbiome community structure, and (5) a review of relevant model systems and gnotobiotic growth platforms. Chapter Two describes the development of a novel "FlowPot" growth system: a peat-based platform conducive to axenic (microbe-free), gnotobiotic (defined microbiota), and holoxenic (undefined, complex microbiota) *Arabidopsis thaliana* growth. This system provides the ability to maintain control of abiotic parameters and exogenous microbiota, thus providing a valuable platform for

discovery for plant microbiome research. The FlowPot system and offers a substantial improvement over alternative growth systems regarding plant health, tractability to greenhouse conditions, and maintenance of bacterial alpha diversity upon inoculation with soil-derived microbiota. An implementation of the growth system is detailed in Chapter Three, featuring a comparative analysis of the axenic vs. holoxenic Arabidopsis transcriptome, metabolome, and immunocompetence. Axenic Arabidopsis has a reduced level of defense- and immunity-associated gene expression and the defense hormone salicylic acid (SA). We report that axenic Arabidopsis is compromised in defense against the foliar pathogen *Pseudomonas syringae* pv. tomato DC3000 (*Pst*). Immune elicitation experiments revealed that axenic Arabidopsis is also compromised in the ability to recognize and/or mount normal defense responses to the microbe-associated molecular pattern flg22. Axenic susceptibility to *Pst* is partially explained by defective innate immunity. Finally, we report the identity of differentially abundant metabolites and transcripts in axenic and holoxenic Arabidopsis that may be involved in microbiome-influenced host phenotypes. Collectively, research described in this dissertation provides new tools and a discovery platform to empirically characterize function of plant microbiota, as well as detailed characterization of axenic phenotypes and axenic immune deficiency.

This dissertation is dedicated to you, the reader. I hope that this will inspire new hypotheses and future experiments.

ACKNOWLEDGEMENTS

First and foremost, I must acknowledge my mentors, Dr. Sheng Yang He and Dr. Jim Tiedje. Without your support and guidance, none of this work would have been possible. I was originally drawn to Michigan State University because of your respective expertise in microbial suppression of plant innate immunity and the ecology of root-associated microbiota. Over time, as I rotated in your respective laboratories and eventually began to focus my dissertation research, I realized how unique and incredible an opportunity it was to be at a place like Michigan State University; nowhere else in the world could I have such brilliant advisors with different but complementary perspectives. Both of you identify and cultivate passion in your students towards the pursuit of areas of research that inspire us individually. When the inevitable roadblocks and experimental setbacks arise, you inspire tenacity. Sheng Yang, you were incredibly patient when I was resistant to attempt certain experiments, and likewise, you trusted me to conduct many experiments that did not necessarily coincide with the primary objectives of this dissertation. By doing so, you have provided me with the incredible opportunity to make mistakes and retrospectively evaluate depth versus breadth of research, and that has taught me more than you can imagine.

Jim, you inspire me to think globally, and taught me the importance of creativity and communication in science. Your academic legacy and success of past students are inspirational, and I cannot understate how honored and fortunate I am to receive your mentorship. Both of you have provided me with numerous opportunities to present at

academic conferences and build a network of collaborators, and for this, I thank you. I also must acknowledge Dr. Paul Schulze-Lefert. You extended an invitation to visit your group in Cologne, and immediately embraced me as one of your own students. I learned a tremendous amount from your pragmatic approach to tackle highly complex scientific questions. Drs. Amine Hassani, Sergio de los Santos, JP Jerome and Brian Kvitko, you have all been incredible scientific collaborators and lifelong friends. To Kevin Nehil, Alec Bonifer, Kristin Drumheller, Alan Mundakkal, David Rhodes, Caleigh Griffin, and all of the other undergraduate researchers I have had the pleasure of advising, you have taught me more than you can realize and I cannot wait to see where life takes you. I also want to acknowledge Brad Paasch for being a great colleague and friend in our mutual pursuit to characterize microbiome function. To the many members of the He and Tiedje lab groups, you have been tremendous colleagues and thank you for all the discussions and input. It has been a pleasure to work with many of you, and I look forward to crossing paths again in the future.

Finally, I must acknowledge my family. Without the tremendous unwavering love and support from my parents, I never would have finished this dissertation. To my aunt Dr. Margaret McLaughlin, your help and wisdom during the “final push” made this all possible. And to my grandmother Bernice Neuberg, your creativity and love for plants no doubt inspired me to pursue this Ph.D.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

History and review of plant microbiome research

Introduction

Agriculture is arguably the single-most important human innovation. It is an essential element of civilization. Crop diseases pose an enormous threat to agricultural productivity and stability, and epidemics have resulted in economic devastation and countless deaths due to famine. Effective strategies to reduce the incidence and severity of crop disease are greatly needed. Thus, the majority of plant-microbe research focuses on model plant-pathogen interactions. One must not discount that multicellular life is inevitably intertwined with a microbial world, and a vast spectrum of host-microbe and microbe-microbe relationships exist within a plant microbiome. In nature, a plant provides numerous distinct habitats colonized by diverse microbiota throughout each phase of development. Our understanding of plant-microbe interactions is biased towards a disproportionate amount of in-depth research on model pathosystems in a laboratory setting. While such studies provide a fantastic foundation for the molecular underpinnings of important pathogenic interactions, there are billions of microscopic “elephants” in the room that are often ignored.

Like mammals, plants have an innate immune system to recognize and respond to microbial pathogens, thereby defending against pathogenic invasion and disease. Pattern recognition receptors (PRRs) offer surveillance for certain microbe-associated molecular patterns (MAMPs), and upon ligand binding, an immune response ensues. Interestingly, certain symbionts rely on immune responses by the plant host to successfully colonize. Many, if not all, microorganisms, have MAMPs that a plant can recognize. What are the phenotypic impacts of perpetual exposure to MAMPs? While

there are some reports of PRR-triggered immunity (PTI) sculpting plant microbiome community composition, environmental and ecological complexity, make such questions difficult to address. Axenic (germfree) animal models have proven valuable to determine that endogenous microbiota are necessary for the development of a fully functional immune system and nutrition, but due to lack of infrastructure, analogous conclusions cannot be drawn for plants.

The plant microbiome, defined here as the collective host-associated environments naturally colonized by microbiota, is dynamic and compartment-specific. The major plant microbiomes are the root surface (rhizoplane) and immediately surrounding soil (rhizosphere), the surface of leaves (phyllosphere), and within tissue beneath the epidermis (endosphere). Microbiota perpetually colonize these habitats in nature, but the benefit or detriment of a colonized microbiome is largely unknown. Most microbiota do not confer obvious phenotypic impacts on their respective plant hosts and historically have long been regarded as commensal.

With new and accessible '-omics' technologies, along with the foresight to develop effective agricultural strategies to meet future food demands of a growing global population, a multidisciplinary plant microbiome research field is emerging. Over the past 20 years, innovations in DNA sequencing technologies have allowed us to explore new horizons with culture-independent methodology to characterize microbial environments. We now know from the Human Microbiome Project that the cells of a typical human are outnumbered by associated microbial cells ten-fold, and the number of genes from any given human microbiome outnumbers human genes by approximately 100:1 (Goodman et al., 2011). Studies over the past decade provide us

with insight into plant-associated communities as well, revealing that astounding microbial diversity is present on (and within) the roots, vascular tissue, and foliage in nature (Buée et al., 2009; Delmotte et al., 2009; Ikeda et al., 2009). Historically, plant microbiome research has been approached from two different angles: a mechanistic/reductionist approach (plant pathology and plant molecular biology) and an ecological/systems approach (microbial ecology). In this chapter, I will highlight important technological breakthroughs and discoveries in the fields of molecular plant pathology/biology and microbial ecology that establish the foundation of plant microbiome research. This chapter also provides a review of key recent publications to provide the reader with a current status of the field. The gradual convergence of molecular plant biology and microbial ecology marks the beginning of a plant microbiome renaissance. A nearly infinite reservoir of genetic functional potential within the plant microbiome conceptually extends the phenotypic capacity of the plant. A better understanding of microbiome structure and function will bestow the power to intelligently manipulate plant-associated microbial communities to control disease, increase crop yield, and achieve optimal phenotypes.

Early plant pathology

Crop diseases have had a profound impact on civilizations over the course of history. Global potential yield loss due to disease is estimated to be approximately 40%, and regional epidemics pose a significant threat to food security and economic stability (Savary et al., 2012). Plant pathogenic microorganisms, including certain bacteria, fungi, oomycetes, and nematodes, are responsible for many crop diseases. The Great Potato Famine of the 1840s, caused by the oomycete pathogen *Phytophthora infestans*, scourged the country of Ireland and resulted in a 25% population drop (Agrios, 2005). A fungal pathogen responsible for brown spot disease in rice, *Bipolaris oryzae*, contributed to the Bengal Famine of 1943, resulting in loss of an estimated 4 million lives (Dasgupta, 1984). In the 1970s, the Southern Corn Leaf Blight of North America caused by the fungal pathogen *Cochliobolus heterostrophus* resulted in over $1 billion USD in crop losses (Tatum, 1971). Citrus greening (‘Ca. Liberibacter spp.'), the most devastating disease of citrus, cost the Florida citrus industry $3.63 billion and 6,611 jobs from 2006-2012(Alvarez et al., 2016). A number of bacterial pathogens, particularly of the genera *Pseudomonas*, *Burkholderia*, *Xanthomonas*, *Ralstonia* and *Erwinia*, are of significant economic importance (Mansfield et al., 2012; Pfeilmeier et al., 2016). Worldwide, potential yield losses caused by microbial phytopathogens, insects, and invasive weeds, are estimated to account for up to $500 billion USD (Agrios, 2005). Early plant-microbe interaction research and the beginnings of plant pathology were driven by the desire to understand and combat crop disease. Plant pathology was a field driven by necessity, and some would argue, the need to control crop diseases set it apart from basic science (Sequeira, 1988). The enigma of the potato late blight was

solved in 1853 when Anton de Bary, widely regarded as the father of plant pathology, discovered the disease was caused a microorganism (later classified as the oomycete *Phytopthora infestans*) (Agrios, 2005). It is worthwhile noting that in addition to his research into plant-pathogen interactions, de Bary was the first to coin the term "symbiosis", which he used to describe the symbiotic relationship of fungi and algae that comprise lichen (de Bary, 1879). One of many noteworthy alumni of de Bary's laboratory is Pierre-Marie-Alexis Millardet. As an early pioneer of translating plant pathology research to the field, Millardet is accredited with the first successful chemical intervention to control downy mildew, by use of the "Bordeaux Mixture" fungicide, consisting of copper (II) sulfate and slaked lime (Swingle, 1896). Another early pioneer of plant pathology was T.J. Burrill, who discovered that the bacterium *Erwinia amylovora* is the causative pathogen of fire blight, a destructive disease to many fruit trees (Agrios, 2005). Analogous to medical pathology, plant pathology was driven by the paradigm that a plant disease was the result of infection of a susceptible host by a compatible pathogen, and the objective of a plant pathologist was to identify said pathogen and characterize the binary interaction to elucidate the basis of the disease.

Eventually, the concept of a disease-conducive environment came to light. This expanded the historical notion that the mere presence of a pathogen on susceptible host tissue was sufficient to cause disease. Certain environmental variables, such as temperature and humidity, were important factors for the establishment of disease-conducive or disease-suppressive environments. A conceptual "plant disease triangle" was proposed to emphasize the importance of the environment in disease, stating that

the three prerequisites for plant disease epidemics are 1) a susceptible host, 2) a compatible pathogen, and 3) a conducive environment (Figure 1; Stevens, 1960). Yet still to this day, plant-microbe interaction research is overwhelmingly dominated by a handful binary pathogenic interactions.

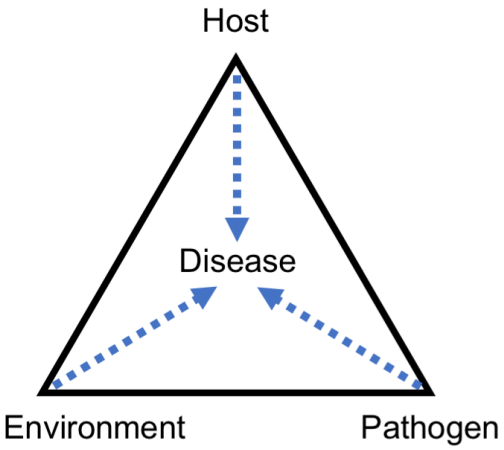


Figure 1.1. Plant disease triangle (adapted from Stevens 1960), depicting the factors necessary for a crop disease to occur and develop to an epidemic. A host plant must be developmentally and genetically susceptible to the pathogen. The environment must be conducive to pathogen proliferation, dispersal, and must confer disease-conducive host physiology to accommodate pathogen proliferation. The pathogen must have appropriate virulence factors to replicate to high levels within the host plant.

Microbial ecology of plant and soil environments

Diverse microorganisms colonize soil in great abundance and provide ecosystem services upon which all life depends. An early pioneer of microbial ecology was Beijerinck. Among his many important discoveries in the field of microbiology, Beijerinck was the first to document biological nitrogen fixation and sulfate reduction, thus demonstrating that microorganisms play a critical role in soil fertility. Beijerinck further reported that N-fixing bacteria were capable of colonizing and forming nodules in the roots of leguminous plants (Chung and Ferris, 1996). Winogradsky, another pioneering soil biologist and microbial ecologist, made major contributions to the field by isolating the first nitrifying bacteria and demonstrating critical roles for microorganisms in the sulfur and nitrogen cycles. Beijerinck and Winogradsky were among the first to promote the ecological concept of microorganisms. The ecology of microorganisms was not considered by most of their contemporaries in the latter half of the 19th century, who instead focused on individual pathogenic microorganisms. Interestingly, Winogradsky began his career as a plant biologist. He later conducted research under the supervision of Anton de Bary, the aforementioned father of plant pathology, but focused on sulfur cycling in soil and pioneered microbial isolation by enrichment and dilution (Dworkin, 2012). Winogradsky was also an influential proponent of environmental stabilization and reduction of complexity for reproducible experimentation. As recounted by Selman Waksman, "...experiments carried out by the youthful Winogradsky may well be considered as among the first careful investigations ever made on the influence of controlled environment on the growth of microorganisms in pure culture, under well-defined experimental conditions" (Dworkin, 2012). Needless

to say, although he did not study the plant microbiome in the context of microbial communities interacting with a host, a strong case can be made that plant microbiome research was first conceived in the laboratory of de Bary.

Winogradsky's careful technique to establish axenic cultures, free of contamination, was also exemplified in the work of the French scientist and close colleague of Louis Pasteur, Émile Duclaux. In addition to his many seminal contributions to medicine, microbiology and the Germ Theory of Disease, a lesser-known contribution by Pasteur that helped to shape the future of microbiome research came from a challenge he posed to the French Academy of Sciences. Perhaps inspired by the concept of pasteurization, Duclaux reported that peas and beans could not thrive in thoroughly sterilized soil and postulated that plants rely upon microbiota for adequate nutrition (Fairman, 1887). Pasteur addressed the French Academy to recognize Duclaux's work, and during his speech, digressed to express doubt that complex multicellular life was capable of survival under axenic, or germ-free, conditions (Pasteur, 1885). Maybe this was true skepticism, or perhaps it was a call to challenge dogma, but nonetheless, Pasteur's skepticism catalyzed the emergence of a new field – Gnotobiology. Animal researchers were not far behind the plant researchers, and systems to rear various animals under effectively germfree conditions were developed by the 1960s (Luckey, 1965). In the words of Duclaux, indeed, "the most fruitful periods of Science are those in which dogmas are shaken" (Duclaux, 1920).

If plant microbiome research was conceived in Anton de Bary's laboratory in Strasbourg, surely it was born in the laboratory of Lorenz Hiltner. Based on culture-based methodology and microscopic investigation, Hiltner postulated that a microbial

community distinct from soil colonizes the roots of plants, likely feeding off root exudate material (Hiltner, 1904). He used the now-colloquial term “rhizosphere” to describe the root-associated microbiome. He further postulated that rhizosphere composition substantially influences plant nutrition (Hartmann et al., 2008). Perhaps Hiltner’s most prophetic insight was the idea that the taxonomic composition of rhizosphere microbiota directly influences plant resistance to pathogens, a concept that is still an active topic of research more than a century later (Hartmann et al., 2008). Interestingly, the entrepreneurial Hiltner shares the first patent for agricultural usage of a microorganism on the use of microbiota for *Rhizobium* inoculants (Hartmann et al., 2008). To put this in perspective, in 2016, the agricultural biostimulants industry (not including biofertilizers and bioinsecticides) is valued at approximately $3 billion USD with an annual projected growth of greater than 10% through 2022 (Raj Khosla, personal communication).

It took nearly 50 years for the microbial communities associated with the above-ground portion of plants to gain attention by the scientific community. In terms of surface area, the plant leaf surface is reported to be the largest biological interface (Delmotte et al., 2009). This environment, including the air-filled apoplastic space of leaf interiors, is termed the “phyllosphere” (Ruinen, 1961). The early history of phyllosphere research is not as rich as that of the rhizosphere, and most literature on leaf-associated microbiota prior to 1960 is almost exclusively focused on individual pathogenic microorganisms. However, by 1930, reports had been published about nutrient exudation and water exchange via the leaf surface, and that leaf-associated microbiota colonized heterogeneously. Microscopic and culture-based evaluations of phyllosphere

microbiota can be found, claiming significantly different fungal and bacterial composition than root-associated communities (Ruinen, 1961). As early as the 1960s, N-fixing bacteria had been isolated from the leaves of tropical plants, but whether or not they contributed to plant nutrition or phenotypes was unknown (Ruinen, 1961). Ruinen further reported that precipitation and moisture content of leaves profoundly impacted microbial abundance and diversity (Ruinen, 1961). Among collections from diverse plants in tropical forests of Indonesia, *Beijerinckia* spp were the most ubiquitous N-fixing bacteria, but other unidentified bacteria, fungi, algae, and protists could be spotted via microscopic evaluation (Ruinen, 1961). Notably, in Rubin's 1961 report, he claims that under normal circumstances, nutrients on the water droplets on a leaf surface are only sufficient for a limited number of cellular divisions, and that vigorous microbial growth only occurs during a fully mature leaf robust photosynthetic activity, and only within pools of water on the leaf; the waxy cuticle is not conducive to much microbial replication.

The *Arabidopsis thaliana* model

Domesticated crops and their pathogens are often difficult to research for various reasons, prompting the adoption of model systems with more powerful tools and fewer logistical constraints. It is important to have multiple diverse models to account for genetic variation across the plant kingdom, but the clear poster child of plant model systems is mouse-ear cress (*Arabidopsis thaliana*, hereafter Arabidopsis). Arabidopsis has proven to be an outstanding model organism for plant research, and it has been used to make seminal discoveries in hormone signaling, stress tolerance, growth, and

development, and disease resistance (Provart et al., 2016). Arabidopsis is an herbaceous flowering dicot of the Brassicaceae family (mustard family), with a relatively short generation time and a fully sequenced genome of 135 Mbp (Arabidopsis Genome Initiative, 2000). By the 1980s, *Arabidopsis* began to rapidly gain popularity for plant researchers around the world (reviewed by Meyerowitz, 2001). After the initial discovery that Arabidopsis was conducive to transformation by *Agrobacterium tumerfacians*, early methods began to emerge to generate selectable Arabidopsis mutants via T-DNA insertion of antibiotic resistance genes (Lloyd et al., 1986). Subsequently, the genome of *Arabidopsis* has been sequenced, and several large and well-characterized T-DNA mutant libraries are now available to the public. Arabidopsis is physically small relative to most domesticated crops, has a short generation time, grows under relatively simple environmental conditions, can be crossed, and has high fecundity. Translation of research is important for model systems, and indeed, key discoveries applicable to other plants have been made in Arabidopsis (Chang et al., 2016). Perhaps most importantly, there is a massive community of Arabidopsis researchers, thus allowing for experimental replication across different labs and a plethora of tools. Arabidopsis is a well-established organism for molecular host-microbe interaction research. One such example is the interaction between Arabidopsis and the pathogen *Pseudomonas syringae* pv tomato DC3000 (*Pst*), the causative agent of bacterial speck disease (Xin and He, 2013). In the early 1990s, it was discovered that *Pst* had the capacity to cause disease in Arabidopsis, and the *Pst*-Arabidopsis pathosystem quickly gained popularity as a platform to dissect bacterial virulence and plant defense (Whalen et al., 1991). A number of plant-microbe model systems are now established that utilize Arabidopsis

as the host, many of which have been used to report fundamental mechanisms of pathogenesis and eukaryotic innate immunity.

Arabidopsis innate immunity

Beyond physical barriers such as the waxy leaf cuticle, the first layer of defense against subversion by microorganisms in Arabidopsis is the innate immune response. The cost of pathogenic relationships and benefits of symbiosis provide adequate selective pressure for the evolution of defense adaptations and intricate signaling networks collectively termed innate immunity. Successful organisms allocate resources towards growth and defense in a niche- and stress-dependent fashion. Defense overspending can be detrimental. Thus, mechanisms to control immune response sensitivity and specificity are important. Despite the fact that host organisms are perpetually colonized in a microbial milieu, very little is known about host-microbiome dynamics or the impact of innate immunity on associated microbiota. Recognition and response to non-self molecular patterns by pattern recognition receptors (PRRs) is collectively termed innate immunity. To initiate a pattern-triggered immunity (PTI) response, a non-self-ligand is first bound by the extracellular domain of a PRR, resulting in autophosphorylation of the cytosolic ectodomain and heterodimerization with other. One such example is recognition of a 22AA motif of the flagellin monomer of certain bacteria (flg22) by the PRR FLAGELLIN SENSITIVE 2 (FLS2) (Gómez-Gómez and Boller, 2000) (Figure 2).

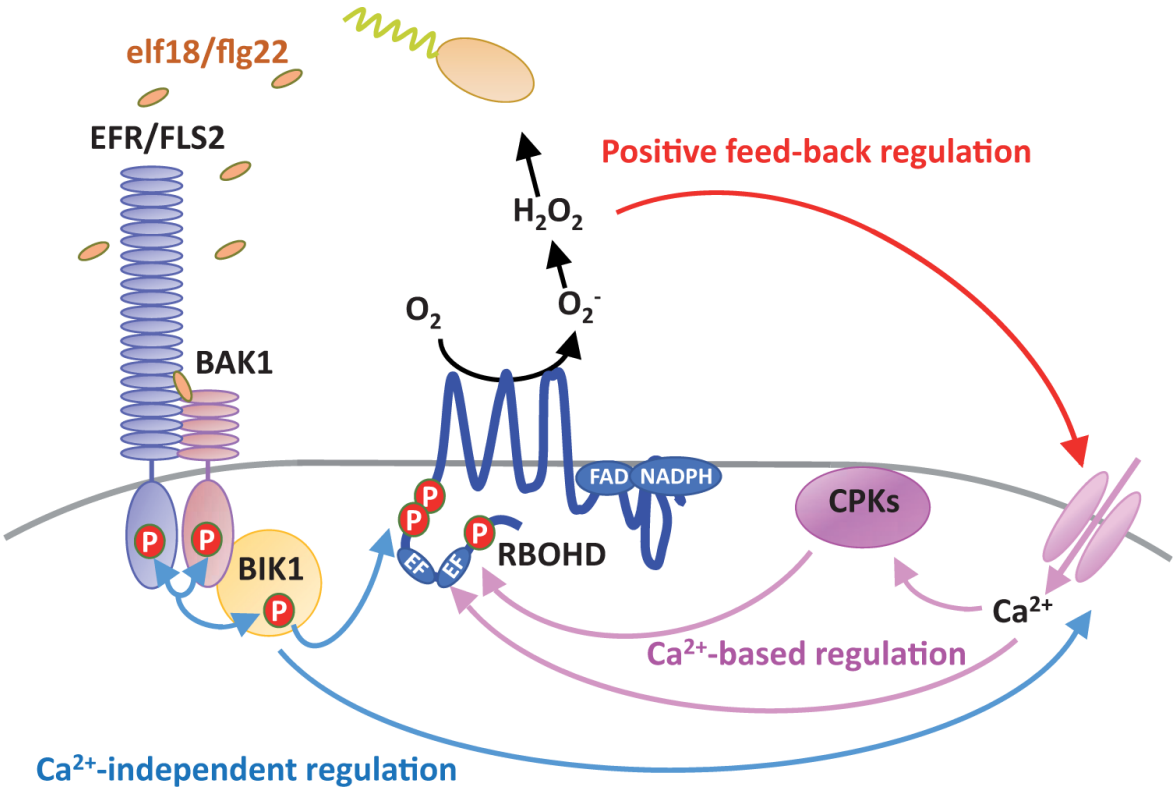


Figure 1.2. Immune responses at the plasma membrane in Arabidopsis. The extracellular leucine-rich repeat domain of a receptor-like kinase (LRR-RLK) pattern recognition receptor (PRR) protein (for example, EFR or FLS2) bind the foreign ligand (EFR binds 18AA 'elf18' from certain bacterial EF-tu; FLS2 binds 22AA 'flg22' from certain bacterial flagellin), resulting in autophosphorylation of the ectodomain and the ectodomain of the co-receptor BAK1. Activated BAK1 directly phosphorylates the NADPH oxidase RBOHD, and also phosphorylates BIK1, which also phosphorylates RBOHD but at a higher affinity. RBOHD reduces dioxygen to superoxide, and in the apoplastic intercellular space between cells, superoxide dismutases (SOD) convert the free superoxide to hydrogen peroxide. This process is called the reactive oxygen species (ROS) burst. Increased concentrations of apoplastic H_2O_2 results in positive feedback and influx of calcium into the cytosol. CPK enzymes interact with RBOHD to increase its activity. (reproduced from Kadota et al., 2015).

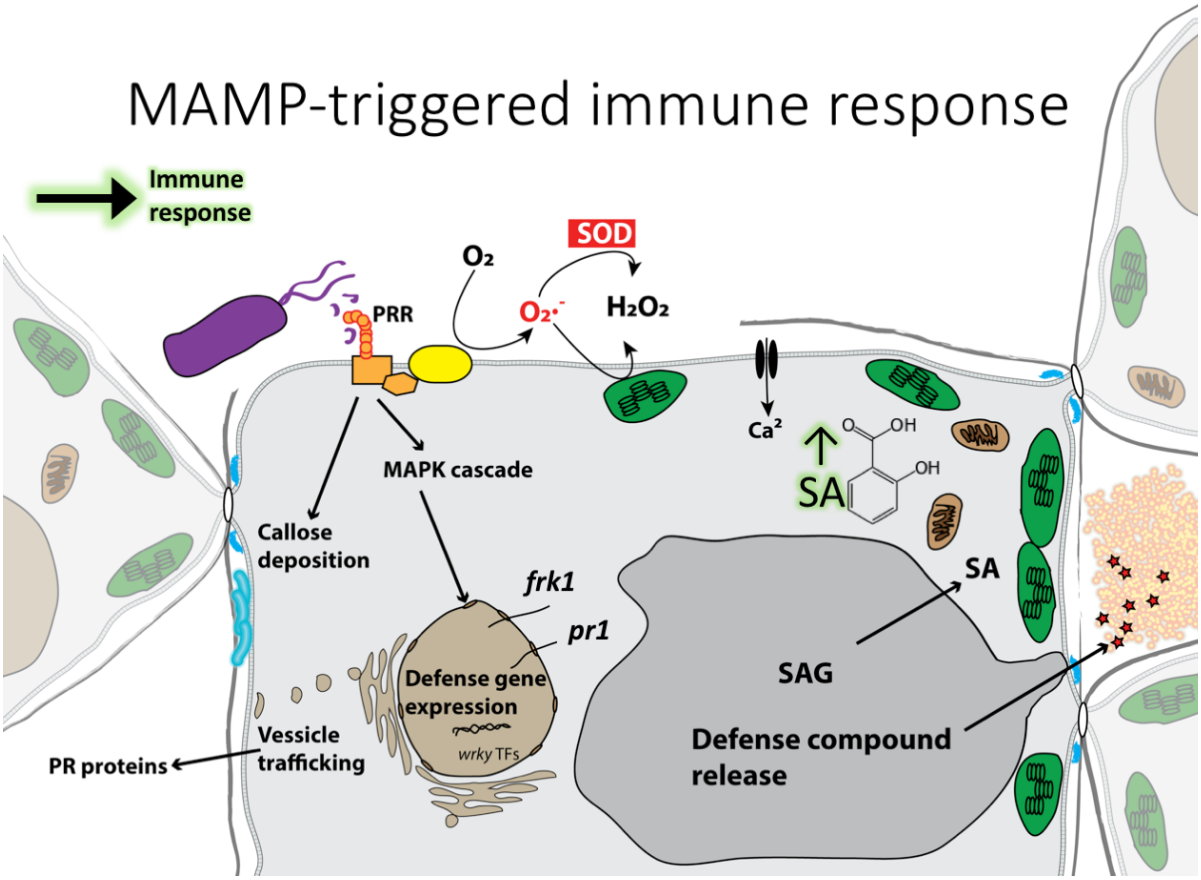


Figure 1.3. MAMP-triggered immune response. PTI activation results in activation of the mitogen-associated protein kinase (MAP kinase) cascades. Activated MAP kinase proteins (MEKK1, MKK5, MPK3/6, and/or MEKK1, MKK1/MKK2, MPK4, MKS1) subsequently localize to the nucleus (Couto and Zipfel, 2016; Xin and He, 2013). Activated MAP kinase proteins bind defense-associated transcription factors, for example, certain WRKY transcription factors), resulting in defense associated gene expression (for example *Frk1* and *Pr1*) (Boudsocq et al., 2010). Additional responses include biosynthesis of Pathogenesis Related proteins (PR proteins); changes in vesicle trafficking, change in redox homeostasis, and increased levels of cytosolic free salicylic acid (SA) (Chisholm et al., 2006). A later response is cell wall reinforcement with callose.

Plant PRRs are the convergently-evolved counterparts to Toll-like receptors in animals (Gómez-Gómez and Boller, 2000). For example, some plants perceive a particular motif of bacterial flagellin through the PRR FLS2 (FLAGELLIN SENSITIVE 2). Similarly, mammals use the Toll-like receptor TLR5 to perceive bacterial flagellin and mount a downstream immune response (Hayashi et al., 2001), albeit a different motif.

During PTI, plants undergo multiple physical and chemical changes. For example, papillae (callose- containing cell-wall appositions) form at infection loci. Callose is a glucan polymer that serves as a matrix in which plant-produced antimicrobials are localized and suspended (Figure 3; Gómez-Gómez et al., 1999). Activated PRRs also initiate downstream defense responses, such as anion fluxes, phosphorylation cascades (including the MAP-kinase cascade), reactive oxygen species accumulation, and defense gene activation (reviewed by (Luna et al., 2011). As demonstrated by Gomez-Gomez and Boller in 2000, exogenous application of MAMP is sufficient to activate MTI. The three best characterized MAMPs/PRRs for *Arabidopsis thaliana* Col-0 are: the semi-conserved N-terminus of bacterial flagellin by FLS2, an 18-AA motif of bacterial EF-Tu by EFR (Kunze et al., 2004), and various forms of fungal chitin and chitosan by CERK1 (reviewed by Chisholm et al., 2006). Peptidoglycan is also known to be a MAMP. A number of other plasma membrane-bound LRR-RLKs also exist in *Arabidopsis*, to which no respective MAMPs have been identified. It is known that MAMP-containing microorganisms regularly inhabit plant tissue, but basal activation of MTI has not been reported. Studies have failed to directly address the effect of exogenous MAMP application on plant-associated microbial communities. Although putative genes from some plant-associated bacteria suggest the potential for reactive

oxygen species tolerance and alkalization tolerance, it is unknown whether some organisms are more tolerant to MTI than others, and likewise, it is not known whether the plant has differential responses to different MAMPs perceived by the same coreceptor complex, or whether this is a generalized response.

Defense hormone signaling

Salicylic acid and jasmonic acid are the two major plant defense hormones and key regulators of plant innate immunity. Jasmonic acid signaling pathway has a significant amount of overlap with the ethylene hormone signaling pathway. The jasmonic acid and salicylic acid hormone signaling pathways are regarded as antagonistic (Bari and Jones, 2009). Salicylic acid signaling is implicated in defense against (hemi-) biotrophic pathogens, while defense against necrotrophic pathogens, including many fungal pathogens, oomycetes, and insects, are dependent on jasmonic acid signaling. Pathogen-induced SA biosynthesis occurs in the chloroplast and is dependent upon isochorismate synthase enzymes, most prominent of which is ICS1 (ISOCHORISMATE SYNTHASE 1). Although salicylic acid can be synthesized in planta via at least two independent pathways, the isochorismate (IC) and the phenylalanine ammonia-lyase (PAL) pathways, the IC pathway is reported to be the most prominent for pathogen-induced production.

The concentration of salicylic acid increases subsequent to an interaction with a pathogenic microorganism. Despite considerable interest, all the steps of salicylic acid biosynthesis have not been elucidated fully. Isochorismate synthase is a critical enzyme in the IC pathway and mutants impaired in isochorismate synthase are more

susceptible to pathogen invasion (Dempsey et al., 2011; Wildermuth et al., 2001)**.** Various forms and modifications of salicylic acid exist in plants. In Arabidopsis, salicylic acid can be methylated, sulfonated, hydroxylated, conjugated with amino acids, and glucosylated. Potentially salicylic acid can be modified to other forms, and further investigation is necessary to make such claims. SA increases the production of reactive oxygen species such as hydrogen peroxide and in general creates a more oxidative environment in the cytosol (Tata, 2005). This will increase the concentration superoxide dismutases. Directly relevant to defense against pathogens, SA prompts the expression of Pathogenesis related (PR) proteins, some of which are reported to have antimicrobial activities. Many SA-inducible genes perpetuate a positive feedback loop of SA-associated defense responses. For example, the SA-inducible protein PAD4 promotes SA biosynthesis in coordination with a similar lipase-like protein EDS1 (Zhou et al., 1998). The master regulator of SA- and JA mediated transcriptional reprogramming during immunity is NPR1. Sweeping REDOX changes upon SA accumulation leads oligomerization of cytosolic NPR1 oligomers, and monomers are released into the nucleus. Numerous post-translational modifications occur to defense-associated proteins after pathogen encounter and in response to elevated SA levels (Withers and Dong, 2017)**.**

Jasmonic Acid and Salicylic acid have been implicated in influencing microbiome community composition in Arabidopsis. In a field study, JA/ET and SA defense-compromised mutants (in the same *A. thaliana* Col-0 background), as well as *Arabidopsis* with exogenous MeJA or SA application were planted among wildtype controls, and culturable leaf-epiphytic and leaf-endophytic bacterial populations were

monitored (Kniskern et al., 2007). Major findings from this study are 1) foliage-associated epiphytic diversity is significantly greater than endophytic diversity in WT plants; 2) induction of SA-mediated defenses reduced endophytic bacterial community diversity; 3) JA-mediated defenses deficiencies confer greater epiphytic bacterial diversity. The small sample size, use of semi-selective culture media, decision to ignore fungi, failure to examine non-culturable populations, single sample site without controls for microhabitat variation, and failure to simultaneously examine root-associated communities leaves us with many unanswered questions regarding the impact of plant defense hormones on associated communities. Another study focused on fitness effects of activated defense signaling (Traw et al., 2007). The Traw *et al.* study reports that induction of SAR, but not JA/ET related defenses, increases the fitness of *Arabidopsis in situ*, despite the (untested) costs that accompany SAR induction. Fitness in the Traw *et al.* study is defined as seed production (seed viability not addressed), and there was a lack of controls to determine the effects of SA and/or JA application have on a germ- free plant, and therefore the fitness benefit conferred by exogenous SA application cannot be attributed to the correlated loss of bacterial diversity. Another study utilized culture-dependent techniques, as well the culture-independent community fingerprinting technique denaturing gradient– gel electrophoresis (DGGE) targeted to the 16S rRNA gene, to examine the effects of JA/ET and SA on the *Arabidopsis* rhizosphere (Doornbos et al., 2011). The Doornbos study was able to identify significantly different rhizosphere community fingerprints for each of the different *Arabidopsis* genotypes and control plants of other species, only three biological replicates were used for each genotype, multiple knock-out lines were

not used to determine if indeed the “fingerprint” can be attributed to the mutant allele. More recently, a reverse genetics approach was published by Lebeis et al. reported that SA-signaling is required for the homeostasis of root-associated microbiota and that the taxonomic composition of the root microbiome is sculpted in a SA-dependent manner. Lebeis et al hypothesize that SA-mediated modulation may occur by a conceptual “gating” of particular bacterial taxa as a in a manner of tolerance to SA-mediated defense, but this raises questions regarding whether previously uncharacterized branches of immune regulatory networks play a substantial role in microbiome governance (Lebeis et al. 2015).

Systemic acquired resistance and induced systemic resistance

Systemic acquired resistance (SAR) and induced systemic resistance (ISR) are two important microbe-induced phenomena with long-term systemic effects on plant immune response and physiology. After infection, endogenous salicylic acid levels rise throughout the plant, resulting in systemic “priming” against infection by some, but not all, microbial pathogens (Durrant and Dong, 2004; Jung et al., 2009). On one hand, plants with activated SAR have higher basal levels of PR proteins (for example, PR-1) and other not fully characterized proteins with hypothetical roles in defense. On the other hand, ISR activated by non-pathogenic rhizobacteria is jasmonate-signaling dependent, and functions independent of SA signaling (Van Wees and De Swart, 2000). Both ISR and SAR are dependent on the function of NPR-1, a master defense response regulator protein (Dong, 2004). Although many intricacies of plant defense signaling and priming have yet to be unraveled, it appears that SAR is most effective

against biographic (derive nutrients from living cells) pathogens and ISR is most effective against necrotrophic pathogens (Durrant and Dong, 2004).

Chapter 2

Construction and optimization of a soil-based growth system conducive to axenic, gnotobiotic, and holoxenic Arabidopsis growth

Acknowledgements: Western blot contributed by Dr. John Froehlich. Some Plant images contributed by Cait Thireault. Experiments in the calcined clay system conducted in collaboration with Dr. Amine Hassani.

Abstract

The presence of endogenous microbiota on and inside plants is hypothesized to influence many phenotypic attributes of the host. Likewise, host factors and microbe-microbe interactions are believed to influence microbiome community assembly. Rigorous testing of these hypotheses necessitates the ability to grow plants in the absence or presence of endogenous microbiota. We introduce the FlowPot axenic growth platform and report vigorous Arabidopsis growth in the presence or absence of microbiota in the system's peat-based substrate. Mechanically, the FlowPot system is unique in that it allows for total-saturation "flushing" of sterile substrate with water and/or nutrient solution, contributing to the healthy appearance of axenic plants *in situ*. We compare bacterial community establishment in two easily-constructed axenic growth systems on the model plant *Arabidopsis thaliana*: an established system with calcined clay substrate and a novel system with peat substrate. We report bacterial community establishment trends in both systems after inoculation with directly-extracted soil-derived microbiota and the established "*At*-SPHERE" community that collectively represents Arabidopsis-associated bacteria cultured from nature. We report system-specific community shifts, as well as commonalities and discrepancies in communities by natural and defined input communities. Relative to the calcined clay growth system, the peat-based FlowPot system retains greater alpha diversity of root-associated communities, and in particular, allows for more proteobacterial growth. In contrast, the calcined clay system is conducive to an actinobacterial-dominated community and increased granularity when comparing rhizosphere versus bulk substrate bacterial communities. This standardized suite of tools and colonization

protocols empowers the plant microbiome research community to conduct harmonized experiments to reveal mechanistic bases of microbial community succession and the impact of microbiota on host phenotypes.

Introduction

In an 1885 address to the French Academy of Sciences, Louis Pasteur expressed doubts that axenic (i.e., germ-free) animals were capable of survival (Pasteur, 1885). Pasteur went on to credit the work of Duclaux who reported that peas and beans could not thrive in thoroughly sterilized soil and that plants likely require nutritive assistance from bacteria (Fairman, 1887). Pasteur's skepticism catalyzed the emergence of a new field – Gnotobiology. Animal researchers were not far behind the plant researchers, and systems to raise mice, chickens and guinea pigs devoid of conventionally culturable microorganisms were established before the first half of the twentieth century (Luckey, 1965). The axenic mouse model, in particular, has proven to be a tremendously valuable tool for microbiome research. Seminal studies demonstrate that enteric microbiota is essential for proper nutrition, enteric physiology, cardiovascular function, and immune function (Clemente et al., 2012; Gensollen et al., 2016; Smith et al., 2007). Two major motivations for establishing axenic experimental systems that also accommodate gnotobiotic (defined microbiota) or holoxenic (undefined microbiota) hosts are 1) to elucidate effects of host-associated microbiota and 2) to identify factors that influence microbiome structure and function. Careful consideration must be taken when designing an axenic growth system to minimize artifacts and sample variability, but not at the expense of versatility.

Environmental factors are the major drivers of differential taxonomic composition and diversity among host-associated microbiomes (Adair and Douglas, 2017; Peiffer et al., 2013; Spor et al., 2011). Thus, abiotic factors must be accounted for when performing a microbial colonization experiments for reproducibility of host phenotypes and

community dynamics. Variation in the food source from commercial suppliers of near-isogenic mice can result in different microbiome composition, and as a result, altered host immune phenotypes (Celaj et al., 2014). In the model plant Arabidopsis, humidity influences microbiome composition and the plant's ability to defend against foliar pathogens (Xin et al., 2016). Ecological drivers of soil microbiomes are associated with physical and chemical factors, including, but not limited to porosity and water retention, C/N/P availability, pH, percent organic content, and cation exchange capacity (Andrew et al., 2012; Fierer and Jackson, 2006). When conducting a microbiome colonization experiment, and comparing gnotobiotic or holoxenic plants to axenic plants, it is essential to consider abiotic factors of the substrate as well as the growth environment, particularly if the intention is to recapitulate a microbial community reminiscent of a "natural" soil microbiome.

Relative to the mechanical, technical, and financial requirements to establish and maintain an axenic mouse colony, an axenic system for a model plant such as Arabidopsis is straightforward (Nicklas et al., 2015). Axenic Arabidopsis growth can be accomplished using routine tissue culture methodology, which in brief entails aseptic placement of a decontaminated seed on a phytonutrient agar substrate (or similar) contained within a sterile, gas-permeable, transparent vessel. Tissue culture systems do not provide a soil-like scaffold for microbial colonization. Furthermore, agar-based systems are notorious for non-uniform nutrient and O_2 delivery over time (Gunning and Cahill, 2009). Hydroponic and aeroponic systems can alleviate the uniformity of nutrient and O_2-delivery by agitation and media replenishment, but do not provide a soil-like scaffold for microbial colonization (Jackson et al., 1991). Furthermore, it can be

challenging to maintain axenic conditions or prevent cross-contamination in common-reservoir hydroponic systems.

Non-soil substrates such as sand, quartz, vermiculite and calcined clay are frequently used in gnotobiotic systems (Bai et al., 2015; Henry et al., 2006; Kloepper et al., 1981; Lebeis et al., 2015). These substrates are porous, thus providing surface area for microbial colonization and root penetration. However, batch-to-batch variation of ceramic substrates can result in a wide range of labile ions (Henry et al., 2006). Calcined clay, for example, has sorptive properties that can result in reduced labile concentrations of P, Fe, Cu and Zn, and desorptive properties that can result in excess labile B, Mg, Ca, S, K, and sometimes to toxic levels, e.g. Mn (Adams et al., 2014). While thorough washing or soaking of the non-soil substrate can reduce the initial excess of labile ions, flow and drainage are important to reduce significant changes in chemistry over time. Furthermore, the aforementioned non-soil substrates lack significant organic carbon typical of soil, unless supplemented.

Soil has also been used as substrate in axenic systems, although it can present challenges due to contamination from insufficient sterilization and hindered plant growth due to suboptimal sterilization protocols. Numerous sterilization methods have been used with soil, including: autoclaving, dry heat, irradiation, microwave, fumigation by gaseous chemicals, saturation with various sterilants (i.e., mercuric chloride, sodium azide, formaldehyde) (Trevors, 1996). All methods of soil sterilization alter, to some extent, physical and/or chemical properties of soil. Chemical sterilization methods are not appropriate for plant growth systems due to phytotoxic effects. Autoclaving soil has been shown to increase levels of water-soluble carbon and reduce pH (Shaw et al.,

1999), but not significantly alter ion exchange capacity. Gamma-irradiation has been reported to minimally disrupt the physical attributes of soils but can result in the generation of reactive oxygen species, capable of depolymerizing C-C of polysaccharides (Bank et al., 2008). Both autoclaving and gamma irradiation can result in changes of the physical structure of the soil, exposing more surface area and thus altering sorptive properties. Complete sterilization of some soils can be achieved with minimal chemical alterations by autoclaving a thin layer of soil for three short (<45 min) autoclave cycles with day intervals (Berns et al., 2008; Lotrario et al., 1995; Wolf et al., 1989). Subsequent flushing of the sterile substrate can increase plant performance, presumably by rinsing away soluble phytotoxic byproducts.

Here, we compare bacterial community establishment in two easily-constructed axenic Arabidopsis growth systems: an established system with calcined clay substrate (Bulgarelli et al., 2012; Lundberg et al., 2012) and a novel system with peat substrate. The peat-based FlowPot system features an inoculation port on each individual vessel that allows for substrate flushing to remove soluble byproducts of sterilization, provides drainage, and accommodates homogenous inoculation with microbiota and/or nutrients. We inoculate both systems with natural soil-derived microbiota from a field abundant with Arabidopsis, and the established "*At*-SPHERE" synthetic community that collectively represents Arabidopsis-associated bacteria (Bai et al., 2015). This set of experiments seeks to empirically determine whether growth system influences microbial community succession in synthetic and natural communities. This standardized suite of tools and colonization protocols empowers the plant microbiome

research community to conduct experiments to reveal mechanistic bases of microbial community succession and the impact of microbiota on host phenotypes.

Materials and Methods

Sterilization of the FlowPot peat-based substrate

Peat potting mix (Sunshine Mix (LG3), SunGro Horticulture, USA) was blended with horticulture-grade medium vermiculite (1:1). The substrate mixture was sprayed with distilled water to achieve ~20% moisture content. Substrate was evenly spread at a depth of 4 cm onto polypropylene or glass trays and covered with aluminum foil in such a way to prevent condensate from pooling and unevenly saturating the substrate, and autoclaved for 30 min on liquid cycle (121.1°C, 15 PSI, slow exhaust with forced liquid cooling). After autoclaving, the substrate was stored for 24 to 48 h at room temperature. The autoclave cycle was repeated, and the substrate was stored for an additional 24 to 48h. The substrate was then aseptically homogenized and ready for use. The substrate was prepared immediately before FlowPot construction. In our experience, depending on the initial moisture content of the substrate, relative humidity, and the calibration of the autoclave, autoclave parameters may need to be optimized empirically.

FlowPot growth system assembly

FlowPots were constructed and assembled using easily-obtained laboratory materials (for supplier details, see Table 1). Luer taper syringes (50 ml, Polypropylene, non-lubricated) were sectioned using a mitre saw with a reversed fine-tooth blade at the "20 ml" mark, retaining only the portion with the Luer connector. Residual debris was removed, and syringes were soaked for 20 min in Multi-Terge ionic detergent (2% v/v), followed by thorough rinsing with distilled water and sterilization by autoclave. Fiberglass screen (Phiferglass, Pfifer Inc., USA) was sectioned into 5 cm squares and

autoclaved. To create the FlowPot platforms, autoclave-compatible plastic (polypropylene or polycarbonate, typical for pipet tip boxes) was sectioned into 12 cm × 8 cm × 1 cm blocks and four holes (8.8 mm diameter) were drilled. Each stand was fastened to the inside bottom of Eco2 micro box (model TPD1600 (XXL), Combiness, USA), secured with filament tape.

Sterile 3 mm soda-glass beads (n=10) were aseptically added to each syringe top and placed on a culture tube rack. The syringe tips were gently packed with the twice-autoclaved substrate mixture until slightly heaping (Figure 1). In our experience, consistency of substrate packing density and avoidance of overly-compressing the substrate were important for optimal plant growth results. The FlowPot construction was completed by covering the substrate-filled syringe tips with a fiberglass screen (one per unit), and securing the assembly with a polycarbonate cable tie (Figure 1). Completed FlowPots were placed on a culture tube racks, put into covered polypropylene bins, covered with autoclave paper or foil, and autoclaved for 45 min on liquid cycle (121.1°C, 15 PSI, slow exhaust with forced liquid cooling). Immediately after autoclaving, the FlowPots were placed in a sterile laminar flow hood or biosafety cabinet to a sterile laminar flow hood. Constructed microboxes were placed in polypropylene bins and covered with autoclave paper. Micro boxes were not stacked, and lids were not secured to allow for thorough autoclave sterilization.

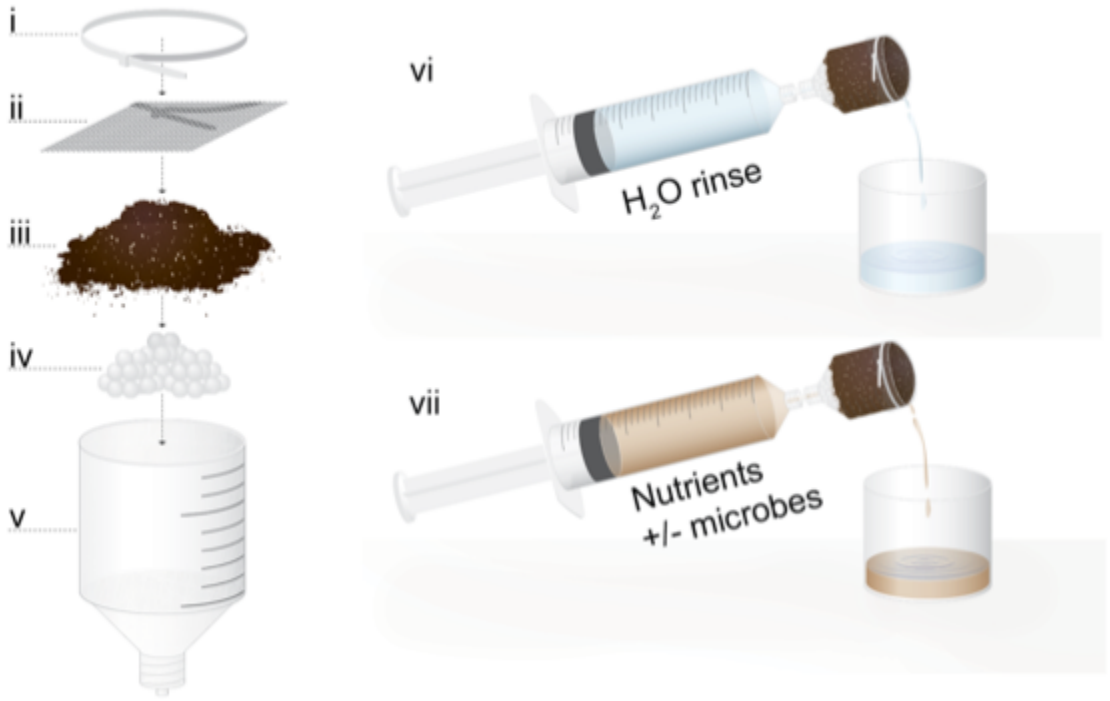


Figure 2.1. Schematics of FlowPot construction and irrigation. Each FlowPot was prepared by fastening a (i) zip-tie to a (ii) mesh retainer to secure (iii) sterilize peat/vermiculite substrate and (iv) glass drainage beads within the (v) 20 ml portion of a 50 ml Luer-lock syringe barrel. Assembled FlowPots were autoclaved, (vi) aseptically rinsed with sterile water, and (vii) irrigated with nutrient solution and (optionally) microbiota.

Calcined clay system assembly

The calcined clay growth system was assembled according to a previously published protocol (Bai et al., 2015; Lebeis et al., 2015). In brief, the substrate (Diamond Pro Calcined Clay Drying 570 Agent, Diamond Pro, Arlington, TX, USA) was thoroughly rinsed with sterile ddH_2O, then autoclaved for 45 min and incubated at 65°C until leather-dry. Calcined clay was aseptically added to Magenta boxes at 100 g per box (approximately half-filled).

Soil collection

German Soil (CAS10) was collected during Spring 2014 from an agricultural field in Cologne, Germany (50.958°N, 6.865°E), at the Max Planck Institute for Plant Breeding Research. The soil has not been subjected to cultivation or chemical treatments for a minimum of 10 years. Large particulate debris was removed from the soil using a steel sieve (3 mm). This identical soil has been used in previously a published study (Zgadzaj et al. 2016).

Michigan Soil (MS15MSU) was collected during Fall 2015 from a sandy loam agricultural field at Michigan State University, East Lansing, Michigan, USA (42.709°N, -84.466°E). The field was being used for *Miscanthus* cultivation and had not been subjected to tilling, fertilization, or any intervention for a minimum of 10 years. The soil was collected from 5 to 15 cm below the surface. Upon collection, the soil was spread out on tables and allowed to sit for one week at room temperature with ~50% relative humidity. Large debris was removed using a 3 mm galvanized metal screen, and 50 g aliquots were prepared in Whirl-Pak bags and stored at 4°C in the dark.

Preparation of input community microbiota

The soil was suspended in sterile distilled water (50 g·L^{-1}) in a covered 2L Erlenmeyer flask and agitated on a rotary shaker for 20 min at 22°C. Afterwards, the soil slurry was allowed to settle for 5 min and filtered through a 40 µm cell strainer into a sterile medium bottle. When required for axenic growth system preparation, a portion of the soil slurry was autoclaved for 45 min. Immediately prior to growth system inoculation, Murashige Skoog nutrient solution (pH 5.7, MES-buffered, Gamborg vitamins) was added to the slurry, bringing the final concentration to either 1/2X or 1/4X, as specified (Murashige and Skoog, 1962). The culture-derived input community *At*-SPHERE was prepared as published previously and resuspended in Murashige Skoog nutrient solution (pH 5.7, MES-buffered, Gamborg vitamins) at 10^6 CFU·ml^{-1} concentration.

Growth system inoculation

Using a sterile test tube clamp or equivalent holder, each FlowPot was aseptically inverted over a sterile funnel placed atop a waste flask. While inverted, each FlowPot was aseptically irrigated by upward flow from the bottom over the course of 30 s with sterile distilled water (50 ml, 65C) using a female-female Luer coupler on a 50 ml syringe (Figure 1). To reduce the risk of contamination, infiltrations were performed in a biosafety cabinet cleaned with Spor-Klenz (STERIS, USA), instruments (Supplemental Figure 1) were flame-sterilized between each unit, and infiltrated FlowPots were placed on a sterile culture tube rack between infiltrations. After irrigation with water, FlowPots were allowed to sit for 30 min. Subsequently, FlowPots were irrigated with the prepared input microbiota. Inoculated FlowPots were then aseptically placed in sterile micro boxes, totaling four per micro box.

Arabidopsis growth conditions

Arabidopsis seeds (250-300 µm) were decontaminated using a chlorine gas sterilization protocol for 6-8 h (Clough and Bent, 1998). Seed-borne contamination and germination efficiency were evaluated by incubating an aliquot of seeds on R2A agar for one week at 22°C. Prior to sowing, seeds were suspended in sterile distilled water at 4°C in the dark for 48-72 h. Per FlowPot or Magenta Box, eight seeds were sown. Plants were grown in growth cabinets set to the following parameters: 23°C with 12/12 h day/night light cycle at an intensity of ~80 µE·m–2·s^{-1}. Approximately two weeks after germination, seedlings were thinned from each FlowPot and Magenta Box, leaving three plants per FlowPot. For an additional contamination check, removed seedlings were incubated on R2A agar for seven days.

Protein extraction and quantification of photosynthesis-associated proteins

Briefly, single *Arabidopsis* leaves (3 weeks old) from FlowPot-grown axenic or holoxenic plants (MS13MSU community) were collected and homogenized using Minute-Chloroplast Isolation Kit (Invent Biotechnology, Inc. Plymouth, MN, USA) according to manufacturer protocol. Chlorophyll content of total leaf homogenate was determined by the method of Arnon (Arnon, 1949) to give a final concentration of 1mg chlorophyll/ml for all samples. All total leaf homogenate (5 mg chlorophyll total) samples were solubilized in Laemmli buffer (Laemmli, 1970) and resolved by SDS-PAGE and either stained with Coomassie Blue or transferred onto polyvinylidene difluoride (PVDF) membrane (Invitrogen) and probed with the following antibodies: anti-D1 Protein (Agrisera, AS01 016), OEC33 (Agrisera, AS06 142-33), OEC23 (Agrisera, AS06 167) or antibodies produced in-house: Anti-Toc159, Anti-Toc75, anti-Tic110,

anti-ClpC. All primary antibodies were incubated using a 1:4,000 dilution. The detection method employed used a secondary anti-rabbit conjugated to alkaline phosphatase (KLP, Inc. Gaithersburg, MD) at 1:5,000 dilution for 1 hour in 5% DM/TBST. The blots were developed using a standard Alkaline Phosphatase (AP) detection system with BCIP/NBT as substrates (Sigma-Aldrich, St. Louis, MO, USA). Quantification of photosynthesis-associated proteins was performed by Dr. John Froehlich, Michigan State University.

Sample collection and DNA extraction

At-SPHERE input microbiota aliquots (3 x 15 ml) were collected at the time of each inoculation, concentrated by centrifugation (15 min at 4,000 x *g*). Bacterial pellets were resuspended in nuclease-free water (Qiagen) stored until Lysing Matrix E tubes at -80C until further processing. German soil input microbiota aliquots (3 x 500 mg) were collected in Lysing Matrix E (MP Biomedicals), and stored at -80C until further processing.

After six weeks of co-incubation in a light cabinet, roots and substrate samples were collected separately from each growth vessel. Roots from each magenta box or micro box were pooled per container and washed twice in sodium phosphate buffer containing Silwet L-77 (0.02% v/v) to remove adherent debris. Within 2 hours, roots were placed in a Lysing Matrix E tube (MP Biomedicals), immediately frozen in liquid nitrogen, and stored at -80°C until further sample processing.

Microbiota from clay substrate were suspended by vigorously shaking in sodium phosphate containing Silwet L-77 (0.02% v/v) and concentrated by centrifugation (10 min at 4,000 x *g*). Bacterial pellets were resuspended in nuclease-free water and

transferred to Lysing Matrix E to be stored at -80°C until further sample processing. For peat substrate, a 500 mg aliquot was collected per FlowPot in a Lysing Matrix E tube and stored at -80C until further processing.

For DNA extractions, samples were homogenized twice by a Precellys 24 tissue lyser (Bertin Technologies), and DNA was extracted according to the manufacturer's protocol (FastDNA SPIN kit for Soil; MP Biomedicals). DNA was eluted in water and stored at -20°C.

16S rRNA gene fragment amplification and MiSeq Library preparation

The 16S rRNA gene fragment DNA libraries for Illumina sequencing were prepared using a previously published two-step amplification protocol (Bai et al., 2015). In brief, DNA concentrations of gel-purified amplicon DNA were adjusted to approximately 3.5 $ng \cdot \mu l^{-1}$. We targeted the V5-V7 region of the 16S rRNA gene using PCR primers 799F (5'-AACMGGATTAGATACCCKG) and 1193R (5'-ACGTCATCCCCACCTTCC). Thermocycling conditions and reagent concentrations were performed according to a previously published protocol (Bai et al., 2015). All reactions were performed in triplicate, and products subsequently pooled. Enzymatic cleanup of the PCR reactions to remove residual nucleotides was performed using 20U Exonuclease I (20U) and antarctic phosphatase (5U) for a 15 min incubation (37°C), followed by a 15 min heat-inactivation (85°C) (New England, BioLabs). The enzymatically processed DNA amplicons were then PCR-barcoded with 12 bp GoLay tags (Caporaso et al., 2012) and Illumina sequencing adapters according to the manufacturer's protocol. Barcoded amplicons were gel-purified according to the manufacturer's protocol (QIAquick Gel Extraction Kit; Qiagen). DNA samples were quantified, pooled, and equilibrated to

~7 ng·μl^{-1}. The amplicon libraries were then subjected to AMPure XP bead purification, according to the manufacturer's protocol (Agencourt, Beckman Coulter). Sequencing was performed using the Illumina MiSeq platform with the MiSeq Reagent kit (version 3 chemistry) and 2 x 300 bp paired-end sequencing protocol.

Preprocessing of 16S rRNA gene fragment amplicons

Quality-filtered reads were demultiplexed by barcode, allowing for one mismatch per barcode. Paired-end reads were merged using PANDAseq (version 2.11; (Masella et al., 2012) using the default parameters for the Ribosomal Database Project pipeline and a minimum PHRED score of 25 (Cole et al., 2014). The PANDAseq parameters were used: "-A rdp_mle -C min_phred:25 -d rbfkms". Reads were further processed using the previously published unsupervised UPARSE method, using the USEARCH software package (version v10.0.240; Edgar, 2013). OTUs were clustered at 97% identity into operational taxonomic units (OTUs), and were classified using the RDP reference database to the highest taxonomic level with >70% confidence (release 11.5; (Cole et al., 2014). Bray-Curtis and weighted UniFrac distances were computed using the R package "phyloseq". Weighted UniFrac distances were calculated between each sample's OTU assemblages (Lozupone et al., 2011). For alpha-diversity analysis of the bacterial communities, read counts were rarefied to an even sequencing depth of 5,000 reads per sample. For beta-diversity analysis of the bacterial communities, read counts were normalized by cumulative sum scaling normalization factors.

Results

Arabidopsis growth in axenic FlowPots

For all experiments presented herein, axenic plants were grown in tandem. A common problem with axenic growth systems that use soil substrate is the release of phytotoxic byproducts during sterilization, resulting in stunted growth and chlorosis or anthocyanin accumulation (Jackson et al., 1991; Trevors, 1996). To assess whether Arabidopsis grown under axenic conditions in the FlowPot system, which involves a critical flushing step to remove possibly phytotoxic byproducts, exhibited any symptoms of phytotoxicity, we carefully examined whole rosettes up to 9 weeks post germination in parallel with Arabidopsis grown in holoxenic FlowPots inoculated with agricultural soil collected from Michigan. Over the course of 9 weeks, we failed to identify any chlorotic tissue or visible anthocyanin accumulation (Figure 2). There were no visible signs of stress in Arabidopsis in the calcined clay system. Microbial contamination was checked using R2A medium at the time of sowing and at the time of collection, and any micro boxes that contained contaminated samples were discarded. To further assess the health of axenic Arabidopsis from the FlowPot system, we quantified seven different photosynthesis-associated proteins from total protein extracted from whole rosette tissue. The abundance of axenic Arabidopsis photosynthesis-associated proteins was quantified relative to simultaneously-grown Arabidopsis in holoxenic substrate (inoculated with MS15MSU Michigan soil microbiota) via Western Blot using native antibodies (Figure 2). The proteins quantified were two outer envelope chloroplast translocases (Toc159, Toc75), an inner envelope chloroplast translocase (Tic110), ATP-Dependent Chloroplast Protease (ClpC), and

three photosystem II complex proteins (D1 protein, OEC33, OEC23). All seven proteins were equivalently abundant in holoxenic and axenic rosettes.

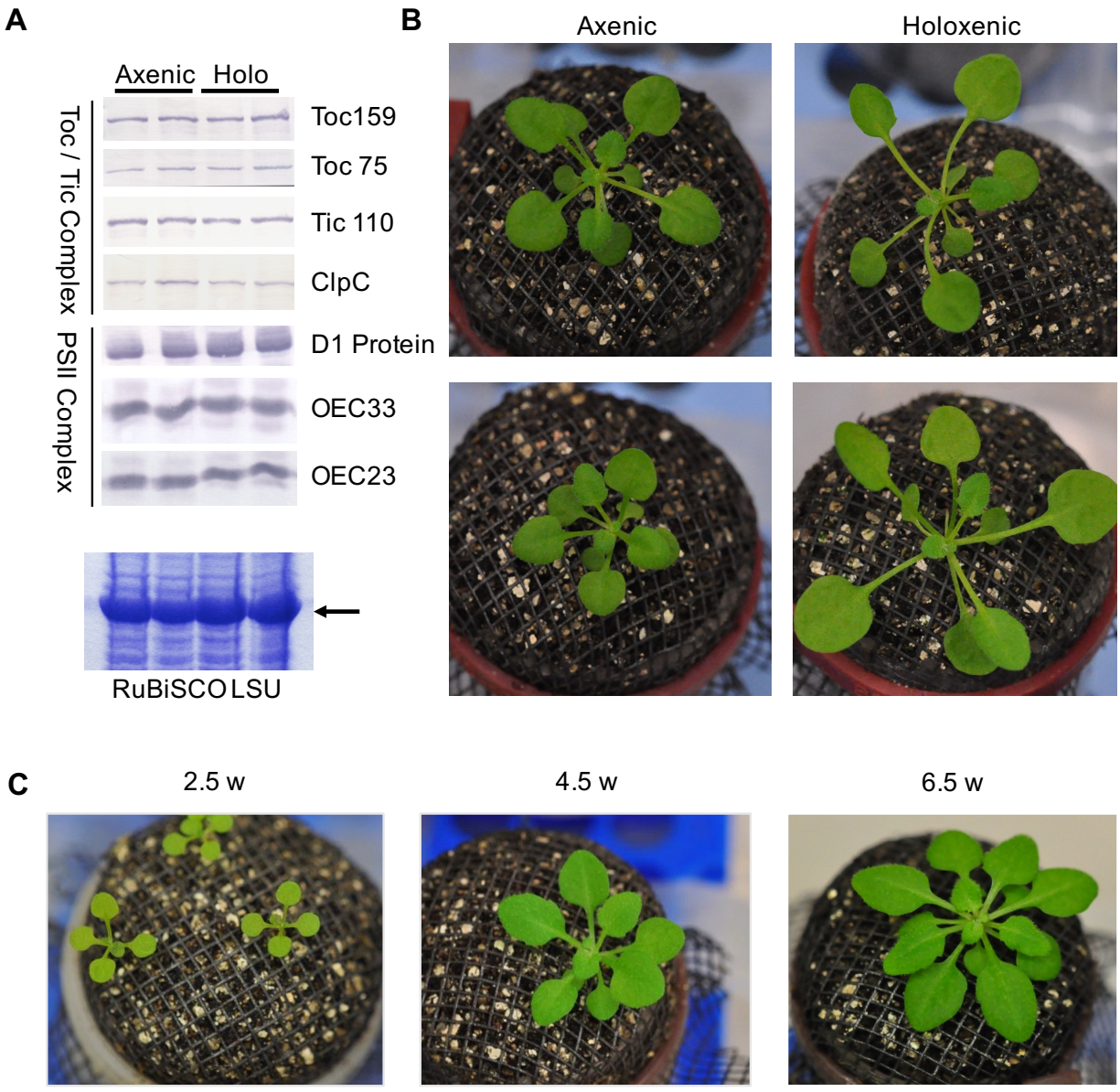


Figure 2.2. *Arabidopsis thaliana* grown in FlowPots with axenic or holoxenic substrate, and photosynthetic protein detection from leaf tissue. (A) Photosynthesis-associated protein quantification from total protein extracts of rosette tissue 3 w post germination. (B) Arabidopsis in FlowPots 4 weeks post germination. Holoxenic substrate was inoculated with a natural soil from Michigan (MS15MSU). (C) Axenic Arabidopsis growth, photographed at 2.5 weeks, 4.5 weeks and 6.5 weeks post germination. Rosette images and the protein gel is representative of at least three replicated experiments. Western blot contributed by Dr. John Froehlich (Michigan State University) and plant photographs contributed by Cait Thireault (Michigan State University).

Beta-diversity

We hypothesized that substrate from different plant growth platforms will influence the microbial community differentiation, resulting in distinct root-associated microbiomes. To test this hypothesis, we inoculated the Magenta box calcined clay system and the FlowPot peat system with established *At*-SPHERE bacteria (Bai et al., 2015), or with complex microbiota directly extracted from German soil (Supplemental Table). We sequenced and performed community analysis on 16S rRNA gene fragment amplicons from DNA collected from (1) *At*-SPHERE and soil input communities, (2) bulk substrate six weeks after inoculation and (3) from whole Arabidopsis roots at six weeks. This resulted in a total of 1,830 OTUs across 90 samples. Samples clustered together by growth system (calcined clay or peat substrate), and by compartment (root or bulk substrate) based on principle coordinate analysis (PCoA) of weighted UniFrac distances (Figure 3). This observation was consistent for samples derived from systems inoculated with German soil microbiota (Figure 3A) and the *At*-SPHERE defined bacterial collection (Figure 3B). For soil-inoculated samples, sample type (root or bulk substrate) was explained by axis PCoA2 (16.0% of variation), and growth system (clay versus peat FlowPot) was explained by axis PCoA2 (60.0% of variation). Trends across axes were less clear for the defined *At*-SPHERE inoculated samples. Irrespective of input community, clusters from clay root and substrate samples had greater distinct separation than FlowPot samples.

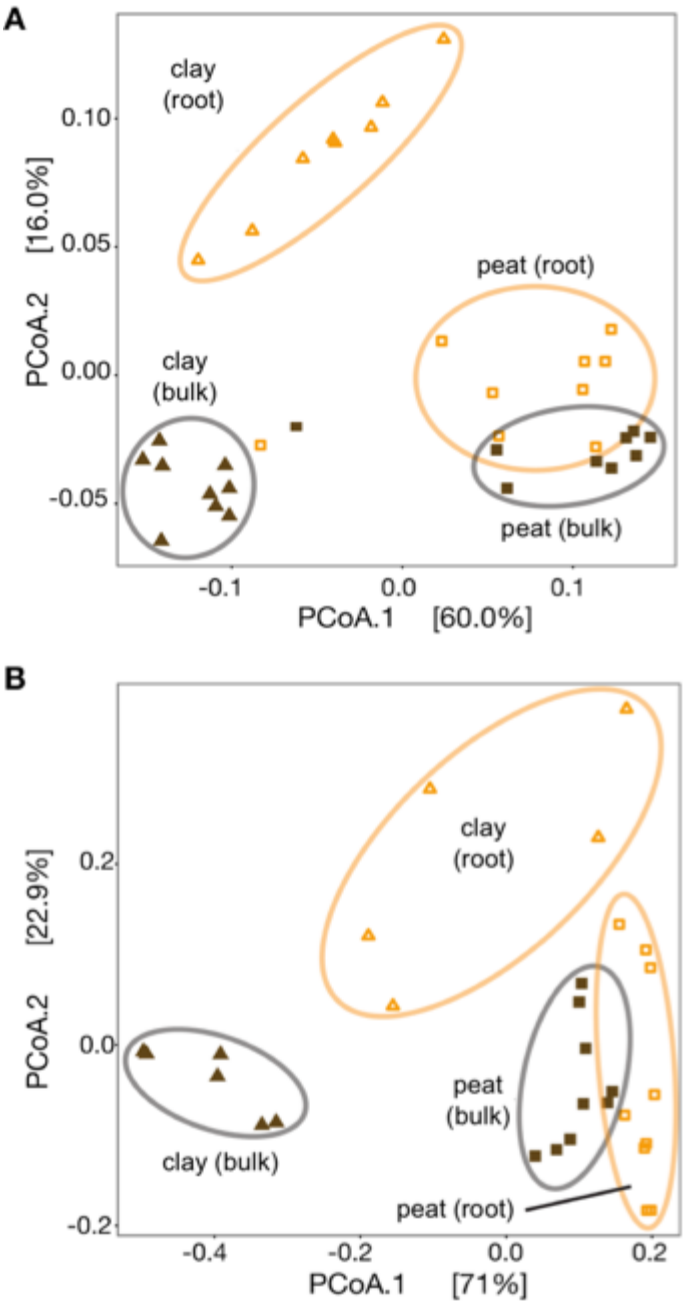


Figure 2.3. Principle coordinates analysis (PCoA) on weighted UniFrac distances of OTU assemblages by growth system (clay or peat FlowPot) and by sample type (root or bulk). Analysis was performed for samples collected from growth systems inoculated with (A) microbiota from German soil or (B) the defined *At*-SPHERE bacterial culture collection. This figure was created in collaboration with Dr. Amine Hassani, Max Planck Institute for Plant Breeding.

Alpha-diversity

We hypothesized that alpha diversity would be highest for the input communities, followed by corresponding bulk substrates, and the selective root environment would produce samples with the lowest alpha diversity. Furthermore, we wanted to determine whether the growth systems accommodated equivalent alpha diversity in their respective substrates. Alpha diversity was measured for mean OTU values from each sample using the Inverse Simpson and Shannon indices, both of which are frequently used for 16S community analysis (McMurdie and Holmes, 2013). Community profiling was performed on fewer than four of each input community (as opposed to n=8 or more for all other samples), and therefore significance testing was not performed on input community data, however, mean alpha diversity measurements with a standard error of the mean were calculated and plotted (Figure 4).

Consistent with our hypothesis, the bulk substrate from both systems inoculated with natural soil microbiota had relatively higher alpha diversity than root samples from the corresponding system. Bulk substrate from the peat FlowPot system with natural soil microbiota had significantly higher alpha diversity than bulk substrate from the clay system treated with the same soil extract inocula ($p<0.0005$ for both indices). Root samples from the FlowPot system had greater alpha diversity than the bulk substrate of the calcined clay system. However, root samples collected from the FlowPot system inoculated with defined *At*-SPHERE bacteria had equivalent alpha diversity to the corresponding substrate. In contrast, a reduction of alpha diversity was detected in when comparing root and bulk samples in the calcined clay system inoculated with *At*-SPHERE bacteria. The Shannon diversity index (but not Inverse Simpson) indicates

that *At*-SPHERE inoculated FlowPots support greater bacterial diversity in the substrate than the calcined clay system (Figure 4).

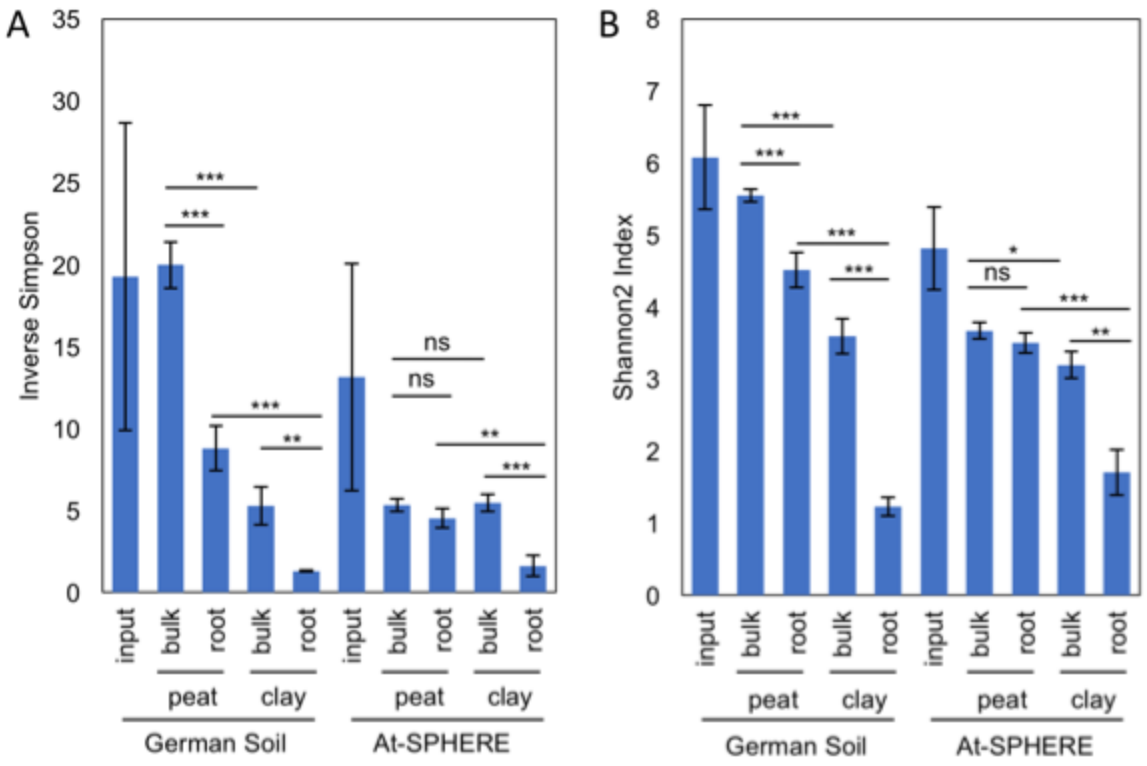


Figure 2.4. Alpha diversity indices for each sample type: (A) inverse Simpson and (B) Shannon (base 2). Significance determined by two-tailed pairwise t-test. Error bars represent standard error of the mean. Alpha diversity indices are calculated from rarefied count table to 5000 reads per sample. All bulk and root samples have a minimum of n=8 samples and consist of all data points from three independent experiments. Input samples are n=2 or n=3 and were not included in significance testing. * $p<0.05$, ** $p<0.005$, *** $p<0.0005$.

Bacterial community differentiation at the Phylum level

Based on principal coordinate and alpha diversity analyses, we have determined that both the peat-based FlowPot system and the calcined clay system support axenic Arabidopsis growth, and with either complex or defined community inocula. Furthermore, we observed a rhizosphere effect in both systems, characterized by a reduction in root-associated bacterial diversity relative to the bulk substrate and distinct clustering of root and substrate samples upon principal coordinate analysis. We hypothesized that phylogenetic signals could explain community differences between the inoculated peat and calcined clay substrates. To address this hypothesis, we calculated average read distribution at the phylum/subphylum level for the German soil and *At*-SPHERE input communities and corresponding bulk substrate and root samples from the FlowPot and calcined clay systems (Figure 5, Supplemental table). Reads from the soil input community were predominantly distributed across seven phyla and subphyla: Actinobacteria, Bacteroidetes, Firmicutes, Acidobacteria, Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria. Actinobacteria reads were most abundant (33.35 ± 6.36%), and reads classified to rare phyla accounted for 7.72 ± 2.8% of soil input (Figure 5, Supplemental Table). Actinobacteria was enriched 0.96-fold in calcined clay bulk substrate and depleted 2.76-fold in peat bulk substrate relative to the soil input (Figure 5, Table 1). Although Actinobacteria was the most abundant phylum in calcined clay bulk samples (65.64 ± 4.74% of reads) and only the fourth-most abundant phylum in FlowPot bulk samples (10.52 ± 0.83% of reads), Actinobacteria abundance in root samples from both systems were similar (FlowPot: 7.00 ± 1.25%, calcined clay: 6.08 ± 1.20%). Betaproteobacteria was the

most abundant phylum for root samples from the both FlowPot and calcined clay systems (FlowPot: 50.58 ± 3.83%, calcined clay: 88.75 ± 1.69%). Root and bulk samples from the soil-inoculated FlowPot system had similar Alphaproteobacteria abundances (bulk substrate: 16.17 ± 1.12% and root: 11.00 ± 1.32%). However, Alphaproteobacteria in samples from the soil-inoculated calcined clay system were 20-fold less abundant than the soil input (bulk: 0.38 ± 0.12% and root: 0.29 ± 0.05%). Strikingly, Alphaproteobacteria abundance did not exceed 0.5% in any sample collected from the soil-inoculated calcined clay groups. Gammaproteobacteria was equivalently abundant (± 2%) in FlowPot bulk substrate and root samples, but more than 30-fold depleted in root samples from the soil-inoculated calcined clay system relative to abundance in respective bulk substrate (bulk: 16.7 ± 3.28%, root: 0.36 ± 0.03%). Bacteroidetes-classified reads were enriched in the FlowPot system relative to the input (input: 2.08 ± 0.84%, bulk: 7.1 ± 1.01%, root: 20.37 ± 2.72%), accounting for a 3.1-fold enrichment in bulk substrate relative to soil input, and a 2.3-fold root enrichment relative to bulk substrate. In contrast, Bacteroidetes represented less than 0.2% of reads from soil-inoculated calcined clay bulk substrate and root samples. Firmicutes were enriched in the FlowPot bulk substrate relative to the soil input community, and very rare in the calcined clay bulk substrate (input: 7.25 ± 2.23%, FlowPot bulk: 22.23 ± 2.25%, calcined clay bulk: 0.13 ± 0.08%). Acidobacteria was found a very low abundance or undetected in all systems (<0.1%), in contrast to 2.66 ± 1.11% of reads from the soil input community.

At phylum-level resolution, bulk substrate of FlowPot samples inoculated with the *At*-SPHERE bacteria were similar to the input (Figure 5, Table 1). Collectively, four phyla

accounted for 95.79% of input community reads Actinobacteria, Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria. Individually, none of the four dominant phyla had greater than 0.11-fold enrichment or depletion in FlowPot substrate relative to the input. The calcined clay substrate, however, had much larger differences in the read distribution amongst the four dominant phyla relative to the *At*-SPHERE input: 4.5-fold enrichment of Actinobacteria, 2.8-fold depletion of Betaproteobacteria, 22.8-fold depletion of Alphaproteobacteria, and 0.6-fold depletion of Gammaproteobacteria. Trends observed for Actinobacteria read distribution from soil-inoculated systems were strikingly similar to trends for *At*-SPHERE inoculated samples. For both calcined clay and FlowPot systems, Actinobacteria preferentially colonized the substrate relative to the root. Relative to Actinobacteria abundance in bulk substrate, FlowPot root samples had 1.6-fold less (bulk: 17.07 ± 1.26%; root: 7.04 ± 0.86%), and calcined clay root samples had 10.6-fold less (68.22 ± 6.74% versus 7.1 ± 1.76%). Notably, the total abundance of Actinobacteria in root samples from both systems was approximately 7%. Betaproteobacteria were root-enriched 14.9-fold in calcined clay relative to bulk, but less than 0.1-fold differentially abundant in FlowPot samples. A subtle enrichment of Gammaproteobacteria in FlowPot root samples was observed relative to the bulk substrate, and collectively Gammaproteobacteria abundance in *At*-SPHERE FlowPot roots was very high (49.17 ± 3.09%), in contrast to *At*-SPHERE calcined clay root samples (5.83 ± 2.52%). High variability of Gammaproteobacteria abundance could possibly be explained by opportunistic colonization of one or several OTUs, or a founder's effect. Firmicutes accounted for 2.78 ± 1.63% of *At*-SPHERE input reads, in

contrast to soil-inoculated FlowPots, the highest average abundance of Firmicutes in *At*-SPHERE inoculated bulk or root samples was 0.53 ± 0.17%.

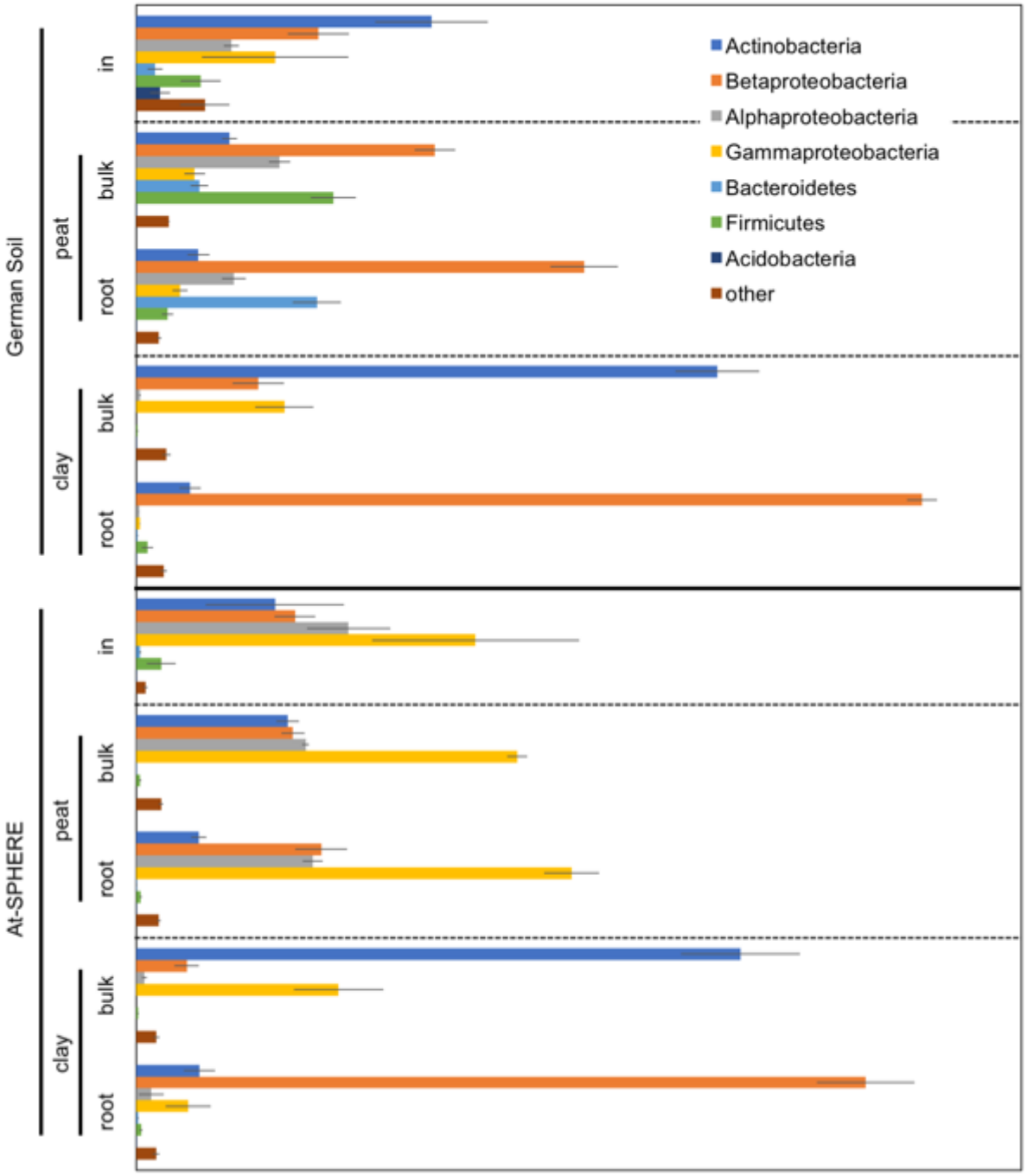


Figure 2.5. Mean percent abundance of OTU-classified reads by phylum. The right bound maximum across the x-axis represents 100%. Error bars represent standard error of the mean for each sample type.

Table 2.1. Enrichment of phylum by sample type relative to input and to bulk.

	natural soil						***At*-SPHERE**					
	enrichment over input				over bulk		enrichment over input				over bulk	
	FlowPot		clay		FP	clay	FlowPot		clay		FP	clay
	bulk	root	bulk	root	root	root	bulk	root	bulk	root	root	root
Act	-1.66	-2.25	0.98	-2.46	-0.59	-3.43	0.13	-1.15	2.12	-1.14	-1.28	-3.26
BP	0.71	1.3	-0.58	2.11	0.59	2.69	-0.02	0.22	-1.66	2.2	0.24	3.86
AP	0.59	0.04	-4.82	-5.21	-0.56	-0.39	-0.33	-0.27	-4.77	-3.83	0.06	0.93
GP	-1.25	-1.67	0.09	-5.44	-0.41	-5.54	0.17	0.36	-0.75	-2.72	0.19	-1.97
Bct	1.77	3.29	-inf	-3.89	1.52	n/a	-inf	-4.29	-inf	-0.96	n/a	n/a
Frm	1.62	-1.05	-5.8	-2.51	-2.66	3.29	-2.83	-2.53	-3.95	-2.39	0.3	1.56
Acd	-5.73	-5.73	-inf	-6.47	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Values expressed as $\log_2(x/y)$. Act= Actinobacteria, BP= Betaproteobacteria, AP= Alphaproteobacteria, GP= Gammaproteobacteria, Bct= Bacteroidetes, Frm= Firmicutes, Acd= Acidobacteria

Bacterial community differentiation at the OTU level

We hypothesize that there are root-enriched OTUs shared among the growth systems, as well as growth system-specific root microbiota. Furthermore, we hypothesize that some root-enriched OTUs will be shared among *At*-SPHERE and German soil communities.

Root-enriched OTUs were identified according to three criteria (1) minimum average abundance on the root sample of 5/5,000 (0.1%) reads, (2) minimum fold-enrichment of the OTU abundance over the corresponding bulk substrate sample of $\log_2(\text{root/bulk})>0.5$, and (3) significance of root-associated enrichment of $p < 0.1$, determined by pairwise two-tailed t-test. The number of OTUs with greater than five reads for each root sample type is 22 for clay with natural soil, 80 peat FlowPots with natural soil, 26 for clay with *At*-SPHERE, and 39 for peat FlowPots with *At*-SPHERE (Table 2). For each sample, OTUs were ranked in order of abundance, and the number of OTUs to reach 90% of the total reads was assessed. The clay system had very few OTUs account for 90% of total reads, relative to the peat FlowPot system (Table 2).

In the clay system with natural soil microbiota, 9 OTUs classified as seven genera met the root enrichment criteria: *Burkholderia* (n=2), *Ralstonia, Herbiconiux, Staphylococcus, Streptococcus*, *Propionibacterium*, unclassified Oxalobacteraceae, unclassified Proteobacteria (Figure 6, Figure 8, and Table 3). The *Ralstonia* OTU1 is 18.49-fold root enriched ($p < 10^{-13}$). The peat FlowPot system with natural soil microbiota had 11 OTUs root-enriched: *Hoeflea*, *Ensifer*, *Rhizobium* (n=2), unclassified Rhizobiaceae, *Flavobacterium*, *Ralstonia*, unclassified Oxalobactereaceae (n=2) *Massilia*, and unclassified Sinobacteraceae. Of the root-enriched OTUs from the

calcined clay and FlowPot systems with natural soil microbiota, only two are shared: *Ralstonia* (OTU1) and an unclassified member of Oxalobacteraceae (OTU10). In the FlowPots with natural soil microbiota, *Ralstonia* (OTU1) was root-enriched 3.61-fold ($p<0.032$), accounting for 5.3% of reads, while in the clay samples with natural soil microbiota, OTU1 is 18.49-fold root enriched ($p < 10^{-13}$) and accounted for >87% of reads.

Samples inoculated with the *At*-SPHERE bacteria were subjected to the same root-enrichment criteria as the holoxenic samples (Figure 7, Figure 9, and Table 3). *At*-SPHERE calcined clay samples yielded eight root-enriched OTUs, classified as the following genera: *Burkholderia* (n=3), *Ralstonia, Staphylococcus*, *Propionibacterium*, *Flavobacterium*, and unclassified Proteobacteria. Six of the eight root-enriched OTUs are shared with the holoxenic calcined clay samples (Table 3).

Table 2.2. Absolute OTU abundance by abundance qualifiers.

	natural soil					***At*-SPHERE**				
		FlowPot		clay			FlowPot		clay	
qualifier	input	bulk	root	bulk	root	input	bulk	root	bulk	root
# > 5	138	104	80	47	22	48	39	39	25	26
# > 1	482	325	279	149	162	187	133	134	114	145
top 90%	167	88	58	20	4	23	20	19	10	9

Table 2.3. Classification of root-enriched OTUs and magnitude of enrichment.

			natural soil				***At*-SPHERE**			
			FlowPot		clay		FlowPot		clay	
OTU	Classification		%R	R/B	%R	R/B	%R	R/B	%R	R/B
100	Act	*Propionibacterium*	-	-	0.6	4.4	-	-	0.5	21.2
320	Act	*Herbiconiux*	-	-	0.4	18.5	-	-	-	-
39	AP	*Devosia*	-	-	-	-	2.6	5.8	-	-
47	AP	*Microvirga*	-	-	-	-	0.9	0.5	-	-
55	AP	*Hoeflea*	0.1	7.8	-	-	0.8	0.6	-	-
51	AP	*Ensifer*	0.1	0.6	-	-	0.6	1.0	-	-
63	AP	*Rhizobium*	0.6	5.3	-	-	0.5	0.6	-	-
122	AP	(Rhizobiaceae)	0.4	1.2	-	-	-	-	-	-
352	AP	*Rhizobium*	0.3	0.6	-	-	-	-	-	-
19	Bct	*Flavobacterium*	16.8	2.3	-	-	-	-	0.2	10.9
1	BP	*Ralstonia*	5.3	3.6	87.3	18.5	10.4	1.4	73.1	19.4
88	BP	*Burkholderia*	-	-	0.7	10.2	-	-	0.7	26.0
1175	BP	*Burkholderia*	-	-	0.2	12.3	-	-	0.2	13.0
46	BP	*Burkholderia*	-	-	-	-	-	-	0.1	7.8
37	BP	*Methylophilus*	-	-	-	-	2.6	15.2	-	-
10	BP	(Oxalobacteraceae)	8.0	7.3	0.1	4.0	-	-	-	-
208	BP	*Massilia*	7.6	11.6	-	-	-	-	-	-
382	BP	(Oxalobacteraceae)	1.6	3.2	-	-	-	-	-	-
179	Frm	*Staphylococcus*	-	-	0.2	10.9	-	-	0.3	16.0
181	Frm	*Streptococcus*	-	-	0.2	13.0	-	-	-	-
23	GP	*Rhizobacter*	-	-	-	-	2.4	7.8	-	-
48	GP	*Lysobacter*	-	-	-	-	1.7	2.3	-	-
226	GP	(Sinobacteraceae)	0.2	7.3	-	-	-	-	-	-

%R= mean percent of root sample reads of the sample group, R/B= root enrichment over bulk;
Act= Actinobacteria, BP= Betaproteobacteria,
AP= Alphaproteobacteria, GP= Gammaproteobacteria, Bct= Bacteroidetes,
Frm= Firmicutes

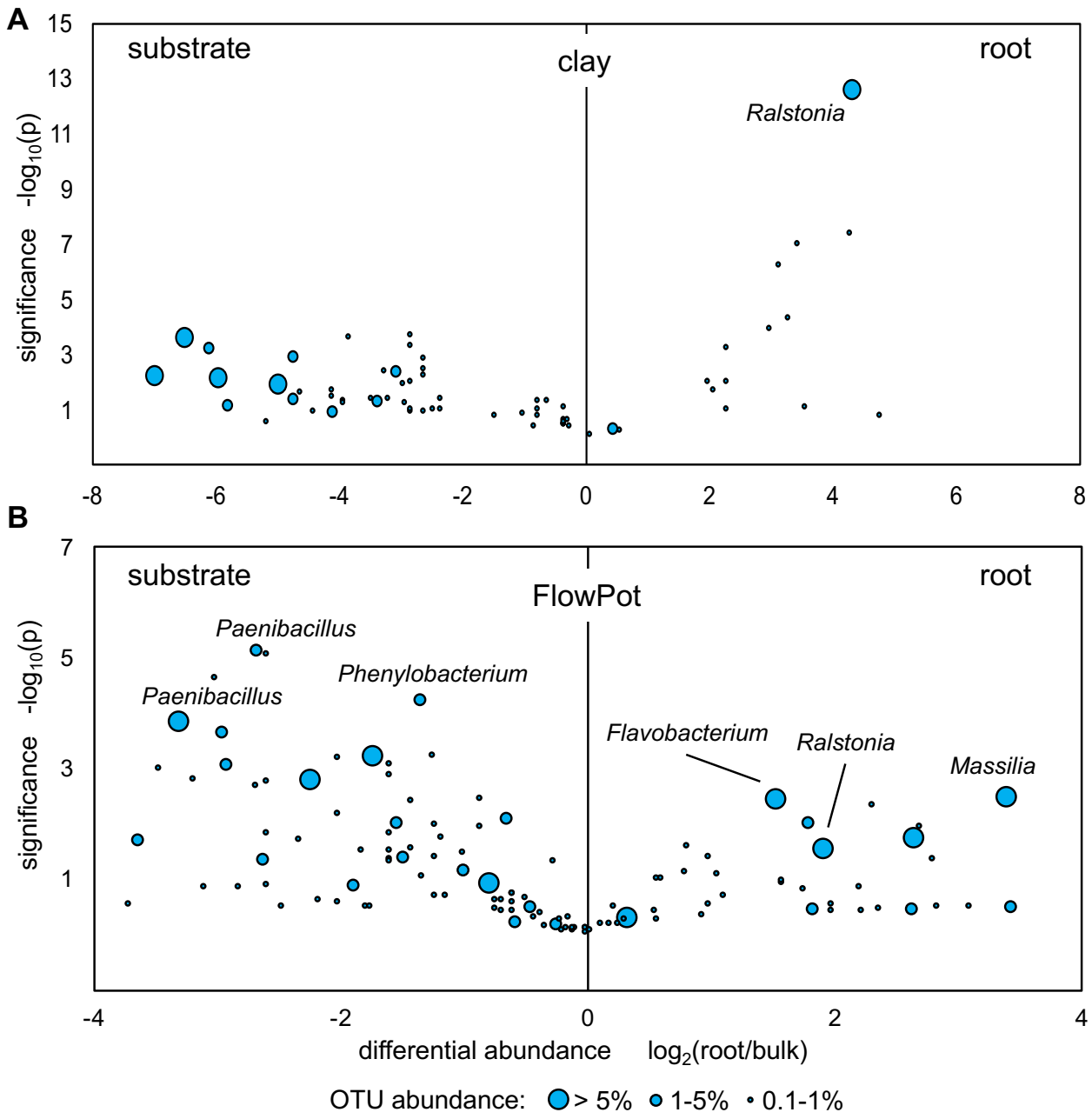


Figure 2.6. Root-associated OTU enrichment for samples with natural soil microbiota (German CAS10) from the (A) calcined clay system and the (B) FlowPot system. Statistical confidence was calculated for each OTU comparison by pairwise t-test and plotted on the y-axis as $-\log_{10}$ (p).

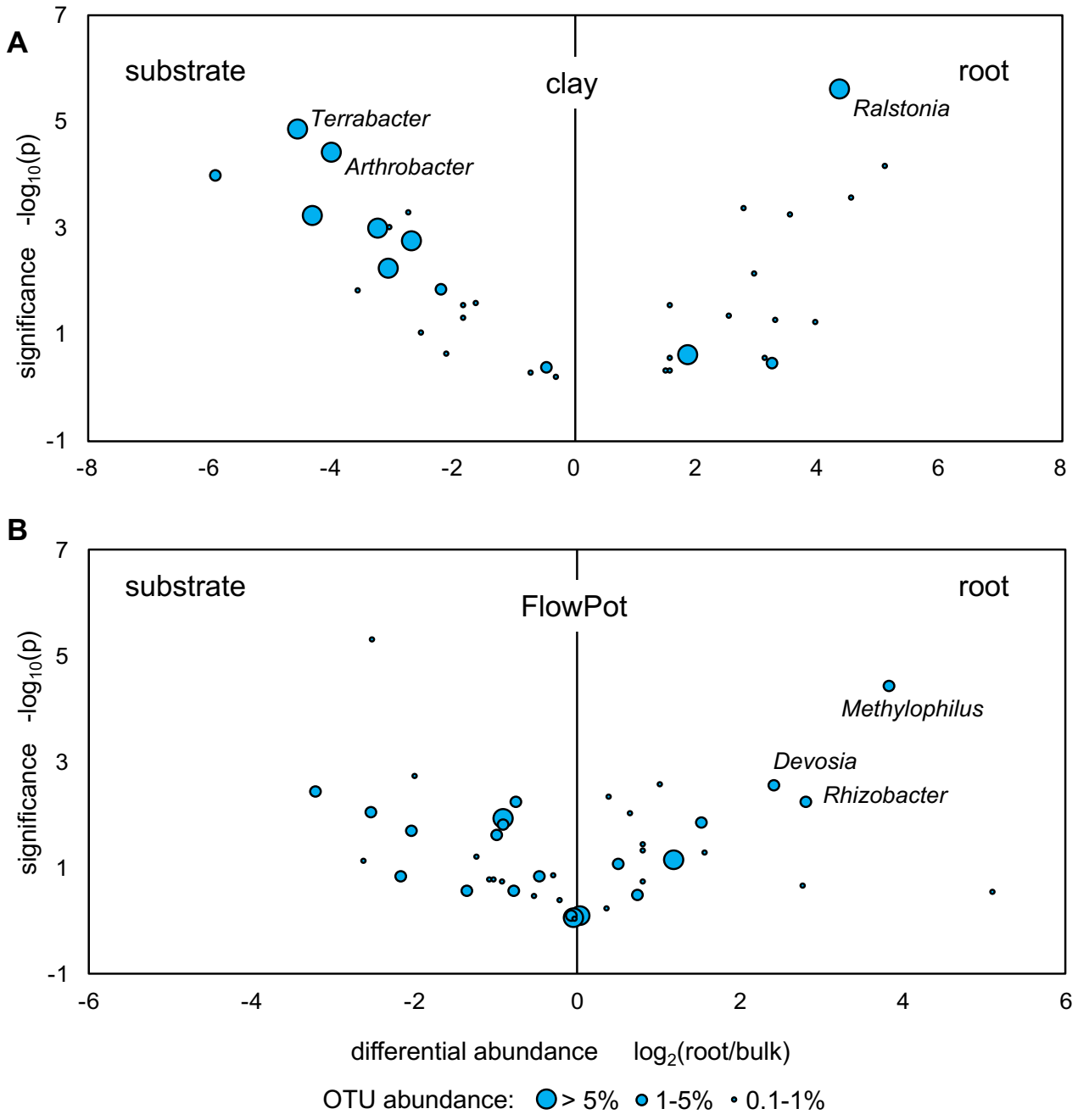


Figure 2.7. Root-associated OTU enrichment relative to abundance in bulk substrate for samples with *At*-SPHERE microbiota from the (A) calcined clay system and the (B) FlowPot system. Statistical confidence was calculated for each OTU comparison by pairwise t-test and plotted on the y-axis as $-\log_{10}(p)$.

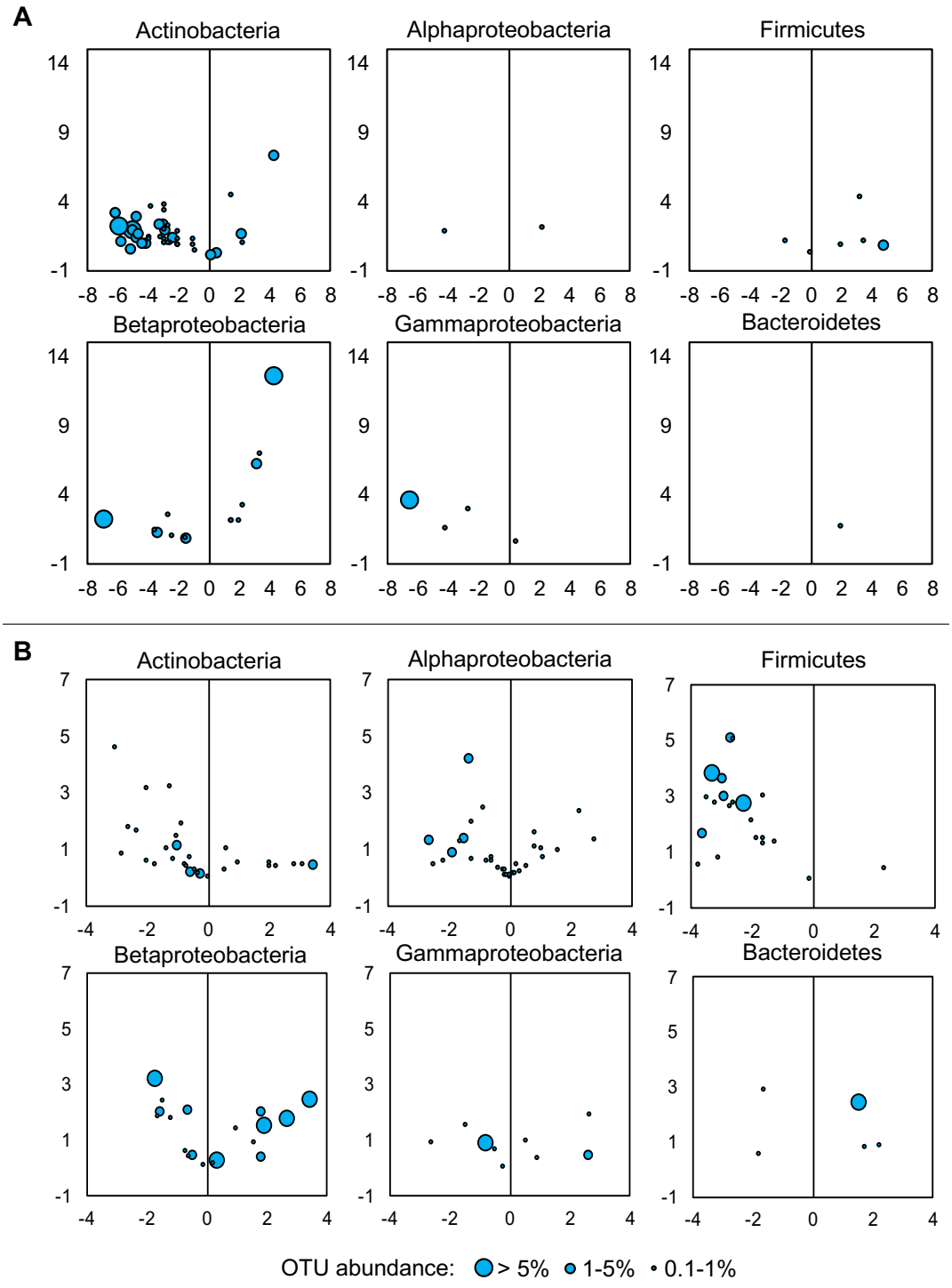


Figure 2.8. Phylum-level resolution of root-associated OTU enrichment relative to abundance in bulk substrate for samples with natural soil microbiota from the (A) calcined clay system and the (B) FlowPot system. Statistical confidence was calculated for each OTU comparison by pairwise t-test and plotted on the y-axis as $-\log_{10}(p)$.

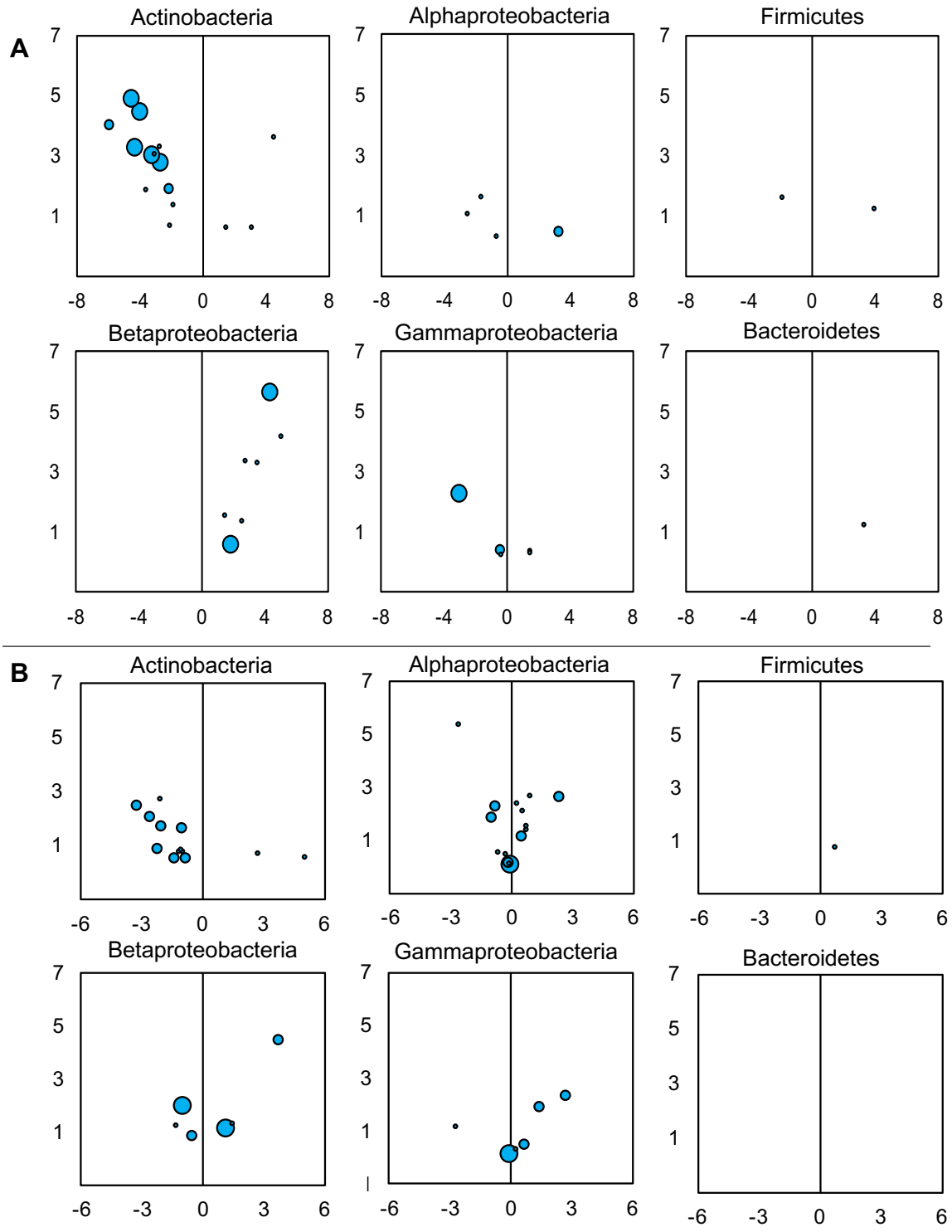


Figure 2.9. Phylum-level resolution of root-associated OTU enrichment relative to abundance in bulk substrate for samples with *At*-SPHERE microbiota from the (A) calcined clay system and the (B) FlowPot system. Statistical confidence was calculated for each OTU comparison by pairwise t-test and plotted on the y-axis as $-\log_{10}$ (p).

Discussion

Over the past decade, great progress has been towards identifying the core microbiota of Arabidopsis using culture-independent methods (Lundberg et al., 2012; Bulgarelli et al., 2012; Delmotte et al., 2009; Kniskern et al., 2007). More recently, microbiome colonization experiments on Arabidopsis under gnotobiotic conditions with synthetic communities have been used to identify host factors that influence the microbiome (Vogel et al., 2016; Lebeis et al., 2015; Bai et al., 2015; Bodenhausen et al., 2014). The trend is shifting from hypothesis-generating descriptive studies to hypothesis-testing mechanistic laboratory studies (Busby et al., 2017). Just as gnotobiotic research in mice has successfully led to discoveries of medical importance (Ley et al., 2006; Kau et al., 2011; Hooper et al., 2012), a major goal of plant microbiome recolonization research in a laboratory setting is to discern plant-microbiome interactions that occur in nature or in an agricultural setting with reduced complexity. It is crucial to know if a growth system can capture and support the microbiota from a natural setting. Based on many studies aiming to identify host factors that influence microbiome assembly, environmental factors have been noted to exert profound impact (McCoy et al., 2017; Goodrich et al., 2014). Ideally, multiple experimental systems with different characteristics should be used to capture different aspects of natural plant-microbiome interactions, resulting in a more robust and realistic general principle.

Here, we introduce the FlowPot: a peat-based growth system that supports vigorous Arabidopsis growth under the following conditions: axenic (total absence of detectable microbiota), gnotobiotic (defined or synthetic microbiota), and holoxenic (direct transfer of a soil microbial community). In this study, we performed a soil microbiome

transplant and characterized microbial community feedback after this perturbation, as well as inoculation with members of the defined *At*-SPHERE Arabidopsis bacterial microbiota, many of which were isolated from plants growing in the same soil (Bai et al., 2015). Relative to the natural soil input community, the bulk substrate of a calcined clay system, and to a lesser extent the FlowPot system, has a reduction in detectable OTUs, consistent with previous reports (Howard et al., 2017; Lau and Lennon, 2012). FlowPot system maintains a similar bacterial phylogenetic profile to a natural soil collected from an agricultural field at the phylum level. Relative to the established the calcined clay growth system, the FlowPot system maintains higher alpha diversity, more closely simulating an agriculturally relevant microbial community. This is probably because calcined clay lacks robust, complex carbon, nitrogen and/or certain micronutrient sources for microorganisms. Although in this study both systems were supplemented with a plant nutrient solution, the carbon source (other than CO2) in the calcined clay growth system likely comes from photosynthates exuded by the plant. This is in contrast to the FlowPot system, in which the substrate is mostly organic material. We speculate that after inoculation, some populations of microorganisms in the calcined clay system may be in a state of metabolic dormancy, some may not even survive. It would be interesting to test the recoverability of viable taxa over time in calcined clay and FlowPots after inoculation. The strong enrichment of Actinobacteria in the calcined clay substrate relative to the root is interesting and suggests that populations of certain taxa may be more capable of surviving the calcined clay environment than other taxa.

The “rhizosphere effect” is the term to describe a root-associated microbiome taxonomically and functionally distinct from surrounding roots (Hiltner, 1904). Previously, the rhizosphere effect in Arabidopsis has been observed in Magenta boxes with calcined clay substrate inoculated with mixtures of defined cultures of Arabidopsis-associated bacteria (Bai et al., 2015; Lebeis et al., 2015). Relative to the FlowPot system the calcined clay system showed a stronger rhizosphere effect as evident from the beta-diversity analysis. We speculate that multiple nutrient sources exist in the FlowPot system, and that metabolically flexible microorganisms can establish in the bulk substrate or the plant root environment. In the calcined clay system, many microorganisms may be forced to compete for the root environment, become/remain metabolically dormant, or die. Some OTUs were present in the soil input community, but not detected at significant abundance the substrate or root. For example, six Acidobacteria OTUs are detected in the soil input at 0.1-0.62% abundance, but not in the calcined clay or FlowPot systems. One hypothesis is that the Acidobacteria are slow-growing in this system, and although they may be alive in the new substrate, they fall below the detection limit due to other faster-growing bacteria. Another possibility is that the substrate lacks physical or chemical properties to sustain the survival of these particular Acidobacteria, although Acidobacteria has been reported to be a dominant phylum in natural Sphagnum peat habitats (Pankratov et al., 2008). A third possibility to consider is that some bacteria maybe be non-viable dead at the time of inoculation. A shortcoming of culture-independent (shotgun and amplicon) community profiling from environmental DNA is the inability to discern live versus dead microorganisms. It is also possible that the perturbation event of inoculation prompted

populations of microorganisms to shift from a dormant to an active lifestyle, and outcompete other microorganisms. In this study, we classified a microorganism as "root-enriched" if it is more relatively abundant in a root sample than the corresponding soil substrate. Some organisms, however, are highly abundant in both the substrate and the root. Just because an organism is not more relatively abundant in a root as compared to the bulk substrate does not exclude the fact that it may have a functional interaction with the plant.

Some OTUs are present in the substrate of one system, but not the other. For example, *Pseudonocardia* (Actinobacteria, Otu17 and Otu22) is >2.5% abundant in soil-inoculated clay, but not detected in FlowPots. Perhaps there are circumstances where *Pseudonocardia* is capable of survival in the FlowPot substrate as well, but the FlowPot substrate also sustains populations of microorganisms that outcompete *Pseudonocardia*. It is critical to consider the dynamics of microbe-microbe interactions and their role in the ecology of microbiomes. The lack of proliferation of a microorganism in a growth system may be due to microbial antagonism and antibiosis (i.e., siderophores, type-VI secretion, bacteriocin production, antibiotics, etc.) rather than abiotic factors (Ma et al., 2014; Loper et al., 1991; Jones et al., 2007). An OTU is root-enriched in one system, and although it is detectable in the other system, it is not root-enriched. *Herbiconiux* (Actinobacteria, Otu320) is root-enriched in the soil-inoculated calcined clay system, and detected in the FlowPot system but at equal equivalent abundance in the root and substrate. This OTU is very low abundance, and increased sequencing depth could potentially solve this problem. Otu320 is very low abundance in both soil-inoculated systems: ~0.4% in clay roots, 0% in clay substrate,

~0.18% in FlowPot roots, ~0.1% in FlowPot substrate. Notably, all root-enriched OTUs specific to a system were less than 0.2% abundant in the other system (substrate and root samples). Six Acidobacteria OTUs are detected in the German soil input at 0.1-0.62% abundance, but not in the calcined clay or FlowPot systems.

One must consider that in nature, a plant microbiome is populated by many non-bacterial microorganisms, including fungi, oomycetes, and protists. Furthermore, macroscopic organisms in the soil such as nematodes and springtails may affect a plant as well. The 16S rRNA gene amplicon approach to microbiome analysis only reveals a portion of the microbiome. Non-bacterial microorganisms may influence the bacterial constituency of a plant's microbiome via competition for nutrients, predation, influencing host metabolism, changing soil features, and providing a new repertoire of metabolites in the ecosystem. The German soil inoculum likely contains viable fungi; traces of fungal and protist mitochondrial DNA were detected in the 16S amplicon reads. The extent to which non-bacterial microorganisms influenced the bacterial interaction in this set of experiments is unknown. Follow-up experimentation is required to address these questions.

The complexity of the plant microbiome is not well understood. In the agricultural biotechnology industry, microbial products have been applied to the field –to stimulate plant growth (biostimulants), providing additional nutrient sources (biofertilizers), and controlling pests (biopesticides). The agricultural biologicals industry is rapidly emerging, but a better understanding of plant-microbe associations in a community context would most certainly lead to innovations (Ciancio et al., 2016; Busby et al., 2017; Mueller and Sachs, 2015). A better grasp of the complexities of the plant

microbiome could lead to new and effective technologies that can enhance crop growth in what is considered marginal land, improve plant nutrient use efficiency, and confer tolerance to both biotic and abiotic stress tolerance. A future direction for the FlowPot system and other axenic technologies is to develop its use for agriculturally relevant model systems such as maize and soybean. Furthermore, expansion of reference culture collections such as *At*-SPHERE to include isolates from other environments, or to build collections based on function rather than 16S, will increase the power of recolonization experiments. One application of an axenic colonization system is to determine the most successful plant colonizers because presumably, they form a stronger association with the plant and/or effectively out-compete other microorganisms. Coupled with the power of fully-sequenced culture collections, one can hypothesize which genomic features are necessary for successful colonization of a plant host. We can also introduce microorganisms into established microbiomes on a large-scale and determine which isolates can invade and persist. As we move forward in the plant microbiome field an important question is now possible to address: how does the microbiome influence host phenotype? Healthy and robust plant growth in a nutrient-replete axenic FlowPot system allows us to study a plant in an axenic state, and directly compare this to colonized plants. Is there anything wrong with an axenic plant? Although we could not detect or identify any signs of stress in our axenic Arabidopsis, a further in-depth study of the physiology of axenic Arabidopsis relative to Arabidopsis colonized by conventional soil community is of interest.

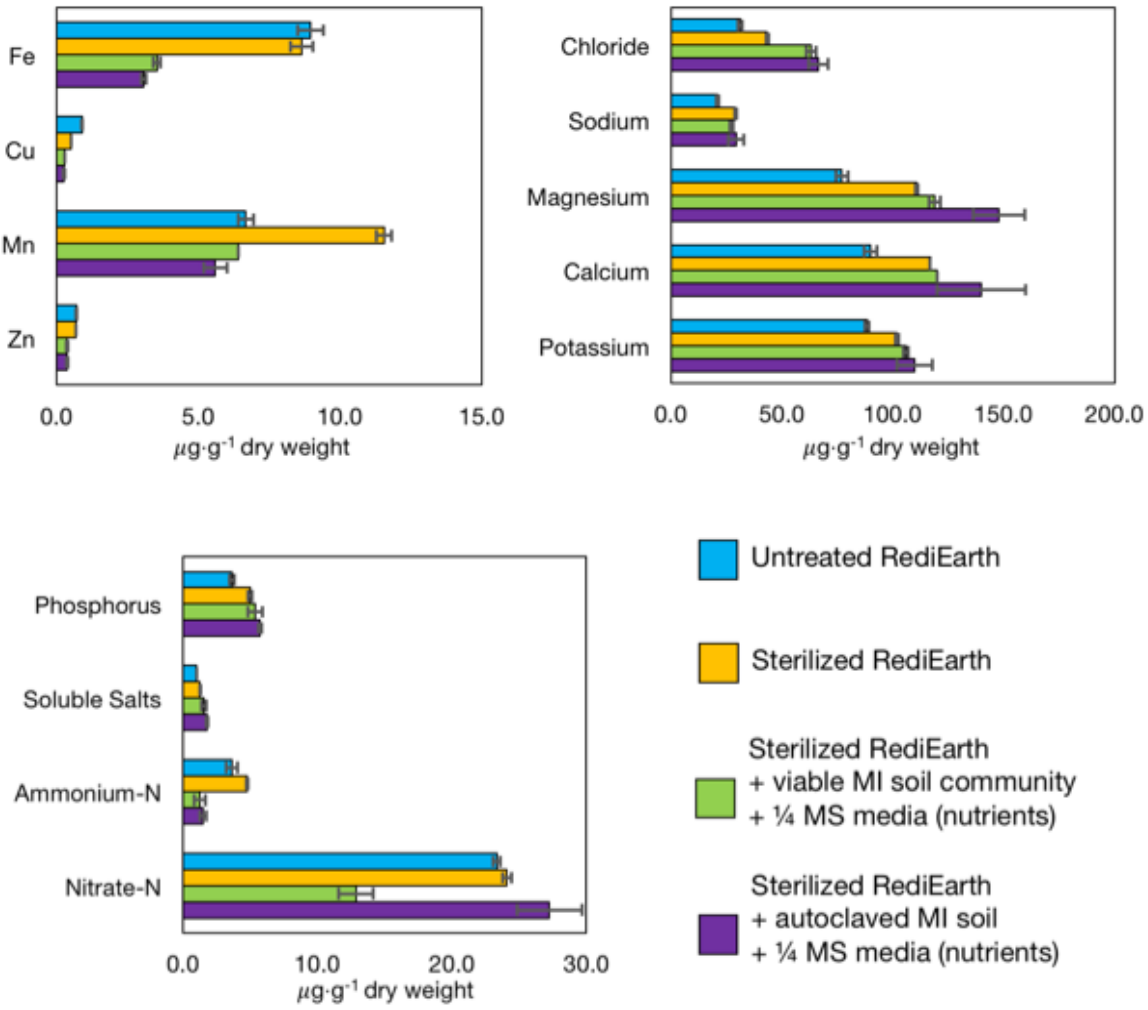


Figure 2.10. Changes in substrate chemical properties after treatments. Chemical analysis of peat substrate after four treatments: as prepared by the manufacturer (no treatment), 7 d after autoclave regime (no flushing), 7 d after autoclave regime and flushing regime with ddH2O followed by Michigan soil microbiota suspended in 1/4X MS media (liquid), 7 d after autoclave regime and flushing regime with ddH2O followed by autoclaved Michigan soil microbiota suspended in 1/4X MS media (liquid).

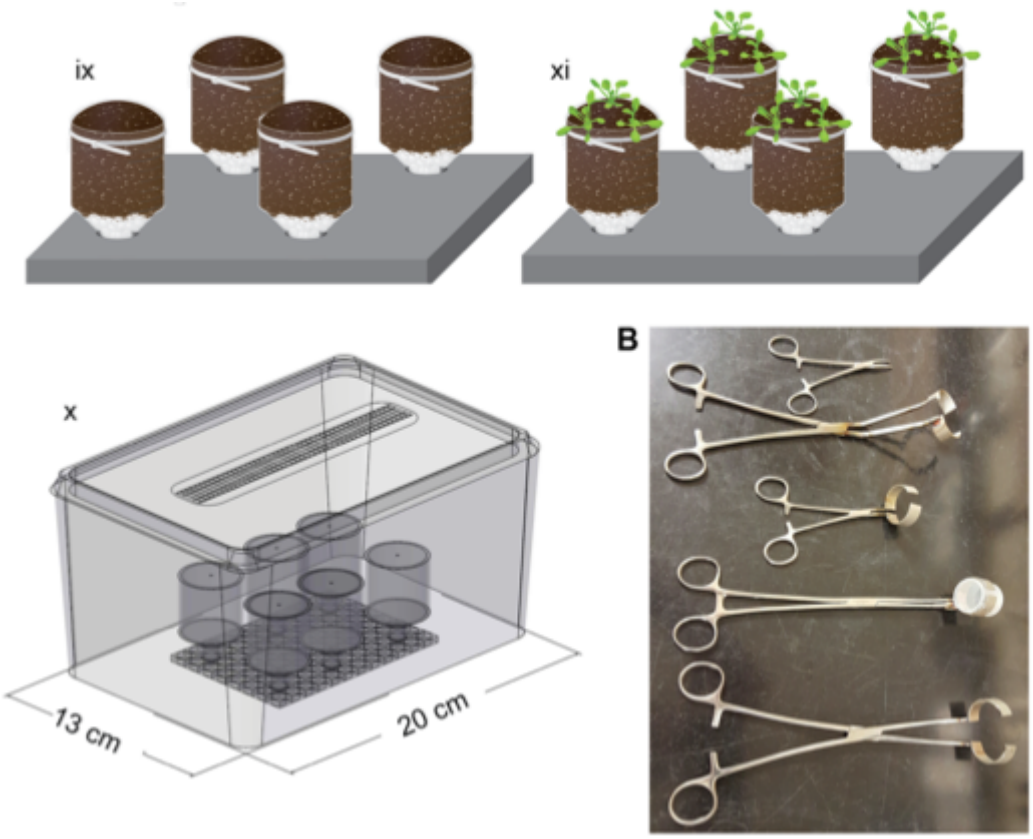


Figure 2.11. Instruments for the aseptic manipulation of FlowPots, and illustration of FlowPots on a stand and within the Eco2 microbox.

Table 2.4. Abundance of reads by phylum.

	natural soil					***At*-SPHERE**				
		FlowPot		clay			FlowPot		clay	
	in	bulk	root	bulk	root	in	bulk	root	bulk	root
Act	33.4 ± 6.4	10.5 ± 0.8	7.0 ± 1.3	65.6 ± 4.7	6.1 ± 1.2	15.7 ± 7.8	17.1 ± 1.3	7.0 ± 0.9	68.2 ± 6.7	7.1 ± 1.8
BP	20.6 ± 3.5	33.7 ± 2.3	50.6 ± 3.8	13.8 ± 2.9	88.8 ± 1.7	17.9 ± 2.3	17.7 ± 1.3	20.9 ± 2.9	5.7 ± 1.4	82.4 ± 5.5
AP	10.7 ± 0.9	16.2 ± 1.2	11.0 ± 1.3	0.4 ± 0.1	0.3 ± 0.1	24.0 ± 4.7	19.1 ± 0.4	19.9 ± 1.1	0.9 ± 0.3	1.7 ± 1.4
GP	15.7 ± 8.3	6.6 ± 1.2	4.9 ± 0.9	16.7 ± 3.3	0.4 ± 0.0	38.3 ± 11.7	43.0 ± 1.1	49.2 ± 3.1	22.8 ± 5.1	5.8 ± 2.5
Bct	2.1 ± 0.8	7.1 ± 1.0	20.4 ± 2.7	< 0.1	0.1 ± 0.0	0.4 ± 0.2	< 0.1	< 0.1	< 0.1	0.2 ± 0.1
Frm	7.3 ± 2.2	22.2 ± 2.5	3.5 ± 0.6	0.1 ± 0.1	1.3 ± 0.6	2.8 ± 1.6	0.4 ± 0.1	0.5 ± 0.1	0.2 ± 0.1	0.5 ± 0.2

Values represent mean of samples, ± standard error of the mean.
Act= Actinobacteria, BP= Betaproteobacteria,
AP= Alphaproteobacteria, GP= Gammaproteobacteria, Bct= Bacteroidetes,
Frm= Firmicutes

Table 2.5. Phylum level classifications of most abundantly ranked OTUs that account for 90% of reads

	natural soil					***At*-SPHERE**				
		FlowPot		clay			FlowPot		clay	
	in	bulk	root	bulk	root	in	bulk	root	bulk	root
Act	51	21	12	16	2	8	8	3	6	4
BP	29	16	15	3	2	3	3	4	2	2
AP	30	26	17	-	-	6	6	8	-	1
GP	15	6	4	1	-	5	3	4	2	2
Bct	4	3	3	-	-	-	-	-	-	-
Frm	12	16	7	-	-	1	-	-	-	-
Acd	10	-	-	-	-	-	-	-	-	-

Act= Actinobacteria, BP= Betaproteobacteria,
AP= Alphaproteobacteria, GP= Gammaproteobacteria, Bct= Bacteroidetes,
Frm= Firmicutes, Acd= Acidobacteria

Table 2.6. Natural soil, clay system. Top 10 OTUs by rank abundance.

Sample type	OTU	reads/5000	Phylum	Genus
bulk	3	1028 ± 404	Actinobacteria	Arthrobacter
	2	809 ± 160	Gammaproteobacteria	Pseudomonas
	7	527 ± 172	Actinobacteria	Patulibacter
	8	371 ± 110	Betaproteobacteria	Burkholderia
	5	369 ± 111	Actinobacteria	Nocardioides
	14	237 ± 55	Actinobacteria	Solirubrobacter
	1	217 ± 148	Betaproteobacteria	Ralstonia
	11	204 ± 45	Actinobacteria	
	17	190 ± 47	Actinobacteria	Pseudonocardia
	22	110 ± 55	Actinobacteria	Pseudonocardia
root	1	4367 ± 85	Betaproteobacteria	Ralstonia
	38	117 ± 52	Actinobacteria	Streptomyces
	88	35 ± 3	Betaproteobacteria	Burkholderia
	86	30 ± 14	Actinobacteria	Streptomyces
	100	29 ± 7	Actinobacteria	Propionibacterium
	309	27 ± 21	Firmicutes	Streptococcus
	17	22 ± 9	Actinobacteria	Pseudonocardia
	320	19 ± 2	Actinobacteria	Herbiconiux
	7	16 ± 16	Actinobacteria	Patulibacter
	181	11 ± 7	Firmicutes	Streptococcus

Table 2.7. Natural soil, FlowPot system. Top 10 OTUs by rank abundance.

Sample type	OTU	reads/5000	Phylum	Genus
bulk	599	524 ± 72	Betaproteobacteria	Massilia
	42	372 ± 67	Firmicutes	Paenibacillus
	40	356 ± 95	Betaproteobacteria	Massilia
	19	291 ± 47	Bacteroidetes	Flavobacterium
	34	258 ± 53	Firmicutes	Bacillus
	2	246 ± 58	Gammaproteobacteria	Pseudomonas
	16	183 ± 29	Betaproteobacteria	Massilia
	28	179 ± 13	Alphaproteobacteria	Phenylobacterium
	81	146 ± 29	Betaproteobacteria	Massilia
	53	110 ± 42	Alphaproteobacteria	Brevundimonas
root	19	842 ± 158	Bacteroidetes	Flavobacterium
	40	447 ± 107	Betaproteobacteria	Massilia
	10	402 ± 129	Betaproteobacteria	
	208	382 ± 102	Betaproteobacteria	Massilia
	1	263 ± 82	Betaproteobacteria	Ralstonia
	4	194 ± 185	Actinobacteria	Arthrobacter
	12	193 ± 178	Gammaproteobacteria	Dyella
	89	159 ± 128	Betaproteobacteria	Herminiimonas
	599	157 ± 49	Betaproteobacteria	Massilia
	2	142 ± 28	Gammaproteobacteria	Pseudomonas

Table 2.8. *At*-SPHERE, clay system. Top 10 OTUs by rank abundance.

Sample type	OTU	reads/5000	Phylum	Genus
bulk	3	1168 ± 183	Actinobacteria	Arthrobacter
	2	974 ± 259	Gammaproteobacteria	Pseudomonas
	5	857 ± 185	Actinobacteria	Nocardioides
	6	394 ± 58	Actinobacteria	Terrabacter
	4	344 ± 72	Actinobacteria	Arthrobacter
	93	259 ± 53	Actinobacteria	
	1	177 ± 55	Betaproteobacteria	Ralstonia
	15	170 ± 45	Actinobacteria	Nocardioides
	12	135 ± 18	Gammaproteobacteria	Dyella
	9	103 ± 32	Betaproteobacteria	Variovorax
root	1	3657 ± 453	Betaproteobacteria	Ralstonia
	9	382 ± 226	Betaproteobacteria	Variovorax
	2	116 ± 53	Gammaproteobacteria	Pseudomonas
	12	100 ± 40	Gammaproteobacteria	Dyella
	3	73 ± 25	Actinobacteria	Arthrobacter
	25	66 ± 63	Alphaproteobacteria	Rhizobium
	4	53 ± 22	Actinobacteria	Arthrobacter
	5	43 ± 10	Actinobacteria	Nocardioides
	15	37 ± 14	Actinobacteria	Nocardioides
	320	36 ± 21	Actinobacteria	Herbiconiux

Table 2.9. *At*-SPHERE, FlowPot system. Top 10 OTUs by rank abundance.

Sample type	OTU	reads/5000	Phylum	Genus
bulk	2	1990 ± 74	Gammaproteobacteria	Pseudomonas
	10	441 ± 57	Betaproteobacteria	
	13	320 ± 15	Alphaproteobacteria	Caulobacter
	1	228 ± 14	Betaproteobacteria	Ralstonia
	27	219 ± 67	Actinobacteria	Oerskovia
	4	186 ± 52	Actinobacteria	Arthrobacter
	9	181 ± 23	Betaproteobacteria	Variovorax
	32	157 ± 18	Alphaproteobacteria	Sphingopyxis
	115	117 ± 19	Alphaproteobacteria	Caulobacter
	12	104 ± 28	Gammaproteobacteria	Dyella
root	2	2065 ± 214	Gammaproteobacteria	Pseudomonas
	1	521 ± 151	Betaproteobacteria	Ralstonia
	13	313 ± 41	Alphaproteobacteria	Caulobacter
	10	238 ± 43	Betaproteobacteria	
	12	177 ± 67	Gammaproteobacteria	Dyella
	9	132 ± 21	Betaproteobacteria	Variovorax
	37	129 ± 21	Betaproteobacteria	Methylophilus
	27	128 ± 46	Actinobacteria	Oerskovia
	39	128 ± 29	Alphaproteobacteria	Devosia
	23	120 ± 32	Gammaproteobacteria	Rhizobacter

Chapter 3

Characterization of axenic *Arabidopsis thaliana*

Acknowledgements: MAPK and UPLC/MS experiments were contributed by Brad Paasch. Non-targeted GC/MS was performed by Dr. Tim Tschliplinsky (Oak Ridge National Lab). Assistance with plant growth and FlowPot assembly from David Rhodes and Caleigh Griffin were greatly appreciated. Thoughtful discussions and guidance from Dr. JP Jerome and Dr. Brian Kvitko were instrumental.

Abstract

Plants and animals are naturally colonized by diverse microbiota, and although mechanistic studies on select host-microbe model systems have revealed important molecular underpinnings of symbiosis and pathogenesis, little is known about the dynamic relationship between a host and its associated microbiota in a community context. Utilizing the FlowPot growth system described in Chapter 2, here we characterize differential immune-associated phenotypes in holoxenic and axenic Arabidopsis, revealing functional attributes of plant microbiota by proxy of host phenotypic characterization. Despite similar appearances and growth rates in nutrient-replete substrate, normal basal abundance of immunity-associated transcripts are lacking in axenic plants. Likewise, metabolic profiling and phytohormone quantification reveal a deficiency of defense and immune-associated metabolites in axenic plants. We further report that axenic plants are compromised in transcriptional, translational, and post-translational responses to the immune elicitation and priming capacity. Axenic plants are compromised in defense against the foliar bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000. Diseases assays using a type-III secretion system deficient virulence-compromised DC3000 mutant strain and an immune-deficient Arabidopsis polymutant reveal that axenic susceptibility is partially explained by a compromised innate immune system. The identity of differentially abundant transcripts and metabolites from axenic and holoxenic Arabidopsis are reported herein, which may contribute to microbiome-influenced host phenotypes. These results provide the first direct evidence that endogenous microbiota are required for the development of a normal level of immunocompetence in soil-grown plants.

Introduction

Despite the vast microbial diversity of the plant microbiome, only a select few pathogenic and symbiotic interactions have been studied in depth. The majority of plant-microbe interactions and their phenotypic repercussions to the host are poorly understood, particularly in a community context. Dysbiosis of a plant's microbiome can occur when a virulent pathogenic microorganism infects a plant host, resulting in disease manifestation. The dogmatic Plant Disease Triangle has been the most popular model to explain prerequisites for plant disease to occur: (1) a disease-conducive environment, (2) a virulent pathogen, and (3) a susceptible host (Scholthof, 2007). Indeed, these elements are necessary for disease to occur, but a potentially neglected factor that influences disease susceptibility is the microbiome.

Previous research indicates that the plant microbiome is intimately linked to plant nutrition. For example, phosphorus solubilization, nitrogen fixation, and carbon conversion are all ecosystem services provided by soil and plant-associated microbiota, for which the health of plants is dependent upon (Berendsen et al., 2012; DeAngelis et al., 2008; Jackson et al., 2008; Turner and Haygarth, 2001). Given the presumed importance of the plant microbiome towards plant health, numerous studies attempt to identify host factors that influence microbiome composition (Bodenhausen et al., 2014; Horton et al., 2014; Lebeis et al., 2015; Peiffer et al., 2013). The function of the plant microbiome, regarding influence over host phenotype, has been studied from various approaches. A traditional approach, albeit removed from understanding plant-microbe interactions in a community context, is to isolate and exogenously apply microorganisms to a plant and compare the host phenotype to a mock-inoculated

control, either in the lab or the field. Correlation studies have been performed to generate hypotheses about microbiome function, whereupon taxonomic-level microbiome community profiling is correlated with corresponding plant phenotypic measurements (Kembel et al., 2014; Wagner et al., 2014). Another approach is to survey the microbiome in depth using meta-'omics approaches (i.e., metagenomics, metatranscriptomics, meta-proteomics, etc.), and based on database annotations of what is identified, hypothesize what effect the microbiome may have on the host (Delmotte et al., 2009; Garoutte, 2016; Ryffel et al., 2016; Tringe et al., 2005). Another approach is to inoculate a plant or soil with a microbial consortium, either via soil or a mix of microbial cultures and assess the host phenotype (Mueller and Sachs, 2015; Panke-Buisse et al., 2015). Here, we propose an alternative approach-- characterize phenotypic attributes of an axenic (germ-free) plant and contrast them directly with holoxenic (colonized by natural microbiota) plants grown under identical conditions. Correlation studies that link plant phenotypes with associated taxonomic microbiome composition, in both diseased and non-diseased states, are the basis for many hypotheses of microbiome-influenced plant phenotypes. One such example links microbiome composition with flowering time, whereupon researchers identified a correlation between flowering time and reproductive fitness of an Arabidopsis relative to soil microbiome composition by studying natural populations (Wagner et al., 2014). Correlations between flowering phenology and relative abundance of particular taxa enabled Wagner et al. to formulate hypotheses regarding which taxa play a role in flowering phenology. In another study, researchers surveyed phyllosphere bacterial communities from over 50 species of trees in a neotropical forest in Panama and

identified correlations between host attributes (i.e., plant species, nutrient concentrations, leaf mass, host phenology, etc.) (Kembel et al., 2014). Both Wagner et al. and Kembal et al. were able to formulate testable hypotheses regarding which microorganisms may play a role in host phenology.

Numerous in-depth descriptive studies that utilize meta-'omics approaches (i.e., metagenomics, metatranscriptomics, metaproteomics, etc.) report attributes of plant microbiota, likewise providing a foundation for the hypothesis of functional plant-microbiome interactions. With the advent of high throughput sequencing technology, researchers are able to extract environmental DNA and perform shotgun metagenome sequencing and hypothesize about community functional potential based on the database annotations of identified reads. This approach has been used on rhizosphere and soil microbiomes. Comparative metagenomics across different environments allows one to hypothesize about microbial traits necessary for colonization, and identification of genes whose annotations (based on previous plant-microbe interaction mechanistic studies) suggest may play a role in the microbial manipulation of host phenology, one can hypothesize about community functional potential (Bai et al., 2015; Tringe et al., 2005). A major challenge with shotgun metagenomics is the inability to discern which genes are expressed, and under what circumstances. To mitigate this and get a step closer to function, metatranscriptomic and hybrid metagenomic/proteomic approaches have been used to identify the activity of microbial populations in situ (Delmotte et al., 2009; Garoutte, 2016). The hybrid and hybrid metagenomic/proteomic approach, in essence, captures mass spectroscopy signatures on peptide fragments from an environment, and subsequently mapped to

curated databases as well as to the corresponding metagenome (Delmotte et al., 2009; Knief et al., 2012). The community proteogenomics approach may be difficult in the rhizosphere and endophytic environments, due to artifacts from the soil as well as "contaminating" plant proteins.

Reducing environmental complexity is often necessary to identify strong correlations between microbial taxonomic composition and plant phenology, due to stochastic and uncontrollable abiotic parameters. Thus, some researchers have, in a sense, brought the ecology to the lab by stabilizing environmental variables and working with synthetic microbiota or complex microbial consortia. Inoculations of plant material with simplified microbial populations have been performed by many groups, but few have taken the analysis to the next level to assess corresponding host phenology. One group has used pure cultures of abundant phyllosphere bacteria to inoculate leaf material and utilize novel mass spectroscopy approaches to identify and quantify carbohydrate signatures on the leaf surface in the presence/absence of select phyllosphere microbiota (Ryffel et al., 2016). Likewise, host transcriptome profiling has been performed on Arabidopsis inoculated with a few select bacterial cultures isolated from the phyllosphere, but in a highly artificial tissue culture environment (Vogel et al., 2016).

An alternative "top-down" approach with additional complexity is to inoculate a plant or soil with microbial consortia and select for the desired host phenotype (i.e. heat

tolerance, salt tolerance, etc.) over multiple cycles of artificial selection. This has been performed by both industrial (for example, Bioconsortia) and academic groups, and has allowed researchers to identify consortia that directly correlate with a host phenotype (Mueller and Sachs, 2015; Panke-Buisse et al., 2015). There is a lack, however, of high-resolution profiling of host phenology tied to microbiota.

An alternative approach to identifying functional roles for host-associated microbiota is to establish an axenic, or germ-free, system and assess relative phenology in the presence and absence of microbiota. These axenic model systems have been used to form testable hypotheses about microbiome function, by proxy of identifying axenic abnormalities relative to conventionally colonized controls. As previously described in Chapter 2, the germ-free mouse model has been widely used in mammalian research. Here we characterize phenotypic attributes of an axenic (germ-free) plant and contrast them directly with holoxenic (colonized by natural microbiota) plants grown under identical conditions. Although conceptually similar studies have been performed (Badri et al., 2013; Carvalhais et al., 2013), we rigorously optimized our axenic system to isolate the microbiome variable, and normalize abiotic factors between axenic and holoxenic plants. For the purposes of this study, we chose to characterize axenic and holoxenic *Arabidopsis thaliana* (hereafter, Arabidopsis) to identify microbiome-dependent phenotypes. This study was performed under nutrient-replete conditions because pleiotropic effects of nutrient deprivation could obscure the ability to identify additional and potentially more subtle phenotypes, particularly in a non-diseased state. We took a top-down approach and broadly characterized axenic attributes using transcriptomics, metabolomic and phytohormone quantification to generate

hypotheses of microbiome function based upon axenic/holoxenic discrepancies. Multiple indications of immunity-related abnormalities prompted us to focus our efforts and perform, to our knowledge, the first in-depth and focused characterization of innate immunity of axenic Arabidopsis.

Materials and Methods

Soil collection and biochemical analysis

The soils used in this study were chosen from a collection of soils with a presumed high degree of dissimilarity in community composition: sandy loam agricultural soil (hereafter Ag soil) was chosen from a *Miscanthus* plot on Michigan State University campus (42.7089°N, -84.4656°E; altitude 261 M; East Lansing, Michigan, USA), fine loam soil from an arid region of California (hereafter Arid soil) was collected from a backslope near the Cleveland National Forest (33.3037°N, -116.8511°E; altitude 1,604 M; California, USA), and fine silt soil from an undisturbed grassland (hereafter Prairie soil) in Iowa (41.8361°N, -93.0078°E; altitude 271 M; Iowa, USA) (Figure 2). Soil collections were performed by removing the top ~5 cm of soil and sampling ~1 kg of soil in the 5-10 cm depth zone. A three-point linear transect separated by 1 m intervals was implemented for each location. After collection, the soil was homogenized, passed through a 3 mm test sieve, and dried for one week at room temperature under ambient humidity (~40%). Soil was stored in 50 or 100 g aliquots in Whirl-Pak bags (Nasco) at 4°C in the dark. Chemical analyses were performed on soil samples by the Michigan State University Soil and Plant Nutrient Laboratory.

Arabidopsis seeds and growth conditions

Arabidopsis thaliana accession Col-0 was obtained from ABRC (www.arabidopsis.org). The following mutants in Col-0 background were obtained from various laboratories: *fls2-1* (Thomas Boller), *npr1-1* (Xinnian Dong), *coi1-16* (John Turner) and *sid2* (Mary Wildermuth). The *bak1 bkk1-2 cerk1* and *efr fls2 cerk1* mutants were a gift from Cyril Zipfel. All seeds were size-selected (250-300 μm) using stacked test sieves.

Seed-associated microbiota were killed using a vapor-phase sterilization protocol, exposing aliquots of ~400 seeds to 6-8 h of chlorine gas (Clough and Bent, 1998). After vapor-phase sterilization, seed aliquots in 1.5 ml microfuge tubes were placed under sterile laminar airflow for ~10 min to remove any residual gas, and promptly closed and put into a desiccator at 4°C in the dark for storage. Seeds were suspended in 1 ml of sterile ddH2O for 48 - 72 h at 4°C in the dark to imbibe H_2O and synchronize germination prior to sowing. FlowPot assembly and plant growth protocol were performed as previously described in Chapter 2 of this dissertation.

DNA extractions

Whole rosettes were aseptically collected three weeks after germination and placed in 2.0 ml high-density polypropylene tubes with three 3 mm zirconium beads. Samples were frozen in liquid nitrogen and stored at -80°C until further processing. Three rosettes from the same FlowPot were pooled per sample collection tube. Sample tubes were refrozen in liquid nitrogen and contents homogenized using a Tissuelyser II (Qiagen, 30 Hz for 2 x 1 min). DNA collections were performed using a Powersoil kit (MoBio). Lysis solution from the PowerSoil kit was added to the homogenized leaf sample, resuspended by pipetting, and transferred back to the PowerSoil homogenization tube. Samples were then heated to 70C for 5 min, and remaining steps were performed according to the manufacturer's instructions. Matrix and soil DNA extractions were performed according to manufacturer's instructions with one modification: with the additional 5 min 70C incubation step during initial lysis. DNA was quantified and checked for quality using a ND-1000 UV Nanodrop spectrophotometer (Thermo Scientific).

ITS and 16S rRNA gene-targeted amplification

Targeted amplification of 16S rRNA gene fragments (regions V5-V7) was performed using primers 799F and 1192R, designed to exclude *Arabidopsis* mtDNA and pDNA (Bulgarelli et al., 2012). To accommodate compatibility with the Fluidigm Access Array tagging workflow, primers were modified with 5' CS-adapter sequences for tagging: CS1-799F [5'-ACACTGACGACATGGTTCTACA-AACMGGATTAGATACCCKG-3'] and CS2-1192R [5'-TACGGTAGCAGAGACTTGGTCT-ACGTCATCCCCACCTTCC-3']. PCR reactions (25 µl) were performed in triplicate and pooled after gel separation: AccuPrime HiFi Polymerase (0.15 µl, Thermo Fisher), DMSO (1.00 µl), forward and reverse primers (200 nM each), 10X AccuPrime Buffer II (2.50 µl, Thermo Fisher), template DNA (~20 ng), DNA-free water. Thermocycling conditions were as follows: hot start 94°C (1 min); 30 cycles of [94°C (20 s), 53°C (30 s), 68°C (45 s)], and a final extension at 68°C (2 min). No-template control reactions were performed as a contamination check, and entire PCR runs were repeated in the event of contamination. Amplicon DNA was gel purified, pooled by sample, checked for quality using an ND-1000 NanoDrop spectrophotometer (Thermo Scientific). Illumina adapters and barcodes were added during a second PCR according to the Access Array workflow (Fluidigm), priming from the CS1 and CS2 tags. Secondary PCR was performed using CS-16S fusion primers to increase fidelity and improve library quality. Amplicon products were normalized for concentration with SequalPrep DNA Normalization Plates (Invitrogen), and tagged amplicons were then pooled. DNA concentrations were measured by Qubit dsDNA and Kapa qPCR after all the tagged libraries were pooled. The pool was loaded on a MiSeq flowcell (Illumina, version 2

chemistry) according to the manufacturer's specified loading concentrations, supplemented with PhiX to reduce highly similar cluster density, and sequenced in a 2 x 250 bp format using a 500-cycle reagent cartridge. Base calling was done by Real Time Analysis software (Illumina, RTA v1.8.54), and sequencing was performed at the Michigan State University Research Technology Support Facility.

Template DNA from all samples was also processed by the Michigan State University Research Technology Support Facility to amplify ITS1 for fungal community profiling using primers ITS1F1 and ITS2R (Schmidt et al., 2013), modified with 5' Fluidigm CS1 and CS2 oligo sequences: CS1-ITS1F1 [5'-ACACTGACGACATGGTTCTACA-GAACCWGCGGARGGATCA-3'] and CS2-ITS2R [5'-TACGGTAGCAGAGACTTGGTCT-TGTGTTCTTCATCATG-3'].

Bioinformatic analysis of ITS1 and 16S rRNA gene fragment amplicons

Raw Illumina output files were checked for quality using FastQC and preprocessed using the Ribosomal Database Project pipeline (Cole et al., 2014). In brief, sequences were demultiplexed and trimmed for length, adapter removal, and quality (Q-score cutoff of 26). Paired-end reads from the 16S amplicons were assembled using the RDP-modified version of PANDAseq (Masella et al., 2012), and all remaining singleton sequences were discarded. ITS amplicons were quality trimmed, but not assembled. BBDuk was used to remove *Arabidopsis* reads from both 16S and ITS datasets, retaining summary statistics for the abundance of mitochondrial and chloroplast reads per sample (Bushnell et al., 2016). All sequences were rarified and stripped of singletons using USEARCH and UPARSE, respectively (Edgar et al., 2011). Clustering (97% identity cutoff) was independently performed on 16S and ITS sequences using

UPARSE (Edgar, 2013). Chimeric sequences were removed from the 16S and ITS datasets using UCHIME (Edgar, 2013), and ITS/16S OTU counts were determined by mapping reads back to respective '.uc' files. Representative sequences for 16S OTUs were extracted, aligned using the RDP Infernal 1.1 alignment model, and assigned taxonomy using the RDP classifier classification from the RDP database (Cole et al., 2014). Representative sequences for ITS OTUs were assigned taxonomy using the RDP classifier trained to the UNITE fungal database (ITS trainset 07-04-2014). ITS OTUs were defined using the USEARCH pipeline (version 9.2) and refined using a 97% identity cutoff of denoised sequences. Each pooled sample was rarefied to 100,000 reads. ITS OTUs were classified using the UNITE ITS database (version 7.1) trimmed to ITS1 and 50 flanking bases using ITSx (version 1.0.11) and HMMER (version 3.1b2).

RNA extraction and quantitative PCR

Total RNA was collected and purified from whole rosette tissue using the RNeasy Plant Mini Kit (Qiagen) according to manufacturer's protocol with optional on-column DNase digestion. Purified RNA was eluted in TE buffer (Tris-HCl 10 mM, pH 7.5, EDTA 1 mM). RNA concentrations were determined using an ND-1000 NanoDrop spectrophotometer (Thermo Scientific) or by Qubit RNA HS fluorometric assay (Thermo Fisher). Total RNA samples were collected in 2.0 ml nucleic acid LoBind tubes (Eppendorf) and stored at -80°C.

For quantitative PCR analysis, total RNA collections were used for cDNA biosynthesis. In brief, first strand cDNA libraries were generated using M-MLV reverse transcriptase according to previously published protocols (Withers and Yao, 2013). Quantitative PCR was performed in 10 µl reaction volumes with ~0.5 ng·µl^{-1} cDNA template using SYBR

Green reagents (Applied Biosystems) and primers specific to each experiment. All qPCR reactions were performed on an ABI 7500 Fast qPCR instrument (Applied Biosystems) with three technical replicates and a minimum of three biological replicates. A dissociation curve to confirm the presence of a single peak for each primer set was generated after the final cycle of each reaction using standard parameters (15 s at 95°C, 60 s ramp from 60°C to 95°C in 1°C increments, and 15 s at 95°C). All qPCR experiments were repeated a minimum of three times with consistent results.

RNA-seq library preparation and analysis

Total RNA was extracted from whole rosettes of FlowPot grown Arabidopsis with Ag, Prairie or Arid input microbiota. A biological replicate is defined as a pool of eight rosettes collected from four FlowPots within the same micro box. Three biological replicates per condition were collected, totaling nine holoxenic and nine axenic replicates. Total RNA was checked for quality using a Bioanalyzer 2100 (Agilent), and all samples were determined to have an RNA integrity (RIN) score of six or greater. Stranded sequencing libraries were prepared using the NuGEN Ovation RNA-SEQ System for Model Organisms (Arabidopsis) according to manufacturer's protocol (NuGEN). Library preparation and sequencing was performed by the Michigan State University RTSF. Sequencing was performed on the HiSeq 2500 (Illumina) with a 1 x 50 bp single read stranded format using Illumina HiSeq SBS reagents (version 4). Base calling was done by Illumina Real Time Analysis (RTA version 1.18.64). FastQ files were checked for quality using FastQC (version 0.11.5; Andrews, 2010), then processed using the Joint Genome Institute BBMap tool suite (version 36.86; (Bushnell, 2016).

Preprocessing was performed using BBDuk to remove low-quality regions, PhiX reads, Illumina adapters, and poly-A tails. The following parameters were implemented for adapter and poly-A removal: k = 23, mink = 11, hdist = 1, ktrim = r, trimq = 12. BBMap was used to map reads to the *Arabidopsis thaliana* Col-0 reference genome file in SAM format (ARAPORT11 assembly; Krishnakumar et al., 2015). A minimum of 30 million quality-passed reads were used for each sample. Splice variants were collapsed so that a single expression value could be determined for each gene locus. Structural RNAs were excluded from expression value calculations. Ambiguous read mappings were assigned to the "best" match according to default parameters. Expression values were determined from read counts normalized as fragments per kilobase per million reads mapped (FPKM). The significance of differential expression was determined for p values using Benjamini-Hochberg correction.

Tissue collection flg22 elicitation experiments

Plants were grown for three weeks in FlowPots under axenic conditions or with specified microbiota. Whole rosettes were severed at root-shoot junctions and carefully transferred to sterile distilled H_2O (20 ml, 22°C) in 50 ml conical tubes or Petri dishes. Suspended rosettes were gently rocked for 2 h, then bottom-loaded with aqueous flg22 (20 ml, 22°C, 250 nM final concentration, 0.1% DMSO). Samples were gently rocked at a 45-degree angle until the specified post-elicitation time points. For MAPK phosphorylation assays, the final concentration of flg22 was adjusted to 1 μm, rocking was omitted, and individual leaves were collected rather than whole rosettes. Mock-treated samples were subjected to the equivalent volume of 0.1% DMSO. At the time of collection, plant tissue was quickly blotted dry, immediately frozen in liquid

nitrogen, and stored at -80°C until homogenization for RNA or protein extraction. A biological sample is defined as three rosettes pooled from the same FlowPot or 3-4 leaves pooled from the same FlowPot for RNA and protein extraction, respectively.

Protein extraction

Membrane proteins were extracted as previously described. Briefly, tissues were mechanically disrupted and taken up in extraction buffer (50 mM Tris HCl pH 8.0, 150 mM NaCl, 10% glycerol, 1% Igepal CA 630 [Sigma Aldrich] and 0.5% sodium deoxycholate), supplemented with protease and phosphatase inhibitor cocktails (Roche) as needed). Soluble proteins were taken up in a buffer consisting of 50 mM Tris HCl, pH 8.0, and 150 mM NaCl, Samples were cleared of debris, normalized to total protein concentration. The relative abundance of each protein was visualized by western blot using anti AtMPK3 (Sigma Aldrich), anti AtMPK6 (Sigma Aldrich), anti BAK1 (Agrisera) or anti FLS2 (Agrisera) antibodies. Phosphorylation of AtMPK3/6 was measured using an anti-phosphor-p44/42 MAPK.

Phytohormone quantification by UPLC/MS

Plants were grown under the previously described FlowPot regime in the presence or absence of Michigan (Ag), Iowa (Prairie), or California (Arid) soil microbiota for three weeks. Extraction was performed according to previously published methodology, with some modifications (Zeng et al., 2011). In brief, rosette tissue (~50-100 mg) was weighed, harvested, and frozen in liquid nitrogen. Tissue and cells were homogenized using three 3 mm Zirconium beads (Glen Mills) and Tissuelyser II (Qiagen, 30 Hz for 2 x 1 min). Homogenates were vortexed and spun down at 12,000 × *g* for 10 min at 4°C.

Extraction was performed with 1 ml methanol water and taken up in phytohormone extraction buffer (methanol: water (80:20 v/v), 0.1% v/v formic acid, 0.1 $mg \cdot ml^{-1}$ butylated hydroxytoluene (BHT) supplemented with 100 nM deuterated-ABA (ABA-2H6) as an internal standard. Upon overnight incubation at 4°C with gentle agitation, samples were cleared by centrifugation, filtered through a 0.2 µm PTFE membrane (Millipore) and transferred to autosampler vials. Calibration standards (1.56 nM -100 nM) were prepared using SA, IAA, ABA, JA, MeJA (Sigma-Aldrich) and JA-Ile (Paul Staswick) supplemented with 100 nM ABA-2H6. Separation, quantification, and analysis were performed as described (Zeng et al., 2011). Transitions from deprotonated molecules to characteristic product ions were monitored for SA (m/z 137>93), JA (m/z, 209.1>59), ABA (m/z 263.1>153.1), JA-Ile (m/z 322.2>130.2, and ABA-2H6 (m/z 269.1>159.1) in negative ion mode and IAA (m/z 176.1>130.1) and MeJA (m/z 225.1>151.1 in positive ion mode. For SAG analysis, a gradient of 100 mM ammonium formate acidified with 0.05% formic acid (solvent A) and methanol (solvent B) was applied for 5 min over the same column at 0.4 $ml \cdot min^{-1}$. The gradient increased from 98:2 A:B to 50:50 A:B for 180 s, then immediately increased to 100% B for 60 s and returned to 98:2 A:B for the remainder of the run. Transitions from deprotonated molecules to characteristic product ions were monitored for SAG (m/z 299>137) using a cone voltage and collision energy of 22 V and 11 eV, respectively. UPLC/MS work was performed by Brad Paasch with assistance from Dr. Tony Schilmiller and Dr. Dan Jones at Michigan State University RTSF Mass Spectrometry Facility.

Metabolite quantification by GC/MS

Plants were grown under the previously described FlowPot regime in the presence or absence of Ag, Prairie, or Arid soil microbiota for three weeks. Extraction was performed according to previously published methodology (Zeng et al., 2011). In brief, ~200 mg (fresh weight) of rosette tissue were frozen in liquid nitrogen, homogenized upon the addition of three 3 mm Zirconium beads (Glen Mills) using a Tissuelyser II (30 Hz for 2 x 1 min, Qiagen), and extracted with 1 ml methanol : water (1 : 1 v/v) containing formic acid [0.1% (v/v)] and butylated hydroxytoluene (BHT, 0.1 $mg \cdot ml^{-1}$) at 4°C for 24 h. Homogenates were vortexed and spun down at 12,000 × *g* for 10 min at 4°C, and supernatants were collected, filtered through 0.2 µm PTFE membrane (EMD Millipore), and transferred to autosampler vials. Due to ion suppression of detectable signal, sorbitol was added as an internal standard to each of the extracts. Extracts were dried under nitrogen, derivatized, and processed according to previously published methodology (Li et al., 2012; Tschaplinski et al., 2012). Metabolite quantities are reported as $\mu g \cdot g^{-1}$ fresh weight (sorbitol equivalent), derived from the mean of at least three biological replicates from three experimental replicates (nine samples per condition). Statistical confidence is reported as P-values calculated using a student's t-test, and error is reported as standard error of mean values. Sample processing and GC/MS was performed in collaboration with Oak Ridge National Laboratory.

Bacterial growth conditions, sources, and strain construction

All bacterial strains were obtained from the He lab culture collection, with the exception of *Pseudomonas syringae* pv. tomato DC3000 Tn*7*-GFP, which was constructed for

this study by triparental mating according to previously published methods (Teal et al., 2006). The insertion locus of the mini-Tn*7-nptII-gfp* cassette was confirmed by PCR to be in the predicted *att*Tn*7* site downstream of *glmS*, and donor plasmid was removed by supplementing growth media with 10% sucrose for *sacB*-mediated counter-selection. All *Pseudomonas* strains were grown at 28°C, and *E. coli* strains grown at 37°C in LM or LB media with appropriate antibiotics, respectively.

Infiltration of rosettes with bacterial suspensions

Overnight cultures of the specified *Pseudomonas* strains were prepared in LM broth supplemented with rifampicin and incubated at 28°C. The overnight cultures were diluted to OD_{600} = 0.002 (~5 x 10^6 CFU·ml^{-1}) in liquid LM media, spread onto LM agar plates, and incubated for 24 h at 28°C to form lawns. The working infection titre was prepared by gently rinsing the lawns in sterile $MgCl_2$ [1 mM] and diluting to OD_{600} = 0.0002 (~5 x 10^5 CFU·ml^{-1}) in $MgCl_2$ [1 mM] unless specified otherwise. Bacterial concentrations were confirmed by plating a serial dilution of the suspension in triplicate on LM agar supplemented with appropriate antibiotics, and CFUs were counted after 48 h of incubation at 28°C. Silwet L-77 [0.0025% v/v] (Lehle Seeds, Round Rock, TX, USA) was added 10 min prior to infiltrations, and homogenized with a magnetic stir rod at low speed to prevent foaming. Sterile 80 ml beakers were each filled with 60 ml of bacterial suspension. Immediately prior to vacuum infiltration, each FlowPot was immersed, stabilized by a sterile paper clamp fastened to the Luer end. The immersed FlowPots (six at a time) were placed into a glass bell vacuum desiccator (Wheaton). Vacuum was applied until the pressure within the glass bell desiccator reached 70 kPa, whereupon the vacuum was maintained for 1 min and slowly released. The vacuum

step was repeated for a total of two iterations. FlowPots were removed from suspensions and placed on KimWipes. Residual liquid was gently blotted from rosettes, and the FlowPots were placed back into the Eco2 boxes. Miracloth (EMD Millipore) was placed over the opening of the Eco2 box and fastened using the cut-away lip portion of an Eco2 lid. Each covered Eco2 box was placed in a BSL2 hood for ~1 h, or until the rosettes no longer appeared water-soaked. Microbox lid covers were then replaced, and boxes were put back in the growth chamber at 23°C [100 $\mu mol \cdot m^{-2} \cdot s^{-1}$] for 2, 3 or 4 d as specified. Mock inoculations were performed in parallel.

Bacterial multiplication in planta protocol

Whole rosettes of representative plants were collected after vacuum infiltration at the time point specified. Leaf discs (3 mm^2) were punched from representative leaves at approximately the same developmental stage and were placed in 2.0 ml high-density screw-cap microtubes containing three zirconium beads (3 mm). Per biological replicate, a total of four leaf discs were pooled from plants sharing the same FlowPot. To each tube, $MgCl_2$ buffer [1 mM] or phosphate buffer (PB, 10 mM, pH 7) was added (500 µl), and leaf discs were mechanically homogenized using a Tissuelyser II (Qiagen, 30 Hz for 2 x 1 min). After homogenization, an additional 500 µl of MgCl2 or PBS buffer was added to each tube, vortexed, and a 3-step 100-fold serial dilution was performed in a round-bottom 96-well microplate with 500 µl well capacity (Nunc, #267334). For assays in which the infiltrated bacteria constitutively express *gfp* (*Pst*-GFP), cell counts were enumerated using an Accuri C6 flow cytometer (BD Biosciences) equipped with a robotic sampler arm (BD Biosciences). In brief, a 10X dilution of a *Pst*

DC3000-infiltrated leaf homogenate was analyzed, followed by a wash step and a 100X dilution of *Pst*-GFP infiltrated leaf homogenate 3 dpi with estimated bacterial counts of ~$1x10^8$ viable cells·ml^{-1}. A volume of 100 µl was analyzed from each sample at medium fluidics flow rate. The Accuri C6 was set to the following parameters: 6,000 SSC threshold, FL1 filter (510/15), FL2 filter (530/40), FL3 (670LP). A GFP-specific gate was established by removing background events from the *Pst* DC3000 infiltrated leaf homogenate sample viewed on log_{10} scale scatter plots of FL1 by FSC and FL2 by FSC axes (supplemental figure). A gate for the positive control *Pst* DC3000-GFP sample was established by simultaneous visual inspection of the *Pst* DC3000-GFP and *Pst* DC3000 leaf homogenate scatter plots, and events within the GFP-specific gate were determined to be analogous to CFU·cm^{-2} when compared to plate counts of serial dilutions. Experimental cells·cm^2 data was recorded for samples with 100-6,000 GFP events·s^{-1} according to manufacturer's suggestions. Multiple dilutions within the 100-6,000 GFP events·s^{-1} range for the same biological sample are considered technical replicates and data points are reported as the mean of technical replicate values. A minimum of six biological replicates were analyzed for each condition per experiment, and all multiplication assays were repeated a minimum of three times with consistent results. Multiplication assays for non-GFP bacteria were performed according to a previously published protocol (Katagiri et al., 2002).

Oxidative burst assay

Reactive oxygen species generated upon flg22 elicitation were quantified using a luminol-based kinetic assay (Leslie and Heese, 2014). In brief, micro boxes containing 3-4 weeks old *Arabidopsis* were partially opened to allow relative humidity equilibration

for ~30 m. Leaf discs (3 mm^2) from representative plants were transferred to a white flat-bottom 96-well microplate (Berthold Technologies), and suspended in ddH_2O (100 µl per well) with abaxial side facing down. Leaf discs were floated for 12-14 h to attenuate wounding responses. Horseradish peroxidase (HRP, 10 $mg \cdot ml^{-1}$) was dissolved in ddH2O, and luminol [17 $mg \cdot ml^{-1}$] was dissolved in DMSO. From stock solutions, a combined aqueous working solution of HRP [80 $\mu g \cdot ml^{-1}$] and luminol [13.6 $\mu g \cdot ml^{-1}$] was prepared. A separate elicitor solution was prepared by diluting flg22 in ddH2O [800 nM]. The HRP/luminol working solution was added to each well containing a leaf disc (50 µl per well), and the plate was then incubated in the dark for 15 min. The elicitor solution was added to each well (50 µl per well), and the suspended leaf discs were immediately placed in a SpectraMax L Microplate Reader (Molecular Devices) to collect luminescence data. Final flg22 concentration for oxidative burst assays was 200 nm. The SpectraMax L parameters were set to the following: kinetic read mode, photon counting mode, integration time 1 s, target wavelength 470 nm, and assay time 90 min. Leaf discs from multiple plants sharing a common FlowPot were considered technical replicates, and the mean of 4-6 technical replicates is considered a biological replicate. Each experiment used a minimum of three biological replicates from different micro boxes, and the flagellin-insensitive mutant *fls2-1* or *efr fls2-1* was used as a negative control. Data is reported as mean photons per square centimeter of leaf tissue per second, and error bars represent standard deviation of the mean. Experiments were repeated a minimum of three times with consistent results.

Callose deposition assay

Staining for callose depositions was performed according to a previously published protocol with some modification (Nomura et al., 2011). In brief, representative 2-3 w rosettes were vacuum-infiltrated with flg22 (500 nM + 0.0025% Silwet L-77) and placed in micro boxes. After a 24 h incubation period under standard growth conditions, whole rosettes were collected and immersed in a clearing solution (1:3 acetic acid: ethanol) for 1 h at 65°C, then incubated overnight at room temperature with gentle rocking. Destaining solution was replaced if saturated. Clearing solution was aspirated and replaced with rinse buffer (K_2HPO_4 [150 mM]) for ~30 min, and replaced with aniline blue staining solution (K_2HPO_4 [150 mM] pH 8.5 and aniline blue [1 $mg \cdot ml^{-1}$]) for 2 h. Representative leaves were mounted in glycerol [50% v/v] and observed under UV excitation to detect and photograph callose depositions. Depositions were enumerated using ImageJ software and are reported as mean depositions per square centimeter for three leaves (leaf two, three, and four), among eight biological replicates. Error bars represent the standard deviation of depositions per square centimeter. Callose deposition assays were repeated a minimum of three times with similar deposition enumeration and identical conclusions. Flg22-insensitive mutant *fls2-1* was used as a negative control.

Results

Bacterial and fungal compositions of input soil communities are distinct

To minimize the community-specific effects of a microbiome on Arabidopsis phenotypic attribute, I elected to use three microbial input communities from distinct soil types and geographic locations. Chemical analysis of the soils revealed that the Prairie soil was closest to neutral pH (6.6), and that the Ag and Arid soils are more acidic (pH 5.6 and 5.4, respectively), and that the Ag and Arid soils have low buffering capacity (Table 1). Other notable differences between the soils are relatively low concentrations of P (9 $\mu g \cdot g^{-1}$), K (77 $\mu g \cdot g^{-1}$), Ca (186 $\mu g \cdot g^{-1}$), Mg (52 $\mu g \cdot g^{-1}$) and Mn (7.3 $\mu g \cdot g^{-1}$) in Ag soil, a relatively high concentration of P (200 $\mu g \cdot g^{-1}$) and organic content (6.3%) in Prairie soil, and relatively high concentrations of K (379 $\mu g \cdot g^{-1}$), Fe (190 $\mu g \cdot g^{-1}$), Na (28 $\mu g \cdot g^{-1}$), S (357 $\mu g \cdot g^{-1}$), Cl (1,000 $\mu g \cdot g^{-1}$) and soluble salts (3.52 $\mu g \cdot g^{-1}$) in Arid soil (Table 1). Precipitation and seasonal temperature fluctuations also contribute to different environmental factors that likely contribute to composition and diversity of each respective soil microbiome.

Figure 3.1. Map of input community soil collection locations and photographs taken at the time of collection. From Left to Right: a fine loam “Arid” soil from a backslope in southern California (33.3037°N, -116.8511°E; altitude 1,604 M), a sandy loam “Ag” agricultural soil from Michigan from an unamended *Miscanthus* plot (42.7089°N, -84.4656°E; altitude 261 M) and a fine silt “Prairie” soil from an undisturbed grassland in Iowa (41.8361°N, -93.0078°E; altitude 271 M).

Table 3.1. Attributes of soils used in this study

	Agricultural	Prairie	Arid
%Organic	2.1	6.3	2.3
pH	5.6	6.6	5.4
BpH	6.7	6.9	6.9
P	9	200	72
K	77	170	379
Ca	186	1499	1958
Mg	52	296	284
Mn	7.3	49.8	68
Cu	1.1	6.7	6.2
Fe	35	46	190
NO3-	22	7	317.9
NH4+	2.1	8.8	5.6
S	14	12	357
Na	1	2	28
Soluble Salts	0.12	0.23	3.52
Chloride	151	178	1000

All measurements are in $\mu g \cdot g^{-1}$

Bacterial and fungal communities of the Ag, Arid, and Prairie soils were analyzed for microbial community composition by targeted amplicon sequencing of the 16S rRNA SSU gene fragment and the ribosomal intergenic spacer region (ITS1), respectively. Among the three input soil samples, 367 ITS1 OTUs accounted for a minimum of 0.1% of the reads for at least one sample. Although 14 OTUs could not be classified to the phylum level, the majority of OTUs were classified to four known phyla (n=353): Ascomycota (n=257), Basidiomycota (n=64), Zygomycota (n=30), and Glomeromycota (n=2). OTU classifications fell into sixteen different classes, of which 182 were classified to the genus level (>0.5 confidence from RDP classifier). Only four core OTUs were detected, classified as the following genera: *Fusarium* (n=2) and *Trichoderma* (n=2 (Figure 2C). Collectively, these four OTUs accounted for <10% of the total classified ITS reads for each sample (Arid=2.7%, Prairie=3.9%, Ag=9.6%) (Table 3). Two OTUs were exclusively shared among Arid and Prairie, both of which are classified as *Alternaria*. Seven OTUs are exclusively shared among Arid and Ag, five of which are classified as fungal genera (within Ascomycota): *Humicola* (n=2), unidentified Chaetothyriales (n=2), and *Drechslera* (n=1). Notably, Zygomycota was far less abundant in Arid soil relative to Prairie and Ag (Zygomycota reads as percent of total: Arid=0.17%, Prairie=31.53%, Ag=4.94%).

Bacterial community composition analysis of the input soils yielded 298 16S rRNA SSU OTUs (97% cutoff), each accounting for a minimum of 0.1% average of the reads for at least one sample group. Relative to ITS1 OTUs, there is a greater number of core 16S OTUs are shared among all input communities (n=13; Figure 2A), also accounting for a relatively higher portion of shared reads (Arid=18.7%, Prairie=22.9%, Ag=36.6%).

Notably, Prairie soil contains over 10-fold more Firmicutes and Bacteroidetes than Arid or Ag soils (Figure 2B; Figure 5). Agricultural soil had the greatest relative abundance of Acidobacteria, positively correlating with the acidic pH of the soil. The Arid soil had the highest relative abundance of Actinobacteria (Figure 2).

Table 3.2. Classifications of core 16S OTU classifications from input soil

Phylum	Family	Genus
Actinobacteria	Mycobacteriaceae	Mycobacterium
Firmicutes	Bacillaceae	Bacillus
Alphaproteobacteria	Bradyrhizobiaceae	Bradyrhizobium
Alphaproteobacteria	Bradyrhizobiaceae	unclassified
Alphaproteobacteria	Rhizomicrobium	Rhizomicrobium
Alphaproteobacteria	Sphingomonadaceae	Sphingomonas
Alphaproteobacteria	Sphingomonadaceae	Sphingomonas
Alphaproteobacteria	Xanthobacteraceae	Pseudolabrys
Alphaproteobacteria	Hyphomicrobiaceae	Rhodoplanes
Betaproteobacteria	Burkholderiaceae	Burkholderia
Betaproteobacteria	Comamonadaceae	Variovorax
Betaproteobacteria	unclassified	unclassified
Betaproteobacteria	Comamonadaceae	unclassified

Classifications based on RDP taxonomy with at least 80% confidence.

Table 3.3. Classifications of core ITS1 OTU classifications from input soil

Phylum	Family	Genus
Ascomycota	Nectriaceae	Fusarium
Ascomycota	Nectriaceae	Fusarium
Ascomycota	Hypocreaceae	Trichoderma
Ascomycota	Hypocreaceae	Trichoderma

Classifications based on RDP taxonomy with at least 80% confidence.

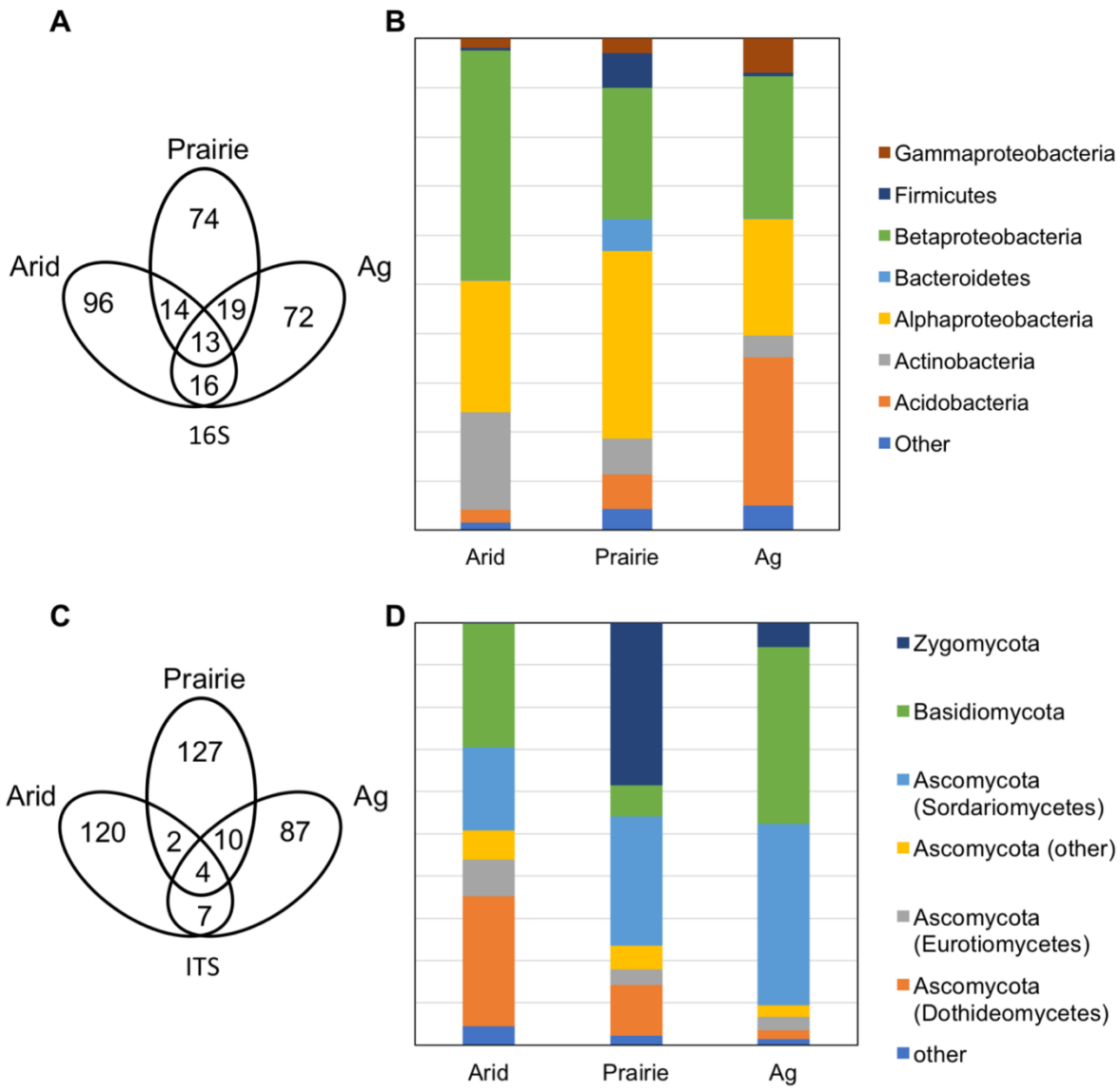


Figure 3.2. Venn diagram of classified bacterial 16S rRNA OTUs from input communities, displaying number of shared OTUs among each sample, along with (B) relative abundance of OTU (>0.1%) classifications at the Phylum level. (C) Venn diagram of classified fungal ITS1 OTUs, displaying number of shared OTUs among each sample, (D) relative abundance of OTU (>0.1%) classifications at the Phylum level.

Fungal and bacterial community differentiation

We have established that the input soil communities are distinct, and in Chapter 2, we determined that both synthetic and naturally-extracted soil communities differentiate to form distinct root-associated and substrate-inhabiting communities six weeks after inoculation in the FlowPot system. As a basis for downstream phenotypic analysis of the host, we further characterized differentiation of the Ag, Arid, and Prairie input communities at week three, comparing within-pot differentiation of rosette-associated microbiota from substrate-associated, as well as comparisons of community composition as the result of different input community inoculum. This experiment was conducted in the FlowPot system according to the schema depicted in Figure 3. Total rosette tissue and bulk substrate was collected three weeks after germination (occurs 2-3 days after sowing).

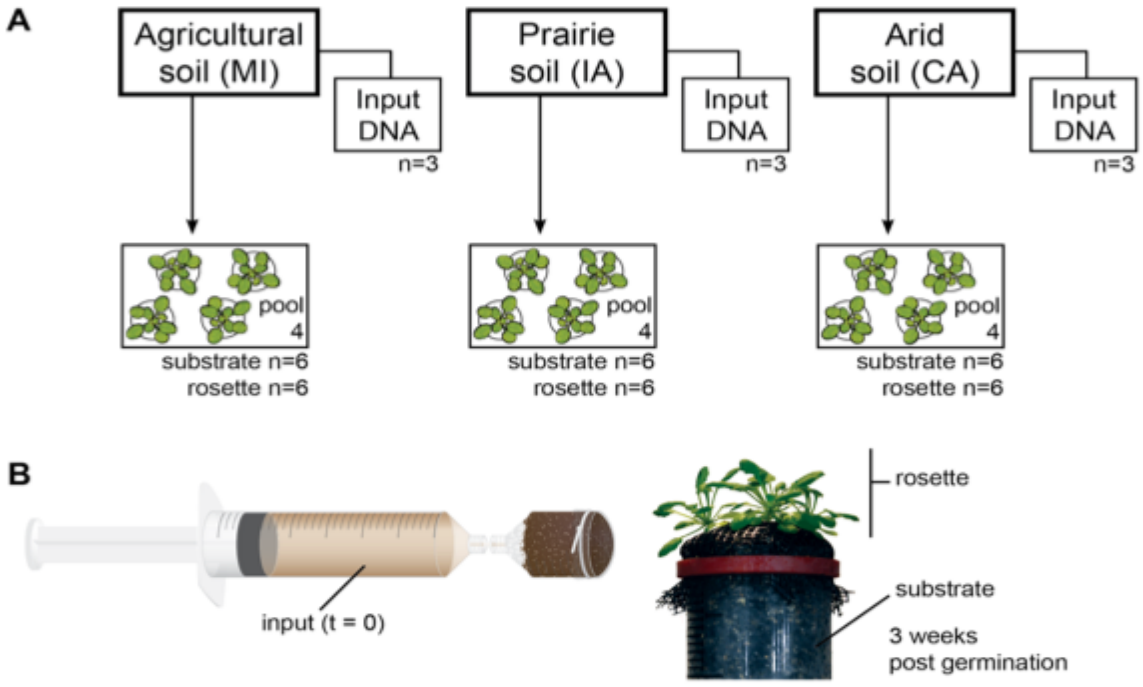


Figure 3.3. Experimental design for 3-community FlowPot inoculation: Agricultural soil, Prairie soil, and Arid soil. (A) Three aliquots of each soil input community are collected at t=0. For each soil input community, 6 micro boxes with 4 FlowPots each are (B) inoculated. At week 3, (B) whole rosettes are collected, along with substrate, and samples are pooled by box. Thus, each micro box containing 4 individual FlowPots is considered a single sample.

In addition to the steps taken for input ITS1 fungal community analysis, for substrate and rosette samples we took an additional step to remove artefactual OTUs resulting from amplification of *Arabidopsis* DNA, with kmer-based decontamination tool BBDuk (version 36.92, kmer length=150). A total of 13.26% of unique sequences prior to denoising and OTU clustering. Additionally, a BLAST search was performed using the TAIR10 assembly of *Arabidopsis thaliana* Col-0 to scan for OTUs with high similarity to *Arabidopsis* DNA. One OTU (OTU1) with 99.1% identity across 100% of the query to a region on *A. thaliana* chromosome 3, annotated as the ITS1 region, was removed. Thus, despite USEARCH classification of OTS1 as *Xylaria* with 98.3% confidence, OTU1 was treated as plant contamination. The reference sequence to which OTU1 has the highest identity is *Xylaria* sp. B101 (genbank accession KJ512146.1). The ITS sequence of *Xylaria* sp. B101 is an unpublished direct submission to NCBI described as a fungal endophyte of *Cleisostoma williiamsonii,* and no additional information is provided regarding whether steps were taken to eliminate plant genomic DNA during template preparation. Among samples with >50,000 high quality read pairs, *Arabidopsis* (OTU1) accounted for an average of 0.001%, 11.762% and 99.419% of classified paired reads for respective soil input, bulk, and rosette DNA samples. Among the detected ITS1 OTUs across all samples, 1413 are classified as Fungi, 15 are classified as Plantae, 72 are classified as Protista, and 20 as unidentified. OTUs are classified at the phylum level for 988/1413 OTUs. Plantae OTUs are unclassified at higher resolution than the domain level. All Protista OTUs are classified as Cercozoa, and 44/72 of the Cercozoa classifications have greater than 90% confidence of classification. Fungal-classified OTUs accounted for an average of 98.644%, 85.413%

and 0.580% of paired reads for respective soil input, matrix, and rosette samples. Although rosette DNA with Ag microbiota has <0.1% read pairs classified as fungal OTUs, rosettes with Arid-derived and Prairie-derived microbiota have >1.1% and >0.6% of read pairs classified as fungal OTUs, respectively. Read pairs classified as Protista account for <1.0% of all samples with only two exceptions: Prairie input (2.05%) and Ag substrate (5.30%). A single core OTU (OTU4) identified as *Fusarium* sp. (74.85% confidence) was detected among rosettes with Arid microbiota. Three core OTUs identified as *Penicillium* sp. (OTU10, 85.29% confidence), *Mucor* sp. (OTU39, 86.78% confidence), and *Umbelopsis* sp (OTU16, 99.16% confidence) were detected among rosettes with Iowa-derived microbiota. Rosettes with Ag microbiota have a single core ITS1 OTU classified as a member of Chaetomiaceae (OTU2, 80.3% confidence).

The bulk substrate bacterial community composition for all three samples was mostly (>80%) composed of Proteobacteria (by decreasing abundance: Betaproteobacteria, Alphaproteobacteria, and Gammaproteobacteria) (Figure 4). While Arid and Prairie groups respectively had 3.7% and 4.1% Bacteroidetes reads, Ag had >0.1% Bacteroidetes. None of the bulk substrate samples had >0.5% Actinobacteria. Beta diversity analysis using UniFrac distances revealed strong clustering by sample type (Figure 5). Despite similar taxonomic profiles at the phylum-level resolution, bulk substrate samples only 16 core OTUs (Figure 6C), and rosettes share only 6 (Figure 6D). Among soil input communities, Ag and Arid had the greatest alpha diversity based on the inverse Simpson index (Figure 6A), but bulk substrate and rosettes with Ag and Arid microbiota had less alpha diversity than samples with Prairie microbiota. Input

communities all had greater alpha diversity than bulk substrate and rosettes. For each corresponding sample, the alpha diversity of bulk substrate was not significantly different from rosette microbiomes (Figure 6AB). The most highly abundant OTUs in rosette samples with Arid microbiota were classified as: *Massilia* Otu2 (30.9±5.6% of reads), *Chryseobacterium* Otu5 (21.5±8.7% of reads), *Achromombacter* Otu8 (17.9±4.3% of reads), and *Pseudomonas* Otu6 (9.8±1.7% of reads). Dominant Prairie rosette microbiota were: *Novosphingobium* Otu4 (20.1±4.3%), *Pseudomonas* Otu6 (15.1±3.8% of reads), *Rhizobium* Otu13 (9.6±4.0%), and *Cupriavidus*/*Ralstonia* Otu11 (5.3±1.6%). For rosettes with Agricultural microbiota were predominantly composed of: *Paraburkholderia* Otu1 (55.7±1.6%), *Rhizobium* Otu3 (16.4±1.4%), *Paraburkholderia* Otu14 (7/0±1.9%). OTUs with greater relative abundance in the rosette that the corresponding bulk substrate were considered rosette-enriched (also >0.1% abundant, pairwise t-test $p<0.05$). Rosettes with Arid microbiota were enriched in 9 OTUs (Figure 8), Prairie with 21 OTUs (Figure 9), and Agricultural with 11 OTUs (Figure 10; Table 4).

Table 3.4. Rosette-enriched OTUs relative to abundance in bulk substrate.

			Arid			Prairie			Agricultural		
OTU	Classification		%Rs	Rs/B	p	%Rs	Rs/B	sig	%Rs	Rs/B	p
42	Ac	Gordonia	1.1	17.2	*	0.1	6.7				
90	Ac	Gordonia	0.3	8.5	*						
64	Ac	Leifsonia				0.5	16	*	0.1	6	*
114	Ac	Nocardioides				0.1	5.4	*			
48	Ac	Rhodococcus				0.6	3	*			
5	Bc	Chryseobacterium	21.5	6.5		1.5	16.2	**			
40	Fr	Paenibacillus	0.6	24.1	*	0.1	10	*	0.6	35.7	
10	AP	Caulobacter	0.4	23.5	*	0.1	1	**			
941	AP					0.2	19.3	*			
1429	AP								0.2	12.9	***
352	AP					0.2	8.3	***			
4	AP	Novosphingobium				20.1	16	***			
13	AP	Rhizobium				9.6	27.7	*	0.5	32.2	
160	AP	Rhizobium				0.4	8.6	*			
27	AP	Rhizobium				2.2	3.1	*			
3	AP	Rhizobium				0.4	2.4		16.4	47.2	***
59	AP	Sphingomonas							0.1	3.1	***
8	BP	Achromobacter	17.9	22	**	2.4	0.7		0.6	0	
1	BP	Paraburkholderia	0.4	0.1		0.5	1.5		55.7	0.4	**
166	BP	Paraburkholderia							0.4	1.1	*
7	BP	Paraburkholderia	0.2	1.7					1.7	10.2	***
1530	BP	Herbaspirillum							0.1	1.4	*
21	BP	Herbaspirillum	1	8.4	**	1.6	0.6				
2	BP	Massilia	31	0.3		2.7	4.8	**	2.3	2.1	*
43	BP	Massilia	0.2	9.9	**	0.1	4.9	**	0.3	29.8	***
61	BP								0.4	28.7	***
11	BP	Ralstonia				5.3	2.9	*			
20	GP	Erwinia	0.8	0.4		1.3	3.2	*			
24	GP	Pantoea	3.1	5.9	*	0.1	0.7				
592	GP	Pseudomonas				0.6	1.5	*			
6	GP	Pseudomonas	9.8	5.7	**	15.1	1.2	*	1.6	1	
99	GP	Stenotrophomonas	0.2	0.9		0.8	5.6	*			

%Rs= mean percent of rosette sample reads of the sample group, R/B= root enrichment over bulk. Only OTUs with >0.1% abundance in one or more rosette samples were included. Ac= Actinobacteria, Bc=Bacteroidetes, Fr= Firmicutes, AP= Alphaproteobacteria, BP= Betaproteobacteria, GP= Gammaproteobacteria, p-values from two-tailed t-test: $p < 0.05$ *, $p < 0.005$ **, $p < 0.0005$ ***.

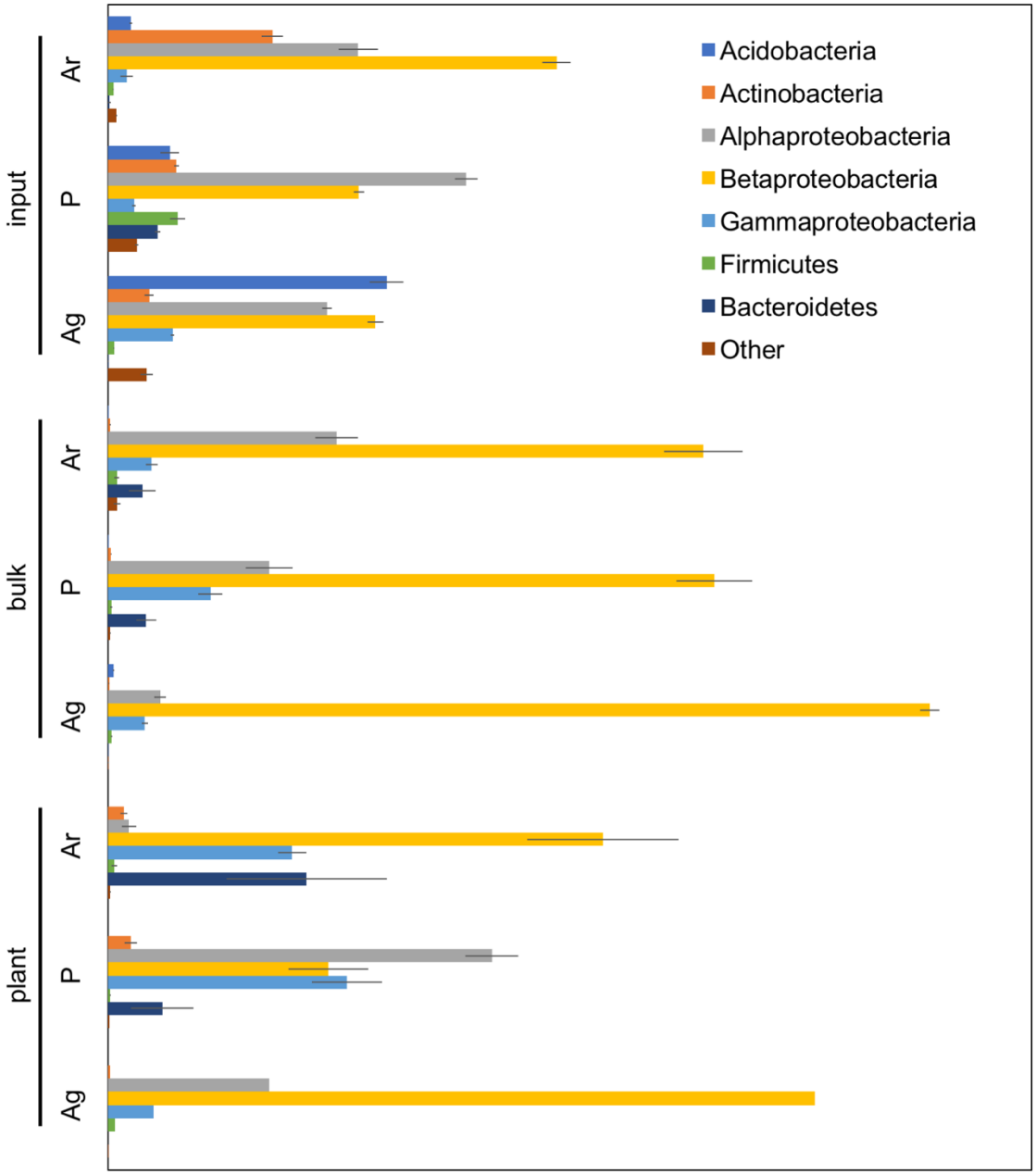


Figure 3.4. Mean percent abundance of 16S OTU-classified reads by phylum. The right bound maximum across the x-axis represents 100%. Error bars represent standard error of the mean for each sample type. Ar= Arid soil; P=Prairie soil; Ag=Agricultural soil.

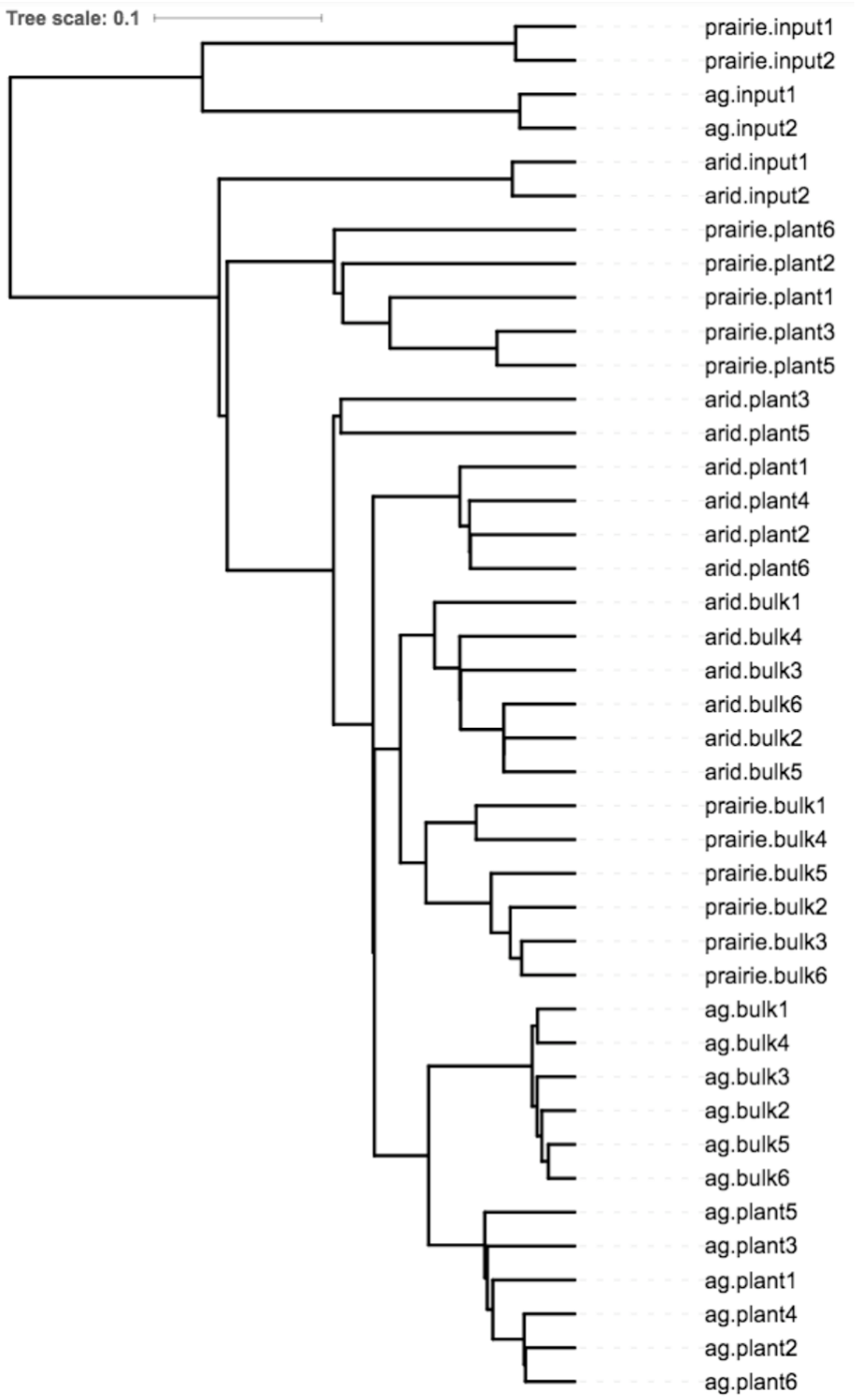


Figure 3.5. UniFrac distances between samples based on 16S OTU clustering.

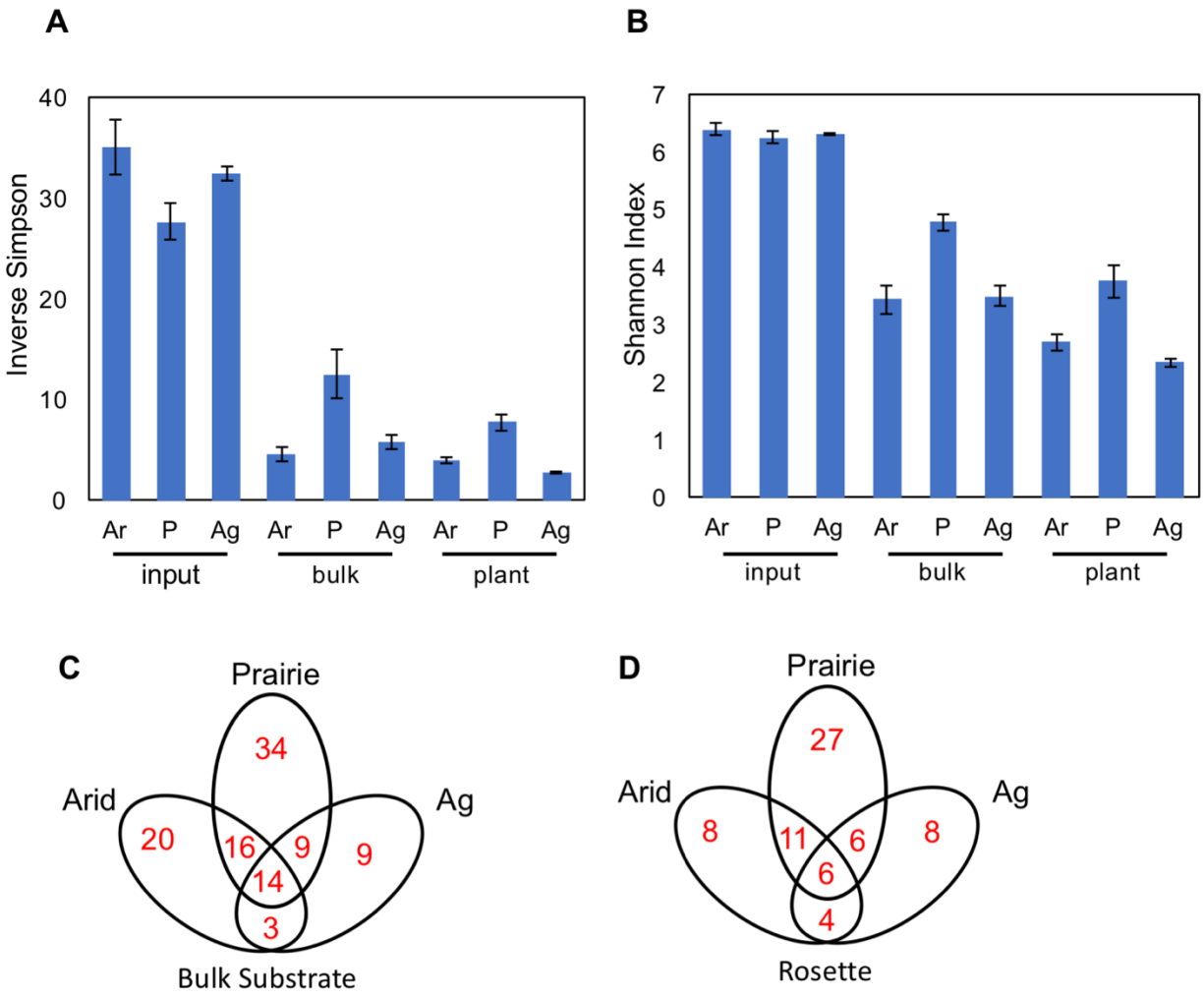


Figure 3.6. Alpha diversity indices for each sample type: (A) inverse Simpson and (B) Shannon (base 2). Error bars represent standard error of the mean. Alpha diversity indices are calculated from rarefied count table to 10,000 reads per sample. All bulk and root samples have a minimum of n=5 samples. The number of shared OTUs with a minimum of 0.1% mean read abundance in at least one sample group was determined for (C) bulk substrate and (D) plant rosettes. Soil input microbiota abbreviations: Ar=Arid, P=Prairie, Ag=Agricultural.

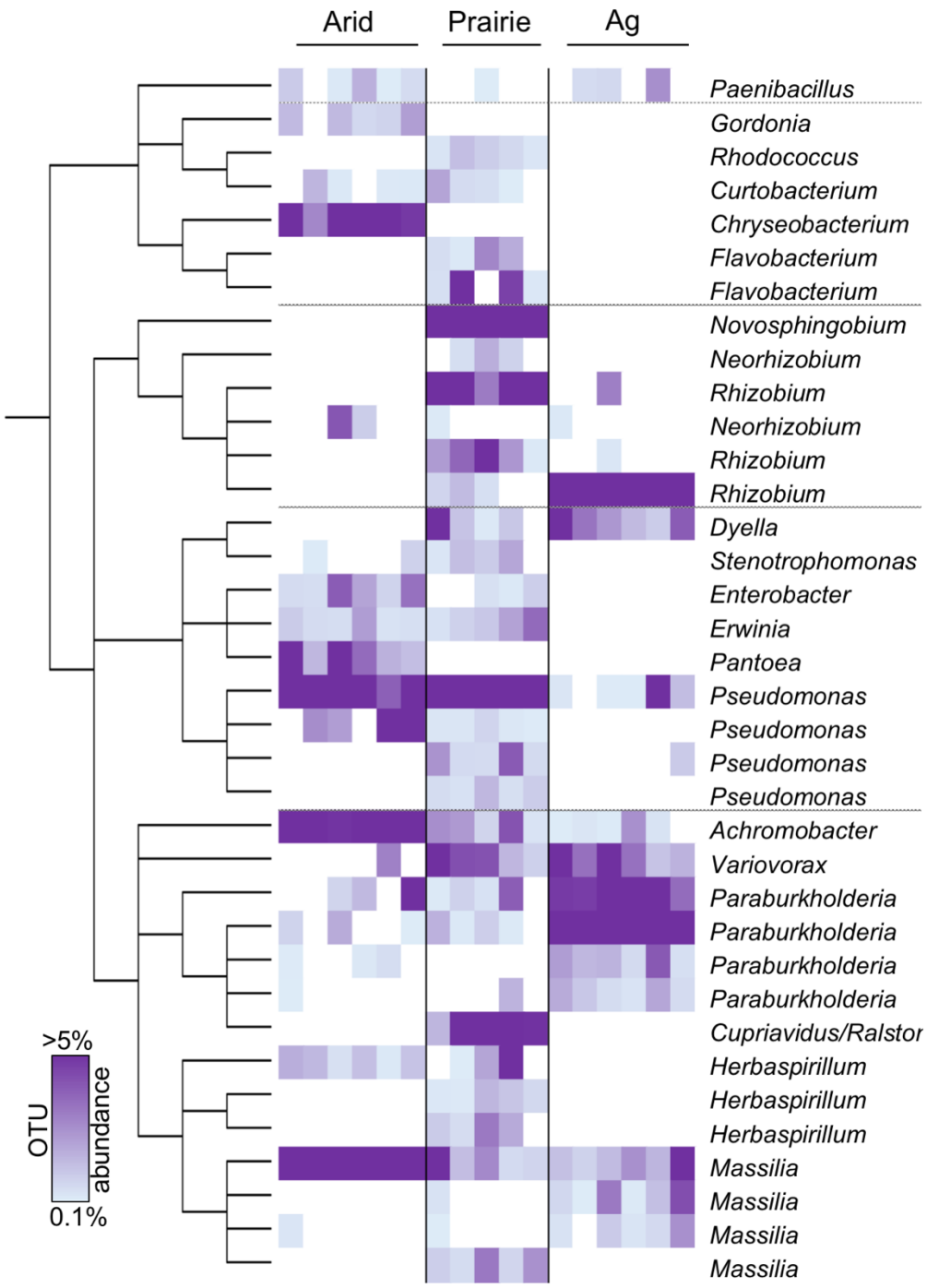


Figure 3.7. Abundance of 16S OTUs for plant (whole rosette) communities with genus classifications. Each of the 36 OTUs individually accounted for >0.5% of the average reads in at least one rosette sample group. Each column represents a sample replicate. Reads are sum scale normalized.

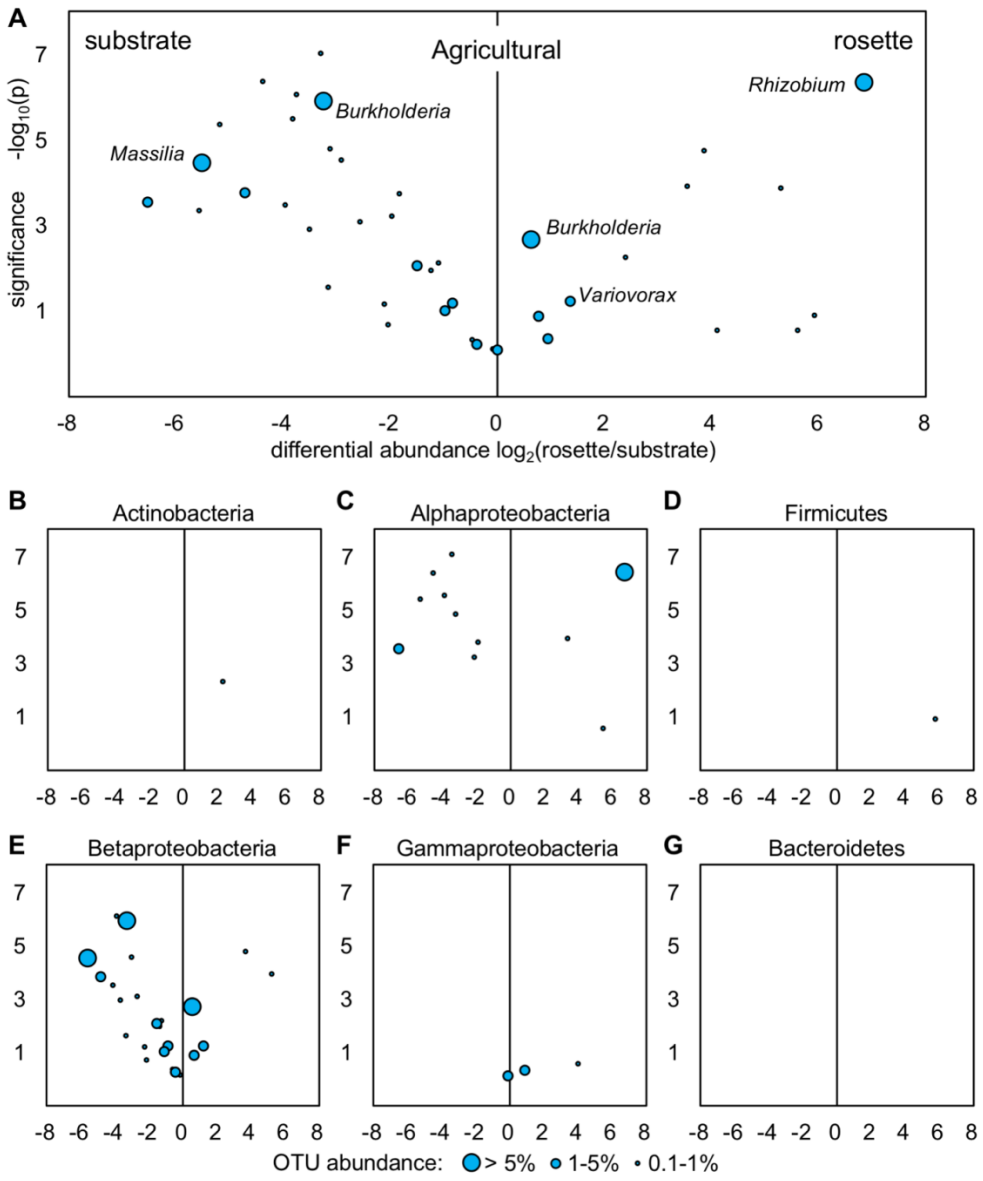


Figure 3.8. Rosette microbiome 16S OTU enrichment relative to abundance in bulk substrate for Agricultural microbiota samples. (A) All 16S OTUs plotted and OTUs filtered by phylum (B-G). Log_2(rosette/bulk) abundance plotted across the x-axis. Statistical confidence was calculated for each OTU comparison by pairwise t-test and plotted on the y-axis as $-log_{10}(p)$. OTU data point symbols are size scaled to reflect abundance in rosette microbiome (if $x>0$) or bulk substrate microbiome (if $x<0$).

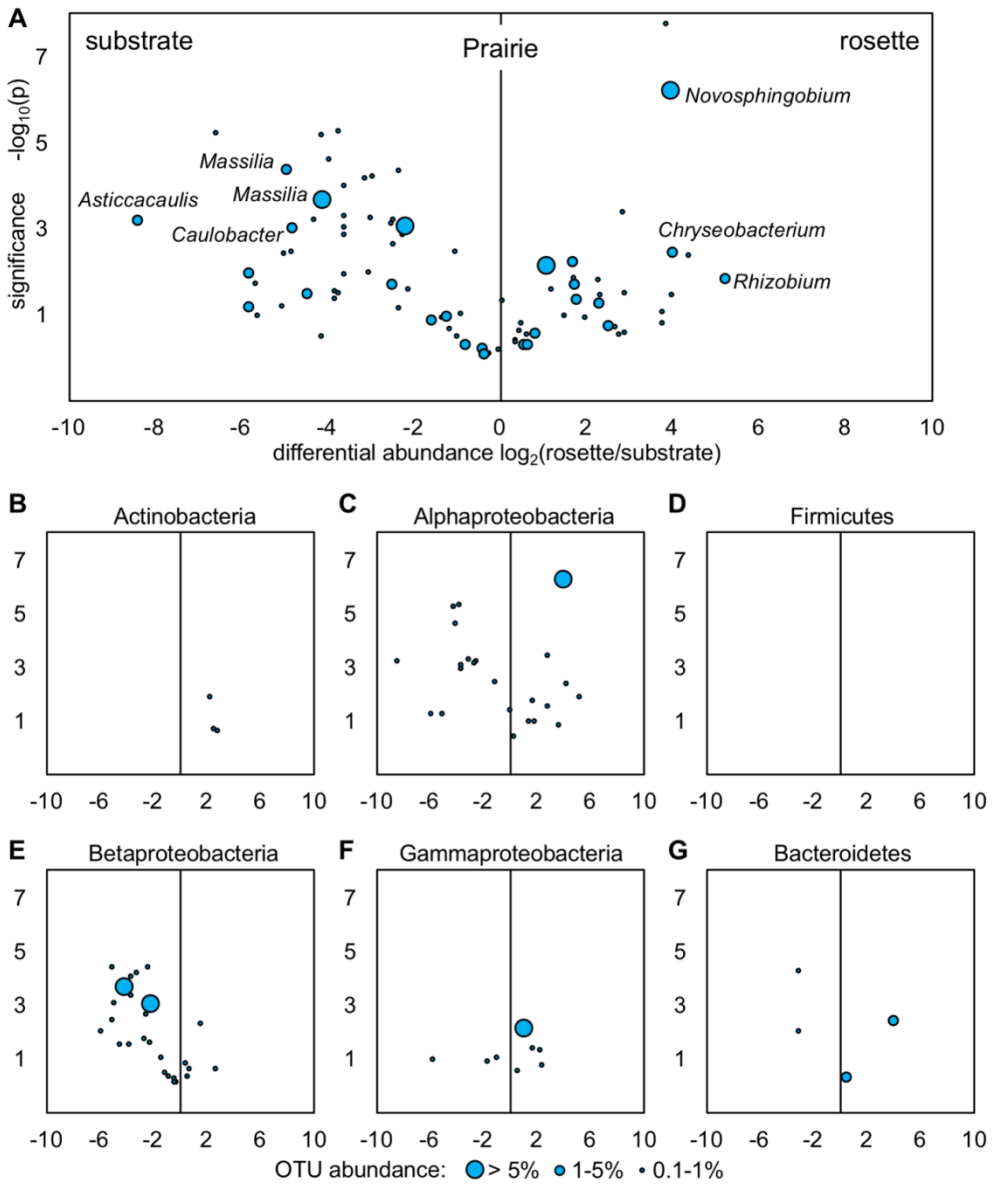


Figure 3.9. Rosette microbiome 16S OTU enrichment relative to abundance in bulk substrate for Prairie microbiota samples. (A) All 16S OTUs plotted and OTUs filtered by phylum (B-G). Log_2(rosette/bulk) abundance plotted across the x-axis. Statistical confidence was calculated for each OTU comparison by pairwise t-test and plotted on the y-axis as $-log_{10}(p)$. OTU data point symbols are size scaled to reflect abundance in rosette microbiome (if $x>0$) or bulk substrate microbiome (if $x<0$).

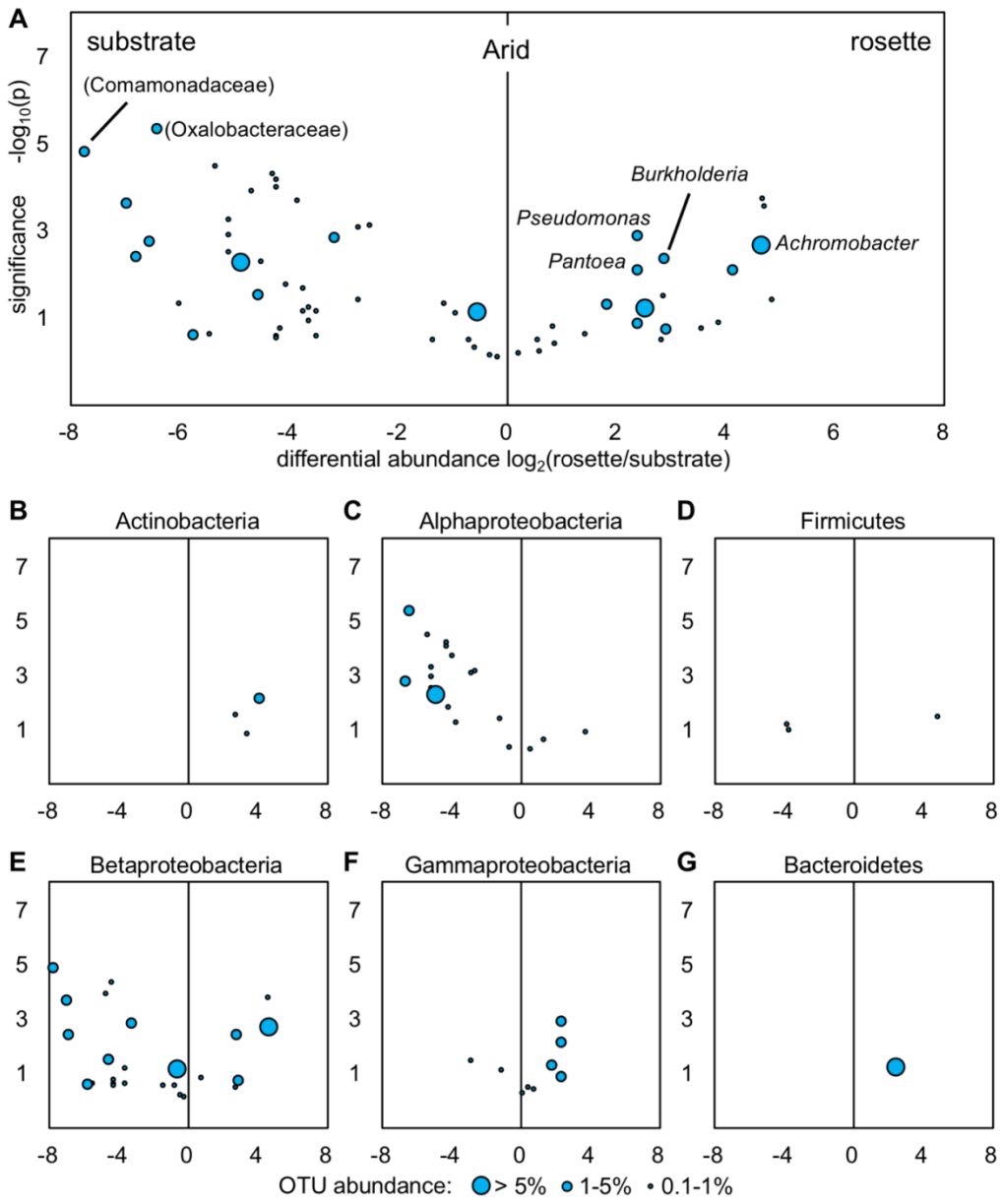


Figure 3.10. Rosette microbiome 16S OTU enrichment relative to abundance in bulk substrate for Arid microbiota samples. (A) All 16S OTUs plotted and OTUs filtered by phylum (B-G). Log_2(rosette/bulk) abundance plotted across the x-axis. Statistical confidence was calculated for each OTU comparison by pairwise t-test and plotted on the y-axis as $-log_{10}(p)$. OTU data point symbols are size scaled to reflect abundance in rosette microbiome (if x>0) or bulk substrate microbiome (if x<0).

Axenic plants lack normal basal expression of immune-associated genes

To achieve a global snapshot of basal transcriptional differences between axenic and holoxenic Arabidopsis, we used the FlowPot system to grow axenic plants in parallel with plants in substrate inoculated with Ag (MI), Arid (CA) or Prairie (IA) soil microbiota (Figure 8), collected whole rosette tissue at week 3, and performed RNAseq on total RNA. Axenic plants were confirmed to be truly axenic, or devoid of culturable microbiota, by incubating sterilized inoculum at t=0 on R2A medium, as well as incubating plant material on R2A medium at the time of tissue collection for RNA extraction. There was no detectable microbial growth from material from any axenic micro boxes used in this experiment.

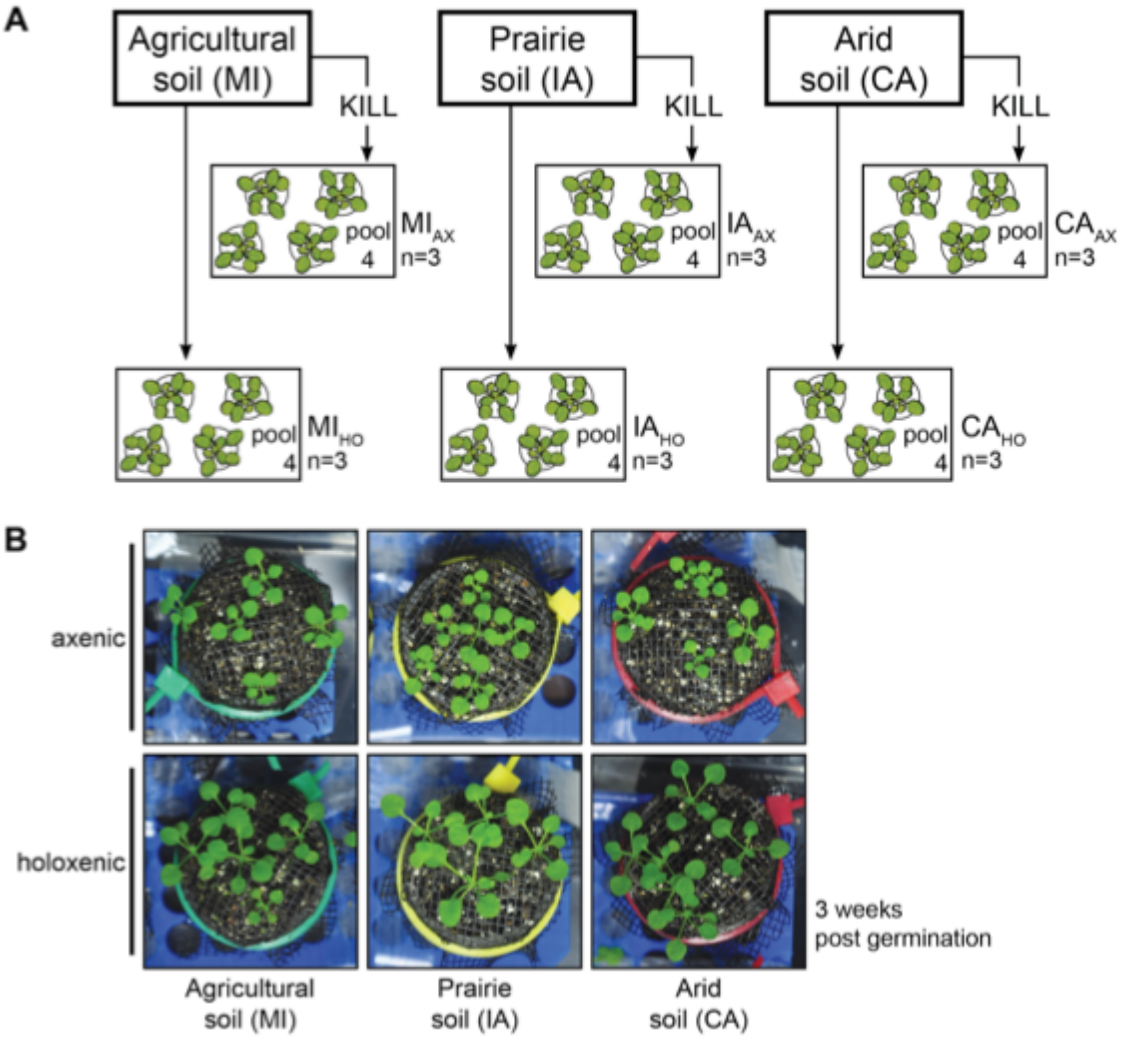


Figure 3.11. RNA extraction experimental design and pipeline. (B) Images of plants at the time of collection for total RNA extraction.

RNA-seq data was analyzed using two separate group definitions: all holoxenic (n=9) vs all axenic (n=9), and input-specific holoxenic (3 groups of n=3) vs. input-specific axenic (3 groups of n=3). In the all vs. all (axenic vs holoxenic) analysis, biological replicates do not account for the input community source (holoxenic n=9, axenic n=9). This allowed for greater statistical power and noise reduction to determine core differentially expressed genes. A gene's expression value is defined as mean reads per million (RPM), averaged across all nine replicates. The input-specific analysis provides granularity to determine whether certain genes are differentially expressed when colonized by certain taxa, but at the expense of statistical power. This allows for the determination of pan differentially expressed genes.

Using stringent criteria to define an expressed gene (RPM>5), we detected the pan expression of 15,050 genes, accounting for ~54% of annotated protein-coding Arabidopsis genes (Table 5). Differential expression of genes was analyzed at three levels of stringency: $|\log_2(\text{RPM/RPM})|>2.0$ (Table 5B), $|\log_2(\text{RPM/RPM})|>1.0$ (Table 5C), and $|\log_2(\text{RPM/RPM})|>0.5$ (Table 5D). Collectively, there are only three 'core' differentially expressed genes at the $|\log_2(\text{RPM/RPM})|>2.0$ stringency, all of which are less abundant in axenic: a peroxidase superfamily protein-coding gene (*AT5G64120*), a MLO-family protein-coding gene (*MLO12*), and a chitinase family protein-coding gene. Notably, all three of these genes are defense-associated based upon gene ontology. The peroxidase superfamily protein is involved in lignification of cell walls and induced during fungal defense, the MLO12 protein is a calmodulin-binding and is annotated as a mildew resistance locus, and the chitinase family protein is also involved in fungal defense (Supplemental Table).

Of the 30 core differentially expressed genes (i.e., differentially expressed axenic vs. all holoxenics) using moderate stringency criteria, 29 are less abundant in axenics (Table 6). The only gene expressed at relatively higher levels in axenic is a putative sugar-phosphate exchange protein (*AT4G32480*) (mean $\log_2$(ax/holo)=1.5). Gene ontology analysis (GO:BP), based on direct experimental evidence, indicates that 11 of the axenic-depleted genes are categorized as "defense response to other organism [GO:0098542]", ($p<0.0008$, Benjamini-Hochberg corrected). Notably, this list includes two defense-associated WRKY transcription factors (*WRKY48*, *WRKY53*), REDOX-associated glutathione S-transferase (*GSTF6*), genes involved in glucosinolate metabolism (*CYP81F2*, *IGMT2*), a cell wall-associated peroxidase (AT5G64120), receptor-like kinases with leucine-rich repeat domains (*IOS1*, *MLO12*, *CRK14*), and a chitinase family protein (*AT2G43620*).

Table 3.5. Differentially expressed genes (DEG) with stringent criteria.

	ax enriched		ax depleted		differential (DEG)	
	genes	%exp	genes	%exp	genes	%exp
Agricultural (Ag)	1	0.01	46	0.3	47	0.31
Ag+Pr	1	0.01	23	0.16	24	0.16
Prairie (Pr)	5	0.03	31	0.2	36	0.24
Arid+Pr	0	0	3	0.02	3	0.02
Arid	0	0	15	0.1	15	0.1
Arid+Ag	0	0	4	0.03	4	0.03
Arid+Ag+Pr (core)	0	0	3	0.02	3	0.02
Pan (holo pooled, ax pooled)	0	0	22	0.15	22	0.15

Criteria: $|\log2(RPM/RPM)| \geq 2.0$ $p < 0.05$, $RPM \geq 5$

Table 3.6. Differentially expressed genes (DEG) with moderate criteria.

	ax enriched		ax depleted		differential (DEG)	
	genes	%exp	genes	%exp	genes	%exp
Agricultural (Ag)	22	0.15	188	1.24	210	1.39
Ag+Pr	5	0.03	97	0.66	102	0.69
Prairie (Pr)	36	0.24	164	1.08	200	1.32
Arid+Pr	1	0.01	33	0.22	36	0.24
Arid	28	0.18	173	1.14	201	1.32
Arid+Ag	2	0.01	55	0.37	57	0.39
Arid+Ag+Pr (core)	1	0.01	29	0.2	30	0.21
Pan (holo pooled, ax pooled)	9	0.06	136	0.9	145	0.96

Criteria: $|\log2(RPM/RPM)| \geq 1.0$ $p < 0.05$, $RPM \geq 5$

Table 3.7. Differentially expressed genes (DEG) with sensitive criteria.

	ax enriched		**ax depleted**		**differential (DEG)**	
	genes	%exp	genes	%exp	genes	%exp
Agricultural (Ag)	248	1.64	634	4.18	882	5.82
Ag+Pr	37	0.25	207	1.40	246	1.66
Prairie (Pr)	281	1.85	535	3.53	816	5.39
Arid+Pr	21	0.14	97	0.66	140	0.95
Arid	246	1.62	605	3.98	851	5.59
Arid+Ag	41	0.28	174	1.18	219	1.48
Arid+Ag+Pr (core)	10	0.07	70	0.48	81	0.55
Pan (holo pooled, ax pooled)	104	0.69	501	3.33	605	4.02

Criteria: $\log2(RPM/RPM)| \geq 0.5$ $p < 0.05$, $RPM \geq 5$

Table 3.8. Core differential expression meeting significance criteria at moderate stringency.

Name	Locus	Arid	Ag	Prairie	Description
AT5G64120	AT5G64120	-2.69	-3.98	-4.45	Peroxidase superfamily protein
MLO12	AT2G39200	-2.25	-3.93	-3.78	Seven transmembrane MLO family
AT2G43620	AT2G43620	-2.00	-3.56	-3.70	Chitinase family protein
IOS1	AT1G51800	-1.55	-4.08	-2.81	Leucine-rich repeat RLK
IGMT2	AT1G21120	-2.34	-2.07	-3.12	O-methyltransferase family protein
AT1G30720	AT1G30720	-1.60	-2.82	-2.51	FAD-binding Berberine family
GSTF6	AT1G02930	-1.15	-2.34	-2.95	glutathione S-transferase 6
AT1G30730	AT1G30730	-1.35	-2.71	-2.43	FAD-binding Berberine family
AT2G27660	AT2G27660	-1.52	-2.96	-1.74	Cysteine/Histidine-rich C1 domain
AT1G26420	AT1G26420	-1.44	-2.24	-2.45	FAD-binding Berberine family protein
WRKY48	AT5G49520	-2.42	-1.97	-1.68	WRKY DNA-binding protein 48
BCB	AT5G20230	-2.17	-1.78	-1.72	blue-copper-binding protein
AT1G72060	AT1G72060	-1.63	-1.92	-1.91	serine-type endopeptidase inhibitor
AT2G18690	AT2G18690	-1.73	-1.87	-1.91	transmembrane protein
FRB1	AT5G01100	-2.25	-1.80	-1.45	O-fucosyltransferase family protein
CRK14	AT4G23220	-1.69	-1.80	-1.46	cysteine-rich receptor kinase
CYP81F2	AT5G57220	-1.15	-1.47	-2.19	cytochrome P450
AT5G67340	AT5G67340	-1.02	-1.99	-1.73	ARM repeat superfamily protein
AT5G35735	AT5G35735	-1.39	-1.77	-1.63	Auxin-responsive family protein
WRKY53	AT4G23810	-1.46	-1.59	-1.43	WRKY family transcription factor
AT1G49000	AT1G49000	-1.37	-1.54	-1.36	transmembrane protein
AT1G65845	AT1G65845	-1.33	-1.51	-1.34	transmembrane protein
EXO70B2	AT1G07000	-1.44	-1.59	-1.15	exocyst subunit exo70 family
CNI1	AT5G27420	-1.07	-1.25	-1.73	carbon/nitrogen insensitive 1
PMT6	AT4G36670	-1.08	-1.59	-1.15	Major facilitator superfamily protein
AT5G52750	AT5G52750	-1.23	-1.47	-1.23	Heavy metal transport/detoxification
AT1G61360	AT1G61360	-1.28	-1.16	-1.53	S-locus lectin protein kinase family
AT4G39830	AT4G39830	-1.20	-1.39	-1.37	Cupredoxin superfamily protein
DIC2	AT4G24570	-1.36	-1.02	-1.16	dicarboxylate carrier 2
AT4G32480	AT4G32480	+1.75	+1.51	+1.10	sugar phosphate exchanger

Criteria (for all sample groups): $|\log2(RPM/RPM)| > 1.0$, $RPM > 5$, $p < 0.05$ (Benjamini-Hochberg corrected)

At high-sensitivity differential gene expression criteria, there are 605 differentially-expressed genes (4.02% of all protein-coding genes): 501 axenic-depleted and 104 axenic-enriched. To simultaneously visualize significance and magnitude of differential expression, a volcano plot was generated, with sensitive parameters for differential gene expression indicated by solid black lines (Figure 11). Gene ontology enrichment analysis (GO:BP biological process) was performed on these significantly enriched and significantly depleted regions. Overwhelmingly, the most strongly enriched category with the highest confidence is GO:BP involved in defense, with a p-value (Benjamini-Hochberg corrected) of $p < 3.62 \times 10^{-42}$. Of the 501 axenic-depleted transcripts, 137 are annotated as defense-associated (Figure 12). Collectively, at stringent and moderate differential expression criteria, there are more differentially-expressed genes in common among holoxenic plants grown in Ag (Michigan) and Prairie (Iowa) inoculated substrate. A distance matrix was generated based on gene expression values (RPM), and principal component analysis was performed and plotted to visualize whether community-specific gene expression profiles cluster together (Figure 11). Indeed, while all axenic samples cluster together, Ag (Michigan) and Prairie (Iowa) samples cluster away from the axenic, and gene expression profile differences are explained entirely by the x-axis (PC1). The Arid (California) holoxenic plants had expression profiles that clustered together, but apart from the Arid and Ag holoxenic plants, forming a third discrete cluster.

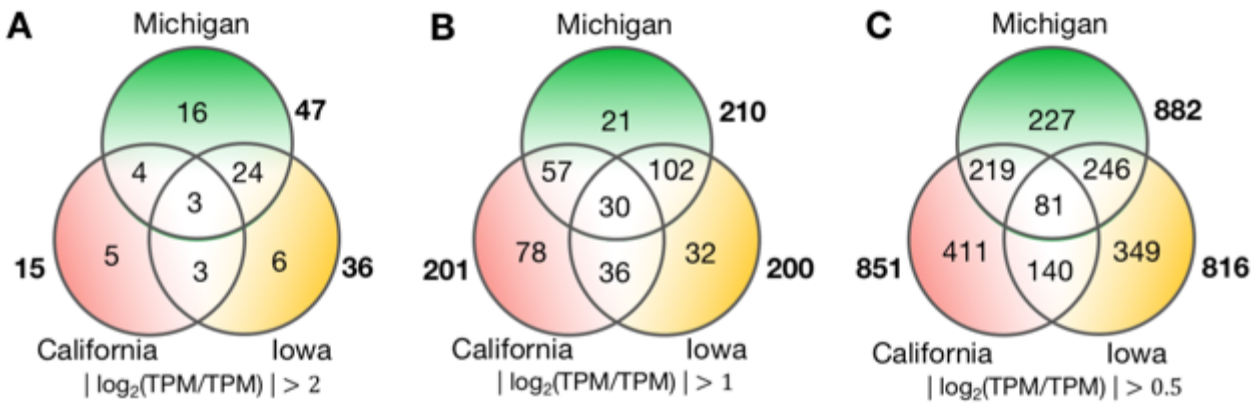


Figure 3.12. Total differentially expressed genes considering input community (sample size = 3 for each), displayed with at three stringency levels: (A) strict, (B) moderate, and (C) sensitive.

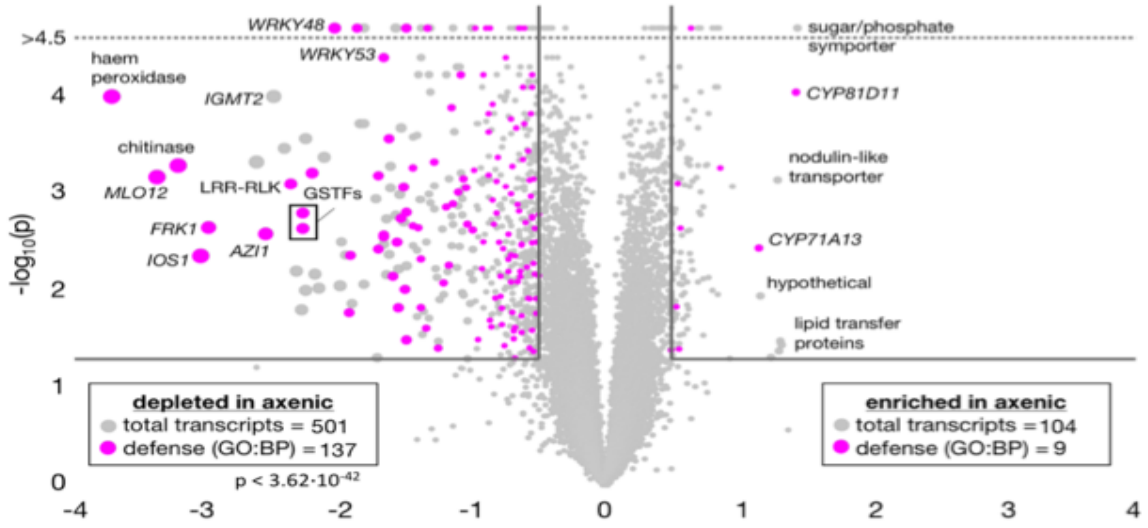


Figure 3.13. Volcano plot of pan differential gene expression (all axenic/holoxenic). Positive values across the x-axis indicate the gene is more highly expressed in axenic relative to holoxenic ($\log_2(RPM_{AX}/RPM_{HO})$. The boxed-out regions on the left and right correspond to the most sensitive criteria for differential expression ($\log_2(RPM_{AX}/RPM_{HO})|>0.5$). Gene ontology analysis (GO:BP biological process) was performed on these significantly enriched and significantly depleted regions. Annotations that indicate the gene is involved in biotic defense are highlighted in pink. Of the 501 axenic-depleted transcripts, 137 are annotated as defense-associated. The p-value (Benjamini-Hochberg corrected) of this enrichment analysis is $p < 3.62 \times 10^{-42}$. Gene names or corresponding brief gene descriptions are annotated as labels next to the symbol representing the corresponding gene.

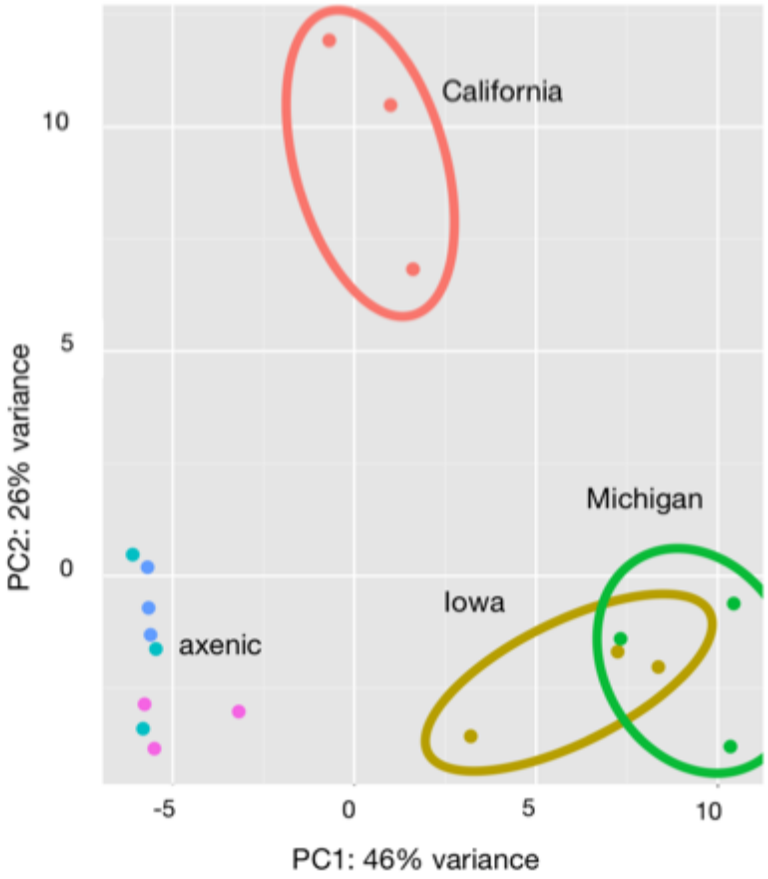


Figure 3.14. Principal component analysis for transcriptome profiles in this experiment. Profiles for axenic plants all cluster together and separate from holoxenic samples. Iowa and Michigan holoxenic cluster together, while holoxenic California clusters away from the other two groups. Michigan=Ag; Iowa=Prairie; California=Arid.

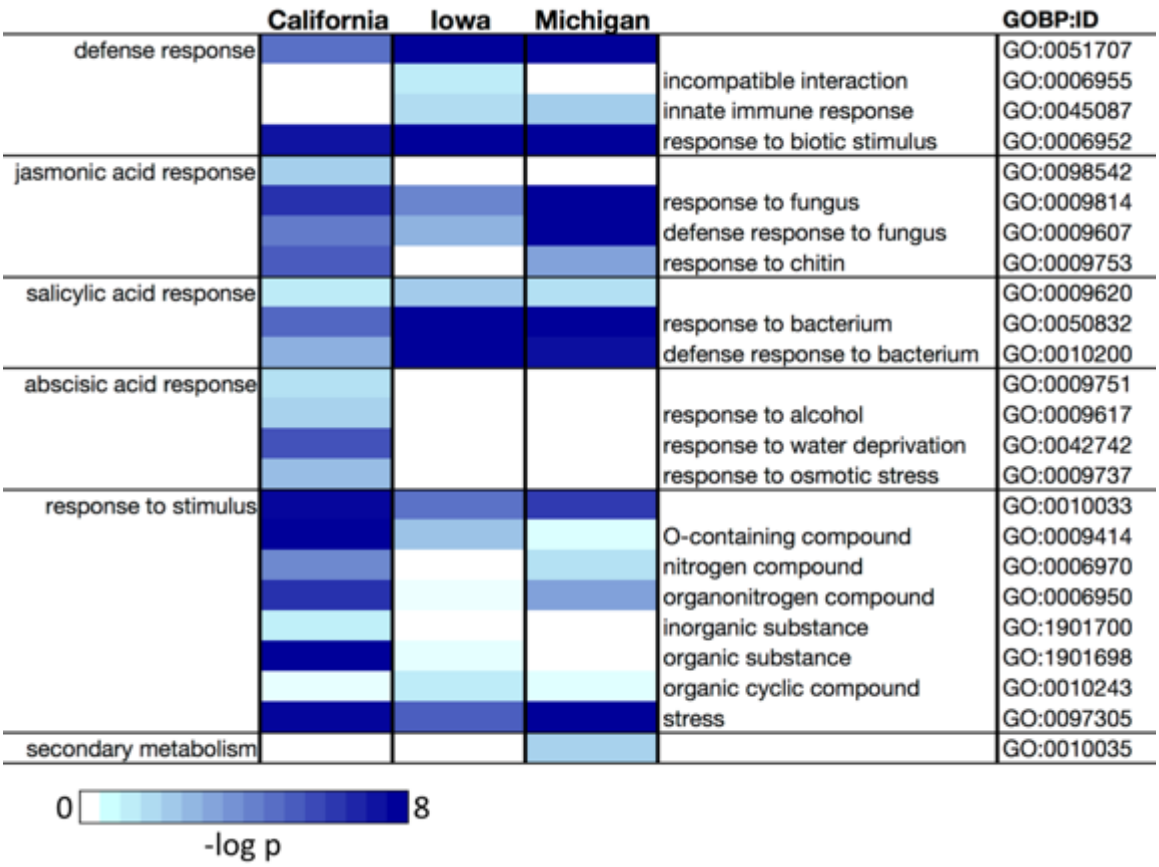


Figure 3.15. Gene set enrichment analysis of microbiome composition-dependent changes in host transcriptome, relative to corresponding axenic samples. Confidence is expressed as $-\log_{10}$ of a p-value (Benjamini-Hochberg corrected).

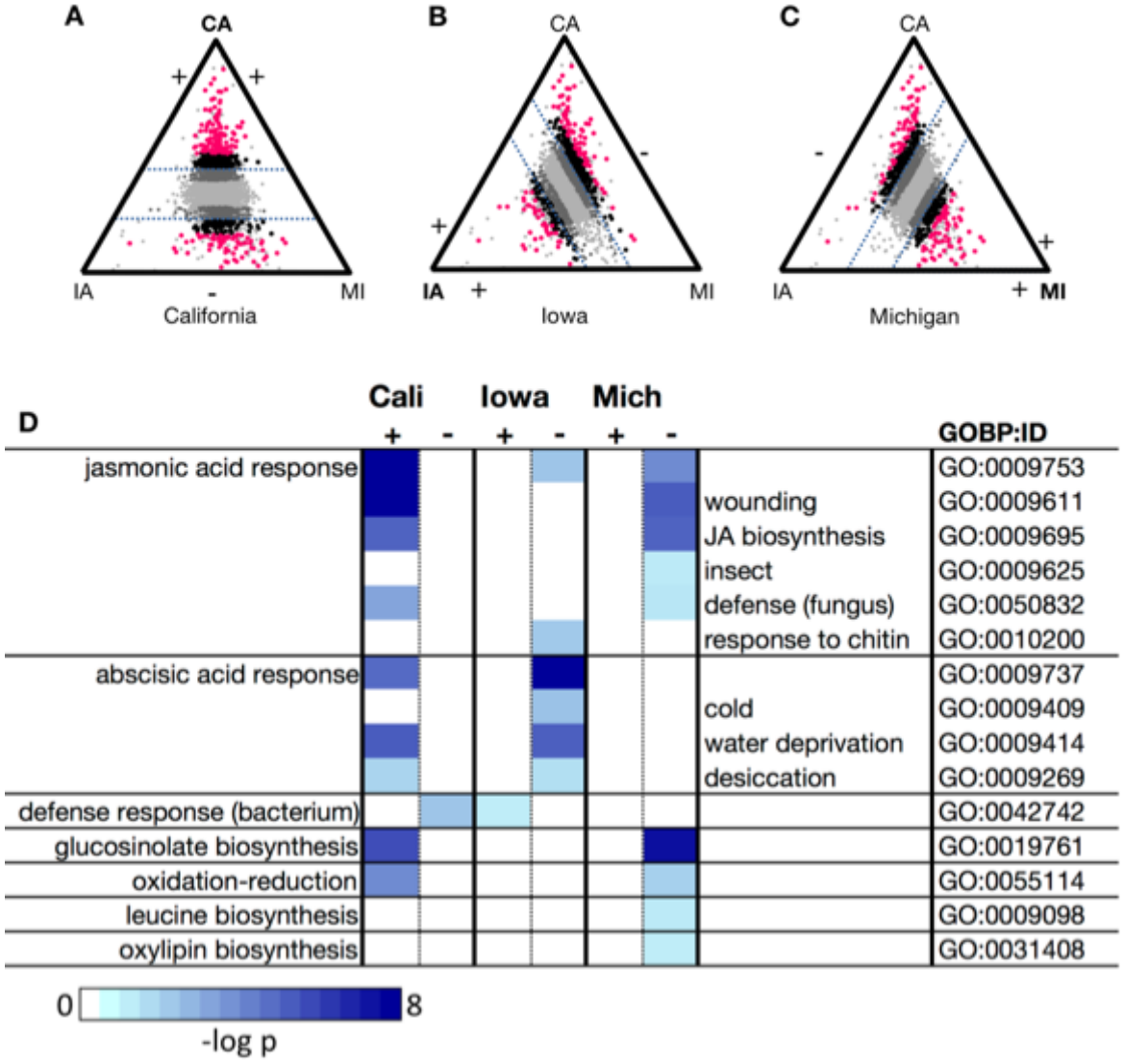


Figure 3.16. Community-specific gene set enrichment analysis, contrasting holoxenic samples by input community. (A-C) Ternary plots where each gene is represented by a symbol, and deviation from the mean RPM expression value is represented by skewing towards the input community for which the expression is relatively higher. The dotted line represents one standard deviation from the mean; the dark gray symbols represent two standard deviations from the mean, and the pink symbols represent statistical significance by 2-way ANOVA. Light gray symbols indicate the p-value for enrichment is $p>0.05$ and/or that the mean RPM for that sample group is within 0.5 standard deviations of total mean RPM among all holoxenic groups. CA=Cali=Arid, Iowa=IA=Prairie, MI=Mich=Ag.

		log2(ax/ho)					log2RPM		
	Gene	Ar	Ag	Pr	all	p	ax	ho	Description
BAK1	IOS1					**			impaired oomycete susceptibility; LRR-RLK
	BIR1					**			BAK1-interacting receptor-like kinase
	FLS2					**			immune receptor for bacterial flagellin
	RBOHD					***			NADPH/respiratory burst oxidase protein
	EFR					*			immune receptor for bacterial EF-TU
	BAK1					**			BRI1-associated receptor kinase
	PEPR1					*			amplifies immune response to pathogens
	BRI1					*			brassinosteroid insensitive
CERK1	BKK					**			BRI1-associated kinase
	CERK1					***			chitin elicitor receptor kinase
	LIK1					**			involved in chitin perception
	LYM1								peptidoglycan sensing and response
	LYM3								peptidoglycan-binding
PTI	FRK1					**			MAPK-specific flg22-induced RLK
	FOX1					*			FAD-linked oxidoreductase
	AT5G39580					*			haem peroxidase
	PER4					*			peroxidase
	YLS9					*			similar to NDR1; senescence-associated
	WAK2					**			wall-associated kinase; SA-inducible

Figure 3.17. Differential and absolute expression of PTI recognition and response genes, coding for proteins of the BAK1 pattern recognition complex, CERK1 pattern recognition complex, and pattern-triggered immune (PTI) response marker genes. Differential expression is defined as $\log_2(RPM_{AX}/RPM_{HO})$; absolute expression is $\log_2(RPM)$. Significance determined by t-test, p-values: * p<0.5, ** p<0.005, *** p<0.0005, **** p<0.00005 (Benjamini-Hochberg corrected). Soil input microbiota abbreviations: Ar=Arid, P=Prairie, Ag=Agricultural.

		log2(ax/ho)					log2RPM		
	Gene	Ar	Ag	Pr	all	p	ax	ho	Description
JA	JAZ1					*			jasmonate-zim-domain protein
	OPR3								required for JA biosynthesis
	JAZ5					*			jasmonate-zim-domain protein
	MYC2								flavonoid biosynth and REDOX response
	MYC4					***			bHLH transcription factor
	JAM2					*			JA associated bHLH transcription factor
	JAZ8								jasmonate-zim-domain protein
	AOS								Involved in JA biosynthesis
	LOX2								lipoxygenase required for JA accumulation
	JAR1								catalyzes the formation of jasmonyl-Ile
	MYC3					*			bHLH transcription factor
	MES16					***			methyl esterase
PR	PR4					*			pathogenesis-related; chitin-binding
	PR2					*			beta 1,3-glucanase
	PR1					*			pathogenesis-related gene; SA-responsive
	PR3								basic chitinase
	PR5					*			pathogenesis-related gene

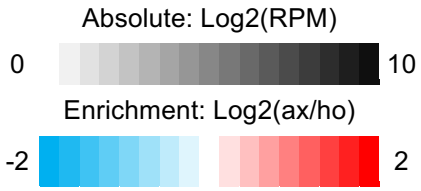


Figure 3.18. Differential and absolute expression of jasmonic acid-associated (JA) and pathogenesis-related (PR) marker genes. Differential expression is defined as log2(RPMAX/RPMHO); absolute expression is log2(RPM). Significance determined by t-test, p-values: * p<0.5, ** p<0.005, *** p<0.0005, **** p<0.00005 (Benjamini-Hochberg corrected). Soil input microbiota abbreviations: Ar=Arid, P=Prairie, Ag=Agricultural.

		log2(ax/ho)					log2RPM		
	Gene	Ar	Ag	Pr	all	p	ax	ho	Description
SA	AZI1					**			azelaic acid induced; involved in SAR
	ANK					*			ankryn protein; SA-inducible
	BCS1					***			amplifies SA signaling
	SYP122					***			syntaxin, neg regulation of SA/JA signaling
	CBP60G					**			MAMP-inducible; involved in SA
	PEN3					***			SA-dependent non-host resistance
	WAK1					*			wall-associated kinase
	SARD1								SAR deficient; SA biosynthesis regulator
	WAK2					**			wall-associated kinase; SA-inducible
	PAD4					*			phytoalexin deficient, lipase; SA signaling
	NPR1					***			key regulator of SA-dep SAR
	TGA5					*			NPR1-interacting B-ZIP protein
	EDS1								enhanced disease susceptibility
	BSMT1								methylates salicylic acid and benzoic acid
	PBS3								AvrPphB susceptible; auxin-responsive
	ICS1								isochorismate synthase; SA accumulation
	CPR1								constitutive PR genes, neg reg of defense
	UGT75B1								UDP-glucosyltransferase; callose
	EPS1								pathogen-induced SA biosynthesis
	PAL3					*			phenyl alanine ammonia-lyase

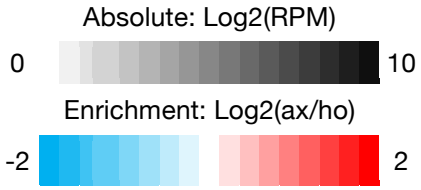


Figure 3.19. Differential and absolute expression of salicylic acid-associated genes. Differential expression is defined as log2(RPM_{MAX}/RPM_{HO}); absolute expression is log2(RPM). Significance determined by t-test, p-values: * p<0.5, ** p<0.005, *** p<0.0005, **** p<0.00005 (Benjamini-Hochberg corrected). Soil input microbiota abbreviations: Ar=Arid, P=Prairie, Ag=Agricultural.

	Gene	log2(ax/ho) Ar	Ag	Pr	all	p	log2RPM ax	ho	Name
WRKY	WRKY48					***			+ reg of defense against *Pst*
TFs	WRKY18					**			+ reg of ETI; complex with WRKY40/60
	WRKY30					***			
	WRKY60					**			pathogen-responsive
	WRKY53					***			+ reg of senescence and defense
	WRKY47					***			
	WRKY46					***			WRKY53-mediated defense against *Pst*
	WRKY6					***			
	WRKY8					**			
	WRKY33					**			- reg of SA signaling
	WRKY58					*			
	WRKY25					**			
	WRKY15					***			
	WRKY40					*			+ reg of ETI; confers *Pst* defense
	WRKY54								
	WRKY17					*			- reg of *Pst* resistance
	WRKY3					*			
	WRKY4					**			
	WRKY7								- reg of *Pst* resistance
	WRKY22					*			mediates dark-induced leaf senescence
	WRKY45								
	WRKY69								
	WRKY29								
	WRKY51								JA responsive
	WRKY28								interacts with TCP20 to regulate SA
	WRKY21								
	WRKY70								- reg of defense

Figure 3.20. Differential and absolute expression of WRKY transcription factors. Differential expression is defined as log2(RPMAX/RPMHO); absolute expression is log2(RPM). Significance determined by two-tailed t-test, p-values: * p<0.5, ** p<0.005, *** p<0.0005, **** p<0.00005 (Benjamini-Hochberg corrected).

Basal metabolite analysis (GC/MS)

To further characterize the basal state of the axenic plants and contrast with the basal state of holoxenic plants grown in holoxenic substrate (with the Ag, Arid or Prairie microbiota), we performed metabolic profiling in collaboration with Oak Ridge National Lab. Relative to all holoxenic communities, phenyl-B-D-glucopyranoside was ~100-fold more abundant in axenic plants (Figure 21, Supplemental Table). Four additional metabolites were significantly more abundant in axenic plants, two of which were identified as sucrose and the polyamine putrescine, and the other two were not fully identified beyond classification as a feruloyl glycoside and an N-containing metabolite, respectively (Figure 21). Based on the GC/MS spectral peak library of Oak Ridge National Lab, 93 metabolites were above the lower threshold of detection. Each metabolite was quantified relative to a sorbitol standard, and differential abundance of each metabolite was assessed by comparing community-specific holoxenic abundance to the corresponding axenic sample in triplicate and tested for significance using a t-test. The number of differentially abundant metabolites is community-specific: 18/93 for Arid (n=14 axenic-enriched), 25/93 for Ag (n=20 axenic-enriched), and 39/93 for Prairie (n=8 axenic-enriched). Ascorbic acid is significantly enriched in Ag and Prairie axenic, but not in Arid axenic. The compound “13.79 204 231 glycoside” is significantly depleted in axenic plants relative to holoxenic-Arid and holoxenic Prairie holoxenic, but not for holoxenic-Ag. Only two metabolites were consistently lacking in axenic plants relative to all holoxenic plants: citric acid and 2-ethylhexanoic acid. Based on our transcriptome data, we hypothesized that azelaic acid and salicylic acid would be reduced basal levels in axenic plants. However, azelaic acid differential

abundance was either not significant (Arid, Prairie), or slightly more abundant in axenic (Ag: axenic 9.4±1.0, holoxenic 5.6±0.5 $\mu g \cdot g^{-1}$ FW).

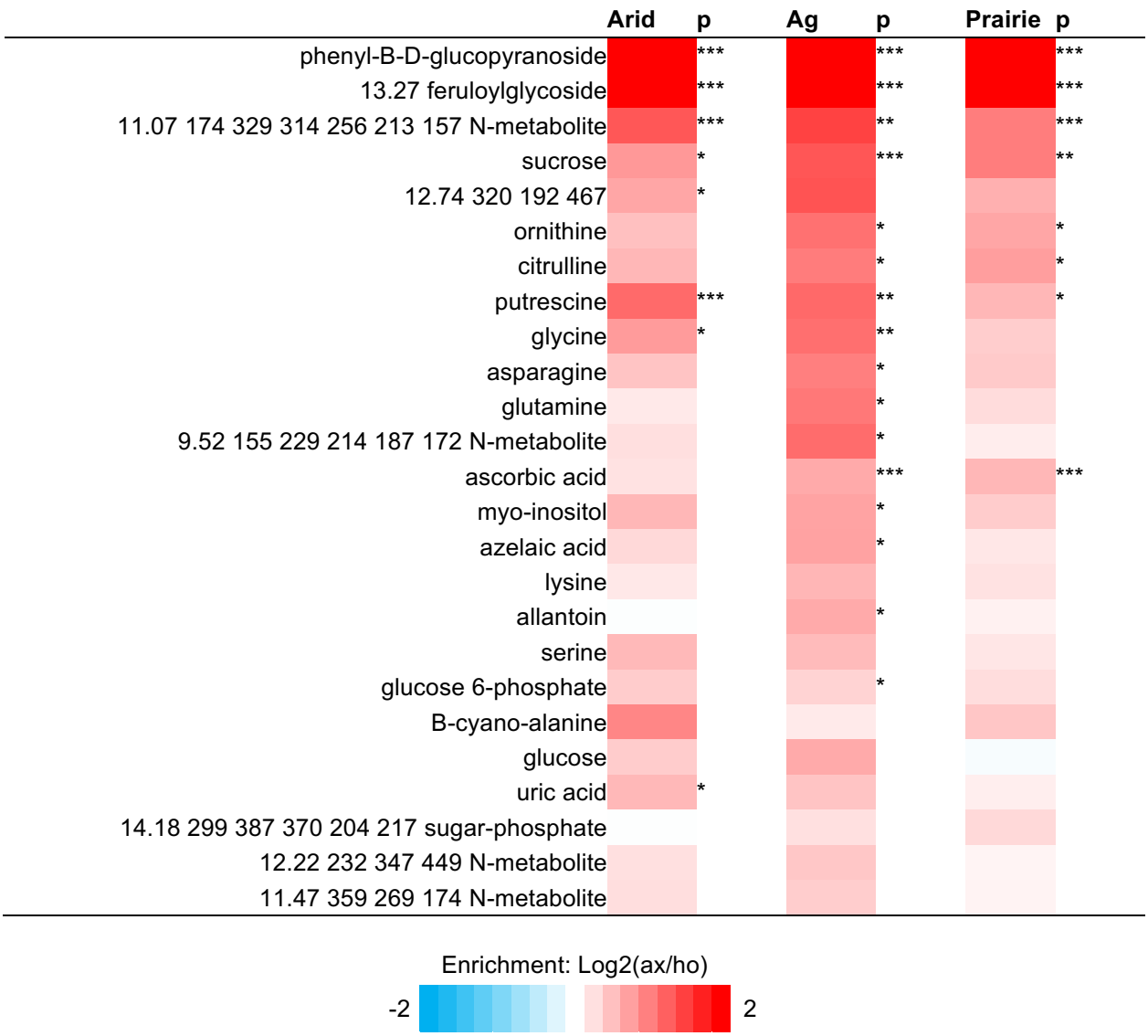


Figure 3.21. Relative abundance of metabolites extracted from whole rosette tissue and quantified by GC/MS (against a sorbitol standard). Differential metabolite abundance is defined as log2(Axenic/Holoxenic) in µg·g-1 FW; corresponding raw concentrations are in Table S3-2. Significance determined by t-test, p-values: * p<0.5, ** p<0.005, *** p<0.0005, **** p<0.00005. (continued on next page [1/4])

Figure 3.21 (cont'd)

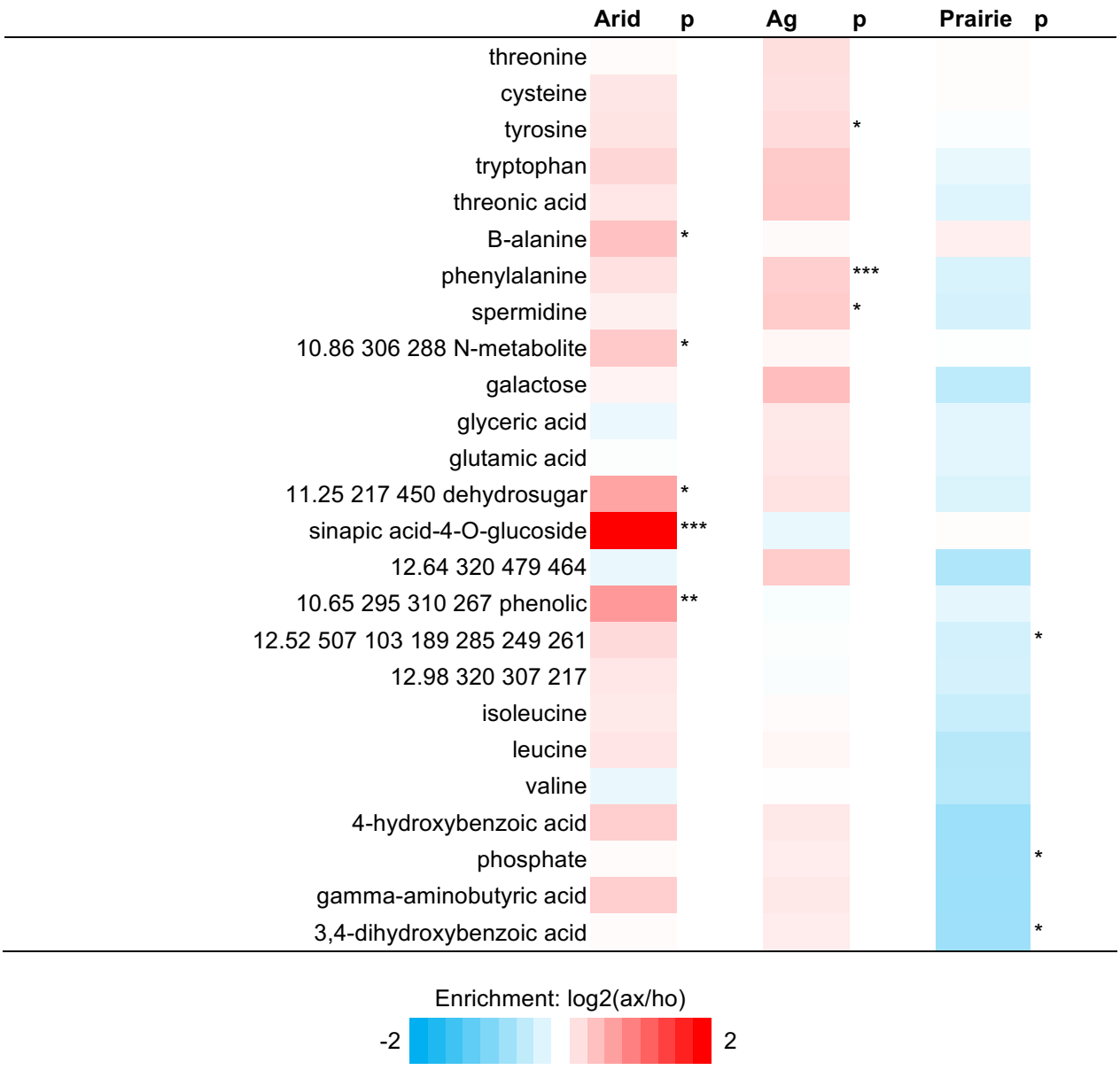


Figure 3.21 (cont'd)

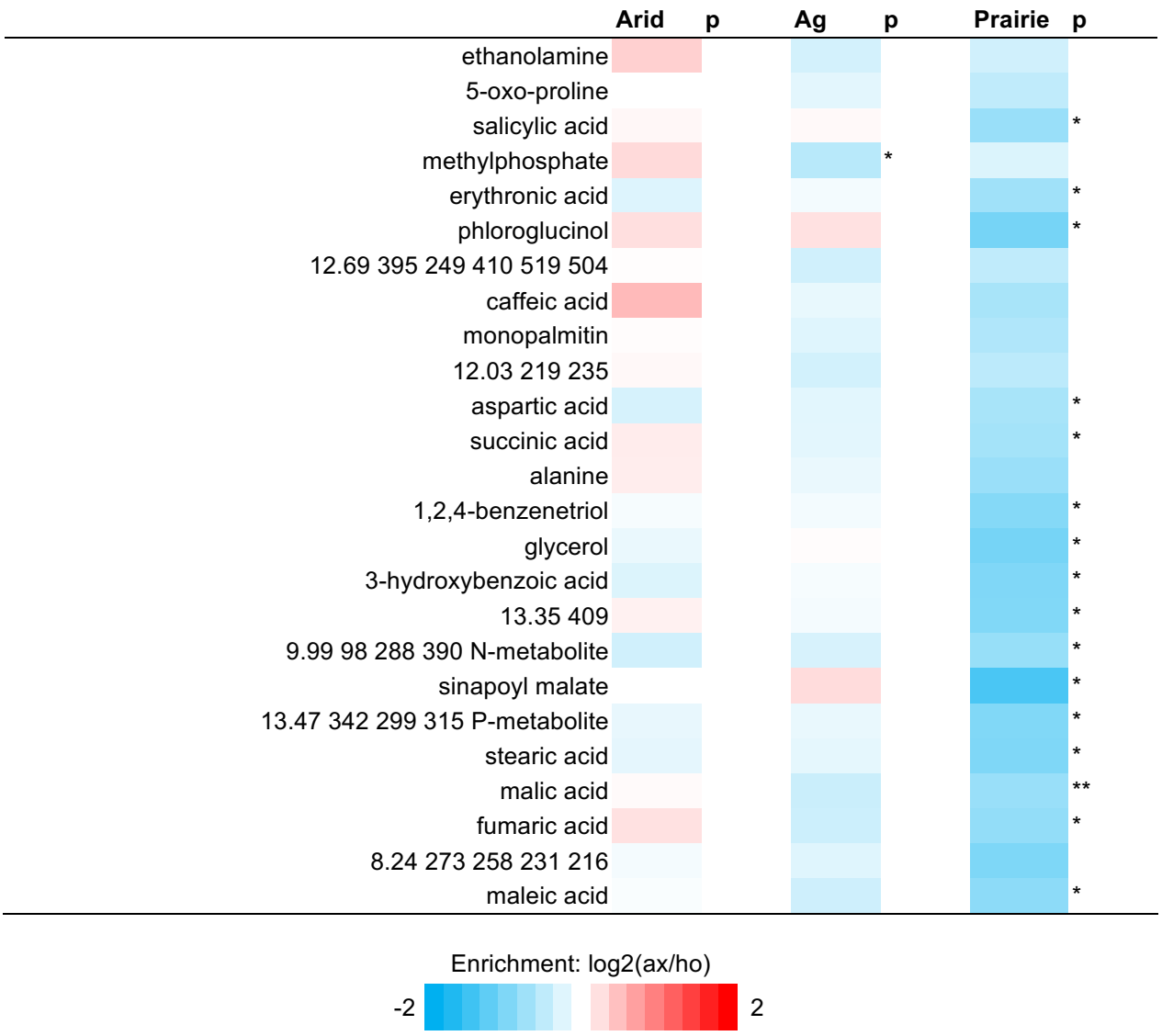
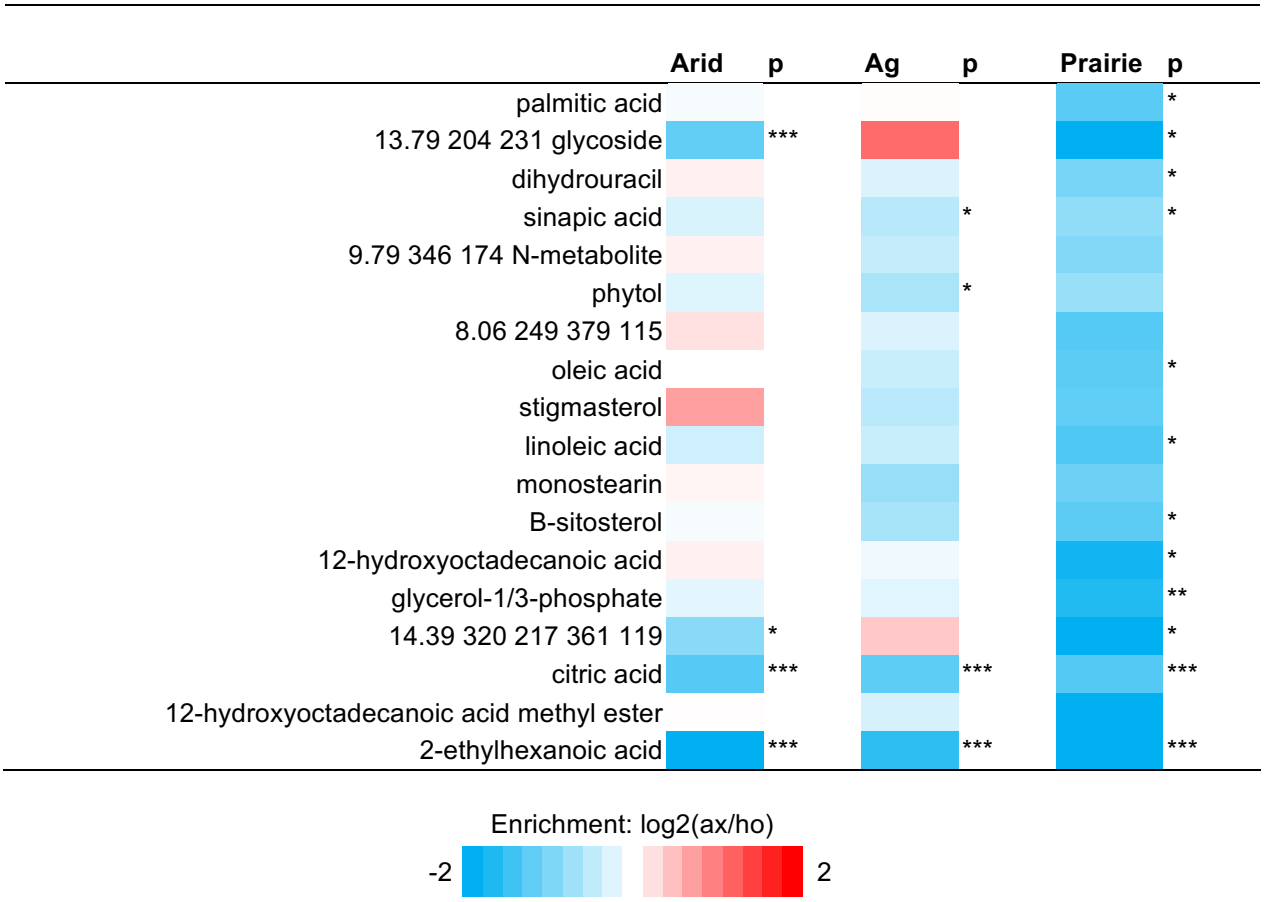


Figure 3.21 (cont'd)



Basal: Salicylic acid quantification (UPLC/MS)

To further characterize the basal status of axenic and holoxenic Arabidopsis, we followed the clues that there may be defense associated phenotypic differences. Although GC/MS did not reveal a significant different in salicylic acid accumulation, we elected to quantify using a more sensitive UPLC approach with a salicylic acid standard. Additionally, we quantified the glucosylated form of salicylic acid (SA-glucosyl). Relative to Arabidopsis grown in any holoxenic substrate tested, corresponding axenic plants were deficient in basal SA levels (Figure 22). Although rosettes from plants grown in holoxenic Prairie substrate had similar levels of SA-glucosyl to plants grown in corresponding axenic substrate, holoxenic Prairie rosettes had the highest relative and absolute level of free SA (Figure 22).

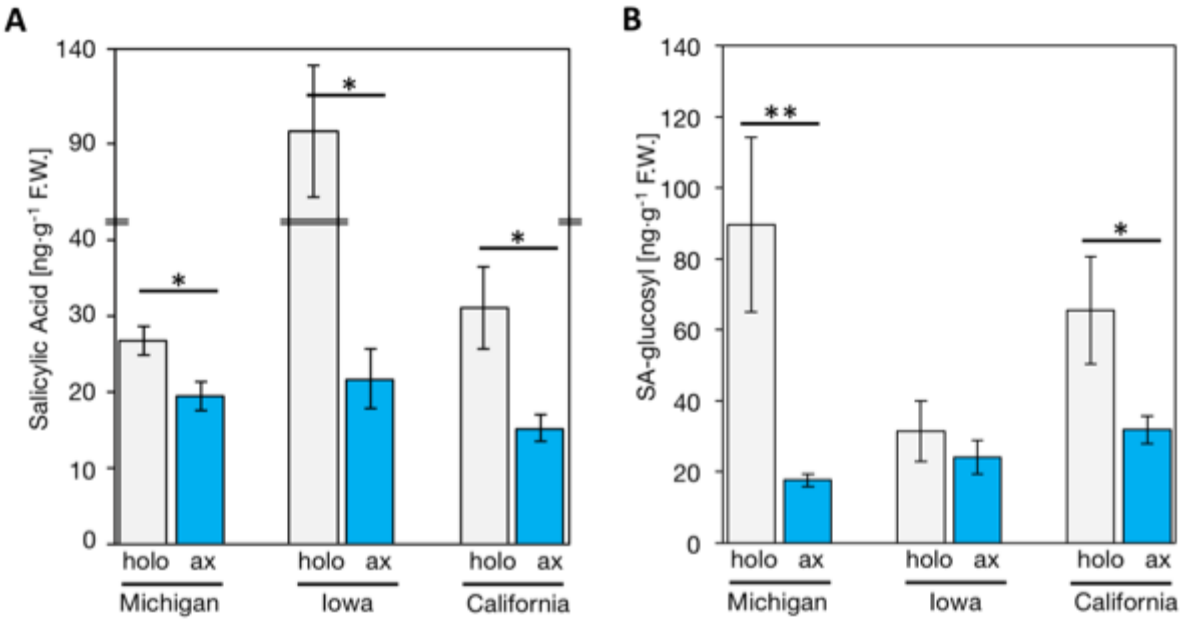


Figure 3.22. Defense hormone salicylic acid quantification as (A) free SA and (B) SA-glucosyl. Quantifications were performed with eight biological replicates per condition, and with three experimental replicates all showing the same trends. Significance determined by two-tailed t-tests; (Michigan=Ag, Iowa=Prairie, California=Arid);
* $p<0.05$, ** $p<0.005$;

Transcriptional response to immune elicitation is defective in axenic plants

Based on abnormally low levels of defense-associated gene expression and reduced basal salicylic acid levels, we hypothesized that axenic plants may be immunodeficient. To further interrogate immune-associated phenotypes of axenic Arabidopsis, we performed a series of immune elicitation experiments using a fragment of bacterial flagellin protein "flg22" (Felix et al. 1999). Initially, comparatively analyzed holoxenic and axenic flg22-inducible expression of the following immune-associated genes were quantified by qRT-PCR: *ICS1* (*ISOCHORISMATE SYNTHASE 1*), *FLS2* (*FLAGELLIN-SENSITIVE 2*), and *FRK1* (*FLG22-induced receptor-like kinase 1*). Although the induction of the expression of all three genes was detected in holoxenic and axenic plants, the absolute expression levels of all three genes was higher in holoxenic, except for *FLS2* in Prairie(Iowa) holoxenic, which was not significantly different from the corresponding axenic.

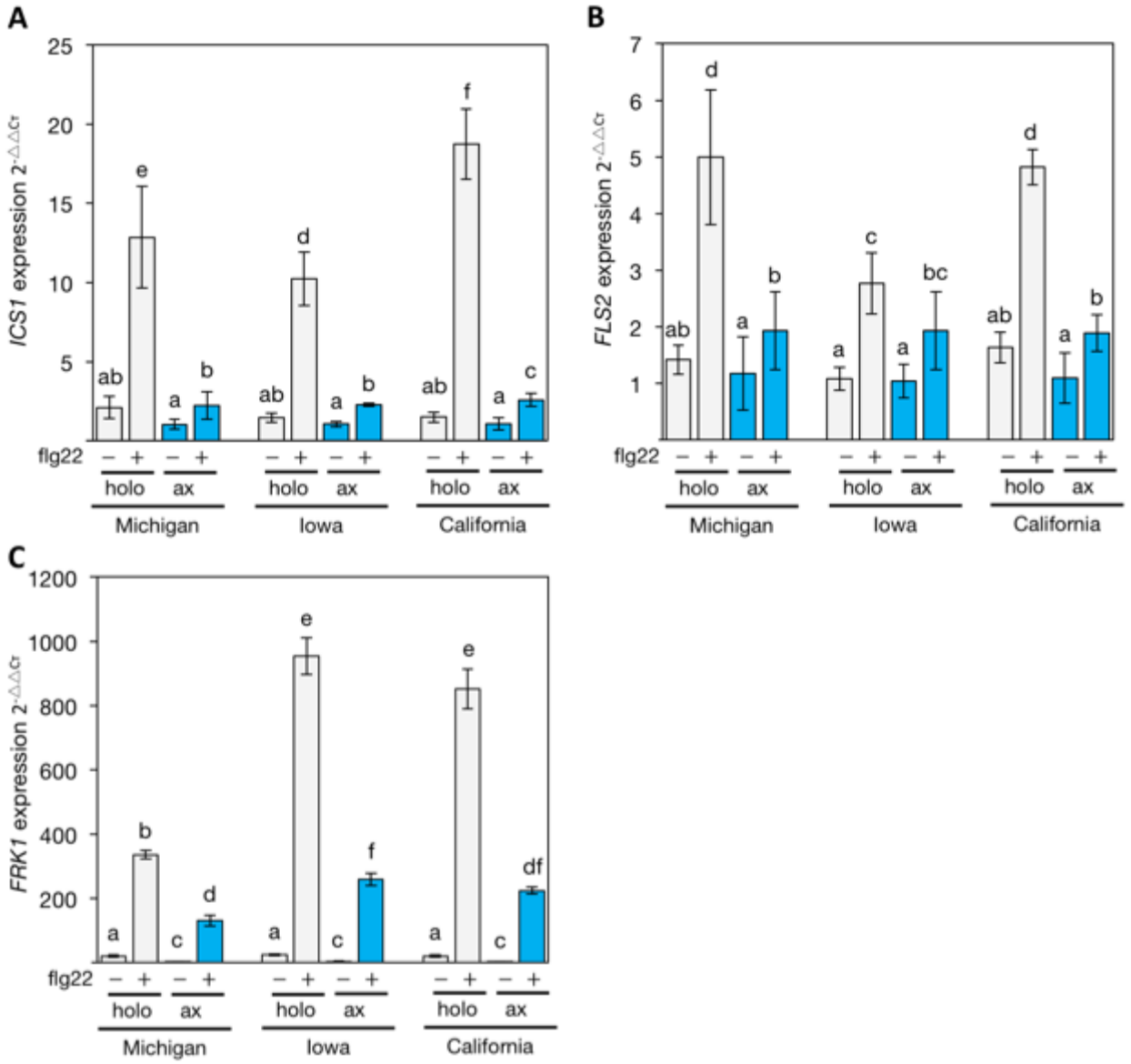


Figure 3.23. Expression of immunity and defense associated genes 2 hours after treatment with flg22. (Michigan=Ag, Iowa=Prairie, California=Arid). Significance testing by one-way ANOVA with post-hoc Tukey-Kramer test.

Compromised axenic posttranslational immune response

An early event that occurs during PTI after immune elicitation by flg22 is phosphorylation of mitogen activated protein kinase (MAPK) proteins (Ligterink et al. 1997). First, we interrogated whether axenic and holoxenic plants had equivalent basal levels of MPK3/6 and then whether flg22-induced phosphorylation of MAPK3/6 differentially occurs in axenic plants. The Axenic abundance of both MPK3 and MPK6 are equivalent to Ag-holoxenic in a basal state (Figure 24B). However, upon flg22 elicitation, axenic plants are partially compromised in phosphorylation of MPK3/6 relative to Ag-holoxenic (Figure 24A). At a high concentration of flg22 [1000 nM], both axenic and holoxenic plants underwent MPK3/6 phosphorylation, but holoxenic plants had a greater relative abundance of p-MPK3/6 (Figure 24A). At lower concentrations of flg22 [100nM], phosphorylated MPK3/6 was not consistently detected in axenic plants (Figure 24A).

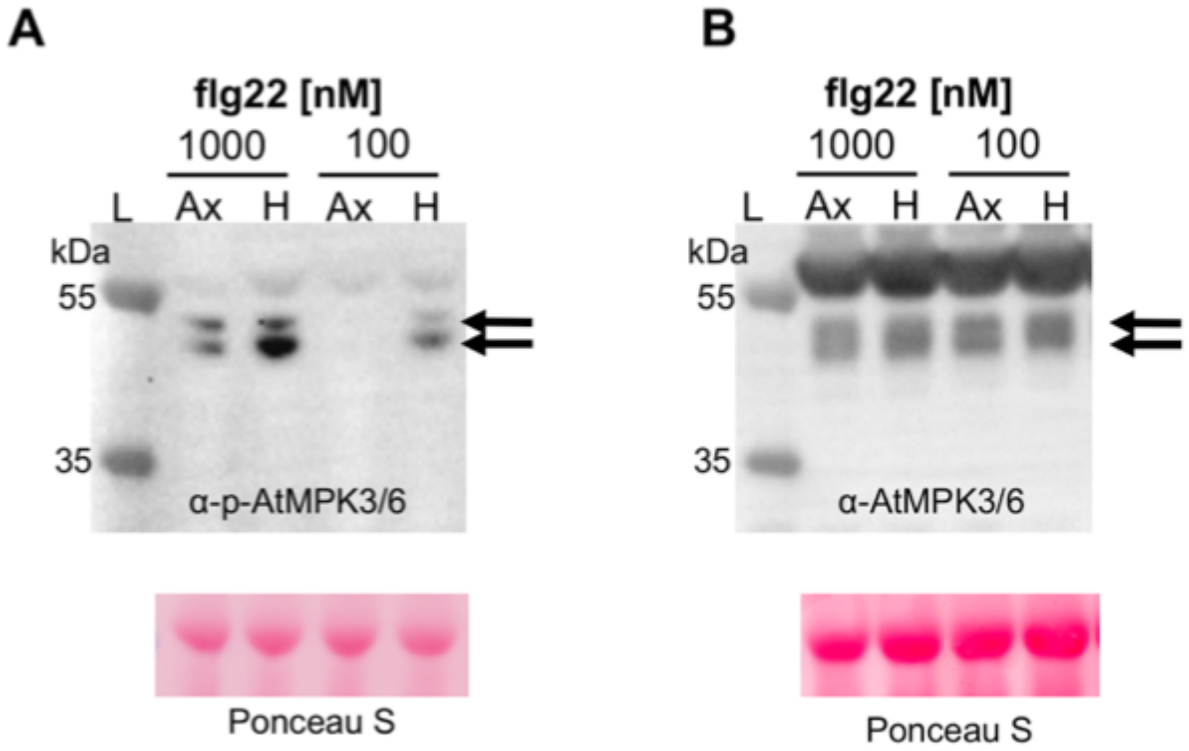


Figure 3.24. MAPK protein detection and quantification by Western blot of (A) Mitogen-Activated Protein Kinase 3/6 (MPK3/6), and (B) phosphorylated Mitogen-Activated Protein Kinase 3/6 (p-MPK3/6) after treatment with 1000nM or 100nM flg22. The visible band in the Ponceau stained gels are from the RuBiSCO LSU. Holoxenic refers to Ag soil inocula.

Oxidative burst and callose deposition are compromised in axenic plants

Foliar treatment of Arabidopsis with flg22 elicits an immune response also characterized by reinforcement of cell walls with callose and a rapid apoplastic “burst” of reactive oxygen species (ROS) (Dixon et al. 1994). Although flg22-induced callose depositions were observed in axenic leaves, they tended to be concentrated in clusters, in contrast to homogenous deposition patterns in holoxenic leaves (Figure 25A). Collectively, axenic Arabidopsis had fewer callose depositions per unit area than holoxenic. Furthermore, a luminol ROS quantification assay revealed that axenic plants are partially compromised in oxidative burst after elicitation (Figure 25D).

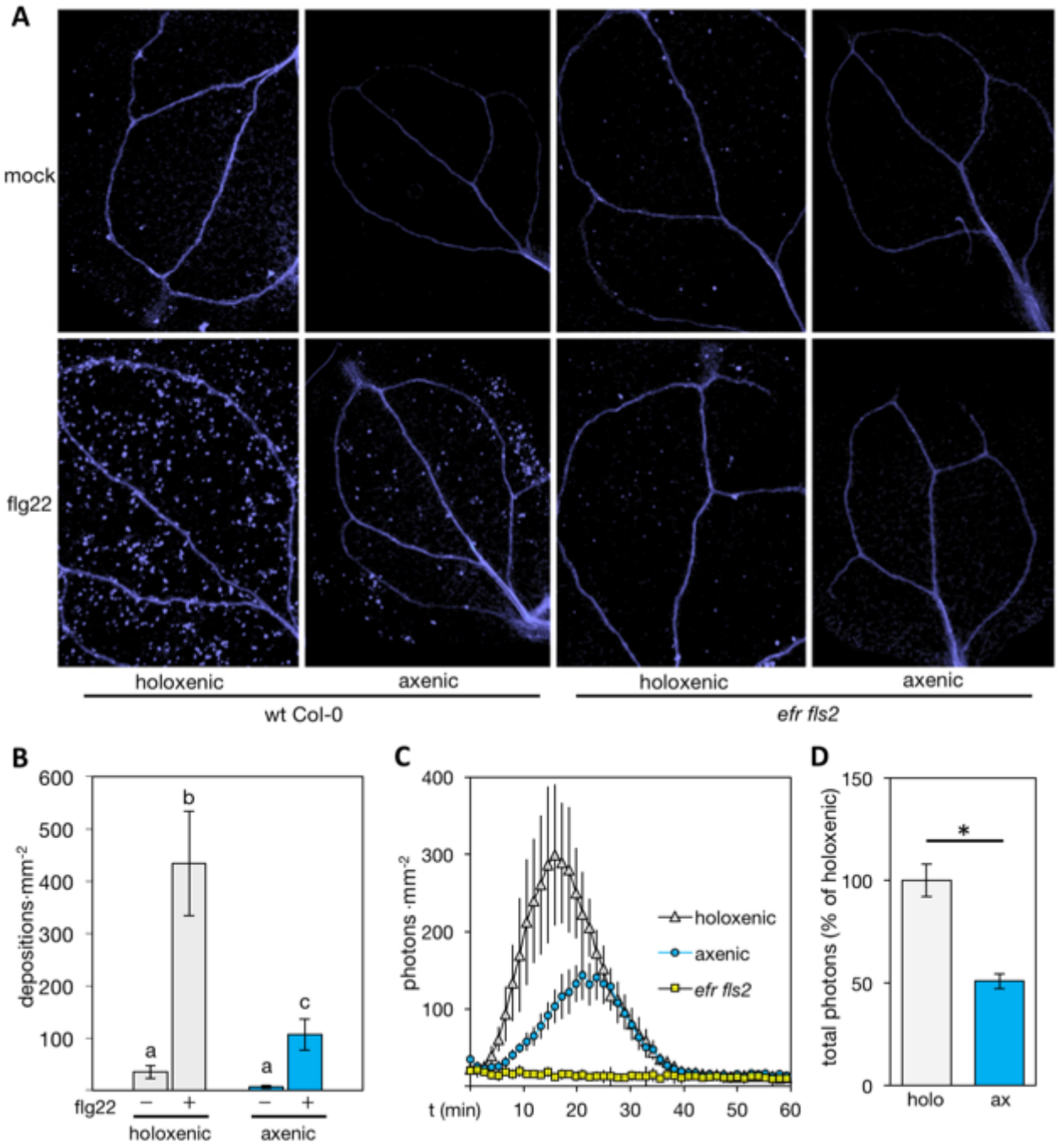


Figure 3.25. Response to immune elicitation by flg22 [500 nM] quantified by (A-B) callose deposition counts, and oxidative burst of reactive oxygen species (C-D). Significance testing by one-way ANOVA with post-hoc Tukey-Kramer test or t-test, * 0.05 > p.

Defective priming and defense against pathogens in axenic plants

Based on numerous indications of defective innate immunity, we hypothesized that axenic Arabidopsis would have decreased capacity to defend pathogen invasion. Indeed, the foliar pathogen *Pseudomonas syringae* pv. tomato DC3000 (*Pst*) was able to grow to a significantly higher population density in axenic plants two days after vacuum infiltration (with 5 x 105 CFUmg-2) relative to holoxenic plants with Ag, Arid or Prairie microbiota (Figure 26A). Notably, among the three holoxenic plants, *Pst* had a significantly lower population in Arabidopsis with Arid microbiota relative to the Ag or Prairie microbiota. Foliar disease symptoms, particularly chlorosis and tissue collapse, were more pronounced in axenic infected plants (Figure 26C). We further tested whether the salicylic acid analog benzothiadiazole (BTH) would prove to be equally efficacious at priming the defenses of axenic and holoxenic plants (Görlach et al. 1996). BTH was efficacious in both axenic and holoxenic plants in regard to reducing the pathogenicity of *Pst*, but the priming was more effective in holoxenic plants (Figure 26B). Notably, BTH pretreatment of axenic plants and mock pretreatment of holoxenic plants resulted in equivalent susceptibility to *Pst*. To interrogate whether indeed innate immunity could explain axenic susceptibility, we performed diseases assays using *Pst* Δ*hrcC*, which has a defective type III secretion system and is unable to deliver immune-suppressing effectors to inhibit host innate immunity. Likewise, we performed vacuum infiltration disease assays using used a PTI-compromised Arabidopsis triple mutant *bak1 bkk cerk1* (*bbc* hereinafter) with wild type *Pst* and *Pst* Δ*hrcC* (Figure 26C). Consistent with wild type Arabidopsis (Col-0), the *bbc* mutant grown in holoxenic

condition was ~10-fold resistant to *Pst*, compared to that grown in axenic condition, and there was not a significant difference between WT and *bbc* mutant Arabidopsis based on this vacuum infiltration assay. However, the magnitude of axenic susceptibility to *Pst* Δ*hrcC* in the *bbc* mutant background was significantly greater (Figure 26C).

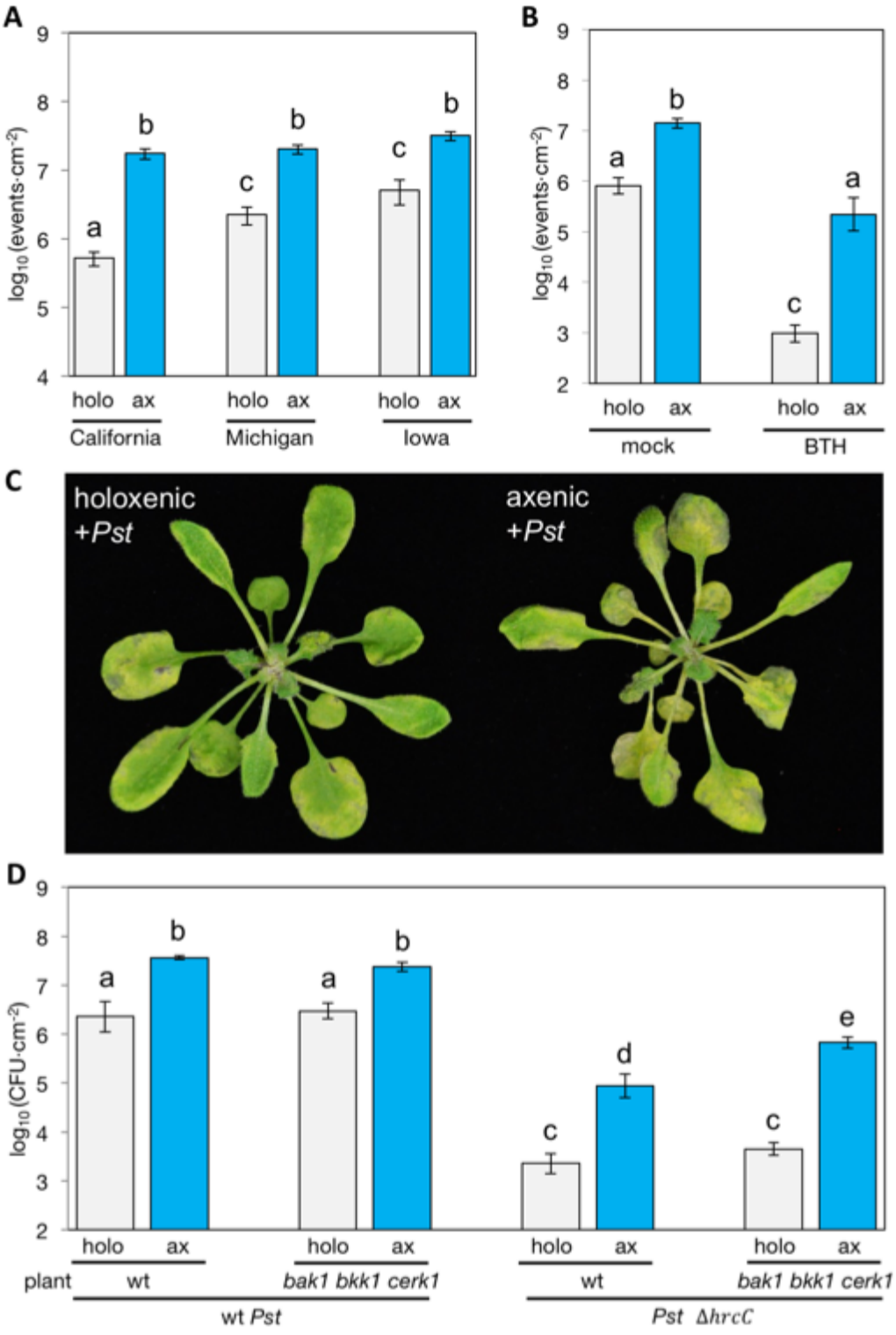


Figure 3.26. Pathogen susceptibility of axenic and holoxenic Arabidopsis to (A) *Pseudomonas syringae* pv tomato DC3000 measured as colony forming units per area at two days post inoculation. (B) BTH protection assay using *Pst* DC3000 in wild-type Arabidopsis (Col-0), two days post infection and three days post BTH pretreatment. (C) Disease symptoms

Figure 3.26. (cont'd)
in five-week-old Arabidopsis, three days after infection with Pst DC3000. (D) Susceptibility of WT Arabidopsis (Col-0) and the *bak1 bkk cerk1* (*bbc*) mutant to *Pst* and *Pst* Δ*hrcC*. Significance testing by one-way ANOVA with post-hoc Tukey-Kramer test. (Michigan=Ag, Iowa=Prairie, California=Arid);

Discussion

Here, we report that soil-grown axenic Arabidopsis has numerous dysfunctions in innate immunity. In this study, our initial objective was to perform an in-depth basal characterization of axenic Arabidopsis relative to holoxenic Arabidopsis with three distinct microbiomes, and our findings inspired us to further characterize the microbiome effect on plant innate immunity. All experiments herein were conducted in the homogenous environment of the FlowPot system under nutrient-replete conditions, enabling us to minimize abiotic stress and look beyond the nutritional services provided by microbiota to identify subtle but significant differences in the basal transcriptome and metabolome. In our hands, the presence/absence of soil microbiota in a non-diseased state results in significant differential expression of merely ~0.2% of Arabidopsis protein-coding genes, based on standard criteria (log2 fold change >2, p-value < 0.05, minimum expression > 5 RPM). However, at higher sensitivity, we discovered axenic plants are defective in the basal expression of a significantly disproportionate amount of defense and immunity-associated genes, including genes associated with the defense hormone Specifically, genes associated with the defense hormone salicylic acid (SA), pattern-triggered immunity (PTI) and systemic acquired resistance (SAR). Targeted phytohormone quantification validated that indeed, axenic plants have reduced levels of SA. Three microbiomes from geographically, physically and chemically distinct soils were used as holoxenic input microbiota, enabling us to identify and report a repertoire of “core” microbiome-modulated plant genes, as well as community-specific influence over the host transcriptome. In-depth basal characterization of axenic and holoxenic Arabidopsis also revealed numerous

differentially abundant transcripts and metabolites that have not been previously reported to be related to biotic interactions. Empirical investigation into axenic immunodeficiency revealed that recognition and/or response to elicitation by flg22, a common bacterial molecular pattern, is less robust. We also report that axenic Arabidopsis is significantly compromised in the ability to defend against a bacterial foliar pathogen. Likewise, protection against infection via pre-activation of SA signaling is less efficacious in axenic plants. Foliar disease assays with a PTI-compromised Arabidopsis *bbc* mutant and virulence-compromised *Pst Δhrc* mutant definitively revealed that immunodeficiency is not sufficient to explain increased axenic susceptibility to infection, but other mechanisms such as microbe-microbe antagonism and niche exclusion are also likely involved. The absence or drastic reduction in basal expression of multiple PTI marker genes in axenic plants relative to all three holoxenic plants include such markers as: *FRK1, FOX1, AT5G39580, PER4, YLS9, and WAK2*. This is not unexpected, because an axenic plant is naive to MAMPs, whereas holoxenic plants constitutively encounter low-level amounts of MAMPs. A low-level activation of PTI may coincide with other defense phenotypes observed in axenic plants. For example, the lower expression of SA-responsive gene expression, including *AZI1, PEN3, PAD4, ACD6* and *NPR1*, are indicative of a compromised SA signaling in axenic plants. There is also a subtle down-regulation of *Pathogenesis Related 1* (*PR1*), a very widely used SA marker (Tsuda et al., 2013), in axenic plants. This was further validated by the finding of a reduced basal concentration of SA in axenic plants. Interestingly, some well-characterized pattern recognition receptors (PRRs) or co-receptors (*IOS1, BIR1, FLS2, EFR, BAK1, PEPR1, BRI1, BKK, CERK1, LIK1*) as well as

several uncharacterized LRR-RLK proteins (*AT1G51890, AT1G51790, AT1G51850*), and *AT4G14370* (TIR-NBS-LRR) are also expressed at a lower basal level than in holoxenic plants. The protein BAK1 is of particular significance because it is a coreceptor of multiple PRRs (e.g., EFR and FLS2). Also interesting is the finding of numerous axenic-downregulated genes are specifically implicated as important for SAR, including *ACD6* and *NPR1*. Based on gene ontology (GO:BP) analysis, "Defense Response" to "Biotic Stimulus", "Response to Bacterium", "Response to Fungi", and "Response to Stress" comprise the core down-regulated categories in axenic plants. Another striking example of depleted gene expression in axenic plants is the family of *WRKY* transcription factor genes. WRKY transcription factors are often associated with biotic interactions and stress response in Arabidopsis (Birkenbihl et al., 2017). Consistent with a previous report that WRKY70 is more abundant in axenic plants (Zhou et al., 2017), this gene was not found in our list of genes with repressed expression in axenic plants.

Three distinct microbiomes, intentionally selected from geographically and physiochemically distinct soils, were chosen for this study to represent diverse holoxenic reconstitutions. As a result, we were able to identify community-specific effects. For example the Arid (California) microbiota conferred a significant relative enrichment of gene expression associated with abscisic acid (ABA). Interestingly, ABA response is important for drought tolerance and desiccation stress, and given the climate in which "Arid" microbiota were collected, it is an attractive hypothesis that microbiota from arid regions may confer drought tolerance. Additionally, JA response was significantly greater in the California holoxenics relative to the Prairie and

Agricultural holoxenics. Furthermore, Arid holoxenic plants had a greater relative enrichment of glucosinolate biosynthesis genes expressed. Glucosinolates are required for innate immune response in Arabidopsis (Clay et al., 2009). Although Ag and Prairie microbiota conferred enrichment in JA subcategories (GO:BP 'defense response to fungus' and GO:BP 'response to fungus'), the overarching JA ontology category was only significantly enriched for in Arid-holoxenic. The Agricultural soil conferred higher relative expression of secondary metabolism associated genes, and the Prairie microbiota conferred the greatest relative enrichment of 'defense response to bacteria' genes. In this study, we also identified microbiome community-dependent effects on the Arabidopsis metabolome. Relative to axenic plants, as well as plants with agricultural (Ag) or arid microbiota, Arabidopsis with prairie (Pr) microbiota have significantly higher levels of stearic acid, malic acid, fumaric acid, maleic acid, palmitic acid, sinapic acid, phytol, oleic acid, and linoleic acid. Prairie microbiota plants also have significantly higher levels of glycerol 1/3-phosphate (G3P), which has been reported as a defense-priming compound (Conrath et al., 2015; Kachroo and Robin, 2013). The community composition of Arabidopsis with Prairie, Ag, and Arid microbiota is likely responsible for the differential phenotypes of holoxenic plants, but specifically which microbiota are responsible is unknown. The Arabidopsis with Prairie microbiota is dominated by the following bacteria, for some of which there are examples of isolates that elicit biocontrol and/or prime defenses: *Chryseobacterium (Domenech et al., 2006), Pseudomonas (Pieterse et al., 2014), Achromobacter,* and *Massilia*. Less abundant, but consistently present microbiota in Arabidopsis with Prairie microbiota are *Enterobacter*, *Erwinia, Pantoea* and *Herbaspirillum*. Compared to other holoxenics,

Arabidopsis with Prairie microbiota also had the greatest abundance of rosette-associated fungal reads relative to chloroplast reads, dominated by four OTUs classified as two genera: *Mucor* and *Penicillium*. Potentially, one or more members of these genera have the ability to further prime defenses in Arabidopsis (Marques et al., 2006; Thuerig et al., 2005). Likewise, it could be rhizosphere-associated microbiota or substrate-colonizing microbiota that emit volatiles that can prime the defenses of the plant (Lee et al., 2012). It is also possible that the increased resistance of Arid holoxenic is entirely attributed to microbe-microbe antagonism.

This study also provides a reference data for the identification of genes relevant to plant-microbiome interactions. For example, the expression of multiple genes encoding FAD-binding berberines are depleted in axenic Arabidopsis. FAD-binding berberines are involved in cell wall modification and REDOX reactions, but not biotic interactions *per se* (Daniel et al., 2015). Likewise, the expression of multiple protease inhibitor/seed storage/lipid transfer protein-coding genes are depleted in axenic plants without any experimental evidence to suggest a function. It would be interesting to further investigate the role of these genes in plant/microbiome interactions.

Arabidopsis has can recognize a 22 AA peptide (flg22) of bacterial flagellin via ligand binding by the transmembrane leucine-rich repeat receptor-like kinase protein FLS2. Upon flg22 binding, FLS2 autophosphorylation and activation of several branches of PTI responses ensue. An early detectable response (<15 mins) is apoplastic burst of reactive oxygen species (ROS burst), mediated by the association of the NADPH oxidase RboHD with the FLS2 receptor complex (Couto and Zipfel, 2016). Later responses for which axenic Arabidopsis is impaired include transcriptional responses (SA and PTI marker genes), post-translational response (MAPK phosphorylation), early REDOX response (apoplastic oxidative burst), and later cell wall reinforcement (callose deposition). We speculate that axenic Arabidopsis may be compromised in PTI responses not due to lack receptors or recognition problems, as justified by an inability to detect differential abundance of FLS2 (Brad Paasch, personal communication). Rather, we speculate there could be a basal REDOX homeostasis issue or simply deficient levels of basal SA could explain axenic supe-susceptibility of axenic plants. REDOX homeostasis has been attributed as an important element for stress tolerance in Arabidopsis, and differential expression of sulfur-containing protein-coding genes and glutathione-related genes could induce a higher redox status in holoxenic plants (Dutilleul et al., 2003), a possibility that needs further investigation.

It is possible that factors in addition to compromised immunity contribute to increased axenic susceptibility. These factors could include but not limited to, defects in PTI, lack of microbe-microbe antagonism, and the increased availability of nutrient sources in the axenic phyllosphere such as sucrose. Metabolic profiling revealed that axenic plants have significantly higher levels of phenyl-B-D-glucopyranoside, sucrose, putrescine (and trending increased abundances of other polyamines), feruloylglycoside. and ascorbic acid. The underlying basis for this is unknown, but warrants further investigation. Likewise, axenic plants a significant reduction in basal levels of citric acid and 2-ethylhexanoic acid, and trending reduced abundance of linoleic acid and glycerol-1/3/phosphate. Citric acid is potentially a REDOX sink, and G3P has been previously implicated in defense, but 2-ethylhexanoic acid is puzzling and the biosynthetic pathway for this compound in Arabidopsis is unclear, although there are reports of its priming potential.

Complex multicellular life is inevitably intertwined with the microbial world. While some host-microbe associations are mutualistic, many microorganisms have evolved mechanisms to exploit eukaryotes for nutrient resources at the expense of the host (i.e., parasitism). Presumably, the cost of parasitism and benefits of mutualism prompt the selection for defense adaptations, ultimately giving rise to the evolution of intricate signaling networks collectively termed *innate immunity*. But despite the fact that all plants and animals have evolved in concert with microbiota, very little is known about host- microbiome dynamics and the functional impact the microbiome has on host phenotypes, particularly in plants. The innate immune system of plants is essential for defense against pathogens; a sessile lifestyle rooted in soil bestows an inability to escape stress in a milieu of potential pathogens. The molecular underpinnings of plant innate immunity have traditionally been studied using model pathosystems (i.e., a well-characterized plant such as *Arabidopsis thaliana* treated with an artificially high concentration of microorganism). Although this approach has built a fantastic framework for understanding plant defense signaling, technological limitations have hindered our ability to study plant immune response to the indigenous *microbiome*, thus leaving a fundamental gap in our understanding of the evolution and dynamic interplays between plants, indigenous microbiota and pathogens. For future directions, I believe that the FlowPot gnotobiotic system and the dataset obtained in this study will provide the foundation for future inquiry into a better mechanistic understanding of functional impact of the plant microbiome on plant-microbe interactions and other phenotypic aspects of plant biology.

Table 3.9. Expressed genes (EG)

Sample grouping	n	%total
Agricultural (Ag)	15162	55.25
Ag+Pr	14794	53.90
Prairie (Pr)	15151	55.20
Arid+Pr	14794	53.90
Arid	15214	55.43
Arid+Ag	14758	53.77
Arid+Ag+Pr (core)	14597	53.19
Pan (holo pooled, ax pooled)	15050	54.84

Criteria: mean RPM ≥ 5, 27445 loci

Table 3.10. Pan- differential expression at moderate stringency

		expression			Log2(axenic/holo)			
Name	**Description**	**ax**	**ho**	**p**	**all**	**Ar**	**Ag**	**Pr**
AT5G64120	Peroxidase superfamily protein	3.4	56.9	***	-3.7	-2.7	-4.0	-4.5
MLO12	Seven transmembrane MLO family protein	1.8	28.2	**	-3.4	-2.3	-3.9	-3.8
AT2G43620	Chitinase family protein	1.4	21.4	**	-3.2	-2.0	-3.6	-3.7
IOS1	Leucine-rich repeat protein kinase	1.9	23.0	**	-3.1	-1.6	-4.1	-2.8
FRK1	flg22-induced receptor-like kinase 1	1.1	15.7	**	-3.0	-1.4	-3.4	-3.3
XBAT34	hypothetical protein	2.3	19.4	***	-2.6	-2.2	-3.4	-2.2
AZI1	azelaic acid induced 1; SAR component	1.6	14.3	**	-2.6	-1.5	-3.2	-2.6
IGMT2	O-methyltransferase	2.9	21.1	***	-2.5	-2.3	-2.1	-3.1
AT1G30720	FAD-binding Berberine	2.8	19.4	***	-2.4	-1.6	-2.8	-2.5
AT1G51890	Leucine-rich repeat protein kinase	1.2	10.4	**	-2.4	-1.2	-2.8	-2.8
AT3G46280	kinase-like protein	1.2	10.1	*	-2.3	-1.2	-3.2	-2.0
AT4G22470	Protease inhibitor/seed storage/LTP	1.8	12.7	*	-2.3	-0.8	-3.2	-2.1
GSTF6	glutathione S-transferase 6	3.0	18.4	**	-2.3	-1.2	-2.3	-3.0
AT4G12490	Protease inhibitor/seed storage/LTP	1.1	9.2	**	-2.3	-1.0	-2.8	-2.6
AT1G30730	FAD-binding Berberine	2.1	13.9	***	-2.3	-1.4	-2.7	-2.4
AT1G51790	Leucine-rich repeat protein kinase	1.4	10.5	*	-2.3	-0.2	-3.2	-2.4
GSTF7	glutathione S-transferase 7	2.4	14.7	**	-2.2	-1.4	-2.2	-2.9
AT2G27660	Cysteine/Histidine-rich C1 domain	2.1	13.1	*	-2.2	-1.5	-3.0	-1.7
AT1G26380	FAD-binding Berberine	1.1	8.4	*	-2.2	-0.7	-2.1	-3.0
AT1G26420	FAD-binding Berberine	1.4	9.4	***	-2.1	-1.4	-2.2	-2.5
WRKY48	WRKY transcription factor	3.9	19.2	****	-2.0	-2.4	-2.0	-1.7
ACS7	Enzyme involved in ethylene biosynthesis	1.4	8.6	*	-2.0	-0.8	-2.7	-2.2
FLOT1	SPFH/Band 7/PHB domain membrane-associated	1.8	10.1	**	-2.0	-0.9	-2.5	-2.3
ERD7	Senescence/dehydration-associated protein	19.0	76.8	**	-2.0	-1.8	-3.0	-0.8
EARLI1	Protease inhibitor/seed storage/LTP	1.2	7.4	*	-1.9	-0.7	-2.6	-1.9

Criteria: log2(RPM/RPM)| ≥ 1.0, $p < 0.05$, RPM ≥ 5. Sorted by ascending average enrichment. Statistical confidence was determined by pairwise t-test (all holoxenic vs all axenic), with Benjamini-Hochberg correction. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, **** $p < 0.00005$

Table 3.10. (cont'd)

Name	Description	expression ax	ho	p	Log2(axenic/holo) all	Ar	Ag	Pr
PLA2A	phospholipase A 2A	5.4	23.2	**	-1.9	-1.0	-2.4	-2.0
AT5G48657	defense protein-like protein	1.9	9.9	*	-1.9	-1.0	-2.7	-1.7
BCB	blue-copper-binding protein; oxidative stress resp.	12.9	49.7	****	-1.9	-2.2	-1.8	-1.7
AT1G72060	serine-type endopeptidase inhibitor	10.6	40.7	***	-1.9	-1.6	-1.9	-1.9
AT1G51850	Leucine-rich repeat protein kinase	1.4	7.5	*	-1.8	-0.3	-2.4	-2.2
AT2G18690	transmembrane protein	2.0	9.6	***	-1.8	-1.7	-1.9	-1.9
FRB1	O-fucosyltransferase family protein	3.9	16.3	****	-1.8	-2.3	-1.8	-1.5
CNGC3	cyclic nucleotide gated channel 3	4.3	16.6	**	-1.7	-1.5	-2.2	-1.3
AT1G13470	hypothetical protein (DUF1262)	1.4	6.9	*	-1.7	-0.1	-2.7	-1.5
AT4G14370	Disease resistance protein (TIR-NBS-LRR)	2.3	9.8	**	-1.7	-0.9	-2.1	-2.1
CRK5	cysteine-rich receptor-like kinase	2.8	11.4	**	-1.7	-0.4	-2.1	-2.3
AT2G36690	2-oxoglutarate and Fe(II)-dependent oxygenase	1.3	6.3	**	-1.7	-0.5	-2.0	-2.0
CYP71B23	cytochrome P450	3.0	11.7	**	-1.7	-0.8	-2.0	-1.7
CRK14	cysteine-rich receptor-like kinase	3.4	13.0	****	-1.7	-1.7	-1.8	-1.5
WRKY18	WRKY transcription factor	10.7	36.3	**	-1.7	-2.6	-2.0	-0.5
AT1G05675	UDP-Glycosyltransferase	2.9	11.2	**	-1.7	-1.4	-1.3	-2.2
AT3G18250	Putative membrane lipoprotein	1.3	6.2	*	-1.7	-0.5	-1.8	-2.2
AT1G56660	MAEBL domain protein	31.0	98.2	**	-1.6	-1.2	-2.4	-1.1
CYP81F2	cytochrome P450	9.6	31.7	***	-1.6	-1.2	-1.5	-2.2
STZ	salt tolerance zinc finger	4.3	15.3	**	-1.6	-2.0	-1.0	-1.7
AT5G67340	ARM repeat superfamily protein	4.5	15.9	**	-1.6	-1.0	-2.0	-1.7
CRK6	cysteine-rich receptor-like kinase	1.8	7.5	*	-1.6	-0.8	-2.3	-1.3
AT1G10340	Ankyrin repeat family protein	1.7	7.1	**	-1.6	-0.9	-1.9	-1.6
AT5G35735	Auxin-responsive family protein	8.3	26.9	****	-1.6	-1.4	-1.8	-1.6
ARCK1	receptor-like cytosolic kinase; ABA neg regulator	1.5	6.4	**	-1.6	-1.2	-2.4	-1.1

Criteria: $\log_2(RPM/RPM)| \geq 1.0$, $p < 0.05$, $RPM \geq 5$. Sorted by ascending average enrichment. Statistical confidence was determined by pairwise t-test (all holoxenic vs all axenic), with Benjamini-Hochberg correction. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, **** $p < 0.00005$

Table 3.10. (cont'd)

		expression			Log2(axenic/holo)			
Name	**Description**	**ax**	**ho**	**p**	**all**	**Ar**	**Ag**	**Pr**
AT1G27020	plant/protein	3.3	11.7	*	-1.6	-0.9	-2.1	-1.2
AT5G10760	Eukaryotic aspartyl protease family protein	1.2	5.5	*	-1.6	-0.6	-2.0	-1.8
AT5G37600	cytosolic glutamine synthetase with NH3 affinity	10.8	33.6	**	-1.6	-0.8	-2.1	-1.5
LOX4	PLAT/LH2 domain lipoxygenase	19.4	58.3	**	-1.5	-2.1	-1.4	-0.9
WRKY30	WRKY transcription factor	1.2	5.4	***	-1.5	-1.2	-1.6	-1.8
AT2G39210	Major facilitator superfamily	16.1	47.9	**	-1.5	-0.9	-2.0	-1.5
ABCG42	ABC-2 and Plant PDR ABC-type transporter	7.6	23.7	**	-1.5	-0.9	-2.0	-1.3
JAZ1	jasmonate-zim-domain 1	14.6	43.4	*	-1.5	-2.2	-1.1	-0.8
WRKY60	WRKY transcription factor	5.6	17.7	**	-1.5	-0.7	-2.1	-1.6
RLP21	receptor like protein 21	1.4	5.8	*	-1.5	-0.8	-2.2	-1.2
PCC1	pathogen and circadian controlled 1	3.3	11.2	*	-1.5	0.0	-2.5	-1.4
WRKY53	WRKY transcription factor	14.1	41.8	****	-1.5	-1.5	-1.6	-1.4
ERD10	Dehydrin; bacteriostatic	93.5	257	**	-1.5	-2.3	-1.7	-0.2
AT1G49000	transmembrane protein	1.9	6.9	***	-1.5	-1.4	-1.5	-1.4
PP2-A5	phloem protein 2 A5	6.6	19.8	**	-1.5	-1.5	-1.6	-1.1
WAKL10	Leucine-rich repeat protein kinase, wall-associated	1.3	5.3	**	-1.5	-0.9	-1.4	-2.0
UCP5	mitochondrial dicarboxylate carriers	8.3	23.8	***	-1.4	-1.7	-1.6	-1.0
PYL6	PYR1-like 6; ABA-responsive	2.3	7.8	*	-1.4	0.4	-2.2	-1.7
ATPMEPCRB	pectin methylesterase inhibitor; BR regulator	11.2	31.6	**	-1.4	-1.1	-2.1	-1.0
AT1G65845	transmembrane protein	22.6	61.8	***	-1.4	-1.3	-1.5	-1.3
ACD6	ankyrin repeat family protein; *Pst* resistance	35.6	96.1	**	-1.4	-0.7	-2.0	-1.3
EXT4	Extensin: cell-wall hydroxyproline-rich glycoproteins	106	279	**	-1.4	-0.9	-1.6	-1.6
EXO70B2	exocyst subunit exo70 family protein B2	12.1	33.4	****	-1.4	-1.4	-1.6	-1.2
PP2-B13	phloem protein 2-B13	2.1	7.1	**	-1.4	-0.9	-1.4	-1.6
GLR2.8	glutamate receptor 2.8	1.6	5.8	****	-1.4	-1.1	-1.8	-1.3

Criteria: log2(RPM/RPM)| ≥ 1.0, $p < 0.05$, RPM ≥ 5. Sorted by ascending average enrichment. Statistical confidence was determined by pairwise t-test (all holoxenic vs all axenic), with Benjamini-Hochberg correction. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, **** $p < 0.00005$

Table 3.10. (cont'd)

		expression			Log2(axenic/holoxenic)			
Name	**Description**	**ax**	**ho**	**p**	**all**	**Ar**	**Ag**	**Pr**
AT3G16530	Legume lectin; chitin-inducible	12.7	35.0	**	-1.4	-0.7	-1.0	-2.3
ANK	ankyrin	1.6	5.8	*	-1.4	-0.4	-1.8	-1.5
AT3G22235	cysteine-rich TM module stress tolerance protein	4.9	14.4	*	-1.4	0.1	-2.4	-1.4
AT3G61280	O-glucosyltransferase rumi-like	2.8	8.8	*	-1.4	-0.7	-2.1	-1.0
RLP23	receptor like protein 23	1.9	6.4	*	-1.4	0.3	-1.9	-2.0
RLK1	receptor-like protein kinase 1, mannose binding	2.3	7.4	**	-1.4	-0.3	-2.0	-1.6
CNI1	carbon/nitrogen insensitive 1	4.2	12.2	****	-1.3	-1.1	-1.3	-1.7
FBS1	F-box family protein	2.5	7.8	*	-1.3	-2.3	-0.7	-0.5
PMT6	Major facilitator superfamily protein	12.7	33.5	**	-1.3	-1.1	-1.6	-1.2
AT5G52750	Heavy metal transport/detoxification protein	2.5	7.8	***	-1.3	-1.2	-1.5	-1.2
AT1G11960	ERD (early-responsive to dehydration stress)	16.4	42.3	**	-1.3	-1.1	-2.0	-0.8
AT1G61360	S-locus lectin protein kinase	6.0	16.5	****	-1.3	-1.3	-1.2	-1.5
AT4G39830	Cupredoxin superfamily protein	3.2	9.5	***	-1.3	-1.2	-1.4	-1.4
CHX16	cation/H exchanger 16	3.2	9.3	**	-1.3	-0.5	-1.7	-1.4
BCS1	cytochrome BC1 synth; amplifies SA signal	3.8	10.7	***	-1.3	-1.8	-0.7	-1.3
CYP707A3	cytochrome P450; ABA catabolism	9.1	23.7	**	-1.3	-1.7	-1.5	-0.7
VSR7	Vacuolar sorting receptor 7	2.0	6.3	*	-1.3	-0.2	-1.8	-1.5
AT3G28540	P-loop with nucleoside triphosphate hydrolases	9.9	25.3	*	-1.3	-0.2	-1.8	-1.5
EXT3	extensin 3	25.5	62.3	*	-1.3	0.4	-2.2	-1.8
MC2	metacaspase 2	2.1	6.4	*	-1.3	0.2	-2.0	-1.4
AT1G51820	LRR repeat protein kinase	2.7	7.8	*	-1.3	-0.3	-1.7	-1.6
AT1G52200	PLAC8 family protein	3.6	9.8	**	-1.2	-0.9	-1.5	-1.2
AT1G30700	FAD-binding Berberine	2.0	6.0	*	-1.2	-0.2	-1.5	-1.8
PTR3	peptide transporter 3; pathogen response	1.7	5.3	*	-1.2	-0.8	-1.8	-1.0
DIC2	dicarboxylate carrier 2	39.3	91.9	***	-1.2	-1.4	-1.0	-1.2

Criteria: log2(RPM/RPM)| ≥ 1.0, $p < 0.05$, RPM ≥ 5. Sorted by ascending average enrichment. Statistical confidence was determined by pairwise t-test (all holoxenic vs all axenic), with Benjamini-Hochberg correction. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, **** $p < 0.00005$

Table 3.10. (cont'd)

Name	Description	expression ax	ho	p	Log2(axenic/holoxenic) all	Ar	Ag	Pr
AT5G41750	Disease resistance protein (TIR-NBS-LRR)	9.5	23.1	**	-1.2	-0.9	-1.4	-1.2
AT4G38560	phospholipase-like protein PEARLI 4	2.6	7.2	*	-1.2	-0.2	-1.6	-1.5
PR4	pathogenesis-related 4	12.8	30.3	*	-1.2	-0.2	-1.3	-1.7
AT5G28630	glycine-rich protein	10.1	24.2	**	-1.2	-1.2	-1.5	-0.8
ORP1A	OSBP (oxysterol binding protein)	2.5	6.8	**	-1.2	-0.7	-1.7	-0.9
AT3G07195	RPM1-interacting protein 4 (RIN4) family	1.9	5.5	**	-1.2	-0.6	-1.4	-1.3
RLP52	disease resistance protein induced by chitin	2.4	6.6	***	-1.2	-0.8	-1.3	-1.4
AT1G18390	Serine/Threonine kinase family catalytic domain	6.5	15.6	****	-1.2	-0.9	-1.1	-1.5
AT2G37130	Peroxidase superfamily protein	9.6	22.5	**	-1.2	-0.7	-1.2	-1.5
PIRL2	plant intracellular ras group-related LRR 2	5.2	12.8	**	-1.2	-0.6	-1.0	-1.9
AT5G25930	LRR protein kinase	10.9	25.2	***	-1.1	-0.7	-1.2	-1.5
AT2G26440	Plant invertase/pectin methylesterase inhibitor	11.8	26.9	*	-1.1	-0.2	-1.6	-1.2
WRKY47	WRKY transcription factor	34.8	77.0	***	-1.1	-1.0	-1.4	-0.9
RLP33	receptor like protein 33	3.6	8.9	**	-1.1	-0.5	-1.3	-1.6
AT4G29780	nuclease	16.5	36.7	**	-1.1	-1.5	-0.6	-1.0
ACA.I	autoinhibited Ca2/ATPase II	57.7	125.0	***	-1.1	-1.0	-1.1	-1.2
NAC036	NAC domain containing protein	5.2	12.3	**	-1.1	-1.4	-1.4	-0.5
MSS1	high affinity, hexose-specific/H symporter	5.3	12.5	**	-1.1	-1.4	-0.9	-1.0
IGMT3	O-methyltransferase	7.1	16.3	**	-1.1	-0.7	-0.8	-1.7
SYP122	Syntaxin; regulator of JA/SA and apotosis	18.3	40.2	***	-1.1	-1.6	-1.2	-0.7

Criteria: log2(RPM/RPM)| ≥ 1.0, $p < 0.05$, RPM ≥ 5. Sorted by ascending average enrichment. Statistical confidence was determined by pairwise t-test (all holoxenic vs all axenic), with Benjamini-Hochberg correction. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, **** $p < 0.00005$

Table 3.10. (cont'd)

Name	Description	expression ax	ho	p	Log2(axenic/holoxenic) all	Ar	Ag	Pr
AT1G35710	Leucine-rich repeat protein kinase	38.6	82.6	*	-1.1	-0.1	-1.6	-1.1
AR781	pheromone receptor-like protein (DUF1645)	8.3	18.7	**	-1.1	-1.6	-0.9	-0.8
UGT72E1	UDP-glucosyl transferase for lignin metabolism	30.2	65.0	**	-1.1	-0.4	-1.4	-1.3
AT5G08760	transmembrane protein	8.6	19.3	**	-1.1	-0.9	-1.6	-0.7
AT3G47090	LRR protein kinase	3.2	7.8	**	-1.1	-0.5	-1.4	-1.4
GSTF2	glutathione S-transferase PHI 2	41.0	87.3	**	-1.1	-0.5	-1.1	-1.4
AT4G35985	Senescence/dehydration-associated	5.5	12.6	*	-1.1	-1.4	-1.7	-0.1
BIR1	BAK1-interacting leucine-rich repeat protein kinase	16.3	34.7	**	-1.1	-1.0	-1.5	-0.6
AT1G36060	Integrase-type DNA-binding	2.5	6.2	*	-1.0	0.0	-1.3	-1.4
AT5G05600	2-oxoglutarate and Fe(II)-dependent oxygenase	13.5	28.9	*	-1.0	-1.8	-1.3	0.4
ERF5	ethylene responsive element binding factor 5	6.9	15.2	**	-1.0	-1.2	-0.6	-1.4
RGXT2	rhamnogalacturonan xylosyltransferase 2	2.5	6.1	*	-1.0	-0.1	-1.6	-1.1
LHT1	lysine histidine transporter 1	3.1	7.3	*	-1.0	-0.3	-1.7	-0.9
AT4G38550	phospholipase-like protein (PEARLI 4)	256.6	517.4	****	-1.0	-1.0	-1.1	-0.9

Criteria: log2(RPM/RPM)| ≥ 1.0, $p < 0.05$, RPM ≥ 5. Sorted by ascending average enrichment. Statistical confidence was determined by pairwise t-test (all holoxenic vs all axenic), with Benjamini-Hochberg correction. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, **** $p < 0.00005$

Table 3.10. (cont'd)

Name	Description	expression ax	ho	p	Log2(axenic/holoxenic) all	Ar	Ag	Pr
CYP71A13	cytochrome P450 family 71 polypeptide	5.5	1.9	**	1.2	1.4	1.2	0.9
AT5G05250	hypothetical protein	19.5	8.1	*	1.2	1.0	1.5	0.9
AT4G22495	hypothetical protein	297.4	124.6	*	1.3	-0.2	2.2	3.1
UMAMIT33	nodulin MtN21 /EamA-like transporter	9.8	3.4	**	1.3	1.4	1.5	1.0
AT4G22505	Protease inhibitor/seed storage/LTP	372.2	149.7	*	1.3	-0.2	2.3	3.2
AT4G22475	transmembrane protein	475.0	189.9	*	1.3	0.0	2.1	2.9
AT4G22485	Protease inhibitor/seed storage/LTP	452.4	179.8	*	1.3	0.0	2.1	2.9
CYP81D11	Cytochrome P450; cis-JA responsive	83.9	30.2	***	1.4	0.6	2.7	2.8
AT4G32480	sugar phosphate exchanger, putative (DUF506)	136.0	49.2	****	1.5	1.8	1.5	1.1

Criteria: log2(RPM/RPM)| ≥ 1.0, $p < 0.05$, RPM ≥ 5. Sorted by ascending average enrichment. Statistical confidence was determined by pairwise t-test (all holoxenic vs all axenic), with Benjamini-Hochberg correction. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, **** $p < 0.00005$

Table 3.11. Metabolite quantification by GC/MS

	Arid		Ag		Prairie	
	axenic	holo	axenic	holo	axenic	holo
phenyl-B-D glucopyranoside	47.1±4.6	1.6±0.4	50.2±8.6	0.7±0.4	57.3±3.6	0.5±0.3
13.27 feruloylglycoside	6.0±0.5	0.5±0.2	7.5±1.0	1.2±0.2	6.8±0.5	1.4±0.4
unk. N-metabolite	49.7±4.1	19.8±3.3	46.5±6.1	16.6±1.3	53.2±4.1	26.2±3.2
sucrose	118.0±17.5	67.1±7.3	336.1±25.9	133.5±15.9	233.7±25.5	115.2±11.1
12.74 320 192 467	3.9±0.3	2.4±0.3	4.3±1.3	1.7±0.2	4.2±0.5	2.7±0.5
ornithine	173.6±20.5	123.2±34.6	152.2±27.1	70.0±7.2	194.2±19.1	119.2±17.6
citrulline	121.0±12.9	81.4±23.8	102.4±18.9	50.2±5.6	135.2±13.3	79.6±12.9
putrescine	48.3±3.9	21.4±3.0	40.7±5.1	18.0±1.4	46.4±3.9	31.2±3.4
glycine	38.2±4.2	22.1±3.6	49.4±7.0	22.5±2.5	40.3±3.8	30.6±4.8
asparagine	316.8±33.2	229.3±54.1	266.8±45.6	132.6±14.3	328.9±37.1	245.6±28.2
glutamine	951.8±101.3	844.4±200.4	667.5±124.7	319.6±39.5	908.2±107.3	747.9±78.8
unk. N-metabolite	177.7±14.7	149.2±40.2	257.0±40.8	115.1±10.5	179.7±26.1	162.5±19.9
ascorbic acid	206.9±12.9	176.5±10.8	319.4±14.9	201.1±8.0	274.6±9.3	185.0±15.7
myo-inositol	18.6±2.5	12.5±1.7	26.3±3.6	15.9±1.7	22.6±2.6	17.1±0.9
azelaic acid	12.3±1.6	10.0±2.3	9.4±1.0	5.6±0.5	15.4±1.5	13.5±1.7
lysine	27.2±2.4	24.0±4.6	25.1±3.9	16.8±1.2	26.1±2.4	22.2±3.6
allantoin	28.3±3.1	28.7±5.6	24.6±3.1	15.5±1.1	28.4±2.2	26.3±5.6
serine	44.3±5.9	30.1±5.7	33.3±4.1	23.0±2.9	48.8±11.4	42.5±11.9
glucose 6-phosphate	5.5±0.5	4.2±0.7	6.3±0.4	4.9±0.3	6.9±0.6	5.7±0.6
B-cyano-alanine	11.5±3.3	5.9±1.9	6.2±2.2	5.5±1.7	9.6±4.5	7.0±1.9
glucose	24.1±3.0	18.2±2.3	27.4±4.5	17.2±1.1	27.6±2.7	28.8±3.9
uric acid	17.2±2.2	11.7±0.9	19.0±1.8	13.7±2.2	15.8±2.0	14.3±2.7
unk. sugar-phosphate	2.6±0.2	2.6±0.4	3.0±0.2	2.5±0.2	3.3±0.2	2.7±0.3
unk. N-metabolite	51.8±6.1	43.7±9.2	32.3±6.8	23.7±2.8	49.4±4.5	46.4±5.7
unk. N-metabolite	4.3±0.5	3.6±0.8	2.7±0.4	2.1±0.2	4.0±0.4	3.7±0.5

RT (min) key m/z; Values are μg·g-1 FW (sorbitol equivalent) ± standard error of the mean.

Table 3.11. (cont'd)

	Arid axenic	holo	Ag axenic	holo	Prairie axenic	holo
threonine	56.5±8.7	39.4±4.6	39.8±4.8	33.8±3.3	50.6±7.7	46.1±6.8
cysteine	4.8±0.3	3.7±0.4	4.6±0.4	3.3±0.3	4.3±0.3	4.7±0.5
tyrosine	10.4±0.8	10.2±1.5	10.0±1.4	8.4±0.8	10.9±1.2	10.7±1.9
tryptophan	3.4±0.2	2.9±0.6	3.4±0.5	2.9±0.3	3.6±0.3	3.5±0.4
threonic acid	2.9±0.3	2.5±0.3	3.3±0.2	2.7±0.2	3.1±0.3	3.1±0.3
B-alanine	9.4±0.8	7.5±1.1	7.7±1.1	5.8±0.7	8.2±0.9	9.2±1.7
phenylalanine	32.1±2.8	28.1±4.1	32.0±4.0	23.7±1.6	29.8±3.0	35.4±5.5
spermidine	52.0±2.8	37.0±4.0	41.2±4.3	40.0±2.9	46.1±3.7	42.2±3.2
10.86 306 288 N-metabolite	4.8±0.3	4.0±0.4	5.9±0.1	4.5±0.3	5.5±0.3	6.7±0.9
galactose	5.2±0.8	4.8±0.8	6.2±0.5	4.7±0.4	6.4±0.6	8.0±1.0
glyceric acid	5.2±0.3	3.9±0.5	4.2±0.2	4.0±0.2	4.8±0.3	4.8±0.5
glutamic acid	230.7±28.2	215.5±39.2	247.4±32.6	172.8±18.1	238.4±25.3	338.1±55.9
11.25 217 450 dehydrosugar	51.7±3.5	57.6±6.0	74.2±5.7	65.3±5.2	63.6±3.2	73.5±5.2
sinapic acid-4-O-glucoside	1.5±0.2	1.6±0.3	1.5±0.3	1.3±0.3	1.4±0.1	1.7±0.3
12.64 320 479 464	3.1±0.4	1.9±0.2	1.7±0.2	1.4±0.1	2.2±0.3	2.7±0.5
unk. phenolic	8.9±1.5	1.0±0.3	3.7±0.9	4.2±0.5	5.3±0.5	5.2±1.1
12.52 507 103 189 285 249 261	28.2±3.2	31.5±3.8	19.9±1.9	14.9±2.2	19.7±2.4	30.1±3.9
12.98 320 307 217	6.2±0.6	3.6±0.5	3.4±0.2	3.5±0.4	5.1±0.5	5.9±0.9
isoleucine	17.9±2.0	14.6±1.7	11.8±1.8	12.0±1.6	14.7±1.5	18.6±4.0
leucine	13.5±1.7	11.9±1.3	8.8±1.1	9.1±1.3	10.8±1.3	13.6±2.7
valine	41.2±4.5	36.7±4.7	31.1±4.1	30.4±2.8	35.4±3.4	47.6±9.6
4-hydroxybenzoic acid	10.8±1.0	9.4±1.1	12.2±1.6	11.6±1.5	11.4±0.7	16.9±3.1
phosphate	689.6±47.4	771.0±75.2	544.9±51.6	539.0±28.6	579.4±34.8	847.8±91.7
gamma-aminobutyric acid	28.9±2.9	22.1±1.8	19.5±3.0	17.1±1.1	21.7±2.4	37.2±7.6
3,4-dihydroxybenzoic acid	13.2±1.8	12.9±1.5	16.3±2.2	14.7±2.1	14.4±1.3	24.4±4.3

RT (min) key m/z; Values are μg·g-1 FW (sorbitol equivalent) ± standard error of the mean.

Table 3.11. (cont'd)

	Arid axenic	holo	Ag axenic	holo	Prairie axenic	holo
ethanolamine	235.4±21.1	181.3±17.0	118.9±12.5	150.5±13.5	176.7±19.4	227.4±25.0
5 oxo proline	896.7±61.9	895.1±115.9	647.9±81.3	750.5±55.2	687.6±62.5	967.2±115.1
salicylic acid	20.6±2.6	19.6±2.3	29.2±4.0	28.2±3.2	23.6±2.3	41.1±7.4
methylphosphate	36.4±3.0	29.6±3.7	25.2±3.3	36.9±1.7	29.3±3.2	35.5±5.5
erythronic acid	7.5±0.8	9.1±1.2	6.4±0.7	6.8±0.6	6.3±0.8	10.6±1.4
phloroglucinol	2.3±0.6	1.9±0.3	3.0±0.4	2.5±0.3	2.1±0.2	4.5±0.9
12.69 395 249 410 519 504	42.9±5.9	42.4±6.1	17.6±2.1	22.6±3.4	35.1±3.7	49.2±8.8
caffeic acid	1.1±0.3	0.8±0.1	1.2±0.1	1.4±0.1	1.2±0.1	1.9±0.4
monopalmitin	121.7±29.6	119.3±13.8	68.9±10.9	81.9±16.5	57.1±11.6	87.5±19.3
12.03 219 235	154.7±21.8	148.5±28.9	95.6±10.1	121.8±16.3	174.0±19.6	250.3±62.3
aspartic acid	476.0±37.9	593.8±66.1	431.0±51.1	503.8±43.1	418.0±48.9	668.9±82.9
succinic acid	14.8±2.2	13.3±1.8	11.0±1.2	12.8±1.5	13.2±1.5	21.6±3.5
alanine	196.9±20.8	177.8±39.2	213.4±27.4	238.7±23.6	184.5±20.1	318.3±60.7
1,2,4 benzenetriol	4.9±0.6	5.2±0.6	4.5±0.6	4.8±0.6	4.8±0.5	9.3±1.8
glycerol	60.3±7.6	67.5±7.6	57.1±6.1	56.2±7.1	32.0±3.1	67.1±11.9
3 hydroxybenzoic acid	2.8±0.3	3.4±0.4	3.7±0.5	3.9±0.6	3.1±0.3	6.2±1.2
13.35 409	2.3±0.4	2.1±0.3	1.4±0.2	1.5±0.2	1.6±0.2	3.2±0.7
9.99 98 288 390 N metabolite	814.5±148.9	1046.5±179.3	437.8±59.0	542.5±122.9	658.2±78.4	1152.3±188.9
sinapoyl malate	8.4±1.5	8.4±1.8	6.5±1.1	5.4±1.1	6.8±1.2	18.2±3.3
13.47 342 299 315 P metabolite	22.4±2.8	25.3±3.4	18.6±3.3	20.9±3.5	20.7±2.0	40.9±7.8
stearic acid	82.1±11.0	94.5±12.0	66.9±10.6	76.9±12.3	75.4±7.5	150.7±28.8
malic acid	71.1±5.8	68.9±9.0	52.9±6.8	70.4±5.2	60.9±4.5	105.6±12.3
fumaric acid	639.0±30.6	541.2±96.5	471.1±76.3	618.8±36.8	512.3±64.4	919.8±119.4
8.24 273 258 231 216	264.2±43.6	280.0±28.7	191.4±25.7	226.7±30.1	261.5±30.0	526.3±123.5

RT (min) key m/z; Values are μg·g-1 FW (sorbitol equivalent) ± standard error of the mean.

Table 3.11. (cont'd)

	Arid		Ag		Prairie	
	axenic	holo	axenic	holo	axenic	holo
maleic acid	27.7±3.3	28.6±3.7	22.4±3.5	29.2±5.1	23.0±2.8	42.6±7.7
palmitic acid	116.3±16.9	121.7±16.6	98.0±19.0	96.1±15.6	94.0±10.2	231.6±48.6
13.79 204 231 glycoside	25.4±1.8	59.4±5.5	141.2±41.1	62.7±8.1	24.0±1.5	135.2±42.8
dihydrouracil	13.2±1.5	12.3±1.3	10.1±1.3	12.3±1.7	10.4±1.3	21.5±4.3
sinapic acid	654.5±39.5	805.9±79.3	511.1±61.4	748.5±93.0	542.3±55.3	983.0±141.5
9.79 346 174 N metabolite	31.2±5.1	28.8±7.5	13.1±1.9	18.0±1.9	22.7±2.0	44.4±14.0
phytol	424.8±55.4	506.0±52.0	307.5±19.1	490.2±80.6	478.7±54.0	836.5±210.5
8.06 249 379 115	613.6±106.2	519.1±144.0	308.8±89.8	375.6±94.4	344.6±65.7	869.5±262.5
oleic acid	6.1±1.0	6.1±1.0	3.3±0.4	4.5±0.6	5.4±0.5	13.1±2.8
stigmasterol	10.5±3.9	6.3±0.9	2.9±0.2	4.3±1.0	5.9±0.8	13.9±3.8
linoleic acid	13.6±2.1	17.7±2.6	9.0±1.4	12.1±3.4	17.3±2.3	44.9±11.3
monostearin	332.4±80.2	313.7±48.4	100.5±14.1	176.6±46.9	128.4±28.7	283.0±73.0
B sitosterol	33.7±6.2	35.3±4.5	13.5±1.7	21.6±4.1	28.6±3.5	69.2±16.7
12 hydroxyoctadecanoic acid	9.6±3.2	8.9±2.6	5.2±1.1	5.6±2.2	5.5±0.9	19.9±4.9
glycerol 1/3 phosphate	9.4±1.2	10.9±1.6	6.4±0.7	7.6±1.1	8.7±0.8	29.0±6.0
14.39 320 217 361 119	2.0±0.4	3.8±0.4	5.3±1.5	3.9±0.7	1.5±0.2	8.7±2.2
citric acid	56.0±2.3	139.8±6.9	59.4±3.3	142.4±8.7	52.6±4.3	133.8±10.7
12 hydroxyoctadecanoic acid methyl ester	56.8±23.7	56.0±20.3	15.9±4.2	19.9±7.9	20.3±3.9	112.4±49.2
2 ethylhexanoic acid	16.3±2.9	82.7±8.7	25.3±6.1	78.7±6.7	13.1±1.2	188.5±35.8

RT (min) key m/z; Values are μg·g-1 FW (sorbitol equivalent) ± standard error of the mean.

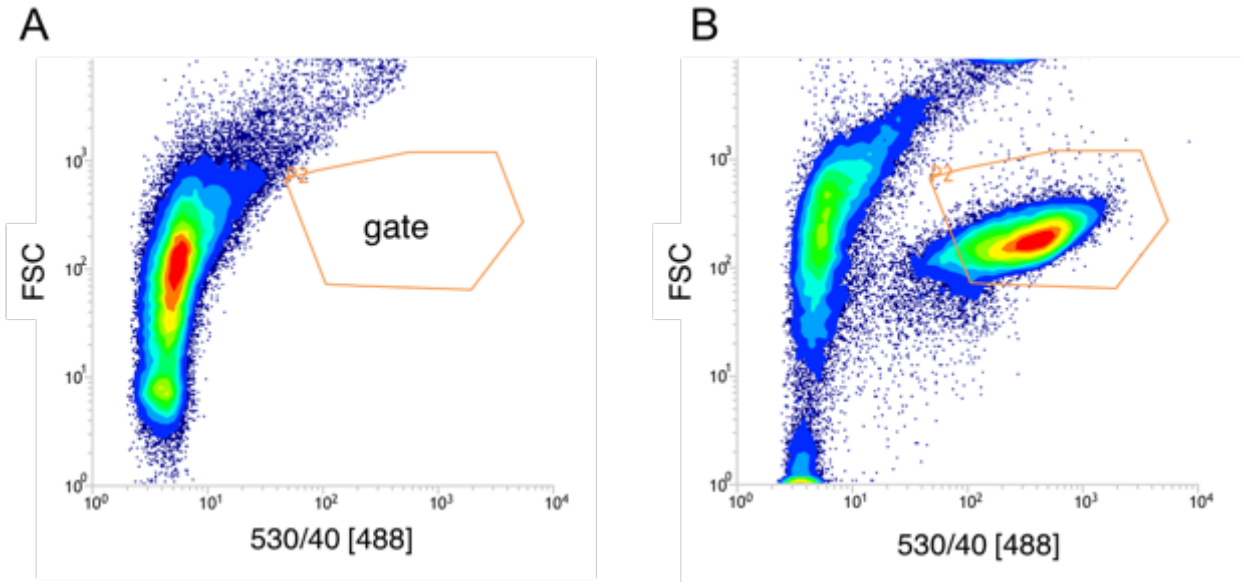


Figure 3.27. Flow cytometry scatter plot of suspended Arabidopsis leaf disc samples from bacterial counts protocol. (A) mock-inoculation control and (B) infected with *Pst*-GFP $5x10^6$ $CFUml^{-1}$. Events were enumerated from the gate (in orange) as a function of volume. Forward scatted (FSC) is on the Y-axis and green fluorescence intensity (530/40 [488]) is on the X-axis.

Chapter 4

Discussion and future directions for plant microbiota research

Plant microbiome research is the product of a multidisciplinary alliance among plant biologists and, enabled by –omics technology, inspired by coordinated human microbiome initiatives, and motivated by food demands of a growing global population. Great strides have been made in the past decade towards a better understanding of which microorganisms inhabit plant microbiomes. Well-designed plant microbiome survey experiments have revealed factors likely to drive microbiome community composition, and cutting- edge -omics technologies (genomics, targeted/shotgun metagenomics, transcriptomics, metabolomics, proteomics, etc.) are powerful for tools for hypothesis generation. Fundamentally, hypothesis-driven empirical research that allows us to interpret these enormous datasets beyond correlations of patterns. The annotation for each gene/transcript/peptide/metabolite and taxonomic classification of each microorganism in a modern plant microbiome survey is based on homology or shared attributes to a biologically characterized example. Perhaps impeded by technological or logistical hurdles, all too often, modern microbiome studies will report correlations and postulate functional repercussions, but fall short of hypothesis testing. Hypothesis testing is essential to translate plant microbiome research into real-world agricultural practice.

A recent summit of leading plant microbiome researchers in Asilomar, California, gathered to discuss the status of the field and to establish research priorities for harnessing plant microbiomes in sustainable agriculture (Busby et al., 2017): (1) define core microbiomes and metagenomes in these model systems, (2) elucidate the rules of synthetic, functionally programmable microbiome assembly, (3) determine functional

mechanisms of plant-microbiome interactions, (4) characterize and refine plant genotype-by-environment-by-microbiome-by-management interactions. In this section, I offer my perspective on and propose experiments and/or endeavors towards achieving some of these objectives, with emphasis on potential applications for the FlowPot system in this pursuit. As is true for most technologies, the FlowPot system design can be refined and improved. Where appropriate, I offer specific design and protocol improvement suggestions for future models. I offer my perspective towards achieving a better mechanistic understanding of axenic immune deficiency.

Development of model host-microbiome systems with associated microbial culture collections and reference genomes

Development of culture collections for plant microbiomes across different model species is no doubt a research community task. Currently, large culture collections exist for Bacteria, fungi, oomycetes, and protists exist for Arabidopsis thaliana in various research laboratories, some of which have been fully sequenced, including: Paul Schulze Lefert (Max Planck, Cologne, Germany), Erick Kemen (Max Planck, Tubingenden, Germany), Detlef Weigel (Max Planck, Tubingenden, Germany), Jeff Dangl (University of North Carolina, Chapel Hill, USA), Julia Vorholt (ETH Zurich, Switzerland) and Joy Bergelson (University of Chicago, USA). I also established a 400-member bacterial culture collection of Arabidopsis phyllosphere-associated bacteria, along with full-length 16S rRNA gene sequences, currently in the laboratory of Sheng Yang He. Ideally, a central repository should be established with each isolate, along with comprehensive associated metadata, following mutually agreed-upon standards and guidelines (Bowers et al., 2017; Chain et al., 2009). It is the opinion of this author that full genome sequencing information will be required for organism identification in the near future, and thus the genomes should be sequenced and annotated for all isolates included in the collection. Along the same framework as the Earth Microbiome Project, the collection should include a diversity of sampling locations to capture increased diversity (Thompson et al., 2017). Associated metadata should be included in a publically accessible database (Arabidopsis accession, isolation location, tissue type, environmental details, growth media, culture conditions, etc.). Such an effort would require substantial funding, organization, and a willingness among participating

laboratories to contribute fairly. Perhaps it would be the highest probability of success would be to coordinate such an effort with separate funding agencies in the EU, United States, Asia, and Australia/New Zealand. An essential element of this endeavor is strict adherence to laboratory and database standards, and even if the strains are not located in a single internationally-accessible location, the database containing genomes and associated metadata should be globally accessible.

Once established for Arabidopsis, this will provide a framework to build additional databases for alternative model systems, expanding upon already initiated efforts for barley (Bulgarelli et al., 2015), wheat (Donn et al., 2015), corn (Aira et al., 2010), rice (Edwards et al., 2015), soybean (Mendes et al., 2014), medicago (Stanton-Geddes et al., 2013), *Populus* (Hacquard and Schadt, 2015), and tomato (Tian et al., 2015). Current culture collections exist for many of the additional proposed models. An international summit among laboratory directors and governors of each respective culture collection would greatly facilitate coordination and centralize what resources are currently available. Participation by industry would also be greatly valuable, as biotechnology companies may have large culture collections.

Define core microbiomes and metagenomes

The concept of an absolute core microbiome for a plant species is controversial, and there are several schools of thought that fundamentally come down to the species concept of microbiota and where thresholds are set for gene sequence clustering. In communities of bacteria, there are reports of very distantly-related microorganisms having functionally redundancy. Thus, a global absolute core community of microbiota inhabiting the Arabidopsis microbiome may not exist at conventional OTU identity cutoffs. Indeed, at low-resolution classification levels, such as phylum-level classification, core microbiomes exist. However, among a single genus such as *Pseudomonas*, there are reports of approximately 70% of gene conservation, despite identical 16S gene sequences (Silby et al., 2009). One of the earliest (now disproven) concepts in microbial ecology is “everything is everywhere, and the environment selects” (de Wit and Bouvier, 2006). Certain organisms have obligate symbioses and vertically-transmitted microbiota. This is a common phenomenon for insects (Baumann et al., 1997; Hansen et al., 2012; Moran, 2001). Seed-borne vertical transmission of microbiota may occur in some plants, as has been reported in maize (Johnston-Monje and Raizada, 2011). However, as demonstrated in this dissertation, axenic Arabidopsis can acquire microbiota from extremely different input communities with minimal overlap of OTUs and grow vigorously. It is the opinion of this author that core microbial species likely exist within a confined geographic area, but biogeography of microorganisms and environmental variability suggest there is not a core microbiome, as defined by 16S rRNA gene sequences. In other words, core is an appropriate term to use in a relative context, but not an absolute context. It is conceivable that global

core microbiome functionality exists, although functions can be convergently evolved and non-orthologous genes can code for very different proteins with similar functionality. Thus, in a shotgun sequencing experiment among many plant microbiome samples, fewer core gene sequences might exist than core functional elements. Based on gene presence/absence, some genes that are common to all bacteria will presumably be “core” metagenomic components, such as *gyrB, rpoE, rpoN, tufAB,* (Parks et al., 2015). It is also possible that core sequence motifs will be identified in such an experiment for which no reliable annotation exists. Thus, extrapolation of biologically interpretable data to formulate testable hypotheses may be limited.

Core taxonomic overlap was identified for input, substrate, and plant-associated communities in FlowPot experiments for both Chapter 2 and Chapter 3, as were core differentially expressed genes. The plant-enriched taxa (relative to substrate abundance) in both the calcined clay system and the FlowPot system may have attributes to establish and maintain a more robust association with the host plant than other input microbiota. This does not necessarily indicate that microbiota with a more robust host association are beneficial to the host, pathogens and parasites have evolved adaptations for close host associations as well. However, it is of interest to identify microbial traits that bestow the ability to form robust host associations in the presence of a diverse microbial community. Such traits could include one or more of the following abilities: to suppress/evade/tolerate host immune responses, to manipulate exudation of nutrient sources to maintain a metabolic relationship with the host, to outcompete other microorganisms for the nutrient-rich host-associated

environment, to form close adhesion structures to maintain a close physical association with the host. From the taxa identified as plant-enriched, it would be interesting to pursue a forward or reverse genetic approach to identify microbial genes of importance to maintain the close association. One candidate for this is OTU1 *Ralstonia* from Chapter 2, which was root-enriched from the natural German soil, as well as the synthetic community. Another commonly-enriched genus with very limited research published on it is *Massilia*. Knowledge of successful plant colonization strategies may help us to predict the success of candidate biostimulants, biopesticides, and biofertilzers in the field based on their genomic features.

Rules of synthetic, functionally programmable microbiome assembly

A fundamental question in plant microbiome research, and an essential prerequisite to translate microbiome research into effective real-world agricultural practice is “how does the microbiome influence host phenotypes of interest?” There are far too many combinations of microbiomes, hosts, and phenotypes to assess by just one research group, and a research community is necessary to address this question. Experimental reproducibility of microbiome experiments is often confounded by environmental complexity and stochasticity, making hypothesis testing and reproducibility major challenges in plant microbiome research. Recognizing a lack of adequate growth systems for microbiome recolonization experiments that simulate agriculturally relevant conditions, I focused on the development of the FlowPot system-- an inexpensive growth platform constructed with common materials and standard greenhouse peat substrate, for highly-controlled plant growth. This FlowPot system was designed to maximize reproducibility by normalization of environmental variables. Microbial communities can be inoculated directly into the soil, allowing for true isolation of the “microbiome” variable. Given that there is no “standard” soil microbiome, I developed the FlowPot system to accommodate axenic (germfree) plant growth, so that the system can be inoculated with microbiota of choice, including but not limited to defined synthetic microbiota, natural soil microbiota, or mock-inoculated with no microbiota. Additional variables can be manipulated, for example: nutrient supplementation, chemical addition, climate simulation, and pathogen introduction. The overall objective and purpose of the FlowPot system is to provide a standardized growth platform for the plant microbiome research community. Data harmonization is

key to drawing accurate comparisons of experimental microbiome results, and to truly leverage the power of a research community, model systems and standards are essential.

Of course, the FlowPot system is not only possible growth platform for plant microbiome research, nor is it the first. One such alternative platform used in several previous high-profile studies, the calcined clay growth system, was compared with the FlowPot system in Chapter 2. The calcined clay system was previously used to grow Arabidopsis in the presence of defined bacterial microbiota to address whether isolation environment (leaf/soil/root) defined where bacterial would preferentially colonize and whether modulation of salicylic acid signaling in the host would predictably influence microbiome composition, respectively(Bai et al., 2015; Lebeis et al., 2015). Using the same synthetic bacterial community from(Bai et al., 2015), we performed recolonization experiments in tandem with the calcined clay system, and report that in fact the FlowPot system accommodates more bacterial diversity more closely recapitulating the microbiome of a greenhouse or agricultural setting. Furthermore, we determined that the enormous microbiota community profile differences between the bulk substrate and the root communities in the calcined clay system we more subtle in FlowPots, perhaps more accurately recapitulating or simulating what would occur in an agricultural setting. It is ideal to have multiple growth platforms for microbiome research, especially since our experiments corroborate the many previous reports that environmental factors are major drivers of community differentiation (Zhou and Ning, 2017).

The breadth of colonization experiments conducted in Chapter 2 could be expanded in a number of interesting ways, depending on the hypothesis to be tested. The objective of the study described in Chapter 2 was primarily to introduce the FlowPot system and validate the health of axenic Arabidopsis. A secondary object was to determine whether bacteria differentially colonize FlowPots and calcined clay, and in what manner. With those objectives accomplished, follow-up on this study could go in various directions

Proposed improvements to FlowPot design and assembly

Current conditions have been optimized for Arabidopsis growth. Arabidopsis is attractive because before bolting, it grows relatively flat. Plants with greater vertical requirements would need an alternative growth vessel. The FlowPots themselves are modular, and will easily fit into containers other than the Eco2 micro boxes. The fiberglass mesh may not accommodate plants with thicker stems than Arabidopsis, so these may need to be adapted with either wider-gage mesh, or removable mesh. Plants with a more elaborate and larger root system than Arabidopsis may require additional growth substrate contained by a larger cylinder than the 50 mL syringe barrel. Companies such as Wilburn Medical USA make a 100 mL syringe that could potentially accommodate for this. As soil volume increases, one must take into consideration the effect this will have on humidity within the chamber and water retention in the substrate. Ideally, empirical experimentation should be conducted on a single lot of sphagnum peat with titrated vermiculite and perlite to determine the water retention. A major bottleneck is the manual labor associated with assembly, along with the risk for contamination. The assembly facility here at Michigan State would immensely benefit from a collaboration with an industrial engineering group and experts from the packaging school.

Despite extensive efforts to optimize batch-to-batch consistency of FlowPots, for unknown reasons, batches occasionally fail, and unpredictably entire boxes of plants will be unhealthy or not survive. In my experience, sometimes this is input community dependent, and other times there is no correlation to microbiome colonization status. The Michigan State University greenhouse staff performs chemical testing on each lot

of Sunshine® Redi-Earth Plug & Seedling mix they receive, and indeed there are variations in nutrient levels. Many factors, however, are not routinely tested. For example, Redi-Earth is blended with an unspecified amount of dolomite limestone, vermiculite, proprietary wetting agent, gypsum, and a proprietary silicon additive. This substrate was used for FlowPots in all experiments conducted at Michigan State because this is the standard greenhouse substrate, but ideally, the substrate should not include proprietary additives. Furthermore, Redi-Earth is pasteurized or steamed by an unspecified protocol. Since there is lot-to-lot variation of nutrient levels, presumably there is lot-to-lot variation of the proprietary additives. It is unknown how the additives may affect microbiota. Greater alpha diversity was present in the Michigan (ag) soil inoculated FlowPots from chapter 2 than the chapter 3 FlowPots inoculated with identical input microbiota. The FlowPots used in chapter 2 were assembled with German peat substrate, while the chapter 3 FlowPots were assembled with Redi-Earth. The extent to which this influences microbial community composition is unknown because the experiment was not conducted simultaneously and time points were not identical (six weeks in Chapter 2, three weeks in Chapter 3).

For future FlowPot optimization, I propose that rather than using the Redi-Earth product, a substrate consisting of pure sphagnum peat blended with medium vermiculite be used. Additional nutrient supplementation and/or wetting agent addition may be necessary to achieve optimal Arabidopsis growth, but this needs to be empirically validated. It is likely that peat from different locations around the world may differ in chemical and physical attributes, so it is important that the peat source is

reported for each experiment. It would be ideal to conduct an experiment testing whether peat from different sources influences the microbial composition.

The Eco2 micro boxes in this set of experiments contain a gas-permeable filter that allows for passive diffusion of air, thus allowing for transpiration, and to some extent, moisture loss. After several autoclave cycles, some filters become discolored and appear partially melted, while others do not. Gas exchange rate and accumulation of volatile organic compounds can affect many aspects of plant physiology, and presumably, the partially melted filters are compromised in full gas exchange capacity. To improve consistency, I suggest that the Eco2 boxes be replaced with similar polypropylene containers, but with an opening port on the side of the box, to which disposable adhesive air filter can be adhered. The side is a better suited place for the filter because if one desires to adapt the system to an overhead optical phenotyping platform, a lid filter may obscure the light and image path. Packing the substrate into the FlowPots by hand is prone to inconsistency. Similar to tamping an espresso, variations in packing intensity greatly influence compaction, and thus may influence consistency.

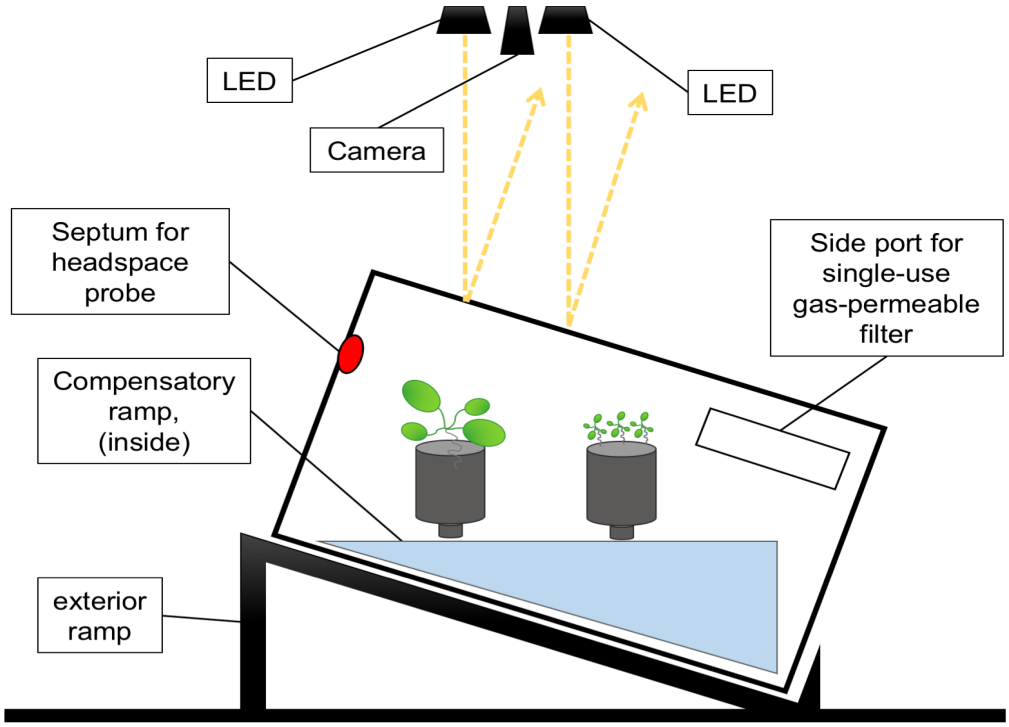


Figure 4.1. Schematic of FlowPot modification to allow it to fit in a DEPI chamber without light reflection interference. Further improvements are illustrated, including filter placement and septum addition for VOC studies and headspace measurements.

The FlowPot system is conducive to high-throughput experimentation given its small form factor, and thus is an attractive platform to perform larger screening experiments. In Chapter 3, we only examined three input communities to determine core microbiome function and core colonization of the plant and substrate. To gain greater statistical power to draw more robust correlations, one could assemble a panel of diverse soil types from a spectrum or gradient of an environmental parameter of choice, and determine to what extend the perturbation of a “transplant” event influences community structure via targeted amplicon sequencing. Presumably, input soil microbiota from peat-like environments would be more amenable to high diversity transfer after colonization in the peat substrate, but this remains to be tested. The FlowPot system, in general, was used as a proof of concept with the Arabidopsis model, and now that this platform has been validated, the research community is encouraged to conduct diverse experiments, cross-check the results from other labs, and establish additional tools to further enhance FlowPots, as well as the field.

Strategies for characterization of axenic immunodeficiency

Based on the experiments conducted in Chapters 3, as well as the dual-RNAseq experiment in axenic and holoxenic Arabidopsis infected with *Pseudomonas syringae*, we can conclude that axenic Arabidopsis is compromised in the ability to defend against *Pst*. The relative magnitude of susceptibility to *Pst* is greater in the *Pst hrcC-* mutant, which is unable to suppress innate immunity. It would be interesting to determine whether the epigenome of axenic plants differs from that of conventionally colonized. There are numerous reports of epigenetic variations in chromatin structure and methylation patterns in DNA that could perhaps be playing a role in host phenotypes. Defense priming and systemic acquired resistance are reported to be, in part, dependent on epigenetic effects on the genome (Conrath et al., 2015; Dowen et al., 2012). It would be interesting to determine if axenic plants have hot spots of relative epigenetic differences as opposed to conventionally colonized plants. An epigenetic-level dissection of microbiota influence over host plants has not been examined yet, and such a study would be a technology-enabled dive into the unknown, perhaps revealing previously undetected layers of plant-microbiome relationships. One could approach this experiment by manipulating the “microbiome variable”, using ~3 reference microbial communities, and perform a time course bs-seq and methyl-seq experiment (Cokus et al., 2008; Hing et al., 2015).

Certain microorganisms will likely have competitive advantages over others, whether they be by antibiosis, niche exclusion, toxin/antitoxin systems, type-VI secretion, direct competition for resources, or other antagonistic characteristics. Rather than starting with a 240 member synthetic community, as is used in Chapter 2, a more practical

approach would be to use a 30-member synthetic community, such as the SynCom used in (Lebeis et al., 2015), but with a high degree of replication to account for stochastic colonization events. Both synCom and At-SPHERE collections are based strictly on taxonomic identification using the 16S rRNA subunit sequences at standard 97% identity cut-offs. One must consider whether these taxa are functionally representative of the Arabidopsis microbiome. Furthermore, the synthetic communities lack fungi, which is appropriate in some experiments, but for prediction of invasion and persistence with robust confidence for success even in a greenhouse, one might consider the addition of fungi or perform the recolonization in an endogenous microbiota background.

An important goal for the plant microbiome research community is to identify characteristics that contribute to or enable a microorganism to invade and persist in a pre-established microbiome. Towards this objective, one could take a microbial genome-wide associated survey approach (Horton et al., 2014) based on many replicates of closely related taxa. The close relatedness of a taxonomic group in question could better equip the scientist to bioinformatically subtract shared among successful and non-successful colonizers, moving towards the objective of identifying functionally relevant genes. It would be prudent to perform these experiments in multiple growth platforms and with multiple background microbiota, because as demonstrated in Chapter 2, substrate and environmental variables influence colonization by certain taxa, and we hypothesize that certain taxa are precluded based on inadequate physical or chemical attributes. The genetic properties of a successful competitive microorganism are valuable information, regardless of whether or not the

microorganism positively influences plant health. Perhaps one has at their disposal a large database of annotated genomes for many microorganisms; multiple selection criteria could be implemented based on HMMs of desired gene sequences. The FlowPot system is conducive to high throughput experimentation with multiple replicates, and would be well suited for a GWAS study to identify functionally relevant genes.

The form factor of the FlowPot system also offers the unique retrofit boxes with septa for headspace measurements, as well as visual growth and photosynthesis measurements. For example, if one were to measure PS-II efficiency fluorometrically as well as growth, the FlowPot system can be used and placed directly in an imaging chamber. The modularity of FlowPots and their containment vessels should maintain sterility or prevent post-inoculation microbiome immigration. The highly controlled and normalized environment of FlowPots and their boxes could potentially enable one to catch more subtle phenotypes.

A major focus of this dissertation was on the characterization of innate immunity deficits and compromised ability to defend against pathogen invasion in axenic conditions, versus colonized by three distinct microbiomes. Although we were able to confirm that all three holoxenic Arabidopsis groups were significantly more immunocompetent than axenic groups, and likewise, inept at defense against *Pst*, it is unknown whether single or multiple members of the microbiome confer this phenotype. For a follow-up experiment, to identify organisms or groups of organisms that confer a desired phenotype, the three-sample panel could be expanded across a gradient or group of similar soils, and introduce only one stress variable. Such an

experiment may provide enough granularity to discern differential effects on pathogen fitness in the host and susceptibility to infection. Through gene expression profiling and metabolic profiling, we were able to identify subtle differences between the colonized plant based upon input community. However since only three input communities were used, the size of this data set is insufficient to draw significant correlations between small taxonomic groups and phenotypic attributes.

In this study, we defined a plant's susceptibility by enumerating precise pathogen population density over time. Pathogen replication in situ was greatly expedited by using the novel flow cytometry enumeration protocol and a th7-GFP-tagged strain. This is an invasive experiment because the leaves we macerated for community extraction. For future experimentation, *Pst* with constitutive lux expression, along with a retrofitted IVIS imager or appropriate CCD camera, one could non-invasively monitor pathogen proliferation by correlating luminescent intensity.

To examine the spectrum of axenic susceptibility, it would be informative to perform disease assays using more pathogens/pests than *Pst*. In particular, it would be informative to perform an insect herbivory assay and disease assays using one or more necrotrophic pathogens. If indeed, axenic are more susceptible to both biotrophs and necrotrophs, this suggests a distinct SA/JA-independent arm of defense signaling is microbiome-influenced. The increased relative abundance of hexanoic acid is interesting in the holoxenic plants. Many of the attributes of axenic susceptibility, or holoxenic resistance, are parallel to literature published about the priming capacity of hexanoic acid (Scalschi et al., 2013). Scalschi et al. demonstrated that hexanoic acid could be used as a biostimulant to prime plants for resistance to necrotrophs and

biotrophs, but not at the expense of growth (Campos et al., 2016). One could further investigate the basis for axenic susceptibility by performing disease assays in JA-insensitive and SA-insensitive backgrounds. Rather than using the wildtype *Pst* DC3000 strain, it would be a cleaner experiment to use a coronatine-deficient *hrcC*- strain. This way the pathogen cannot manipulate hormone signaling directly other than through PTI. I attempted 3-4 disease assay trials in *coi1-16*, *npr1-1* and, *sid2* mutants, but the results were inconclusive. As expected, *coi1-16* was resistant and *npr1-1* and, *sid2* were more susceptible to *Pst,* but the relative magnitude of susceptibility was unhindered (Katsir et al., 2008; Tsuda, 2008; Yu et al., 2001).

One puzzling result that remains unresolved comes from protection assays. BTH protection assays worked very reliable and consistently, always conferring at least 10-fold protection against *Pst*, but flg22 protection assays would not work with the same batch of plants, and a wide range of titrated flg22 concentrations. Finally, after two years of attempts at the hands of 4 different scientists, Dr. Kinya Nomura was able to achieve consistent results in four week old plants, but only with a reduced surfactant concentration (personal communication). This seems to suggest that the original higher concentration of surfactant was masking the effects of PTI priming in the growth system. A previous report ties PTI to apoplastic water availability, and perhaps SA-mediated defense activation (BTH is an SA analog) does not change moisture availability in the apoplast, but PTI does (Xin et al., 2016).

Prior to the development of the FlowPot system, an earlier version of a soil-based axenic growth platform was used to grow and collect 2wo Arabidopsis rosettes for RNA extraction (data not included). We performed mRNA-seq, as well as smRNA-seq,

and identified miR398 as highly abundant in axenic Arabidopsis. This miRNA targets copper superoxide dismutases 2/3, as well as their chaperone. The gene expression from the FlowPot mRNA-seq suggests that indeed, miR398 may be enriched in axenic plants, but this warrants follow up experiments with 398 mutant T-DNA lines (knockout and overexpressor), and qPCR validation of miRNA abundance tracking. miR398 has previously been implicated in innate immunity and defense (Li et al., 2011; Zhu et al., 2011).

A coordinated effort amongst plant microbiome community, as well as adequate financial support from funding agencies and adherence to standards and standard operating protocols will lead to major strides and steps forward and microbiome research. The global population is growing at an alarming rate and our food resources and current agricultural productivity cannot keep pace, especially in areas of the world that need massive technological improvements the most. It is the belief of this author that innovations in microbiome technologies and intelligent application of biostimulants, biopesticides and biofertilizers can offer sustainable efficacious alternatives or supplements to current agricultural practices. With the advent of accessible and inexpensive next-gen sequencing technology platforms at our fingertips and a growing interest in plant microbiome research, perhaps motivated by human microbiome initiatives, we are embarking upon a renaissance of plant microbiome discoveries and innovations.

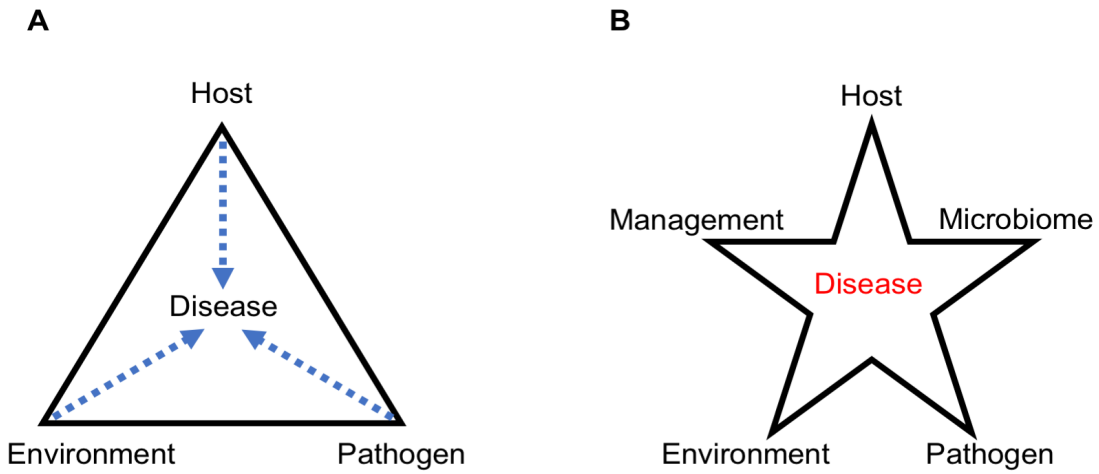


Figure 4.2. Revised disease triangle model, incorporating more vertices that integrate 'management' and 'microbiome'. Plant disease triangle (adapted from Stevens 1960), depicting the factors necessary for a crop disease to occur and develop to an epidemic. A host plant must be developmentally and genetically susceptible to the pathogen. The environment must be conducive to pathogen proliferation, dispersal, and must confer disease-conducive host physiology to accommodate pathogen proliferation. The pathogen must have appropriate virulence factors to replicate to high levels within the host plant. An extension of this model highlights the opportunity for inappropriate management regimes to influence the risk for epidemic, such as import/export of diseased plant material, high-density monoculture cropping systems, and use of one or very few chemicals/traits for pest resistance, thus facilitating the evolution of resistance.

REFERENCES

REFERENCES

Adair KL, Douglas AE. (2017). Making a microbiome: the many determinants of host-associated microbial community composition. Curr Opin Microbiol 35: 23–29.

Adams C, Jacobson A, Bugbee B. (2014). Ceramic Aggregate Sorption and Desorption Chemistry: Implications for Use as a Component of Soilless Media. J Plant Nutr 37: 1345–1357.

Agrios, G.N. (2005). Plant Pathology. 5th eds. Department of Plant Pathology. University of Florida.

Aira, M., Gómez-Brandón, M., Lazcano, C., Bååth, E., and Domínguez, J. (2010). Plant genotype strongly modifies the structure and growth of maize rhizosphere microbial communities. Soil Biol. Biochem. 42, 2276–2281.

Alvarez, S., Rohrig, E., Solís, D., and Thomas, M.H. (2016). Citrus Greening Disease (Huanglongbing) in Florida: Economic Impact, Management and the Potential for Biological Control. Agric. Res. 5, 109–118.

Andrew DR, Fitak RR, Munguia-Vega A, Racolta A, Martinson VG, Dontsova K. (2012). Abiotic factors shape microbial diversity in Sonoran Desert soils. Appl Environ Microbiol 78: 7527–7537.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815.

Arnon DI. (1949). Copper enzymes in isolated chloroplasts. Polyphenoloxidase in beta vulgaris. Plant Physiol 24: 1–15.

Badri, D.V., Zolla, G., Bakker, M.G., Manter, D.K., and Vivanco, J.M. (2013). Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. New Phytol. 198, 264–273.

Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, et al. (2015). Functional overlap of the Arabidopsis leaf and root microbiota. Nature 528: 364–369.

Bai, Y., Müller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., Dombrowski, N., Münch, P.C., Spaepen, S., Remus-Emsermann, M., et al. (2015). Functional overlap of the Arabidopsis leaf and root microbiota. Nature 528, 364–369.

Bank TL, Kukkadapu RK, Madden AS, Ginder-Vogel MA, Baldwin ME, Jardine PM. (2008). Effects of gamma-sterilization on the physico-chemical properties of natural sediments. Chem Geol 251: 1–7.

Baumann, P., Moran, N.A., and Baumann, L. (1997). The Evolution and Genetics of Aphid Endosymbionts. Bioscience 47, 12–20.

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

Berns AE, Philipp H, Narres H-D, Burauel P, Vereecken H, Tappe W. (2008). Effect of gamma-sterilization and autoclaving on soil organic matter structure as studied by solid state NMR, UV and fluorescence spectroscopy. Eur J Soil Sci 59: 540–550.

Birkenbihl, R.P., Liu, S., and Somssich, I.E. (2017). Transcriptional events defining plant immune responses. Curr. Opin. Plant Biol. 38, 1–9.

Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA. (2014). A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. PLoS Genet 10: e1004283.

Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M., and Vorholt, J.A. (2014). A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. PLoS Genet. 10, e1004283.

Boudsocq, M., Willmann, M.R., McCormack, M., Lee, H., Shan, L., He, P., Bush, J., Cheng, S.-H., and Sheen, J. (2010). Differential innate immune signaling via Ca(2+) sensor protein kinases. Nature 464, 418–422.

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 35, 725–731.

Buée, M., Reich, M., Murat, C., Morin, E., Nilsson, R.H., Uroz, S., and Martin, F. (2009). 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. New Phytol. 184, 449–456.

Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, et al. (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 488: 91–95.

Bulgarelli, D., Garrido-Oter, R., Münch, P.C., Weiman, A., Dröge, J., Pan, Y., McHardy, A.C., and Schulze-Lefert, P. (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. Cell Host Microbe 17, 392–403.

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., Rauf, P., Huettel, B., Reinhardt, R., Schmelzer, E., et al. (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 488, 91–95.

Busby PE, Soman C, Wagner MR, Friesen ML, Kremer J, Bennett A, et al. (2017). Research priorities for harnessing plant microbiomes in sustainable agriculture. PLoS Biol 15: e2001793.

Bushnell, B. (2016). BBMap short read aligner. http://sourceforge.net/projects/bbmap

Campos, M.L., Yoshida, Y., Major, I.T., de Oliveira Ferreira, D., Weraduwage, S.M., Froehlich, J.E., Johnson, B.F., Kramer, D.M., Jander, G., Sharkey, T.D., et al. (2016). Rewiring of jasmonate and phytochrome B signalling uncouples plant growth-defense tradeoffs. Nat. Commun. 7, 12570.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6: 1621–1624.

Carvalhais, L.C., Muzzi, F., Tan, C.-H., Hsien-Choo, J., and Schenk, P.M. (2013). Plant growth in Arabidopsis is assisted by compost soil-derived microbial communities. Front. Plant Sci. 4, 235.

Celaj S, Gleeson MW, Deng J, O'Toole GA, Hampton TH, Toft MF, et al. (2014). The microbiota regulates susceptibility to Fas-mediated acute hepatic injury. Lab Invest 94: 938–949.

Chain, P.S.G., Grafham, D.V., Fulton, R.S., Fitzgerald, M.G., Hostetler, J., Muzny, D., Ali, J., Birren, B., Bruce, D.C., Buhay, C., et al. (2009). Genomics. Genome project standards in a new era of sequencing. Science 326, 236–237.

Chang, C., Bowman, J.L., and Meyerowitz, E.M. (2016). Field Guide to Plant Model Systems. Cell 167, 325–339.

Chisholm, S.T., Coaker, G., Day, B., and Staskawicz, B.J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. Cell 124, 803–814.
Chung, K., and Ferris, D.H. (1996). Martinus Willem Beijerinck (1851–1931). J. Am. Med. Assoc.

Ciancio A, Pieterse CMJ, Mercado-Blanco J. (2016). Editorial: Harnessing Useful Rhizosphere Microorganisms for Pathogen and Pest Biocontrol. Front Microbiol 7: 1620.

Clay, N.K., Adio, A.M., Denoux, C., Jander, G., and Ausubel, F.M. (2009). Glucosinolate metabolites required for an Arabidopsis innate immune response. Science 323, 95–101.

Clemente JC, Ursell LK, Parfrey LW, Knight R. (2012). The impact of the gut microbiota on human health: an integrative view. Cell 148: 1258–1270.

Clough SJ, Bent AF. (1998). Transformation of Arabidopsis thaliana. Plant J 16: 735–743.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452, 215–219.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42: D633–42.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., and Tiedje, J.M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42, D633–D642.

Conrath, U., Beckers, G.J.M., Langenbach, C.J.G., and Jaskiewicz, M.R. (2015). Priming for enhanced defense. Annu. Rev. Phytopathol. 53, 97–119.

Couto, D., and Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. Nat. Rev. Immunol. 16, 537–552.

Daniel, B., Pavkov-Keller, T., Steiner, B., Dordic, A., Gutmann, A., Nidetzky, B., Sensen, C.W., van der Graaff, E., Wallner, S., Gruber, K., et al. (2015). Oxidation of Monolignols by Members of the Berberine Bridge Enzyme Family Suggests a Role in Plant Cell Wall Metabolism. J. Biol. Chem. 290, 18770–18781.

Dasgupta, M.K. (1984). The Bengal famine, 1943 and the brown spot of rice--an inquiry into their relations. Hist. Agric. 2, 1–18.

De Bary, A. (1879). Die erscheinung der symbiose.

de Wit, R., and Bouvier, T. (2006). “Everything is everywhere, but, the environment selects”; what did Baas Becking and Beijerinck really say? Environ. Microbiol. 8, 755–758.

DeAngelis, K.M., Lindow, S.E., and Firestone, M.K. (2008). Bacterial quorum sensing and nitrogen cycling in rhizosphere soil. FEMS Microbiol. Ecol. 66, 197–207.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C., and Vorholt, J.A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc. Natl. Acad. Sci. U. S. A. 106, 16428–16433.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C., and Vorholt, J.A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc. Natl. Acad. Sci. U. S. A. 106, 16428–16433.

Dempsey, D.A., Vlot, A.C., Wildermuth, M.C., and Klessig, D.F. (2011). Salicylic Acid biosynthesis and metabolism. Arabidopsis Book 9, e0156.

Dixon, R.A., Harrison, M.J., and Lamb, C.J. (1994). Early Events in the Activation of Plant Defense Responses. Annu. Rev. Phytopathol. 32, 479–501.

Domenech, J., Reddy, M.S., Kloepper, J.W., Ramos, B., and Gutierrez-Mañero, J. (2006). Combined Application of the Biological Product LS213 with Bacillus, *Pseudomonas* or Chryseobacterium for Growth Promotion and Biological Control of Soil-Borne Diseases in Pepper and Tomato. Biocontrol 51, 245.

Dong, X. (2004). NPR1, all things considered. Curr. Opin. Plant Biol. 7, 547–552.

Donn, S., Kirkegaard, J.A., Perera, G., Richardson, A.E., and Watt, M. (2015). Evolution of bacterial communities in the wheat crop rhizosphere. Environ. Microbiol. 17, 610–621.

Doornbos, R.F., Geraats, B.P.J., Kuramae, E.E., Van Loon, L.C., and Bakker, P.A.H.M. (2011). Effects of jasmonic acid, ethylene, and salicylic acid signaling on the rhizosphere bacterial community of Arabidopsis thaliana. Mol. Plant. Microbe. Interact. 24, 395–407.

Dowen, R.H., Pelizzola, M., Schmitz, R.J., Lister, R., Dowen, J.M., Nery, J.R., Dixon, J.E., and Ecker, J.R. (2012). Widespread dynamic DNA methylation in response to biotic stress. Proc. Natl. Acad. Sci. U. S. A. 109, E2183–E2191.

Duclaux, E. (1920). Pasteur: The History of a Mind (W.B. Saunders Company).

Durrant, W.E., and Dong, X. (2004). Systemic acquired resistance. Annu. Rev. Phytopathol. 42, 185–209.

Dutilleul, C., Garmier, M., Noctor, G., Mathieu, C., Chétrit, P., Foyer, C.H., and de Paepe, R. (2003). Leaf mitochondria modulate whole cell redox homeostasis, set antioxidant capacity, and determine stress resistance through altered signaling and diurnal regulation. Plant Cell 15, 1212–1226.

Dworkin, M. (2012). Sergei Winogradsky: a founder of modern microbiology and the first microbial ecologist. FEMS Microbiol. Rev. 36, 364–379.

Edgar RC. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10: 996–998.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27, 2194–2200.

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., Eisen, J.A., and Sundaresan, V. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. Proc. Natl. Acad. Sci. U. S. A. 112, E911–E920.

Fairman CE. (1887). Chemical Relations of Bacteria.

Felix, G., Duran, J.D., Volko, S., and Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J. 18, 265–276.

Fierer N, Jackson RB. (2006). The diversity and biogeography of soil bacterial communities. Proc Natl Acad Sci U S A 103: 626–631.

Garoutte, A. (2016). Identifying the activities of rhizosphere microbial communities using metatranscriptomics.

Gensollen T, Iyer SS, Kasper DL, Blumberg RS. (2016). How colonization by microbiota in early life shapes the immune system. Science 352: 539–544.

Gómez-Gómez, L., and Boller, T. (2000). FLS2: An LRR Receptor–like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis. Mol. Cell 5, 1003–1011.

Gómez-Gómez, L., Felix, G., and Boller, T. (1999). A single locus determines sensitivity to bacterial flagellin in Arabidopsis thaliana. Plant J. 18, 277–284.

Goodman, A.L., Kallstrom, G., Faith, J.J., Reyes, A., Moore, A., Dantas, G., and Gordon, J.I. (2011). Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proc. Natl. Acad. Sci. U. S. A. 108, 6252–6257.

Goodrich JK, Di Rienzi SC, Poole AC, Koren O, Walters WA, Caporaso JG, et al. (2014). Conducting a microbiome study. Cell 158: 250–262.

Görlach, J., Volrath, S., Knauf-Beiter, G., Hengy, G., Beckhove, U., Kogel, K.H., Oostendorp, M., Staub, T., Ward, E., Kessmann, H., et al. (1996). Benzothiadiazole, a novel class of inducers of systemic acquired resistance, activates gene expression and disease resistance in wheat. Plant Cell 8, 629–643.

Gunning T, Cahill DM. (2009). A Soil-free Plant Growth System to Facilitate Analysis of Plant Pathogen Interactions in Roots. Journal of Phytopathology 157: 497–501.

Hacquard, S., and Schadt, C.W. (2015). Towards a holistic understanding of the beneficial interactions across the Populus microbiome. New Phytol. 205, 1424–1430.

Hansen, A.K., Vorburger, C., and Moran, N.A. (2012). Genomic basis of endosymbiont-conferred protection against an insect parasitoid. Genome Res. 22, 106–114.

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

Hayashi, F., Smith, K.D., Ozinsky, A., Hawn, T.R., Yi, E.C., Goodlett, D.R., Eng, J.K., Akira, S., Underhill, D.M., and Aderem, A. (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. Nature 410, 1099–1103.

Henry A, Doucette W, Norton J, Jones S, Chard J, Bugbee B. (2006). An axenic plant culture system for optimal growth in long-term studies. J Environ Qual 35: 590–598.

Hiltner L. (1904). Uber neuere Erfahrungen und Probleme auf dem Gebiet der Bodenbakteriologie und unter besonderer Berucksichtigung der Grundungung and Brache. Arbeiten der Deutschen Landwirtschaftlichen Gesellschaft 98: 59–78.

Hing, B., Ramos, E., Braun, P., McKane, M., Jancic, D., Tamashiro, K.L.K., Lee, R.S., Michaelson, J.J., Druley, T.E., and Potash, J.B. (2015). Adaptation of the targeted capture Methyl-Seq platform for the mouse genome identifies novel tissue-specific DNA methylation patterns of genes involved in neurodevelopment. Epigenetics 10, 581–596.

Hooper LV, Littman DR, Macpherson AJ. (2012). Interactions between the microbiota and the immune system. Science 336: 1268–1273.

Horton, M.W., Bodenhausen, N., Beilsmith, K., Meng, D., Muegge, B.D., Subramanian, S., Vetter, M.M., Vilhjálmsson, B.J., Nordborg, M., Gordon, J.I., et al. (2014). Genome-wide association study of Arabidopsis thaliana leaf microbial community. Nat. Commun. 5, 5320.

Howard MM, Bell TH, Kao-Kniffin J. (2017). Soil microbiome transfer method affects microbiome composition, including dominant microorganisms, in a novel environment. FEMS Microbiol Lett 364. e-pub ahead of print, doi: 10.1093/femsle/fnx092.

Ikeda, S., Kaneko, T., Okubo, T., Rallos, L.E.E., Eda, S., Mitsui, H., Sato, S., Nakamura, Y., Tabata, S., and Minamisawa, K. (2009). Development of a bacterial cell enrichment method and its application to the community analysis in soybean stems. Microb. Ecol. 58, 703–714.

Jackson MB, Abbott AJ, Belcher AR, Hall KC, Butler R, Cameron, J. (AFRC Institute of Arable Crops Research, Long Ashton Research Station, University of Bristol, Bristol BS18 9AF (United Kingdom). (1991). Ventilation in plant tissue cultures and effects of poor aeration on ethylene and carbon dioxide accumulation, oxygen depletion and explant development. Ann Bot 67. http://agris.fao.org/agris-search/search.do?recordID=GB9129134.

Jackson, L.E., Burger, M., and Cavagnaro, T.R. (2008). Roots, nitrogen transformations, and ecosystem services. Annu. Rev. Plant Biol. 59, 341–363.

Johnston-Monje, D., and Raizada, M.N. (2011). Conservation and diversity of seed associated endophytes in Zea across boundaries of evolution, ethnography and ecology. PLoS One 6, e20396.

Jones AM, Lindow SE, Wildermuth MC. (2007). Salicylic acid, yersiniabactin, and pyoverdin production by the model phytopathogen *Pseudomonas syringae* pv. tomato DC3000: synthesis, regulation, and impact on tomato and Arabidopsis host plants. J Bacteriol 189: 6773–6786.

Jung, H.W., Tschaplinski, T.J., Wang, L., Glazebrook, J., and Greenberg, J.T. (2009). Priming in systemic plant immunity. Science 324, 89–91.

Kachroo, A., and Robin, G.P. (2013). Systemic signaling during plant defense. Curr. Opin. Plant Biol. 16, 527–533.

Kadota, Y., Shirasu, K., and Zipfel, C. (2015). Regulation of the NADPH Oxidase RBOHD During Plant Immunity. Plant Cell Physiol. 56, 1472–1480.

Katagiri, F., Thilmony, R., and He, S.Y. (2002). The Arabidopsis thaliana-*pseudomonas syringae* interaction. Arabidopsis Book 1, e0039.

Katsir, L., Schilmiller, A.L., Staswick, P.E., He, S.Y., and Howe, G.A. (2008). COI1 is a critical component of a receptor for jasmonate and the bacterial virulence factor coronatine. Proc. Natl. Acad. Sci. U. S. A. 105, 7100–7105.

Kau AL, Ahern PP, Griffin NW, Goodman AL, Gordon JI. (2011). Human nutrition, the gut microbiome and the immune system. Nature 474: 327–336.

Kembel, S.W., O'Connor, T.K., Arnold, H.K., Hubbell, S.P., Wright, S.J., and Green, J.L. (2014). Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. Proc. Natl. Acad. Sci. U. S. A. 111, 13715–13720.

Kloepper JW, Schroth - Phytopathology MN, 1981. (1981). Plant growth-promoting rhizobacteria and plant growth under gnotobiotic conditions. apsnet.org. https://www.apsnet.org/publications/phytopathology/backissues/Documents/1981Articles/Phyto71n06_642.pdf.

Knief, C., Delmotte, N., Chaffron, S., Stark, M., Innerebner, G., Wassmann, R., von Mering, C., and Vorholt, J.A. (2012). Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. ISME J. 6, 1378–1390.

Kniskern JM, Traw MB, Bergelson J. (2007). Salicylic acid and jasmonic acid signaling defense pathways reduce natural bacterial diversity on Arabidopsis thaliana. Mol Plant Microbe Interact 20: 1512–1522.

Krishnakumar, V., Hanlon, M.R., Contrino, S., Ferlanti, E.S., Karamycheva, S., Kim, M., Rosen, B.D., Cheng, C.-Y., Moreira, W., Mock, S.A., et al. (2015). Araport: the Arabidopsis information portal. Nucleic Acids Res. 43, D1003–D1009.

Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T., and Felix, G. (2004). The N terminus of bacterial elongation factor Tu elicits innate immunity in Arabidopsis plants. Plant Cell 16, 3496–3507.

Laemmli. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Springer. http://link.springer.com/article/10.1038/227680a0.

Lau JA, Lennon JT. (2012). Rapid responses of soil microorganisms improve plant fitness in novel environments. Proc Natl Acad Sci U S A 109: 14058–14062.

Lebeis, S.L., Paredes, S.H., Lundberg, D.S., Breakfield, N., Gehring, J., McDonald, M., Malfatti, S., del Rio, T.G., Jones, C.D., Tringe, S.G., et al. (2015). Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. Science 349, 860–864.

Lee, B., Farag, M.A., Park, H.B., Kloepper, J.W., Lee, S.H., and Ryu, C.-M. (2012). Induced resistance by a long-chain bacterial volatile: elicitation of plant systemic defense by a C13 volatile produced by Paenibacillus polymyxa. PLoS One 7, e48744.

Leslie, M.E., and Heese, A. (2014). A re-elicitation assay to correlate flg22-signaling competency with ligand-induced endocytic degradation of the FLS2 receptor. Methods Mol. Biol. 1209, 149–162.

Ley RE, Turnbaugh PJ, Klein S, Gordon JI. (2006). Microbial ecology: human gut microbes associated with obesity. Nature 444: 1022–1023.

Li, Y., Tschaplinski, T.J., Engle, N.L., Hamilton, C.Y., Rodriguez, M., Jr, Liao, J.C., Schadt, C.W., Guss, A.M., Yang, Y., and Graham, D.E. (2012). Combined inactivation of the Clostridium cellulolyticum lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. Biotechnol. Biofuels 5, 2.

Li, Y., Wang, W., and Zhou, J.-M. (2011). Role of small RNAs in the interaction between Arabidopsis and *Pseudomonas syringae*. Front. Biol. 6, 462–467.

Ligterink, W., Kroj, T., zur Nieden, U., Hirt, H., and Scheel, D. (1997). Receptor-mediated activation of a MAP kinase in pathogen defense of plants. Science 276, 2054–2057.

Lloyd, A.M., Barnason, A.R., Rogers, S.G., Byrne, M.C., Fraley, R.T., and Horsch, R.B. (1986). Transformation of Arabidopsis thaliana with Agrobacterium tumefaciens. Science 234, 464–466.

Loper JE, Buyer - Mol. Plant-Microbe Interact JS, 1991. (1991). Siderophores in microbial interactions on plant surfaces. apsnet.org. https://www.apsnet.org/publications/mpmi/BackIssues/Documents/1991Articles/Microbe04_005.pdf.

Lotrario JB, Stuart BJ, Lam T, Arands RR, O'Connor OA, Kosson DS. (1995). Effects of sterilization methods on the physical characteristics of soil: implications for sorption isotherm analyses. Bull Environ Contam Toxicol 54: 668–675.

Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. (2011). UniFrac: an effective distance metric for microbial community comparison. ISME J 5: 169–172.

Luckey TD. (1965). Effects of Microbes on Germfree Animals11Presented in modified form as the principal talk at the International Symposium on Microecology, Berlin, September, 1964. In: Umbreit WW (ed) Vol. 7. Advances in Applied Microbiology. Academic Press, pp 169–223.

Luna, E., Pastor, V., Robert, J., Flors, V., Mauch-Mani, B., and Ton, J. (2011). Callose deposition: a multifaceted plant defense response. Mol. Plant. Microbe. Interact. 24, 183–193.

Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. (2012). Defining the core Arabidopsis thaliana root microbiome. Nature 488: 86–90.

Ma L-S, Hachani A, Lin J-S, Filloux A, Lai E-M. (2014). Agrobacterium tumefaciens deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. Cell Host Microbe 16: 94–104.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., et al. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. Mol. Plant Pathol. 13, 614–629.

Marques, M.R., Buckeridge, M.S., Braga, M.R., and Dietrich, S.M.C. (2006). Characterization of an extracellular endopolygalacturonase from the saprobe Mucor ramosissimus Samutsevitsch and its action as trigger of defensive response in tropical plants. Mycopathologia 162, 337–346.

Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G., and Neufeld, J.D. (2012). PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics 13, 31.

McCoy KD, Geuking MB, Ronchi F. (2017). Gut Microbiome Standardization in Control and Experimental Mice. Curr Protoc Immunol 117: 23.1.1–23.1.13.

McMurdie PJ, Holmes S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8: e61217.

Mendes, L.W., Kuramae, E.E., Navarrete, A.A., van Veen, J.A., and Tsai, S.M. (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. ISME J. 8, 1577–1587.

Meyerowitz, E.M. (2001). Prehistory and history of Arabidopsis research. Plant Physiol. 125, 15–19.

Moran, N.A. (2001). The Coevolution of Bacterial Endosymbionts and Phloem-Feeding Insects. Ann. Mo. Bot. Gard. 88, 35–44.

Mueller UG, Sachs JL. (2015). Engineering Microbiomes to Improve Plant and Animal Health. Trends Microbiol 23: 606–617.

Nicklas W, Keubler L, Bleich A. (2015). Maintaining and Monitoring the Defined Microbiota Status of Gnotobiotic Rodents. ILAR J 56: 241–249.

Nomura, K., Mecey, C., Lee, Y.-N., Imboden, L.A., Chang, J.H., and He, S.Y. (2011). Effector-triggered immunity blocks pathogen degradation of an immunity-associated vesicle traffic regulator in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 108, 10774–10779.

Panke-Buisse, K., Poole, A.C., Goodrich, J.K., Ley, R.E., and Kao-Kniffin, J. (2015). Selection on soil microbiomes reveals reproducible impacts on plant function. ISME J. 9, 980–989.

Pankratov TA, Serkebaeva YM, Kulichevskaya IS, Liesack W, Dedysh SN. (2008). Substrate-induced growth and isolation of Acidobacteria from acidic Sphagnum peat. ISME J 2: 551–560.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25, 1043–1055.

Pasteur, L. (1885). Observations relatives à la note précédente de M. Duclaux. C.R. Acad. Sci. 100, 68.

Peiffer, J.A., Spor, A., Koren, O., Jin, Z., Tringe, S.G., Dangl, J.L., Buckler, E.S., and Ley, R.E. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. Proc. Natl. Acad. Sci. U. S. A. 110, 6548–6553.

Pfeilmeier, S., Caly, D.L., and Malone, J.G. (2016). Bacterial pathogenesis of plants: future challenges from a microbial perspective: Challenges in Bacterial Molecular Plant Pathology. Mol. Plant Pathol. 17, 1298–1313.

Pieterse, C.M.J., Zamioudis, C., Berendsen, R.L., Weller, D.M., Van Wees, S.C.M., and Bakker, P.A.H.M. (2014). Induced systemic resistance by beneficial microbes. Annu. Rev. Phytopathol. 52, 347–375.

Preston, G., Deng, W.L., Huang, H.C., and Collmer, A. (1998). Negative regulation of hrp genes in *Pseudomonas syringae* by HrpV. J. Bacteriol. 180, 4532–4537.

Provart, N.J., Alonso, J., Assmann, S.M., Bergmann, D., Brady, S.M., Brkljacic, J., Browse, J., Chapple, C., Colot, V., Cutler, S., et al. (2016). 50 years of Arabidopsis research: highlights and future directions. New Phytol. 209, 921–944.

Ruinen, J. (1961). The phyllosphere. Plant Soil 15, 81–109.

Ryffel, F., Helfrich, E.J.N., Kiefer, P., Peyriga, L., Portais, J.-C., Piel, J., and Vorholt, J.A. (2016). Metabolic footprint of epiphytic bacteria on Arabidopsis thaliana leaves. ISME J. 10, 632–643.

Savary, S., Ficke, A., Aubertot, J.-N., and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. Food Sec. 4, 519–537.

Scalschi, L., Vicedo, B., Camañes, G., Fernandez-Crespo, E., Lapeña, L., González-Bosch, C., and García-Agustín, P. (2013). Hexanoic acid is a resistance inducer that protects tomato plants against *Pseudomonas syringae* by priming the jasmonic acid and salicylic acid pathways. Mol. Plant Pathol. 14, 342–355.

Schmidt, P.-A., Bálint, M., Greshake, B., Bandow, C., Römbke, J., and Schmitt, I. (2013). Illumina metabarcoding of a soil fungal community. Soil Biol. Biochem. 65, 128–132.

Scholthof, K.-B.G. (2007). The disease triangle: pathogens, the environment and society. Nat. Rev. Microbiol. 5, 152–156.

Sequeira, L. (1988). On becoming a plant pathologist: the changing scene. Annu. Rev. Phytopathol. 26, 1–14.

Shaw LJ, Beaton Y, Glover LA, Killham K, Meharg AA. (1999). Re-inoculation of autoclaved soil as a non-sterile treatment for xenobiotic sorption and biodegradation studies. Appl Soil Ecol 11: 217–226.

Silby, M.W., Cerdeño-Tárraga, A.M., Vernikos, G.S., Giddens, S.R., Jackson, R.W., Preston, G.M., Zhang, X.-X., Moon, C.D., Gehrig, S.M., Godfrey, S.A.C., et al. (2009). Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas* fluorescens. Genome Biol. 10, R51.

Smith K, McCoy KD, Macpherson AJ. (2007). Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. Semin Immunol 19: 59–69.

Spor A, Koren O, Ley R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. Nat Rev Microbiol 9: 279–290.

Stanton-Geddes, J., Paape, T., Epstein, B., Briskine, R., Yoder, J., Mudge, J., Bharti, A.K., Farmer, A.D., Zhou, P., Denny, R., et al. (2013). Candidate genes and genetic architecture of symbiotic and agronomic traits revealed by whole-genome, sequence-based association genetics in Medicago truncatula. PLoS One 8, e65688.

Stevens, R.B. (1960). Cultural practices in disease control. Plant Pathol.

Swingle, W.T. (1896). Bordeaux Mixture: Its Chemistry, Physical Properties, and Toxic Effects on Fungi and Algae, by Walter T. Swingle.

Tata, J.R. (2005). From oxidative phosphorylation to transcription--a postdoctoral adventure. Trends Biochem. Sci. 30, 529–534.

Tatum, L.A. (1971). The southern corn leaf blight epidemic. Science 171, 1113–1116.

Teal, T.K., Lies, D.P., Wold, B.J., and Newman, D.K. (2006). Spatiometabolic stratification of Shewanella oneidensis biofilms. Appl. Environ. Microbiol. 72, 7324–7330.

Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., Prill, R.J., Tripathi, A., Gibbons, S.M., Ackermann, G., et al. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551, 457–463.

Thuerig, B., Felix, G., Binder, A., Boller, T., and Tamm, L. (2005). An extract of Penicillium chrysogenum elicits early defense-related responses and induces resistance in Arabidopsis thaliana independently of known signalling pathways. Physiol. Mol. Plant Pathol. 67, 180–193.

Tian, B.-Y., Cao, Y., and Zhang, K.-Q. (2015). Metagenomic insights into communities, functions of endophytes, and their associates with infection by root-knot nematode, Meloidogyne incognita, in tomato roots. Sci. Rep. 5, 17087.

Traw, M.B., Kniskern, J.M., and Bergelson, J. (2007). SAR increases fitness of Arabidopsis thaliana in the presence of natural bacterial pathogens. Evolution 61, 2444–2449.

Trevors JT. (1996). Sterilization and inhibition of microbial activity in soil. J Microbiol Methods 26: 53–59.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., Podar, M., Short, J.M., Mathur, E.J., Detter, J.C., et al. (2005). Comparative metagenomics of microbial communities. Science 308, 554–557.

Tschaplinski, T.J., Standaert, R.F., Engle, N.L., Martin, M.Z., Sangha, A.K., Parks, J.M., Smith, J.C., Samuel, R., Jiang, N., Pu, Y., et al. (2012). Down-regulation of the caffeic acid O-methyltransferase gene in switchgrass reveals a novel monolignol analog. Biotechnol. Biofuels 5, 71.

Tsuda, K., Mine, A., Bethke, G., Igarashi, D., Botanga, C.J., Tsuda, Y., Glazebrook, J., Sato, M., and Katagiri, F. (2013). Dual regulation of gene expression mediated by extended MAPK activation and salicylic acid contributes to robust innate immunity in Arabidopsis thaliana. PLoS Genet. 9, e1004015.

Turner, B.L., and Haygarth, P.M. (2001). Biogeochemistry. Phosphorus solubilization in rewetted soils. Nature 411, 258.

Van Wees, S., and De Swart, E. (2000). Enhancement of induced disease resistance by simultaneous activation of salicylate-and jasmonate-dependent defense pathways in Arabidopsis thaliana. Proceedings of the.

Vogei C., Bodenhausen N., Gruissem W., Vorholt J.A. (2016). The Arabidopsis leaf transcriptome reveals distinct but also overlapping responses to colonization by phyllosphere commensals and pathogen infection with impact on plant health. New Phytol 212: 192–207.

Wagner, M.R., Lundberg, D.S., Coleman-Derr, D., Tringe, S.G., Dangl, J.L., and Mitchell-Olds, T. (2014). Natural soil microbes alter flowering phenology and the intensity of selection on flowering time in a wild Arabidopsis relative. Ecol. Lett. 17, 717–726.

Whalen, M.C., Innes, R.W., Bent, A.F., and Staskawicz, B.J. (1991). Identification of *Pseudomonas syringae* pathogens of Arabidopsis and a bacterial locus determining avirulence on both Arabidopsis and soybean. Plant Cell 3, 49–59.

Wildermuth, M.C., Dewdney, J., Wu, G., and Ausubel, F.M. (2001). Isochorismate synthase is required to synthesize salicylic acid for plant defense. Nature 414, 562–565.

Withers, J., and Dong, X. (2017). Post-translational regulation of plant immunity. Curr. Opin. Plant Biol. 38, 124–132.

Wolf DC, Dao TH, Scott HD, Lavy TL. (1989). Influence of Sterilization Methods on Selected Soil Microbiological, Physical, and Chemical Properties. J Environ Qual 18: 39–44.

Xin, X.-F., and He, S.Y. (2013). *Pseudomonas syringae* pv. tomato DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. Annu. Rev. Phytopathol. 51, 473–498.

Xin, X.-F., Nomura, K., Aung, K., Velásquez, A.C., Yao, J., Boutrot, F., Chang, J.H., Zipfel, C., and He, S.Y. (2016). Bacteria establish an aqueous living space in plants crucial for virulence. Nature 539, 524–529.

Yu, D., Chen, C., and Chen, Z. (2001). Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. Plant Cell 13, 1527–1540.

Zeng, W., Brutus, A., Kremer, J.M., Withers, J.C., Gao, X., Jones, A.D., and He, S.Y. (2011). A genetic screen reveals Arabidopsis stomatal and/or apoplastic defenses against *Pseudomonas syringae* pv. tomato DC3000. PLoS Pathog. 7, e1002291.

Zgadzaj R, Garrido-Oter R, Jensen DB, Koprivova A, Schulze-Lefert P, Radutoiu S. (2016). Root nodule symbiosis in Lotus japonicus drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. Proc Natl Acad Sci U S A 113: E7996–E8005.

Zhou, J., and Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? Microbiol. Mol. Biol. Rev. 81.

Zhou, M., Lu, Y., Bethke, G., Harrison, B.T., Hatsugai, N., Katagiri, F., and Glazebrook, J. (2017). WRKY70 prevents axenic activation of plant immunity by direct repression of SARD1. New Phytol.

Zhou, N., Tootle, T.L., Tsui, F., Klessig, D.F., and Glazebrook, J. (1998). PAD4 functions upstream from salicylic acid to control defense responses in Arabidopsis. Plant Cell 10, 1021–1030.

Zhu, C., Ding, Y., and Liu, H. (2011). MiR398 and plant stress responses. Physiol. Plant. 143, 1–9.