DISSERTATION

MANIPULATING THE SOIL MICROBIOME TO INCREASE PLANT HEALTH AND PRODUCTIVITY

Submitted by

Jacqueline Michelle Chaparro

Department of Horticulture and Landscape Architecture

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2015

Doctoral Committee:

Advisor: Jorge M Vivanco

Jan E Leach
Daniel K Manter
Stephen J Wallner

ABSTRACT

MANIPULATING THE SOIL MICROBIOME TO INCREASE PLANT HEALTH AND PRODUCTIVITY

Rhizosphere microbial communities offer immense benefits to plants. The rhizomicrobiome has the ability to help combat numerous biotic and abiotic stresses as well as increase plant health and productivity. In a world where the population keeps increasing at an alarming rate while food is scarce, new alternatives to feed the growing population need to be identified. The answer lies in harnessing and exploiting the beneficial interactions between plants and their rhizosphere microbiome to increase plant health and productivity. An understanding of the mechanisms that govern such interactions is essential to increase plant health and productivity.

Based on this need, an analysis of the interactions between *Arabidopsis thaliana* and its rhizosphere microbial community was undertaken. Initial studies revealed that root exudates serve as a means of initiating, attracting, maintaining, and enhancing rhizosphere microbial community interactions. Furthermore, root exudation changes with development and leads to changes in the functional capacity and the members that make up the rhizosphere microbial community. These changes appear to occur so the plant can recruit specific functions necessary for survival.

Once a framework outlining the importance of root exudation on plant-microbiome interactions was established, compounds from root exudates were added to soil, without the plant, and tested its impact on the soil microbiome. Studies revealed that these compounds when

acting alone do in fact influence the soil microbiome and that distinct chemical classes have a direct influence on the soil microbial community. Most importantly, correlation analysis of microbes and the phytochemicals added to the soil revealed that phenolic compounds appear to predominantly modulate the soil microbial community.

Finally, the knowledge acquired from these studies allowed development of statistical models that could predict the specific influence of root exudate compounds on the soil microbiome. Five statistical models were implemented, tested, and validated. These results identified models based on machine learning to be of great value in their ability to accurately predict the behavior of soil microbial community abundance after exposure to specific compounds.

Overall, the results of this dissertation enable the ability to begin to modulate and manipulate the soil microbial community for increased plant health and productivity.

ACKNOWLEDGEMENTS

I would like to extend my sincerest gratitude to my adviser, Dr. Jorge Vivanco, and the members of my graduate committee whose support throughout this intellectual endeavor has been invaluable: Dr. Daniel Manter, Dr. Jan Leach, and Dr. Stephen Wallner. I am grateful to the National Science Foundation Alliances for Graduate Education and the Professoriate (AGEP) for their funding and support during my graduate career at Colorado State University.

I would also like to thank everyone who has made my time at Colorado State University an exciting and incredible learning adventure which was filled with plenty of twists, turns, and life experiences. Special thanks to the members of the Vivanco lab for all the support, laughs, and constant help. Additionally, I would like to thank Dr. Dayakar Badri for his constant support and assistance. To Brittany Barnett and Amy Sheflin, thank you for each and every happy hour we spent together; it was always a welcome break and made the difficult times bearable. Viviana Betancourt, I want to thank you for all the help, edits, and comedic relief provided during this endeavor.

Finally, to my parents, Martha and Luis Chaparro, and my brother, Kevin Chaparro, thank you for your unwavering support and devotion throughout my life. Your encouragement is what kept me focused on the finish line. To my husband, Steven Russo, thank you for keeping me grounded and knowing just what to say during those moments when things seemed insurmountable. To the entire Russo clan, thank you for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTAL FILES

Supplementary Material 2-1. Hierarchical classification of the functional genes in each sample they are categorized under KEGG orthology.

Supplementary Material 2-2. Pearson correlation analysis of the functional genes involved in Amino Acid Metabolism with root exudate compounds classified as amino acids. Values present significantly correlate at $p<0.05$.

Supplementary Material 2-3. Pearson correlation analysis of the functional genes involved in Metabolism of Secondary Metabolites which includes the Biosynthesis of Other Secondary Metabolites and Metabolism of Terpenoids and Polyketides with root exudate compounds classified as phenolics. Values present significantly correlate at $p<0.05$.

Supplementary Material 2-4. Pearson correlation analysis of the functional genes involved Carbohydrate Metabolism with root exudate compounds classified as sugars. Values present significantly correlate at $p<0.05$.

Supplementary Material 3-1. List of OTUs with the number of sequence reads for each soil microbial community and their taxonomic assignment present at each developmental time point.

Supplementary Material 3-2. Pearson correlation analysis of the bacterial OTUs classified as Acidobacteria, Actinobacteria, Bacteroidetes and Cyanobacteria with compounds released as root exudates. Values shown are statistically significant at $p<0.05$.

Supplementary Material 3-3. Pearson correlation analysis of the statistically significant transcripts (81) with compounds released as root exudates. Values shown are statistically significant at $p<0.05$.

Supplementary Material 4-1. List of OTUs with abundance (number of sequence reads) and their taxonomic assignments present in the controls and treatments.

Supplementary Material 4-2. Correlation analyses of groups of compounds (sugars, sugar alcohols, amino acids and phenolics) present in the whole exudates, ethylacetate fraction, chloroform fraction and water.

Supplementary Material 4-3. Correlation analyses of groups of compounds (sugars, sugar alcohols, amino acids and phenolics) present in the whole exudates, ethylacetate fraction, chloroform fraction and water fraction with soil bacterial sequencing data at the (OTU) genus level. The cells representing values are significant at p value 0.05 and blank cells are non-significant.

Supplementary Material 5-1. List of OTUs with abundance (number of sequence reads) and their taxonomic assignments present in the controls and treatments of *Arabidopsis* co-adapted soil after exposure to the chemical libraries for 2 or 6 weeks.

Supplementary Material 5-2. List of OTUs with abundance (number of sequence reads) and their taxonomic assignments present in the controls and treatments of *Arabidopsis* co-adapted soil after exposure to the artificial chemical blends.

CHAPTER 1 MANIPULATING THE SOIL MICROBIOME TO INCREASE SOIL HEALTH AND PLANT FERTILITY[1]

Synopsis

A variety of soil factors are known to increase nutrient availability and plant productivity. The most influential might be the organisms comprising the soil microbial community of the rhizosphere, which is the soil surrounding the roots of plants where complex interactions occur between the roots, soil, and microorganisms. Root exudates act as substrates and signaling molecules for microbes creating a complex and interwoven relationship between plants and the microbiome. While individual microorganisms such as endophytes, symbionts, pathogens, and plant growth promoting rhizobacteria are increasingly featured in the literature, the larger community of soil microorganisms, or soil microbiome, may have more far-reaching effects. Each microorganism functions in coordination with the overall soil microbiome to influence plant health and crop productivity. Increasing evidence indicates that plants can shape the soil microbiome through the secretion of root exudates. The molecular communication fluctuates according to the plant development stage, proximity to neighboring species, management techniques, and many other factors. This review seeks to summarize the current knowledge on this topic.

Introduction

The relationship between plants and their surroundings is a complex one that, for centuries, has been the focus of much research. For much of this time, the focus was on the

[1] Reprinted from Chaparro, JM;Sheflin, AM, Manter, DK; Vivanco, JM, Manipulating the soil microbiome to increase soil health and plant fertility Biol Fert Soils 2012, 48, 489-499, DOI: 10.1007/s00374-012-0691-4. Reproduced with permission from Springer

plants with little focus on the beneficial plant-microbe interactions (Berg 2009). Plants and microbes have evolved intimate relationships that enable them to coexist (Nihorimbere et al 2011). Many experiments have tried to answer these questions by simplifying the interactions that occur to an individual plant-microbe relationship, but in reality, these interactions are much more complex, involve a vast array of microbes, and often produce synergistic effects (Mendes et al 2011). It is necessary to move away from the potentially simplistic view of individual plant-microbe interactions and take into account all the factors that influence this complex ecosystem. The plant, the soil, and the soil microbes all work together to mediate and influence the various exchanges (see Figure 1-1) that contribute to plant health and productivity. An understanding of how each component manipulates and influences each other is needed. Recent advances in "-omics" research can help us answer these questions and allow us to see how all these interactions relate and influence one another (Morales and Holben 2011). Here we focus on the soil microbiome and its impact on plant health and productivity.

The effect of plants on the soil microbiome

The interaction between plants and their surroundings is a dynamic process in which plants monitor their environment and react to changes. The root system, which was traditionally thought to only provide anchorage and uptake of nutrients and water, is a key element to a plant interacting with its surroundings (Bais et al 2006). Chemical signals emitted by soil microorganisms are received and recognized by plants and then addressed through the release of chemical compounds in the form of root exudates. Secretion of these compounds varies between different plant species (Rovira 1969), ecotypes (Micallef et al 2009), and even distinct roots within a plant (Uren 2007). The diverse compounds released by plants as root exudates create a unique environment in the rhizosphere and include sugars, amino acids, flavonoids, aliphatic

acids, proteins, and fatty acids (Badri et al 2009b). All these different compounds are able to attract and initiate both symbiotic and pathogenic interactions within the rhizosphere (Bais et al 2006). Root exudate composition and concentration change according to the signals received from the environment and the rhizosphere, age of the plant (De-la-Pena et al 2010), soil type (Rovira 1969), and biotic and abiotic factors (Flores et al 1999, Tang et al 1995). For example, De-la-Pena et al. (2010) observed that the protein composition of the root exudates changed when the plant grew alone as compared to when the plant interacted with pathogens or with symbiotic microbes. Root exudates are used as growth substrates (Vandenkoornhuyse et al 2007) by soil microbes and can act as antimicrobials (Bais et al 2006, Perry et al 2007); therefore, as the composition and concentration of the exudates change, so do the microbes that inhabit the rhizosphere (Badri et al 2009a, Micallef et al 2009).

Root exudates are released by a variety of mechanisms. Diffusion, ion channels, and vesicle transport are the primary mechanisms of root exudation and require little to no energy input (Bertin et al 2003, Neuman and Romheld 2007). Recently it has been demonstrated that ATP-binding cassette (ABC) transporters are also involved in root exudation (Badri et al 2008). Micallef et al. (2009) demonstrated that naturally occurring ecotype accessions of *Arabidopsis* exuded a unique suite of compounds into the rhizosphere, with each genotype supporting a different soil bacterial community. This is a clear example of how root exudates can have a significant effect on the soil microbiome. Rhizodeposition, which encompasses border cells, root debris, and root exudates is the major sources of organic C to enter the soil (Uren 2007). It comes at a high C cost to the plant, with young seedlings typically releasing about 30-40% of their fixed C (Whipps 1990). Why would the plant use such a large percentage of its energy to produce and release these rhizodeposits? Perhaps it is to attract microorganisms that service the plant through

secreting growth promoting hormones, preventing disease, or acquiring nutrients via the excretions of a biochemically active root system. For example, Hamilton and Frank (2001) demonstrated that a grazing tolerant grass, *Poa pratensis*, is capable of concentrating microbes that facilitate the uptake of a limiting soil resource needed for growth, in this case N, in its rhizosphere when under herbivore attack. White lupin, on the other hand, is able to discourage microbial growth by drastically decreasing the soil pH in the rhizosphere via the release of organic acids, lowering the competition for P acquisition (Weisskopf et al 2006). At the same time, white lupin prevents microbial degradation of root exudates important for P acquisition (Weisskopf et al 2006). The diversity and relative number of soil microbes was found to decrease with closer proximity to the rhizosphere of the invasive weed *Centaurea maculosa* and changes in the soil microbial community extended to neighboring native grass species (Broz et al 2007). The invasive weed *Chromolaena odorata* has been shown to accumulate high concentrations of native soil pathogenic fungi inhibiting the growth of the native plant species (Mangla and Callaway 2008). Other studies have demonstrated that similar disruptions in the microbial communities of native plants benefits the invasive species while diminishing the success of native plants (Klironomos 2002, Stinson et al 2006). The altered soil microbiome appears to be a significant part of the strategy for invasive weeds to increase its own resources and exploit weaknesses in the native plant.

Soil properties influence microbial diversity

The forces that shape the rhizosphere microbial community cannot be completely understood without a discussion of the influences of the soil environment. Soils are highly diverse allowing for habitation by equally diverse communities of microorganisms with as many as 10,000-50,000 species of microbes existing in 1 gram of soil (Schloss and Handelsman 2006).

Unique bacterial and fungal communities have been associated with soils of varying texture (Girvan et al 2003), N content (Frey et al 2004), P content (Faoro et al 2010), and soil pH (Fierer and Jackson 2006, Lauber et al 2008, Rousk et al 2010). Recent evidence suggests that out of all these factors, soil pH may have the most influence on the bacterial community in the soil (Fierer and Jackson 2006). Rousk et al. (2010) collected soil samples across a long-term liming experiment where the pH varied from 4.0-8.3 while all other factors and variables that compose soil variability were controlled. A strong correlation between soil pH and the diversity and composition of bacterial communities was seen across biomes and was a greater driver of bacterial community composition than dispersal limitations (Rousk et al 2010). It is hypothesized that the reason for this connection between pH and soil bacterial community structure has to do with the sensitivity of bacterial cells to pH, as bacterial taxa exhibit a relatively narrow pH growth tolerance (Rousk et al 2010). Other evidence refutes pH as a driver for soil microbial diversity, indicating that P content, altitude, and the ratio of cations in the soil (Ca^{2+}, Mg^{2+}, and Al^{3+}) are more influential (Faoro et al 2010). Clearly, many influences converge to create the ultimate effect on the soil microbial community and multiple soil factors potentially exhibit synergistic effects.

Although soil factors provide a strong influence on microbial communities, root exudates have been shown to also strongly influence the soil microbial community. Close ties between the composition of soil microbial communities and host plants were found (Broeckling et al 2008). Soil fungal communities changed composition and decreased in total biomass after planting with a non-native model plant or applying the non-native plant's exudates (Broeckling et al 2008). Two model plant species, *Arabidopsis thaliana* and *Medicago truncatula*, were grown in their native soil and in the other plant's soil, non-native. *Arabidopsis* plants or root exudates added

alone maintained the native fungal community in its native soil but not in non-native soil. In non-native soil, some microbial species increased while others diminished. Total fungal biomass was also affected when treated with root exudates alone or grown with *Arabidopsis* plants. The same was observed with *Medicago*. These results strongly suggest that plant root exudates and, therefore, plants themselves are able to affect the composition and total population of soil microflora.

Role of soil microbes in soil health and plant productivity

The purpose of this review is not to cover all of the beneficial effects of soil microbes on plant health and the associated mechanisms of action; these have been covered adequately in other reviews (Babalola 2010, Cummings 2009, Esitken 2011, Lugtenberg and Kamilova 2009, Maheshwari 2011). Instead, the aim is to highlight some of the more recent advances made in this rapidly developing field and emphasize potential practical applications for sustainable and integrated approaches to agriculture. For example, adding beneficial microorganisms to those already present in the soil can maximize plant nutrient uptake (Kirankumar et al 2008), increase plant growth (Cummings 2009, Guiñazú et al 2009), confer resistance to abiotic stress (Selvakumar et al 2012), and suppress disease (De Vleesschauwer and Höfte 2009). These living microorganisms are dynamic and potentially self-sustaining, reducing the need for repeated applications, and can avoid the problem of pests and pathogens evolving resistance to the treatments (Lucas 2011). A possible management technique is to apply plant growth promoting rhizobacteria (PGPRs) as an agricultural treatment to minimize niche vacancy and effectively fill vacant niches. It has been shown that PGPRs colonize particularly and effectively in soils with low microbial biomass (Fliessbach et al 2009) so inoculations are more likely to be successful. Beneficial microorganisms that thrive in this environment can more quickly take up space and

nutrients made available for potential pathogen invaders and assist with achieving sustained niche occupancy (Kaymak and Maheshwari 2010). In addition to "sealing off" open ecological niches and increasing the soil's resistance to pathogen invasion, PGPRs offer benefits of increased yields, nutrient acquisition, stress tolerance, and disease resistance to the plant host (Lugtenberg and Kamilova 2009). As an example of the potential of microbial inoculation, consider the outcome of a greenhouse study using tomato plants inoculated with PGPR and mycorrhizal fungi. It showed that inoculated plants that received less than 75% the full rate of fertilizer had yields identical to uninoculated plants that received full fertilizer treatments (Adesemoye et al 2009). Furthermore, an awareness of the existing soil fertility level is critical to realizing PGPR benefits, as a diminishing effect is seen when starting N, P, and K levels are high (Shaharoona et al 2008).

Recent discoveries have shown that plants also interact with a variety of PGPRs that are capable of increasing photosynthetic capacity (Xie et al 2009, Zhang et al 2008b), conferring drought and salt tolerance (Dimkpa et al 2009, Xie et al 2009, Zhang et al 2008a, 2009a, 2010), increase disease suppression (Chithrashree et al 2011, Jetiyanon and Kloepper 2002, Okubara and Bonsall 2008), plant growth (Hayat et al 2010, Lim and Kim 2009), and improving the effectiveness of the plant's own iron acquisition mechanisms (Zhang et al 2009a). These discoveries may offer potential for PGPR applications to improve agricultural production and sustainability. Currently, producers are faced with a need to reduce inputs like water and fertilizer applications while simultaneously increasing production. In addition, these PGPR traits promise considerable value in biofuel cropping considering the need to produce biofuel crops in areas unsuitable for agricultural production (Tilman et al 2009) where drought and salt tolerance may become especially important. Given the wide variety of effects and mechanisms of action,

it's not surprising that a combination of PGPR treatments has been shown to be even more effective than one treatment alone in suppressing disease (Ahemad and Khan 2011, Berg 2009, Pérez-Piqueres et al 2006, Yang et al 2011). One example of combined inoculations includes the PGPR *Pseudomonas putida* added in combination with nodule-inducing *Sinorhizobium meliloti* in the legume *Medicago sativa*, which resulted in increased nodulation and significantly increased plant biomass (Guiñazú et al 2009). Another study analyzed the benefits of combining PGPR strains. Greenhouse studies showed that the dry weight of tomato transplants were higher when a combination of two PGPR strains, and 75% fertilizer was used when compared to the control (100% fertilizer with no PGPR inoculants) (Hernandez and Chailloux 2004). When these experiments were performed in the field, the treatments with PGPR, mycorrhizal fungi, and 50% fertilizer exhibited a greater yield than the control (100% fertilizer) (Hernandez and Chailloux 2004). This combination of beneficial microbes also had the added effect of stimulating plant N and P absorption (Hernandez and Chailloux 2004). Formulations of compost with beneficial bacteria have also shown the ability to suppress plant pathogens (Pugliese et al 2011, Yang et al 2011). The ability of formulations of multiple beneficial microbes to increase plant productivity and health hint at the potential of the entire microbiome and plants working together with mutually beneficial outcomes.

In some cases, application of a microbial organism that confers benefit may not even be necessary. Sometimes, the same effect can be achieved by applying a microbial elicitor, which is a compound produced by the microorganism that causes the desired effect. For example, exogenous application of the *Bacillus subtilis*-derived elicitor, acetoin (3-hydroxy-2-butanone), was found to trigger induced systemic resistance (ISR) and protect plants against *Pseudomonas syringe* pv tomato pathogenesis (Rudrappa et al 2008). Similarly, adding low doses of

Chryseobacterium balustinum AUR9 cell wall lipopolysaccharides, another bacterial elicitor, to *A. thaliana* reproduced systemic induction (Ramos Solano et al 2008). Determining the precise compounds and dosages necessary for application would allow for commercial development of a non-living application providing the same benefits as the PGPRs themselves. Such treatments could avoid some of the potential complications associated with developing commercial PGPR applications such as low survivability due to competition and adverse environmental conditions (Cummings 2009). While it might be easier to come to market sooner with more consistent results the potential advantages of being self-sustaining and avoiding evolution of resistant super-organisms would probably be lost with such products. Applications of living microbes or their elicitors has potential use for agricultural priming, the induction of ISR (Conrath and Loon 2009), which has been shown as an efficient way to increase pathogen resistance with little cost to the plant (De Vleesschauwer and Höfte 2009). An important addition to strategic management practices will be the development of crop species that are able to accomplish their own priming and ISR induction, which will reduce the use of microbial applications. Although, ideally, adding PGPRs as inoculants into the rhizosphere to exploit the immense benefits they provide is, potentially, an easy fix, there is still much inconsistency in their performance at the field scale (Mark et al 2006, Morrissey et al 2004). Research has begun to focus on how to cater the rhizosphere environment for PGPR rhizosphere colonization by means of rhizosphere engineering (Ryan et al 2009). By understanding which PGPR traits are essential for rhizosphere competence (Barret et al 2011), or by considering which indigenous soil microbial communities respond most favorable to inoculation (Bernard et al 2012).

Role of the microbiome in plant health and productivity

While it's tempting to focus on characterizing microorganisms and their associated functions on the species level, logistically, this proves difficult (Nee 2004). Soil microbes are capable of both directly and indirectly influencing the productivity, diversity, and composition of plant communities (Barea et al 2002, Fitzsimons and Miller 2010, Lau and Lennon 2011, van der Heijden et al 2006, van der Heijden et al 2008). As a result, some characterizations now focus on aspects of community structure that influence plant function. Recently, increasing soil microbial species richness was shown to be a predictor of plant health and productivity (Lau and Lennon 2011, Schnitzer et al 2011, van der Heijden et al 2008, Wagg et al 2011). Plant productivity, diversity, and nutrient acquisition have all been shown to increase with soil fungal diversity (Jonsson et al 2001, Maherali and Klironomos 2007, van der Heijden et al 1998, Wagg et al 2011); however, we are unaware of similar research using soil bacteria alone or in combination with soil fungi. Other studies have indicated that the reduction of microbial diversity (as measured by species richness) does not result in decreased soil ecosystem functions and that other microorganisms can carry out the same function without affecting plant productivity (Nannipieri et al 2003). Perhaps the key aspect determining this relationship is not taxonomic diversity, but rather functional diversity. In other words, it is not who is present but what they are doing that is more informative and revealing (Andren and Balandreau 1999, Bardgett and Shine 1999, Nannipieri et al 2003). Advances in technology have shed light on the importance and need in determining microbial functional diversity along with microbial species diversity in the rhizosphere (Nannipieri et al 2003, 2008). For example, one study found that plant productivity increased only when increased fungal diversity spanned a range of functional groups, not taxonomic groups (Maherali and Klironomos 2007). In support of this possibility,

decomposition rates have been shown to be promoted through increased microbial functional diversity (Balser et al 2002, Bonkowski and Roy 2005). While the soil microbial community exerts changing effects on the plant community, it also changes in response to host plant productivity and community characteristics resulting in a feedback response. For instance, Zak et al. (2003) demonstrated that changes in microbial community biomass, activity, and composition were a direct result of increased plant production. Another study linked plant community evenness with increased microbial biomass which in turn increased microbial functions (Lamb et al 2011). Therefore, it is important to determine microbial functional activity in the rhizosphere. Functional activity has been closely linked to organic C mineralization (Nannipieri et al 2008). Rhizodeposition is the major sources of organic C to enter the soil (Uren 2007). This results in higher enzymatic activity in the rhizosphere soil than the bulk soil, but this increase in enzymatic activity does not always correlate with higher microbial diversity (Nannipieri et al 2008). There is much debate on the actual influence of plant species on microbial diversity (de Ridder-Duine et al 2005, Dennis et al 2010, Garbeva et al 2004). Yet, recent studies using six publicly available rhizosphere microbiomes (Markowitz et al 2008) were used to analyze the functional content of the assembled and unassembled reads from rhizosphere and bulk soil (Barret et al 2011). Comparing these six microbiomes demonstrated that a small percentage of the functions overlapped between the different rhizosphere microbial communities. This suggests that plant species identity is the dominant factor influencing the composition of the rhizosphere microbial communities, as has been previously determined (Berg and Smalla 2009, van Overbeek and van Elsas 2008). Although there was a common core of shared broad functions between the rhizosphere and bulk soil microbiomes, further analysis of the functional traits to the pathway level revealed that certain specific functions are more abundant in the rhizosphere than in the

bulk soil (Barret et al 2011). This added to the importance of these traits to rhizosphere competence.

Along with increasing plant productivity, the soil microbiome also provides an important role in disease-suppressive soils. The ability of a soil to suppress disease is of key importance in measuring soil productivity (Janvier et al 2007). There are many PGPRs that aid in disease suppression via the release of antimicrobial or antifungal compounds that deter plant pathogens (Garbeva et al 2004, Weller et al 2002). For example, fluorescent psuedomonads produce the antibiotic 2, 4-DAPG which has been extensively studied as a protectant against soil-borne diseases and have been directly linked to disease suppression (Raaijmakers et al 1997, Raaijmakers and Weller 1998). *B. subtilis* also releases the antibiotics, surfactin and iturin, into the rhizosphere that play a major role in plant disease suppression (Kinsella et al 2009) while also conferring increased plant growth promotion. Many studies have focused on the disease-suppressive ability of particular taxons or group of microbes but this ability of soils to suppress disease has been linked to the soil community as a whole (Garbeva et al 2004, Malajczuk 1983). Recently, Mendes et al. (2011) determined that the soil microbiome as a whole and not an individual taxon or group of soil microbes is what drives the disease-suppressive ability of the soil.

To achieve healthy and productive plants, soil quality is of great importance. Soil quality has been defined as the "capacity of a soil to function within ecosystem boundaries to sustain plant-animal productivity, maintain or enhance water and air quality, and support human health habitation" (Karlen et al 1997). This definition has been further refined to take into account the dynamic nature of soil as a living system to "sustain biological productivity" (Doran and Safley 1997). The soil microbiome can be used as an indicator of soil quality due to its sensitivity to

small changes in the environment resulting from environmental stresses or natural perturbations (Sharma et al 2010). Elevated levels in species richness and diversity produce high functional redundancy within the soil microbiome, allowing it to quickly recover during stress (Nannipieri et al 2003, Yin et al 2000). The high functional redundancy in soil microbial diversity also confers protection against soil-borne diseases (Brussaard et al 2007, Garbeva et al 2004, Mendes et al 2011, Nannipieri et al 2003). The increase in microbial diversity produces a balanced microbiome that does not allow for pathogens to flourish since the high microbial diversity present in the soil keeps the pathogen "in check" (Garbeva et al 2004, Marrone 1999, Mendes et al 2011, Ochiai et al 2008, Postma et al 2008, Schnitzer et al 2011, Shennan 2008).

There are many key factors involved in soil health. Recently, community evenness has also been identified as an important factor in community functioning, soil health, and plant productivity (Crowder et al 2010, Wittebolle et al 2009). Microbial evenness ensures that no individual microbial taxum is able to take over and flourish, upsetting the ecological balance (Elliot and Lynch 1994). Field studies using potato plants demonstrated that even biocontrol communities among natural enemies of the potato beetle allowed for the improved control of these pests. Treatments where both pathogens and predators of the pestiferous beetle were most evenly distributed also contained plants with the greatest biomass (Crowder et al 2010). Since potato tuber yield is strongly correlated with above-ground productivity, this increase in biomass suggests that natural enemy evenness may also increase crop yield (Crowder et al 2010, Donnelly et al 2001). Increased competition found in diverse and even microbial communities reduces the niche spaces available for potential invaders (Hillebrand et al 2008, Knops et al 1999, Naeem et al 2000), and a lack of community microbial evenness has been associated with reduced plant productivity (Wilsey and Potvin 2000), possibly due to an empty niche effect

leaving some ecosystem services unfulfilled. It is suggested that when environmental fluctuations occur, even communities are quickly able to adapt to the new environment and sustain high productivity over time (Hillebrand et al 2008, Wittebolle et al 2009). These examples highlight the benefits of ensuring even and diverse microbial communities to produce healthy soil, high levels of nutrient cycling (Elliot and Lynch 1994), and to combat stress and disease (van Bruggen and Semenov 2000). In such an ecosystem where the synergistic interactions between the soil, the soil microbiome, and the plant are of great importance, it can be deduced that reduced evenness has potentially negative effects since synergistic interactions fail when one species completely dominates the assemblage (Hillebrand et al 2008).

The strong ability of root exudates to mediate and maintain the soil microbiome allows for the possibilities of exploiting this mechanism. It could be foreseen that plants (i.e., engineered or selected) could cultivate specific soil microbes that are needed or of importance to plant health. Root exudates are a complex mixture of compounds (Uren 2007). If we could tease out which compounds attract which microbes, we could selectively culture beneficial microbes such as PGPRs. Recent studies have demonstrated how specific root exudates can attract specific microbes. Rudrappa et al. (2008) demonstrated that root-secreted malic acid recruits *B. subtilis* to the root. This PGPR is known to be involved in plant growth promotion and plant protection against several plant pathogens. Chemotaxis is another means by which plants recruit PGPRs to the rhizosphere by means of the release of carbohydrates and amino acids (Somers et al 2004). Root exudates have also shown the ability to influence flagellar motility in some rhizosphere bacteria (de Weert et al 2002). The classic and most studied example of how plants are able to culture and attract beneficial microbes comes from the study of the legume *M. truncatula* and its relationship with its symbiont *S. meliloti*. Flavones and flavonols, released as root exudates by

the legume, act as the "ice-breaker" for initiating symbiosis (Zhang et al 2009b). To aid in the attraction of rhizobia to the legume, studies demonstrated that volatile organic compounds, specifically dimethyl sulfide, released by the legume are used to attract nematodes that transport the rhizobia to the legume for the purpose of symbiosis (Horiuchi et al 2005). This example is only one of the potential multitrophic interactions that can exist in the rhizosphere. Such an example illuminates the potential influence plants have in manipulating their environment. For example, *Arabidopsis* mutants lacking an ABC transporter produced changes in root exudation profiles, the ratio of phenolics to sugars changed when compared to wildtype (Badri et al 2009a). This change caused an overhaul of the natural microbial community. The changes in the root exudate chemical composition were able to culture beneficial bacterial communities enriched with PGPRs, N_2-fixing bacteria, and bacteria that are involved in heavy metal remediation (Badri et al 2009a). If we can determine which chemicals are able to attract which microbes we can selectively culture beneficial microbes and concurrently deter pathogenic microbes from colonizing the root.

Management practices influence the soil microbiome

Farm management practices fall into two general categories, organic or conventional; although, specific management objectives and/or styles exist within these categories. The choice of farming practices may lend themselves to different processes or steps to achieve a more diverse and even microbiome. The USDA defines organic farming as "an ecological production management system that promotes and enhances biodiversity, biological cycles, and soil biological activity. Organic farming is based on the minimal use of off-farm inputs and on management practices that restore, maintain, and enhance ecological harmony" (Gold 1995). Whereas organic farming uses no synthetic fertilizers or added inputs to increase productivity,

conventional farming does just the opposite; often using synthetic, chemical fertilizers, and pesticides to benefit crop protection and productivity. Frequently, these treatments are aimed at the microbial "black box" that is the soil microbiome. For example, conventional agriculture may target plant pathogens through the use of pesticides/fungicides, with a potential side effect of reducing soil microbial community diversity and evenness (Crowder et al 2010, Krauss et al 2011, Liu et al 2007, Sugiyama et al 2010). Whereas, organic agriculture may seek to control plant pathogens through competition and/or antagonism by utilizing treatments that promote a more diverse and even microbial community (Sugiyama et al 2010) such as the addition of varying types of organic matter. It is known that the structure of the soil microbiome is influenced by agricultural management practices (Crowder et al 2010, Liu et al 2007, Lumini et al 2011, Reeve et al 2010, Sugiyama et al 2010), land use (Degens et al 2000), and degrees of stress and disturbance (Degens et al 2001). By understanding those influences that combine to create more diverse and even soil microbial communities, fertility, and disease resistance can be inherently restored in depleted, disease-stricken soil environments.

Implications for agriculture

While improving crop productivity is a century-old agrarian goal, high energy prices, globalization and climate change are changing the landscape for seeking solutions. The problem is no longer simply to produce more food, but also to do so in environmentally and socially sustainable ways (Godfray et al 2010, von Braun 2007). As discussed above, agriculture should consider maximizing the coadaptation between plants and microbes in an effort to promote soil microbial diversity. Although, this may reduce short-term productivity, we believe that it will maximize long-term yields while minimizing resource use.

Evidence showing the importance of coadaptation of plant-microbial communities in plant health and productivity continues to mount (Badri and Vivanco 2009, Hierro et al 2005, Lambers et al 2009, van der Putten et al 2009, Wardle 2004). What implications does decoupling the coadapted plant-microbial relationship have on agriculture? The inability of plants to maintain the diversity and evenness of a microbiome that is not co-adaptive has already been described (Broeckling et al 2008). This loss of diversity and evenness is detrimental to ecosystem functioning and plant productivity. In a world where the demand for food increases by the second unhealthy crops with low productivity is unacceptable.

Conclusion

The growing human population, reduction in land and resources, and the need for more environmentally friendly agricultural practices have highlighted the need for sustainable farming. There is evidence showing the close ties between plants and their microbiome. An even and balanced microbiome can be the answer for obtaining healthier more productive plants. Recent studies have begun to hint at the importance of this relationship and have started to examine the system as a whole to better understand the intricacies of the plant-microbiome interaction and its impact on plant health and productivity. There is a complex conversation that occurs between soil microbes and plants, mediated by root exudates, but this conversation still needs a lot more translating. We realize that successful management of soil health and plant productivity is a combination of many factors and individuals coming together to provide optimal conditions for a healthy plant. PGPRs are known to increase plant productivity and health and we need to be able to improve the conversations between plants and those microbes. Root exudates, due to their use as signaling molecules and as substrates by microbes, can be the answer to manipulating this dialogue (Ryan et al 2009). We have seen that ABC transporters play a key role in root

exudation, and the compounds exuded change when they are manipulated. Future studies should determine what key compounds and root exudate compositions will culture these beneficial microbes that produce healthy more productive plants.

Dissertation research objectives

The importance of soil microbial communities to plant health and productivity has been outlined above. The goal of this research was to understand the interactions and mechanisms that enable not only the close associations between plants and the rhizosphere microbial community, but also how these interactions are controlled and modulated with the ultimate purpose of increasing plant health and productivity. This goal was accomplished by evaluating the following hypothesis, utilizing the model organism *Arabidopsis thaliana*:

> The qualitative and quantitative changes in root exudation enables the plant to culture an array of microbes that help in overcoming biotic and abiotic stress so that blends of root exudates can be used to alter soil microbial communities in order to improve and increase plant health and productivity.

In order to comprehend and decipher the mechanisms at play in the rhizosphere we first needed to answer the following questions:

1. Does root exudation change with plant development?
2. Does root exudation influence the rhizosphere metatranscriptome?
3. What changes occur in the rhizosphere microbiome during plant development?
4. Are members in the rhizosphere microbiome influenced by root exudation?
5. In the absence of the plant, do root exudates influence soil microbial communities?

Answers to these questions provide the necessary framework to manipulate the soil microbial community to increase plant health and productivity. Additionally, the techniques and

knowledge gained in these studies could be extended to commercially important crops such as maize, potato, sorghum, rice, etc.

Dissertation overview

This dissertation includes six chapters. Chapter 1 provides the introduction and background information for the overall study. The fundamental results of the study are provided in Chapters 2 through 5. Chapter 6 provides the conclusions and recommendations for future work.

Chapter 2, “Root exudation of phytochemicals in Arabidopsis follows specific patterns that are developmentally programmed and correlate with soil microbial functions,” examines root exudation through development. This chapter focuses on how identified changes in root exudation during development influence the rhizosphere metatranscriptome. Additionally, correlations between *in vitro* collected root exudates and the rhizosphere metatranscriptome suggest that root exudation may be genetically programmed.

Chapter 3, "Rhizosphere microbiome assemblage is affected by plant development,” focuses on the identity of the rhizosphere microbiome during development. The chapter includes an analysis of whether rhizosphere microbial community structure changes with plant development and how this process occurs. The chapter examines how root exudation patterns may play a part in the establishment and maintenance of the rhizomicrobiome. Additionally, the potential benefits that the plant gains from culturing and attracting these microbes to the rhizosphere was examined.

Chapter 4, "Application of natural blends of phytochemicals derived from the root exudates of Arabidopsis to the soil reveal that phenolic-related compounds predominantly modulate the soil microbiome,” analyzes whether the compounds found in root exudation in the

absence of a plant directly influence the soil microbiome. Additionally, root exudates are fractionated in order to identify which classes of compounds readily stimulate soil microbial community dynamics. Outcomes from this study will provide support for the use of natural compounds derived from root exudates as potential soil prebiotics to modulate and help control the soil microbial community.

Chapter 5, "An experimental pipeline for the development of soil prebiotics of agricultural importance," utilizes the information gained in Chapters 2-4 to develop interactive models that could be used to forecast how specific blends of chemical compounds found in root exudates influence specific bacteria in the soil. Furthermore, various statistical modeling approaches were utilized and validated in order to identify which model can be used as a predictive tool to identify compounds that could reliably be used as soil prebiotics.

Finally, Chapter 6 summarizes the overall conclusions for the study, discusses the status of the field, and provides recommendations for future areas of exploration.

Figures

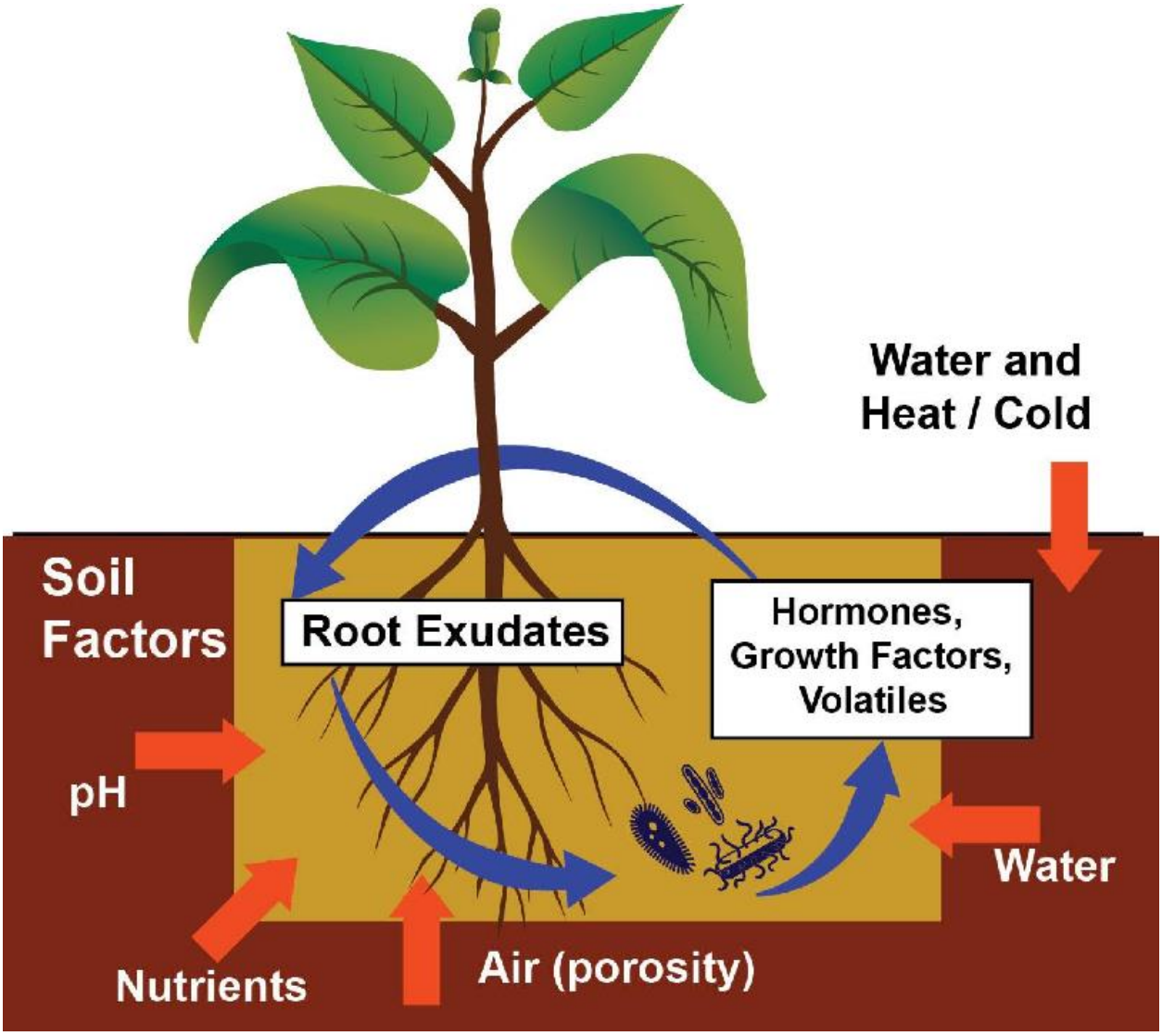

Figure 1-1. Schematic illustration of how soil factors influence both plant roots and soil microbes which in turn reshape the soil environment through a dynamic exchange of chemical responses to living and non-living stimuli

REFERENCES

Adesemoye AO, Torbert HA, Kloepper JW (2009). Plant growth-promoting rhizobacteria allow reduced application rates of chemical fertilizers. *Microb Ecol* **58:** 921-929.

Ahemad M, Khan MS (2011). Plant Growth Promoting Rhizobacteria: Recent Advancements. *Insight Microbiol* **1:** 39-54.

Andren O, Balandreau J (1999). Biodiversity and soil functioning - from black box to can of worms? *Appl Soil Ecology* **13:** 105-108.

Babalola OO (2010). Beneficial bacteria of agricultural importance. *Biotechnol Lett* **32:** 1559-1570.

Badri DV, Loyola-Vargas VM, Broeckling CD, De-la-Pena C, Jasinski M, Santelia D *et al* (2008). Altered profile of secondary metabolites in the root exudates of Arabidopsis ATP-binding cassette transporter mutants. *Plant Physiol* **146:** 762-771.

Badri DV, Quintana N, El Kassis EG, Kim HK, Choi YH, Sugiyama A *et al* (2009a). An ABC transporter mutation alters root exudation of phytochemicals that provoke an overhaul of natural soil microbiota. *Plant Physiol* **151:** 2006-2017.

Badri DV, Vivanco JM (2009). Regulation and function of root exudates. *Plant Cell Environ* **32:** 666-681.

Badri DV, Weir TL, van der Lelie D, Vivanco JM (2009b). Rhizosphere chemical dialogues: plant-microbe interactions. *Curr Opin Biotechnol* **20:** 642-650.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* **57:** 233-266.

Balser TC, Kinzig AP, Firestone MK (2002). Linking Soil Microbial Communities and Ecosystem Functioning. In: Kinzig AP, Pacala SW, Tilman D (eds). *The Functional Consequences of Biodiversity: Empirical Progress and Theoretical Extensions*. Princeton University Press: Princeton, NJ. pp 265-293.

Bardgett RD, Shine A (1999). Linkages between plant litter diversity, soil microbial biomass and ecosystem function in temperate grasslands. *Soil Biol Biochem* **31:** 317-321.

Barea JM, Azcon R, Azcon-Aguilar C (2002). Mycorrhizosphere interactions to improve plant fitness and soil quality. *Anton Leeuw Int J G* **81:** 343-351.

Barret M, Morrissey JP, O'Gara F (2011). Functional genomics analysis of plant growth-promoting rhizobacterial traits involved in rhizosphere competence. *Biol Fertil Soils* **47:** 729-743.

Berg G (2009). Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Appl Microbiol Biotechnol* **84:** 11-18.

Berg G, Smalla K (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* **68:** 1-13.

Bernard E, Larkin RP, Tavantzis S, Erich MS, Alyokhin A, Sewell G *et al* (2012). Compost, rapeseed rotation, and biocontrol agents significantly impact soil microbial communities in organic and conventional potato production systems. *App Soil Ecol* **52:** 29.

Bertin C, Yang X, Weston LA (2003). The role of root exudates and allelochemicals in the rhizosphere. *Plant Soil* **256:** 67-83.

Bonkowski M, Roy J (2005). Soil microbial diversity and soil functioning affect competition among grasses in experimental microcosms. *Oecologia* **143:** 232-240.

Broeckling CD, Broz AK, Bergelson J, Manter DK, Vivanco JM (2008). Root exudates regulate soil fungal community composition and diversity. *Appl Environ Microbiol* **74:** 738-744.

Broz AK, Manter DK, Vivanco JM (2007). Soil fungal abundance and diversity: another victim of the invasive plant Centaurea maculosa. *Isme J* **1:** 763-765.

Brussaard L, de Ruiter PC, Brown GG (2007). Soil biodiversity for agricultural sustainability. *Agr Ecosyst Environ* **121:** 233-244.

Chithrashree, Udayashankar AC, Nayaka SC, Reddy MS, Srinivas C (2011). Plant growth-promoting rhizobacteria mediate induced systemic resistance in rice against bacterial leaf blight caused by Xanthomonas oryzae pv. oryzae. *Biol Control* **59:** 114-122.

Conrath U, Loon LCV (2009). Priming of Induced Plant Defense Responses. In: Van Loon LC (ed). *Advances in Botanical Research*. Academic Press. pp 361-395.

Crowder DW, Northfield TD, Strand MR, Snyder WE (2010). Organic agriculture promotes evenness and natural pest control. *Nature* **466:** 109-112.

Cummings S (2009). The application of plant growth promoting rhizobacteria (PGPR) in low input and organic cultivation of graminaceous crops; potential and problems. *Environ Biotechnol* **5:** 43-50.

De-la-Pena C, Badri DV, Lei Z, Watson BS, Brandao MM, Silva-Filho MC *et al* (2010). Root secretion of defense-related proteins is development-dependent and correlated with flowering time. *J Biol Chem* **285:** 30654-30665.

de Ridder-Duine AS, Kowalchuk GA, Klein Gunnewiek PJA, Smant W, van Veen JA, de Boer W (2005). Rhizosphere bacterial community composition in natural stands of Carex arenaria (sand sedge) is determined by bulk soil community composition. *Soil Biol Biochem* **37:** 349-357.

De Vleesschauwer D, Höfte M (2009). Rhizobacteria-Induced Systemic Resistance. In: Van Loon LC (ed). *Advances in Botanical Research*. Elsevier: Burlington, MA. pp 223-281.

de Weert S, Vermeiren H, Mulders IHM, Kuiper I, Hendrickx N, Bloemberg GV *et al* (2002). Flagella-Driven Chemotaxis Towards Exudate Components Is an Important Trait for Tomato Root Colonization by Pseudomonas fluorescens. *Mol Plant Microbe Interact* **15:** 1173-1180.

Degens BP, Schipper LA, Sparling GP, Vojvodic-Vukovic M (2000). Decreases in organic C reserves in soils can reduce the catabolic diversity of soil microbial communities. *Soil Biol Biochem* **32:** 189-196.

Degens BP, Schipper LA, Sparling GP, Duncan LC (2001). Is the microbial community in a soil with reduced catabolic diversity less resistant to stress or disturbance? *Soil Biol Biochem* **33:** 1143-1153.

Dennis PG, Miller AJ, Hirsch PR (2010). Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *FEMS Microbiol Ecol* **72:** 313-327.

Dimkpa C, Weinand T, Asch F (2009). Plant-rhizobacteria interactions alleviate abiotic stress conditions. *Plant Cell Environ* **32:** 1682-1694.

Donnelly A, Craigon J, Black CR, Colls JJ, Landon G (2001). Elevated CO2 Increases Biomass and Tuber Yield in Potato Even at High Ozone Concentrations. *New Phytol* **149:** 265-274.

Doran JW, Safley W (1997). Defining and Assessing Soil Health and Sustainable Productivity. In: Gupta VVSR, Pankhurst C, Doube BM (eds). *Biological indicators of soil health*. CAB International: Wallingford, UK. pp 1-28.

Elliot LF, Lynch JM (1994). Biodiversity and soil resilience. In: Greenland DJ, Szabolos I (eds). *Soil Resilience and Sustainable Land Use*. CAB International: Wallingford, UK. pp 353-364.

Esitken A (2011). Use of plant growth promoting rhizobacteria in horticultural plants. In: Maheshwari DK (ed). *Bacteria in Agrobiology: Crop Ecosystems*. Springer: Berlin. pp 189-235.

Faoro H, Alves AC, Souza EM, Rigo LU, Cruz LM, Al-Janabi SM *et al* (2010). Influence of Soil Characteristics on the Diversity of Bacteria in the Southern Brazilian Atlantic Forest. *App Environ Microb* **76:** 4744-4749.

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

Fitzsimons MS, Miller RM (2010). The Importance Of Soil Microorganisms For Maintaining Diverse Plant Communities In Tallgrass Prairie. *Am J Bot* **97:** 1937-1943.

Fliessbach A, Winkler M, Lutz MP, Oberholzer HR, Mader P (2009). Soil amendment with Pseudomonas fluorescens CHA0: lasting effects on soil biological properties in soils low in microbial biomass and activity. *Microb Ecol* **57:** 611-623.

Flores HE, Vivanco JM, Loyola-Vargas VM (1999). 'Radicle' biochemistry: the biology of root-specific metabolism. *Trends Plant Sci* **4:** 220-226.

Frey SD, Knorr M, Parrent JL, Simpson RT (2004). Chronic nitrogen enrichment affects the structure and function of the soil microbial community in temperate hardwood and pine forests. *Forest Ecol Manag* **196:** 159-171.

Garbeva P, van Veen JA, van Elsas JD (2004). Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* **42:** 243-270.

Girvan MS, Bullimore J, Pretty JN, Osborn AM, Ball AS (2003). Soil type is the primary determinant of the composition of the total and active bacterial communities in arable soils. *Appl Environ Microbiol* **69:** 1800-1809.

Godfray HC, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF *et al* (2010). Food security: the challenge of feeding 9 billion people. *Science* **327:** 812-818.

Gold MV (1995). Organic production/Organic food: Information Access Tools. USDA National Agricultural Library.

Guiñazú LB, Andrés JA, Del Papa MF, Pistorio M, Rosas SB (2009). Response of alfalfa (Medicago sativa L.) to single and mixed inoculation with phosphate-solubilizing bacteria and Sinorhizobium meliloti. *Biol Fertil Soils* **46:** 185-190.

Hamilton EW, Frank DA (2001). Can plants stimulate soil microbes and their own nutrient supply? Evidence from a grazing tolerant grass. *Ecology* **82:** 2397-2402.

Hayat R, Ali S, Amara U, Khalid R, Ahmed I (2010). Soil beneficial bacteria and their role in plant growth promotion: a review. *Ann Microbiol* **60:** 579-598.

Hernandez M, Chailloux M (2004). Las micorrizas arbusculares y las bacterias rizosfericas como alternativa a la nutricion mineral del tomate. *Cultivos Tropicales* **25:** 5-16.

Hierro JL, Maron JL, Callaway RM (2005). A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range. *J Ecol* **93:** 5-15.

Hillebrand H, Bennett DM, Cadotte MW (2008). Consequences of dominance: a review of evenness effects on local and regional ecosystem processes. *Ecology* **89:** 1510-1520.

Horiuchi J, Prithiviraj B, Bais HP, Kimball BA, Vivanco JM (2005). Soil nematodes mediate positive interactions between legume plants and rhizobium bacteria. *Planta* **222:** 848-857.

Janvier C, Villeneuve F, Alabouvette C, Edel-Hermann V, Mateille T, Steinberg C (2007). Soil health through soil disease suppression: Which strategy from descriptors to indicators? *Soil Biol Biochem* **39:** 1.

Jetiyanon K, Kloepper JW (2002). Mixtures of plant growth-promoting rhizobacteria for induction of systemic resistance against multiple plant diseases. *Biological Control* **24:** 285-291.

Jonsson LM, Nilsson M-C, Wardle DA, Zackrisson O (2001). Context Dependent Effects of Ectomycorrhizal Species Richness on Tree Seedling Productivity. *Oikos* **93:** 353-364.

Karlen DL, Mausbach MJ, Doran JW, Cline RG, Harris RF, Schuman GE (1997). Soil Quality: A Concept, Definition, and Framework for Evaluation (A Guest Editorial). *Soil Sci Soc Am J* **61:** 4-10.

Kaymak HC, Maheshwari DK (eds) (2010) *Microbiology Monographs*. Springer Berlin Heidelberg: Berlin, Heidelberg, 45-79pp.

Kinsella K, Schulthess CP, Morris TF, Stuart JD (2009). Rapid quantification of Bacillus subtilis antibiotics in the rhizosphere. *Soil Biol Biochem* **41:** 374-379.

Kirankumar R, Jagadeesh KS, Krishnaraj PU, Patil MS (2008). Enhanced growth promotion of tomato and nutrient uptake by plant growth promoting rhizobacterial isolates in presence of tobacco mosaic virus pathogen. *Karnataka J Agric Sci* **21:** 309-311.

Klironomos JN (2002). Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature* **417:** 67-70.

Knops JMH, Tilman D, Haddad NM, Naeem S, Mitchell CE, Haarstad J *et al* (1999). Effects of plant species richness on invasion dynamics, disease outbreaks, insect abundances and diversity. *Ecol Lett* **2:** 286-293.

Krauss J, Gallenberger I, Steffan-Dewenter I (2011). Decreased Functional Diversity and Biological Pest Control in Conventional Compared to Organic Crop Fields. *PLoS ONE* **6:** e19502.

Lamb E, Kennedy N, Siciliano S (2011). Effects of plant species richness and evenness on soil microbial community diversity and function. *Plant Soil* **338:** 483-495.

Lambers H, Mougel C, Jaillard B, Hinsinger P (2009). Plant-microbe-soil interactions in the rhizosphere: an evolutionary perspective. *Plant Soil* **321:** 83-115.

Lau JA, Lennon JT (2011). Evolutionary ecology of plant-microbe interactions: soil microbial structure alters selection on plant traits. *New Phytologist* **192:** 215-224.

Lauber CL, Strickland MS, Bradford MA, Fierer N (2008). The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol Biochem* **40:** 2407-2415.

Lim JH, Kim SD (2009). Synergistic Plant Growth Promotion by the Indigenous Auxins-producing PGPR Bacillus subtilis AH18 and Bacillus licheniforims K11. *J Korean Soc Appl Bi* **52:** 531-538.

Liu B, Tu C, Hu S, Gumpertz M, Ristaino JB (2007). Effect of organic, sustainable, and conventional management strategies in grower fields on soil physical, chemical, and biological factors and the incidence of Southern blight. *Appl Soil Ecol* **37:** 202-214.

Lucas JA (2011). Advances in plant disease and pest management. *J Agr Sci* **149:** 91-114.

Lugtenberg B, Kamilova F (2009). Plant-growth-promoting rhizobacteria. *Annu Rev Microbiol* **63:** 541-556.

Lumini E, Vallino M, Alguacil MM, Romani M, Bianciotto V (2011). Different farming and water regimes in Italian rice fields affect arbuscular mycorrhizal fungal soil communities. *Ecol Appl* **21:** 1696-1707.

Maherali H, Klironomos JN (2007). Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science* **316:** 1746-1748.

Maheshwari DK (ed) (2011) *Plant Growth and Health Promoting Bacteria*. Springer: Munster, 431pp.

Malajczuk N (1983). Microbial antagonism to Phytophthora. In: Erwin DC, Bartnicki-Garcia S, Tsao PH (eds). *Phytophthora its biology, taxonomy, ecology, and pathology*. American Phytopathological Society: St. Paul, Minn. pp 197-218.

Mangla S, Callaway RM (2008). Exotic invasive plant accumulates native soil pathogens which inhibit native plants. *J Ecol* **96:** 58.

Mark G, Morrissey JP, Higgins P, O'Gara F (2006). Molecular-based strategies to exploit Pseudomonas biocontrol strains for environmental biotechnology applications. *FEMS Microbiol Ecol* **56:** 167-177.

Markowitz VM, Ivanova NN, Szeto E, Palaniappan K, Chu K, Dalevi D *et al* (2008). IMG/M: a data management and analysis system for metagenomes. *Nucleic Acids Res* **36:** D534-538.

Marrone PG (1999). Microbial pesticides and natural products as alternatives. *Outlook Agr* **28:** 149-154.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332:** 1097-1100.

Micallef SA, Shiaris MP, Colon-Carmona A (2009). Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. *J Exp Bot* **60:** 1729-1742.

Morales SE, Holben WE (2011). Linking bacterial identities and ecosystem processes: can 'omic' analyses be more than the sum of their parts? *FEMS Microbiol Ecol* **75:** 2.

Morrissey JP, Dow JM, Mark GL, O'Gara F (2004). Are microbes at the root of a solution to world food production? Rational exploitation of interactions between microbes and plants can help to transform agriculture. *EMBO Rep* **5:** 922-926.

Naeem S, Knops JMH, Tilman D, Howe KM, Kennedy T, Gale S (2000). Plant diversity increases resistance to invasion in the absence of covarying extrinsic factors. *Oikos* **91:** 97-108.

Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G (2003). Microbial diversity and soil functions. *Eur J Soil Sci* **54:** 655-670.

Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G *et al* (2008). Effects of root exudates in microbial diversity and activity in rhizosphere soils. In: Nautiyal CS, Dion P (eds). *Molecular Mechanisms of Plant and Microbe Coexistence*. Springer-Verlag: Berlin Heidelberg. pp 339-365.

Nee S (2004). More than meets the eye. *Nature* **429:** 804-805.

Neuman G, Romheld V (2007). The release of root exudates as affected by the plant plysiological status. In: Pinton R, Varanini Z, Nannipieri P (eds). *The rhizosphere biochemistry and organic substances at the soil-plant interface*. CRC: Boca Raton, FL. pp 23-72.

Nihorimbere V, Ongena M, Smargiassi M, Thonart P (2011). Beneficial effect of the rhizosphere microbial community for plant growth and health. *Biotechnol Agron Soc* **15:** 327-337.

Ochiai N, Powelson M, Crowe F, Dick R (2008). Green manure effects on soil quality in relation to suppression of Verticillium wilt of potatoes. *Biol Fertil Soils* **44:** 1013-1023.

Okubara PA, Bonsall RF (2008). Accumulation of Pseudomonas-derived 2,4-diacetylphloroglucinol on wheat seedling roots is influenced by host cultivar. *Biol Control* **46:** 322-331.

Pérez-Piqueres A, Edel-Hermann V, Alabouvette C, Steinberg C (2006). Response of soil microbial communities to compost amendments. *Soil Biol Biochem* **38:** 460-470.

Perry LG, Alfored ER, Horiuchi J, Paschke MW, Vivanco JM (2007). Chemical Signals in the Rhizosphere: Root-Root and Root-Microbe Communication. In: Pinton R, Varanini Z, Nannipieri P (eds). *The Rhizosphere Biochemistry and Organic Substances at the Soil-Plant Interface, Second Edition*. CRC Press: Boca Raton, Fl. pp 297-330.

Postma J, Schilder MT, Bloem J, van Leeuwen-Haagsma WK (2008). Soil suppressiveness and functional diversity of the soil microflora in organic farming systems. *Soil Biol Biochem* **40:** 2394-2406.

Pugliese M, Liu BP, Gullino ML, Garibaldi A (2011). Microbial enrichment of compost with biological control agents to enhance suppressiveness to four soil-borne diseases in greenhouse. *J Plant Dis Protect* **118:** 45-50.

Raaijmakers JM, Weller DM, Thomashow LS (1997). Frequency of Antibiotic-Producing Pseudomonas spp. in Natural Environments. *Appl Environ Microbiol* **63:** 881-887.

Raaijmakers JM, Weller DM (1998). Natural plant protection by 2,4-diacetylphloroglucinol - Producing Pseudomonas spp. in take-all decline soils. *Mol Plant Microbe Interact* **11:** 144-152.

Ramos Solano B, Barriuso Maicas J, Pereyra de la Iglesia MT, Domenech J, Gutierrez Manero FJ (2008). Systemic Disease Protection Elicited by Plant Growth Promoting Rhizobacteria Strains: Relationship Between Metabolic Responses, Systemic Disease Protection, and Biotic Elicitors. *Phytopathology* **98:** 451-457.

Reeve JR, Schadt CW, Carpenter-Boggs L, Kang S, Zhou J, Reganold JP (2010). Effects of soil type and farm management on soil ecological functional genes and microbial activities. *ISME J* **4:** 1099-1107.

Rousk J, Baath E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG *et al* (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J* **4:** 1340-1351.

Rovira AD (1969). Plant Root Exudates. *Bot Rev* **35:** 35-57.

Rudrappa T, Czymmek KJ, Pare PW, Bais HP (2008). Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol* **148:** 1547-1556.

Ryan PR, Dessaux Y, Thomashow LS, Weller DM (2009). Rhizosphere engineering and management for sustainable agriculture. *Plant Soil* **321:** 363-383.

Schloss PD, Handelsman J (2006). Toward a census of bacteria in soil. *PLoS Comput Biol* **2:** e92.

Schnitzer SA, Klironomos JN, Hillerislambers J, Kinkel LL, Reich PB, Xiao K *et al* (2011). Soil microbes drive the classic plant diversity-productivity pattern. *Ecology* **92:** 296-303.

Selvakumar G, Panneerselvam P, Ganeshamurthy AN, Maheshwari DK (2012). Bacterial Mediated Alleviation of Abiotic Stress in Crops. In: Maheshwari DK (ed). *Bacteria in Agrobiology: Stress Management*. Springer: New York, NY. pp 205-224.

Shaharoona B, Naveed M, Arshad M, Zahir ZA (2008). Fertilizer-dependent efficiency of Pseudomonads for improving growth, yield, and nutrient use efficiency of wheat (Triticum aestivum L.). *Appl Microbiol Biot* **79:** 147-155.

Sharma SK, Ramesh A, Sharma MP, Joshi OP, Govaerts B, Steenwerth KL *et al* (2010). Microbial Community Structure and Diversity as Indicators for Evaluating Soil Quality. In: Lichtfouse E (ed). *Biodiversity, Biofuels, Agroforestry and Conservation Agriculture*. Springer Netherlands. pp 317-358.

Shennan C (2008). Biotic interactions, ecological knowledge and agriculture. *Philos T Roy Soc B* **363:** 717-739.

Somers E, Vanderleyden J, Srinivasan M (2004). Rhizosphere Bacterial Signalling: A Love Parade Beneath Our Feet. *Crit Rev Microbiol* **30:** 205-240.

Stinson KA, Campbell SA, Powell JR, Wolfe BE, Callaway RM, Thelen GC *et al* (2006). Invasive plant suppresses the growth of native tree seedlings by disrupting belowground mutualisms. *PLoS Biol* **4:** e140.

Sugiyama A, Vivanco JM, Jayanty SS, Manter DK (2010). Pyrosequencing Assessment of Soil Microbial Communities in Organic and Conventional Potato Farms. *Plant Dis* **94:** 1329-1335.

Tang CS, Cai WF, Kohl K, Nishimoto RK (1995). Plant Stress And Allelopathy. In: Inderjit, Dakshini KMM, Einhellig FA (eds). *Allelopathy: Organisms, Processes, And Applications*. Amer Chemical Soc: Washington. pp 142-157.

Tilman D, Socolow R, Foley JA, Hill J, Larson E, Lynd L *et al* (2009). Energy. Beneficial biofuels--the food, energy, and environment trilemma. *Science* **325:** 270-271.

Uren NC (2007). Types, amounts, and possible functions of compounds released into the rhizosphere by soil-grown plants. In: Pinton R, Varanini Z, Nannipieri P (eds). *The Rhizosphere Biochemistry and Organic Substances at the Soil-Plant Interface, Second Edition*. CRC Press: Boca Raton, Fl. pp 1-21.

van Bruggen AHC, Semenov AM (2000). In search of biological indicators for soil health and disease suppression. *Appl Soil Ecol* **15:** 13-24.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T *et al* (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396:** 69-72.

van der Heijden MGA, Bakker R, Verwaal J, Scheublin TR, Rutten M, Van Logtestijn R *et al* (2006). Symbiotic bacteria as a determinant of plant community structure and plant productivity in dune grassland. *FEMS Microbiol Ecol* **56:** 178-187.

van der Heijden MGA, Bardgett RD, Van Straalen NM (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* **11:** 296-310.

van der Putten W, Bardgett R, de Ruiter P, Hol W, Meyer K, Bezemer T *et al* (2009). Empirical and theoretical challenges in aboveground-belowground ecology. *Oecologia* **161:** 1-14.

van Overbeek L, van Elsas JD (2008). Effects of plant genotype and growth stage on the structure of bacterial communities associated with potato (Solanum tuberosum L.). *FEMS Microbiol Ecol* **64:** 283-296.

Vandenkoornhuyse P, Mahe S, Ineson P, Staddon P, Ostle N, Cliquet JB *et al* (2007). Active root-inhabiting microbes identified by rapid incorporation of plant-derived carbon into RNA. *Proc Natl Acad Sci U S A* **104:** 16970-16975.

von Braun J (2007). The World Food Situation: New Driving Forces and Required Actions. In: Institute IFPR (ed).

Wagg C, Jansa J, Schmid B, van der Heijden MG (2011). Belowground biodiversity effects of plant symbionts support aboveground productivity. *Ecol Lett* **14:** 1001-1009.

Wardle DA (2004). Ecological Linkages Between Aboveground and Belowground Biota. *Science* **304:** 1629-1633.

Weisskopf L, Abou-Mansour E, Fromin N, Tomasi N, Santelia D, Edelkott I *et al* (2006). White lupin has developed a complex strategy to limit microbial degradation of secreted citrate required for phosphate acquisition. *Plant Cell Environ* **29:** 919-927.

Weller DM, Raaijmakers JM, Gardener BB, Thomashow LS (2002). Microbial populations responsible for specific soil suppressiveness to plant pathogens. *Annu Rev Phytopathol* **40:** 309-348.

Whipps JM (1990). Carbon Economy. In: Lynch JM (ed). *The Rhizosphere*. J. Wiley: Chichester, UK. pp 59-97.

Wilsey BJ, Potvin C (2000). Biodiversity and ecosystem functioning: importance of species evenness in an old field. *Ecology* **81:** 887-892.

Wittebolle L, Marzorati M, Clement L, Balloi A, Daffonchio D, Heylen K *et al* (2009). Initial community evenness favours functionality under selective stress. *Nature* **458:** 623-626.

Xie X, Zhang H, Paré PW (2009). Sustained growth promotion in arabidopsis with long-term exposure to the beneficial soil bacterium Bacillus subtilis (GB03). *Plant Signaling Behav* **4:** 948-953.

Yang X, Chen L, Yong X, Shen Q (2011). Formulations can affect rhizosphere colonization and biocontrol efficiency of Trichoderma harzianum SQR-T037 against Fusarium wilt of cucumbers. *Biol Fertil Soils* **47:** 239-248.

Yin B, Crowley D, Sparovek G, De Melo WJ, Borneman J (2000). Bacterial functional redundancy along a soil reclamation gradient. *Appl Environ Microbiol* **66:** 4361-4365.

Zak DR, Holmes WE, White DC, Peacock AD, Tilman D (2003). Plant diversity, soil microbial communities, and ecosystem function: Are there any links? *Ecology* **84:** 2042-2050.

Zhang H, Kim MS, Sun Y, Dowd SE, Shi H, Pare PW (2008a). Soil bacteria confer plant salt tolerance by tissue-specific regulation of the sodium transporter HKT1. *Mol Plant Microbe In* **21:** 737-744.

Zhang H, Xie X, Kim MS, Kornyeyev DA, Holaday S, Pare PW (2008b). Soil bacteria augment Arabidopsis photosynthesis by decreasing glucose sensing and abscisic acid levels in planta. *Plant J* **56:** 264-273.

Zhang H, Sun Y, Xie X, Kim M-S, Dowd SE, Paré PW (2009a). A soil bacterium regulates plant acquisition of iron via deficiency-inducible mechanisms. *Plant J* **58:** 568-577.

Zhang H, Murzello C, Sun Y, Kim M-S, Xie X, Jeter RM *et al* (2010). Choline and osmotic-stress tolerance induced in Arabidopsis by the soil microbe Bacillus subtilis (GB03). *Mol Plant Microbe In* **23:** 1097-1104.

Zhang J, Subramanian S, Stacey G, Yu O (2009b). Flavones and flavonols play distinct critical roles during nodulation of Medicago truncatula by Sinorhizobium meliloti. *Plant J* **57:** 171-183.

CHAPTER 2 ROOT EXUDATION OF PHYTOCHEMICALS IN ARABIDOPSIS FOLLOWS SPECIFIC PATTERNS THAT ARE DEVELOPMENTALLY PROGRAMMED AND CORRELATE WITH SOIL MICROBIAL FUNCTIONS[1]

Synopsis

Plant roots constantly secrete compounds into the soil to interact with neighboring organisms presumably to gain certain functional advantages at different stages of development. Accordingly, it has been hypothesized that the phytochemical composition present in the root exudates changes over the course of the lifespan of a plant. Here, root exudates of *in vitro* grown *Arabidopsis* plants were collected at different developmental stages and analyzed using GC-MS. Principle component analysis revealed that the composition of root exudates varied at each developmental stage. Cumulative secretion levels of sugars and sugar alcohols were higher in early time points and decreased through development. In contrast, the cumulative secretion levels of amino acids and phenolics increased over time. The expression in roots of genes involved in biosynthesis and transportation of compounds represented in the root exudates were consistent with patterns of root exudation. Correlation analyses were performed of the *in vitro* root exudation patterns with the functional capacity of the rhizosphere microbiome to metabolize these compounds at different developmental stages of Arabidopsis grown in natural soils. Pyrosequencing of rhizosphere mRNA revealed strong correlations ($p<0.05$) between microbial functional genes involved in the metabolism of carbohydrates, amino acids and secondary

[1] Reprinted from Chaparro, JM; Badri, DV; Bakker, MG; Sugiyama, A; Manter, DK; Vivanco, JM, Root exudation of phytochemicals in Arabidopsis follows specific patterns that are developmentally programmed and correlate with soil microbial functions. PLOS One 2013, 8, (2), e55731, DOI: 10.1371/journal.pone.0055731.

metabolites with the corresponding compounds released by the roots at particular stages of plant development. In summary, our results suggest that the root exudation process of phytochemicals follows a developmental pattern that is genetically programmed.

Introduction

Plants use root exudates as chemical cues to monitor and interact with their surroundings (De-la-Pena et al 2008, 2010). Exudate release is dependent on the needs of a plant (Badri and Vivanco 2009) and exudation can be modified in order to cope with stresses (Selvakumar et al 2012, Zamioudis and Pieterse 2012). For example, to overcome nitrogen deficiency legumes release specific flavonols which attract and initiate symbiotic relationships with rhizobia (Zhang et al 2009). However, when N fertilization is supplemented the symbiotic interaction is halted (Omrane and Chiurazzi 2009). When *Arabidopsis* is attacked by the foliar pathogen *Pseudomonas syringae* pv *tomato*, roots release malic acid which recruits beneficial soil bacteria capable of triggering host defense responses against *P. syringae* (Rudrappa et al 2008). *Zea mays* releases 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one that chemotactically attracts the beneficial rhizobacterium *Pseudomonas putida* KT2440 (Neal et al 2012). Besides such one-to-one interactions, multitrophic interactions also occur in the rhizosphere. For instance, *Medicago truncatula* emits dimethyl sulfide that attracts *Caenorhabditis elegans*, which in turn transports rhizobia close to the legume roots to initiate symbiosis (Horiuchi et al 2005). Evidence is mounting that the cross talk between plants and the soil microbes is largely orchestrated by root exudates, not only at the one compound-one microbe level, but at the community level. For instance, it has been reported that changes in root exudation due to mutation of an ABC transporter gene modulated the soil microbial community composition such that more beneficial

microbes were cultured (Badri et al 2009). Similarly, Micallef et al. (2009a) showed that soil microbial communities are affected by plant age and genotype.

Rhizosphere microbial communities have shown strong ties to root exudate composition (Broeckling et al 2008) and changes in exudate composition result in significant modifications of the soil microbial community (Badri et al 2009). Root exudate composition and concentrations change depending upon the environment in which a plant is growing, including soil edaphic and biological factors (Flores et al 1999, Micallef et al 2009b, Rovira 1969, Tang et al 1995). It has been previously reported that root secretion of proteins changes with plant development, and when challenged by pathogens or symbiotic bacteria (De-la-Pena et al 2008, 2010, Flores et al 1999, Tang et al 1995). Root secretions of some phytochemicals have also been shown to follow a diurnal rhythm (Badri et al 2010). However, there is no information available on how the composition of root secreted primary and secondary metabolites changes over the course of plant development, and how those changes correlate to the functioning of the rhizosphere microbiome.

Soil microbes have been shown to have both negative and positive effects on plant development. For example, *Agrobacterium rhizogenes* influence and manipulate plant development for the formation of hairy roots (Ortiz-Castro et al 2009, Schmulling et al 1988). Similarly, soil microbes such as PGPRs can modulate plant growth through the production of hormones such as auxin and cytokinin or via the release of volatile organic compounds (Ortiz-Castro et al 2009).

The rhizomicrobiome plays an important role in disease suppression by direct antagonism against pathogens (Mendes et al 2011), in overcoming abiotic stress by induced systemic tolerance (Selvakumar et al 2012) and in overcoming biotic stress by increasing the plant's innate immunity (Zamioudis and Pieterse 2012). It has also been documented that phytohormone

production and enhanced access to nutrients due to rhizomicrobiome activity has a positive correlation with plant productivity (Berg 2009) and overall plant health (Berendsen et al 2012, Chaparro et al 2012).

In summary, there is a concerted understanding of the ability of root exudates to influence the structure of rhizosphere microbial communities. Root exudates act as substrates, signals and/or antimicrobials influencing the relative abundance of microbial taxa in the rhizosphere. However, the functional capacity of most of these organisms is unknown and our understanding of the correlation between root exudation and microbiome functioning remains limited. Here, we show how *in vitro Arabidopsis* root exudate composition changes over the course of plant development, and we correlate these patterns with the ability of the soil microbiome to metabolize those compounds under natural soil conditions.

Materials and methods

Plant growth conditions and root exudate collection

Arabidopsis wild type (Col-0) plants were grown and root exudates were collected by using an established protocol as previously described (Badri et al 2008, 2009, 2010, 2012, Biedrzycki et al 2010, De-la-Pena et al 2008, 2010, Micallef et al 2009b), with a few modifications. *Arabidopsis* wild type (Col-0) seeds were surface-sterilized with Clorox® for one minute followed by four rinses in sterile distilled water and plated on Murashige and Skoog medium (MS) (Murashige and Skoog 1962), supplemented with 3% sucrose and 0.9% bactoagar in Petri plates. Petri plates were incubated in a growth chamber (Percival Scientific) at 25°C for seven days, with a photoperiod of 16 h light/8 h dark. To collect root exudates at different developmental time points, seven-day-old seedlings were transferred to Magenta® boxes each containing 10 ml of liquid MS (MS basal salts supplemented with 1% sucrose), incubated on an

orbital shaker at 90 rpm and illuminated under cool white fluorescent light (45 µmol m^{-2} s^{-1}) with a photoperiod of 16 h light/8 h dark at 25°C±2. Prior to exudate collection (7, 14, 21, or 28 days), plants were gently washed with sterile water to remove the surface-adhering exudates and transferred to new Magenta boxes containing 10 ml of sterile water. Growth medium plus dissolved exudates were collected at approximately the same time on the third day, after three days of continuous secretion for each time point (7-10, 14-17, 21-24, and 28-31 days). Each growth stage of the plant was as follows: the 10 day plants consisted of the two leaf growth stage, the 17 day plants were at the 5 leaf rosette stage, the 24 day plants reached the bolting stage and the 31 day plants reached the flowering stage as described by Boyes et al. (Boyes et al 2001), representative pictures of each growth stage can be found in De-la-Pena et al. (De-la-Pena et al 2010). Each time point consisted of three replicates and each replicate consisted of a total volume of 180 ml of exudate-containing medium, from 18 individually-grown *Arabidopsis* plants. The collected root exudates were filtered using nylon filters of pore size 0.45 µm (Millipore, MA) to remove root sheathing and root-border-like cells. After filtration, the exudates were freeze-dried (Labconco, MO) and stored at -20°C for further analyses. Plant root tissues were collected from each replicate of all time points, frozen with liquid nitrogen and stored at -80°C for gene expression analyses. Sterile techniques were used throughout the experiment and there was no evidence of contamination in the media.

Gas chromatography and mass spectrometry (GC-MS) analyses of exudates

Freeze dried root exudates were dissolved in 5 ml of 80% methanol and the supernatant was collected into new glass tubes after centrifugation at 8000 rpm for 15 minutes at room temperature. The supernatants were dried under nitrogen gas and shipped to the Genome Center Core Services at the University of California, Davis for GC-MS analyses. Briefly, the dried

supernatants were derivatized as described by Sana et al. (Sana et al 2010). All samples were spiked with a mixture of fatty acid methyl esters of C8, C9, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28 and C30 linear chain length which served as an internal retention index (Fiehn et al 2008, Sana et al 2010). An Agilent 6890 gas chromatograph (Santa Clara, CA) containing a 30 m long, 0.25 mm i.d. rtx5Sil-MS column with an additional 10 m integrated guard column was used to run the samples. The Agilent 6890 was controlled by the Leco ChromaTOF software version 2.32 (St. Joseph, MI). Resulting text files were exported to a data server with absolute spectra intensities and further processed by a filtering algorithm implemented in the metabolomics BinBase database (Fiehn et al 2005). Quantification was reported as peak height using the unique ion as default. Metabolites were unambiguously assigned by the BinBase identifier numbers using retention index and mass spectrum as the two most important identification criteria. Additional confidence criteria were used by giving mass spectral metadata, using the combination of unique ions, apex ions, peak purity, and signal/noise ratios. All database entries in BinBase were matched against the Fiehn mass spectral library (http://fiehnlab.ucdavis.edu/Metabolite-Library). Data normalization was performed as described in Fiehn et al. (2008), using total metabolite content. The resulting data underwent a log transformation and was subjected to multivariate analyses and significant feature identification using MetaboAnalyst, a web-based metabolomics data processing tool (http://www.metaboanalyst.ca) (Xia et al 2009).

Gene expression analyses from plant root tissue

Total RNA was isolated from frozen root tissues (see above) using TriReagent (Sigma, MO), and was quantified with a Nanodrop ND-1000 Spectrophotometer (Thermo, DE). RNA integrity was checked on a formamide denaturing agarose gel. Two µg of purified total RNA

were reverse-transcribed using Superscript III RT and a poly (T) primer (Invitrogen, CA) at 42°C for one hour according to the manufacturer's instructions. The reaction product was diluted to a concentration of 50 ng μl^{-1} and 1 µl was used for each PCR reaction. The PCR reaction mix (20 µl) contained 0.4 µmoles of each gene-specific primer, 200 µmoles of dNTPs, 1 X reaction buffer and one unit of Taq DNA polymerase (Takara, Japan). PCR included 30 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 2 min in a GeneAmp 2700 thermal cycler (Applied Biosystems, CA). Actin primers were used as a control to determine the uniformity of the concentration of cDNA. The gene-specific primers used for RT-PCR assays are listed in Table 2-1.

Soil experiment

Soil with a history of exposure to *Arabidopsis* was collected in July 2011 from the Michigan Extension Station, Benton Harbor, MI (N42° 05' 34'', W86° 21' 19'' W, elevation 630 feet). The top 5-10 cm of soil was collected from under three patches of *Arabidopsis thaliana* that have been growing naturally in a fallow field for more than 8 years. All the necessary permits were obtained for the described soil. Broeckling et al. (2008) described the soil in detail although from a different collection event. Soil from the same site although collected at other time points was used in previous experiments by Badri et al. (2009) and Broeckling et al. (Broeckling et al 2008). The soil was transported to the laboratory in air tight coolers and stored in a cold room (4°C) until further use. At the time of use, the soil was dried at room temperature, homogenized by hand, and cleaned of plant debris. Pots (2 x 6 x 6-cm) were lined with Whatmann 3MM filter paper to avoid soil loss. The pots were placed in a growth chamber at 25°C with a photoperiod of 16 h light/8 h dark. Six replicate pots were maintained for each of the four developmental time points. Pots without plants served as a bulk soil control that could be

contrasted with rhizosphere communities under the influence of the host plant. *Arabidopsis* seeds were surface-sterilized and grown on MS plates as described above (see plant material and growth conditions). One seven day old seedling was transplanted to each of the six pots. Plants were grown until they were: 17 days, 24 days, 31 days or 38 days old, bulk soil was collected along with the 38 day old sample (Figure 2-1).

Extraction of microbial RNA from soil

For each of the 5 time points (17, 24, 31, 38 days and bulk soil) 6 replicate pots were maintained, rhizosphere soil was collected by obtaining the soil attached to the roots of the plant and bulk soil was collected from the center of the pot. Soil samples were transferred to 2.0 ml tubes, immediately frozen in liquid nitrogen and stored at -80°C until processing. A total of 6 rhizosphere soil samples were collected for each time point (17, 24, 31, 38 days and bulk soil; 30 samples total). Total RNA was extracted from each soil sample using the PowerSoil® total RNA isolation kit (MoBio, CA), with slight modifications to the manufacturer's instructions. The modifications were as follows: after solution SR2 was added to the bead tube, the solution was vortexed at maximum speed for 30 minutes instead of 5 minutes. After the phenol: chloroform: isoamyl alcohol was added to the bead tube, the bead tube was shaken at 200 rpm for 30 minutes instead of being vortexed at high speed for 10 minutes. RNA integrity was checked on a formamide denaturing agarose gel. Microbial RNA was quantified using a Nanodrop ND-1000 spectrophotometer. All RNA samples that had an A_{260}:A_{280} ratio between 1.7 and 2.0 were processed for metatranscriptomics.

Pyrosequencing and analyses

Total RNA collected from the 6 pots per time point were pooled and 15µg of total RNA from each time point and bulk soil were sent to the W.M. Keck Center for Comparative and

Functional Genomics, Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign, where steps from mRNA isolation to pyrosequencing were performed. Briefly, for each of the time points, which consisted of 6 pooled replicates, ribosomal RNA was removed from 5µg of total RNA using the Ribozero rRNA removal Meta-bacteria kit (Epicentre Biotechnologies, WI). The mRNA was converted to cDNA by using barcoded random hexamer primers and nebulized with the nebulization kit supplied with the GS Titanium library preparation kit (454 Life Sciences, CT). Each sample (17, 24, 31, 38 days, and bulk soil) was given a unique 10 bp sequence barcode and the cDNA libraries of each sample were normalized by using the Trimmer Direct kit (Evrogen, Russia) following the manufacturer's instructions and as previously described (Lambert et al 2010). cDNA normalization equalizes the number of gene copies in the library which allows for the discovery of new genes that are transcribed at low levels (Shcheglov et al 2007). The normalized barcoded cDNA libraries were pooled in equimolar concentrations based on average fragment length and concentration. The pooled libraries were quantified using a Qubit fluorometer (Invitrogen, CA) and average fragment sizes were determined by analyzing 1 µl of each sample on a Bioanalyzer (Agilent, CA) using a DNA 7500 chip. The pooled library was diluted to $1x10^6$ molecules μl^{-1}. Emulsion-based clonal amplification and sequencing on the 454 Genome Sequencer FLX+ system was performed according to the manufacturer's instructions (454 Life Sciences, CT). 454 pyrosequencing was performed on 1/8 of a PicoTiter-Plate (454 Life Sciences, CT). Signal processing and base calling were performed using the bundled 454 Data Analysis Software v2.6. 454 sequencing yielded a total of approximately 166,250 sequence reads. MG-RAST (Meyer et al 2008) and Mothur (Schloss et al 2009) were used for quality screening and sequence processing for each of the 5 samples. Sequences were screened on the following criteria: sequences derived from *A.*

thaliana were removed using the Bowtie algorithm (Langmead et al 2009). Sequences were dereplicated and filtered by length to remove sequences that differed by more than two standard deviations from the mean length. Sequences were dropped if they contained five or more ambiguous bases, or appeared to be ribosomal RNA. To equalize sampling effort across time points, a random subset of 14,740 high-quality sequence reads were selected for each time point. A summary of the 454 pyrosequencing data for each sample is found in Table 2-1. Sequence reads were assigned to the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al 2004) subsystem categories using the MG-RAST web-server pipeline. A minimum percent identity cutoff of 70% between our sequences and the sBLAT database and an E-value cutoff of 10^{-5} was used for further quality control (Supplementary Material 2-1).

Correlating rhizosphere microbial function with host plant root exudation

Correlation analyses (SAS ver. 9.3; SAS Institute, NC) were performed with the transformed data of the 17, 24, and 31 day metabolomics data with that of the corresponding metatranscriptomics data as follows: the average of the transformed GC-MS identified compounds were correlated with the overall functional genes identified by Level 2 KEGG orthology as being involved in Carbohydrate Metabolism, Metabolism of Amino Acids, and Metabolism of Secondary Metabolites (which includes the KEGG level 2 categories of Biosynthesis of Other Secondary Metabolites and Metabolism of Terpenoids and Polyketides) with the compounds categorized as sugars, amino acids, and phenolics. We further performed a more in depth correlation of the individual compounds from our metabolomics data through development with the corresponding functional genes at the KEGG functional level. The interactive pathways explorer (iPath2) (Yamada et al 2011) was used to map the functional genes involved in Metabolism, specifically Carbohydrate Metabolism, Amino Acid Metabolism,

Biosynthesis of Other Secondary Metabolites, and Metabolism of Terpenoids and Polyketides along with the root exudate compounds categorized as amino acids, phenolics, and sugars (Figure 2-2).

Results

Arabidopsis root exudation over a developmental time course

The primary and secondary metabolites present in the root exudation profiles of in vitro grown wildtype Col-0 Arabidopsis through a developmental time series were analyzed by GC-MS. After normalization, 107 compounds were detected. Among these, 57 compounds were identified (Table 2-3) based on the mass spectral library database developed by the Fiehn laboratory (University of California, Davis), which includes sugars, sugar alcohols, amino acids, organic acids, fatty acids, phenolics, etc. Hierarchical analysis using a Ward clustering algorithm and Pearson's correlation as a similarity measure revealed that the root exudate profile at each time point clustered separately and that the early (7-10 days and 14-17 days) and later (21-24 days and 28-31 days) developmental time points formed two distinct groups (Figure 2-3A). Principle component analysis (PCA) also showed that the root exudate profiles of early and later developmental time points clustered separately from each other (Figure 2-3B). Most of the variability in the data could be accounted for by component 1 (97.2%), while component 2 accounted for 2% of the variability in the data. The identified compounds contributing most to component 1 in the PCA were glycerol, ethanolamine, fructose, glucose, glycine, alanine, and tagatose. The identified compounds contributing most to component 2 were oxoproline, γ-Aminobutyric acid (GABA), urea, isoleucine, galactose, and tagatose. These data clearly indicate that the quantitative composition of *Arabidopsis* root exudates varies at each developmental stage.

We broadly categorized the 55 identified compounds into four groups: sugars, sugar alcohols, phenolics, and amino acids. In total, we identified nine sugars, seven sugar alcohols, twelve amino acids, and twenty-seven phenolic compounds (Table 2-3). The compounds categorized as phenolics consisted of compounds belonging to organic acids, carboxylic acids, fatty acids, and phenolics. For each group of compounds, cumulative secretion levels where calculated in order to identify potential patterns found throughout development. These cumulative secretion levels did indeed follow a trend depending upon the developmental stage of the plant. For instance, the secretion levels of sugars and sugar alcohols were higher at early developmental time points and gradually lowered at later developmental time points of the plant (Figure 2-4). On the contrary, the cumulative secretion levels of amino acids and phenolics were low during the early developmental time points, but rose at later developmental time points of the plant (Figure 2-4).

Root gene expression analyses

To validate the observed trends in root exudation over developmental time, we examined the gene expression of sugar transporters, ABC transporters and genes involved in secondary metabolite biosynthesis in the root tissues of 10, 17, 24, and 31 day old plants by semi-quantitative RT-PCR. In total, gene expression patterns were analyzed for 43 genes, including twenty-two sugar transporters (Figure 2-5A), six ABC transporters (Figure 2-5B) and fifteen secondary metabolite biosynthesis genes (Figure 2-5C). The expression of 22 sugar transporter genes varied with development. Among those, the expression levels of eight sugar transporters (AtSUC3, AtINT2, AtINT3, AtpGlcT, and four genes belonging to the putative monosaccharide transporter family) were higher in early plant development and gradually decreased at later developmental stages (Figure 2-5A). This result is in agreement with our GC-MS data showing

more sugars secreted at early stages of plant development. The expression of six sugar transporters (AtSUC5, AtPLT4, AtSTP10, AtSUC9 and two genes belonging to the putative monosaccharide transporter family) was not detectable in our RT-PCR analyses and the expression of five sugar transporters (AtSUC1, AtSTP7 and three genes belonging to the putative monosaccharide transporter family) was constant throughout the time points. The remaining three sugar transporters (AtSUC2, AtSUC4 and AtINT1) showed an increase in gene expression until 24 days and then decreased at 31 days.

Among the six ABC transporters (AtPDR2, AtPDR4, AtPDR6, AtPDR7, AtPDR8, and AtPDR9) that were analyzed, none showed a definite pattern of gene expression with respect to the developmental stages of the plant. For the most part the ABC transporter genes showed consistent expression over time (AtPDR2, AtPDR6, ATPDR7, and AtPDR9), with the exception of AtPDR4 which showed high gene expression at 10 days but expression decreasing to undetectable levels by 31 days. AtPDR8 was equally expressed at 17 days and 24 days, but was not detected at the other time points (Figure 2-5B).

We also analyzed the expression of fifteen genes involved in secondary metabolite biosynthesis such as the phenylpropanoid pathway (Figure 2-5C). Several of these genes were only expressed or were more highly expressed at later stages of plant development. These included PAL1 and PAL2 (phenylalanine ammonia-lyase), C4H (cinnamate-4-hydroxylase), 4CL1 and 4CL2 (4-coumarate-CoA ligase), and CYP79B2 (involved in converting tryptophan to indole-3-acetaldoxomine, a precursor of indole glucosinolates and indole-3-acetic acid) (Figure 2-5C). However, the expression of the remaining nine genes (PAL3, 4CL3, FS1, FS2, FS3, F3H, CYP79B3 and CYP71B15) tested was not detectable in our RT-PCR analysis.

Correlations between root exudation patterns and the functional capacity of the soil microbiome

Measuring the patterns of root exudation identified in this study on plants growing in natural soils is problematic due to technical sensitivity limitations and confounding factors such as the release and modification of compounds by other organisms in the soil. Therefore, we correlated the *in vitro* root exudation patterns with the functional capacity of the rhizosphere microbiome to metabolize these compounds at different developmental stages of Arabidopsis grown in natural soils.

We performed a metatranscriptomic analysis to correlate rhizosphere microbial functions with root exudation patterns corresponding to different stages of plant development. A total of 14,740 sequences for each time point (Table 2-2) were uploaded to the Metagenomics-RAST (MG-RAST) server (Meyer et al 2008) and annotated to the Kyoto Encyclopedia of Genes and Genomes (KEGG) hierarchical classification within MG-RAST to assign a putative function to each sequence (KEGG hierarchical annotations for each sample is provided in Supplementary Material 2-1). KEGG divides the functional genes obtained into a four level hierarchy with the first level consisting of five categories: metabolism, genetic information processing, environmental information processing, cellular processes, and human diseases (Supplementary Material 2-1) (Kanehisa et al 2004). In our data, metabolism accounted for over 50% of the functional genes in each sample (Supplementary Material 2-1). Since the metabolic activity of the soil microbes is presumable directly tied to their utilization of the root exudate compounds identified via GC-MS we focused on the metabolic activity of the soil microbes.

We performed Pearson correlation and Spearman rank correlation analyses between the group of exudate components and the abundance and number of related microbial functional

genes assigned to specific KEGG metabolic functions at the different time points. Indeed, we observed that the exudation of phenolics by the plant through development was significantly ($p<0.05$) positively correlated with both the abundance and number of expressed microbial genes involved in secondary metabolism ($r = 0.98$ and 0.95, respectively; Table 2-4). Similarly, there was a positive correlation between the amino acids released as root exudates and the number of expressed microbial genes involved in Amino Acid Metabolism at each time point ($\rho= 1.00$; Table 2-4). Correlations between overall Carbohydrate Metabolism performed by the soil functional microbiome and the sugars released as root exudates were not statistically significant (Table 2-4).

Correlation analyses were also performed between specific compounds present in the root exudates and the functional genes involved in their metabolism. For example, beta-alanine positively correlated to the functional gene adenylosuccinate synthase [EC:6.3.4.4] which is involved in the alanine, aspartate, and glutamate metabolism pathway (Supplementary Material 2-2). Likewise, glycine positively correlated with sarcosine oxidase, subunit alpha [EC:1.5.3.1] and sarcosine oxidase, subunit beta [EC:1.5.3.1] functional genes involved in the glycine, serine, and threonine metabolism pathway (Supplementary Material 2-2). On the other hand, 41 functional genes involved in Amino Acid Metabolism negatively correlated with asparagine. Some functional genes that negatively correlate with asparagine were: shikimate dehydrogenase [EC:1.1.1.25], hydroxypyruvate reductase [EC:1.1.1.81], or anthranilate phosphoribosyltransferase [EC:2.4.2.18] (Supplementary Material 2-2). Similarly, the functional gene farnesyl diphosphate synthase [EC:2.5.1.1 2.5.1.10] involved in the terpenoid backbone biosynthesis pathway positively correlated with six compounds categorized as phenolics (4-hydroxybutyric acid, capric acid, lauric acid, palmitic acid, propane-1,3-diol, and stearic acid;

(Supplementary Material 2-3)). We observed that the sugars arabinose, fructose, and sucrose positively correlated with 27 functional genes involved in carbohydrate metabolism such as chitin synthase [EC:2.4.1.16]; glucan endo-1,3-beta-D-glucosidase [EC:3.2.1.39]; glucan 1,3-beta-glucosidase [EC:3.2.1.58]; or rhamnulose-1-phosphate aldolase [EC:4.1.2.19] (Supplementary Material 2-4).

The iPATH 2 KEGG map (Figure 2-2) visually illustrates that the compounds released as root exudates could be utilized by the soil microbial functional genes. For example, we identified sucrose in the plant root exudates (Table 2-3). Sucrose can be used by two of the identified microbial functional genes alpha-glucosidase [EC:3.2.1.20] and beta-fructofuranosidase [EC:3.2.1.26] for the production of glucose and fructose. Similarly, galactose is used by the functional gene galactokinase [EC:2.7.1.6] to make alpha-D-galactose-1-phosphate. Beta-alanine is used to make 3-oxoproponoate and L-Glutamate by beta-alanine--pyruvate transaminase [EC:2.6.1.18].

Discussion

Soil microbial communities are able to utilize and are impacted by root exudates in a variety of ways. For example, *Bacillus subtilis*, *Rhizobium leguminosarum*, or *Agrobacterium tumefaciens* C58C1 are just a few examples of soil bacterial species that utilize and exhibit chemotaxis towards a wide variety of sugars (Bowra and Dilworth 1981, Loake et al 1988, Ordal et al 1979). Rhizobia use specific flavonoids for initiating symbiosis (Zhang et al 2009). *Agrobacterium* is chemotactically attracted to certain phenolics, such as acetosyringone (Shaw 1991), while *Pseudomonas putida* is able to catabolize flavonoids, such as naringen, p-hydroxybenzoic acid, and quercetin for use as nutritional sources (Pillai and Swarup 2002). Similarly, many gram positive bacteria use amino acids or modified peptides as signal molecules

(Ryan and Dow 2008). Here we present evidence that root exudation of primary and secondary metabolites by *Arabidopsis* changes with plant development and follow specific trends. For instance, the quantity of amino acids and secondary metabolites (phenolics) released from the roots increased over developmental time. On the other hand, sugars were released in the greatest abundance early in the plant's life cycle. These patterns were corroborated by root gene expression analyses, which showed higher expression of the majority of sugar transporter genes tested in the early stages of plant development. Similarly, higher expressions of the genes involved in the phenylpropanoid pathway were seen in later stages of plant development. Sugars serve as ready sources of energy for microbial growth (Behera and Wagner 1974), and secondary metabolites, such as those categorized as phenolics in this study, may function as antimicrobials and signaling molecules in the rhizosphere (de Weert et al 2002, Rudrappa et al 2008, Steinkellner et al 2007, Zhang et al 2009).

Plant secondary metabolites' are known defense signals (Bennett and Wallsgrove 1994) that play important roles in disease resistance (Li et al 2009, Nicholson and Hammerschmidt 1992), in adapting to the changing environment and overcoming stress (Edreva et al 2008). Their increased release at later stages in the plants life cycle is in agreement with De-la-Pena et al. (2010), where defense related proteins showed enhanced secretion during flowering time. The observed increase in the exudation of phenolics at later developmental stages was mirrored by a corresponding increase in microbial functions related to the metabolism of secondary metabolites (Table 2-4). Microbes can quickly evolve a variety of mechanisms to detoxify and overcome the effects of potentially harmful chemicals either by chemically modifying the toxin, metabolizing the toxin, or by extruding the toxin from their cells. For instance, studies analyzing the effect of the toxic compound toluene on the soil bacterial metaproteome have shown an increase in ABC

transporter activity after toluene amendment, presumably due to export of the toxic substance out of the bacterial cell (Volkers et al 2009, Williams et al 2010). Likewise, we observed that as the host plant aged, there was an increase in the expression of ABC transporters and genes involved in membrane transport among soil microbes peaking at 24 days when phenolics are at their most abundant secretion time point (data not shown).

Rates of sugars' root exudation decreased with plant development and this trend did not correlate with the overall functional genes categorized under Carbohydrate Metabolism. This can be due to the fact that many pathways and cycles utilizing sugars such as glycolysis and the citric acid cycle are synergistically regulated by both sugars and amino acids (Blencke et al 2003). Although the correlation of the overall group of compounds categorized as sugars present in the root exudates did not correspond with the whole Carbohydrate Metabolism functional genes, we did see individual sugar compounds positively correlating with particular functional genes involved in Carbohydrate Metabolism (Supplementary Material 2-4). For example sucrose, arabinose, and fructose positively correlated with 27 functional genes involved in carbohydrate metabolism (Supplementary Material 2-4). These results are indicative of the microbial community actively utilizing these specific compounds released by the plant. Studies have shown that *Sinorhizobium meliloti* carries genes necessary for the catabolism of arabinose (Poysti et al 2007). Fructose and alanine have also been shown to produce a positive metabolic priming effect on soil microbes that is manifested in the increased degradation and mineralization of more complex soil organic matter when compared to simple sugars (Hamer and Marschner 2005). The priming effect is thought to be due to the ability of these easily available substrates, i.e. fructose and alanine, to activate microbial metabolism and increase enzyme production (Kuzyakov 2002). The enhanced enzyme production presumably increases the

metabolic capability of the soil microbiome, which in turn benefits the plant as various limiting nutrients can be made available.

Rates of amino acid exudation increased with plant development, and this trend was mirrored by an increase in the number and abundance of expressed functional genes related to the metabolism of Amino Acids (Table 2-4). We observed significant correlations between specific amino acids and functional genes involved in Amino Acid Metabolism (Supplementary Material 2-2). Amino acid availability is necessary for bacterial root colonization. For example, *Pseudomonas fluorescens Pf0-1* shows chemotactic response towards tomato roots due to L-amino acids found in its root exudates (Oku et al 2012). Similarly, *Pseudomonas fluorescens* strain WCS365 colonizes the tomato root in the presence of amino acids such as: aspartic acid, glutamic acid, isoleucine, leucine, and lysine (Simons et al 1997). Our studies identified the release of the amino acid isoleucine (Table 2-3), which has been shown to be one of the major amino acid components required for the colonization of *Pseudomonas fluorescens* on tomato roots (Simons et al 1997). Interestingly, amino acid exudation of rice increased upon the plants exposure to *Cyanobacterium* sp. (Sb26) and *Rhizobium* sp. (Sb16) (Naher et al 2008) presumably due to certain microbial products that are able to trigger amino acid exudation (Phillips et al 2004). Because we collected exudates in an axenic system, this suggests that the amino acid concentration released by the plant in a more natural setting (i.e. when the plant is interacting with its biotic environment, in this case a soil microbial community) may be even higher than we observed here.

Ample evidence demonstrates that plant root exudates mediate the selection of specific rhizosphere microbes. However, there is no information available on how plants and their root exudates influence the rhizosphere microbiome functioning over the course of plant

development. In this study, we suggest that plant root exudates have associations with rhizosphere microbial functions, and that these interactions are dependent upon the plant developmental stage. These observed trends of *in vitro* collected root exudates and their correlation with rhizosphere microbial functions might hint that the qualitative changes in root exudation observed through plant development are genetically regulated and independent of the microbial community. On the other hand, quantitative changes in root exudates might be attributed to the microbial community. In other words, the soil microbial community is able to modify plant root exudation but not control it. For example, the increased exudation of secondary metabolites and amino acids through plant development might be indicative of innate defensive priming by the plant. Studies analyzing the metabolite profiles of potatoes after pathogen inoculation showed that 42 metabolites (consisting of sugars, amino acids, organic acids, and fatty acids) significantly increased or decreased (Abu-Nada et al 2007). Similar to our exudation profiling over a developmental time course, metabolite profiling by Abu-Nada et al. (2007) revealed that pathogen inoculation lead to an up-regulation of amino acids and a down-regulation of sugars. Although our exudation profiles were obtained axenically, we observed the release of compounds that are released as defensive and priming strategies against pathogens or as attractors for beneficial microbes. This may mean that as plants develop and start to set seed, they begin to adopt a more defensive strategy. Incidentally, it was previously reported that defense related proteins are secreted by roots in higher concentrations at flowering time (De-la-Pena et al 2010).

Root exuded metabolites can also have a dual role with respect to their effect on microbes. For example, GABA that in our studies increased following a developmental pattern has been shown to reduce *Agrobacterium tumefaciens* virulence by quenching quorum-sensing

(Chevrot et al 2006); while the beneficial bacteria *Pseudomonas putida* is able to use GABA as its sole nutrient source (Ramos-Gonzalez et al 2005). In contrast, proline nullifies GABAs ability to quench quorum-sensing (Haudecoeur et al 2009). While proline catabolism by the symbiont *Rhizobium meliloti* aids in its ability to colonize the root and establish symbiosis (Jimenez-Zurdo et al 1997). In our study, GABA and oxoproline (which is very similar in structure to proline) increase following plant development (Table 2-3). Experimentally these two signals have been shown to work in opposing fashion with respect to plant *Agrobacterium* infection (Chevrot et al 2006, Haudecoeur et al 2009) yet understanding the interplay of these signals in a complex environment like the rhizosphere still needs to be explored.

In connection to these patterns in root exudation, our rhizosphere metatranscriptomics data showed that fewer functional genes were uniquely expressed at early plant developmental time points compared with later developmental stages (Table 2-2). Thus, it is possible that: 1) high sugar levels exuded in early plant developmental stages may attract a wide range of microbes expressing a limited number of genes (which are similar across taxa) involved in the utilization of sugars as general substrates, and 2) high levels of phenolics exuded in later plant developmental stages might induce the expression of genes belonging to more specialized pathways, where these compounds are used as specific substrates or signaling molecules in ways that vary across taxa. This hypothesis implies that the plant attracts a wide range of microbes in the early stages of development when compared to later developmental stages by secreting sugars which are readily available for metabolism. As the plant develops, it begins to select among rhizosphere inhabitants by releasing phenolics and amino acids. The increase in the number of uniquely expressed microbial functional genes at later plant developmental stages may be indicative of a community-wide microbial response to shifting exudation toward more

recalcitrant or inhibitory compounds. Thus, only those microbes that have evolved means of detoxifying or utilizing these compounds will thrive. A detailed analysis of the rhizosphere microbes associated with Arabidopsis plants at different stages of development may provide a means of answering some of these questions. Micallef et al. (2009a) used denaturing gel gradient electrophoresis to demonstrate that rhizospheric microbial communities change with plant development. Further studies identifying the taxonomic microbial community of the rhizosphere would allow us to identify how specific microbial taxa are influenced by root exudation.

In nature, root exudation is affected by a myriad of factors. Here we analyzed the compounds released as root exudates in a controlled environment and correlated them with the functions present in the rhizosphere. Although we did see correlations between the identified compounds in the *in vitro* studies and the functions carried out in the rhizosphere *in vivo*, additional work is needed to clarify the impacts of root exudation changing not only with plant development, but also in response to the specific microbiomes present at each developmental stage. It is important to note that in our rhizosphere soil community analysis we have not excluded the contributions of components such as proteins and polysaccharides that contribute to root exudation. Our exudate profiling was not exhaustive, and other root exudate constituents, such as proteins and polysaccharides, might also contribute to the changes observed in the functional microbiome at different stages of plant development.

Rhizosphere driven selection of microbial functions has the potential to improve the development and health of plants in a sustainable manner. A deeper understanding of soil microbial functions over plant development can help devise better strategies for disease resistance and thus improving plant and soil health. However, further mechanistic studies are

required to identify specific microbial candidates that perform certain microbial functions of benefit to the plant, and to uncover pathways for inter-species signaling in the rhizosphere.

Tables

Table 2-1. List of the primers used in this study. RT-PCR was used to analyze the expression of sugar transporters, ABC transporters and genes involved in secondary metabolism. Putative M refers to genes involved in the putative monosaccharide transporter family.

	AGI Code	Forward	Reverse
AtSUC1	At1g71880	GGAGTCCAATCTGGTGCAAT	GAATCCTCCCATGGTCGTTG
AtSUC2	At1g22710	AACTTCATCCTCGCCATTTG	GCTTTGAAGGCAGGAGCATC
AtSUC3	At2g02860	TGGGGATCCAACAGGAGATA	CCGGTGGACTTGAAAGAACTC
AtSUC4	At1g09960	TATGGGTGCACTTGGTTTGA	GAGAGGGATGGGCTTCTGAAT
AtSUC5	At1g71890	AATCGATTGGTCGGAAAATG	ATAGCCCCTGACATGGCTGG
Putative M	At1g08920	GTGCGTTGCAAGTTGTGACT	CACCGGTAGAGGCCAATAGA
Putative M	At4g04760	AATCGGATGGTTCGCTATTG	CGGTCATCGATGTCTTGTTG
AtINT1	At2g43330	TTGGTCGGTTTAGGAGTTGG	GGCAGCAACAATGAGAGACA
AtINT2	At1g30220	GGGCATGTTGGATCTCTGAT	CATCGATCTTCGTTCAAGCA
AtINT3	At2g35740	TGGTGATGGTGATTGCTCAT	TATCCTCGTCAGCCGTCTCT
AtPLT4	At2g20780	CGTGAGCTTCTTAGCCCATC	GCACGTGACACAGAGAGGAA
Putative M	At1g79820	CCAAATTGTCGGAGTTGCTT	ATAACCCAGTGAACGGCAAG
AtSTP7	At4g02050	TGAATGCTGGAGCTGTGAAC	TGGCATGCAAATAGCCATAA
AtSTP10	At3g19940	CGAGAGAGGCAAAAATGAGG	TGTCCCGCTGGTCTAATTTC
AtpGlcT	At5g16150	CTCTGGCCAAGGTTCTTCTG	CTCTGATTCGGGATGCAAAT
Putative M	At2g48020	GGGGCTCTAACCACACTGAA	TTCTTCCGGCTCTGTCAACT
Putative M	At3g20460	GTTGAAATCGCTCCCAAAAA	ACGCTTGTCACAGACTGCAC
Putative M	At1g67300	CTGCTGCACTACTCGCTCTG	CGCCGAAAAATGTGGTAAGT
Putative M	At1g05030	GAATCACGAGTGGTGCTCAA	GAGCCAAAGCTGGCATAGAC
Putative M	At1g19450	GAGATCGCTCCACAGACCAT	GAGTAGAAGCCGACGACCTG
Putative M	At5g17010	CCCAACCTATTCCGTTCTGA	GCTTTCAAGCACTTCCCTTG
AtSUC9	At5g06170	AGCCGTTGGTTTCTTCGTTG	TCTTACTAATCACTCCAATAACAAGG
AtPDR2	At4g15230	TGGCAAGAGATGAAGTGTCAGGGAAAG	CTACAGCAGGATCTGGAATGATTTCTTGG
AtPDR4	At2g26910	CACGATTCATCAGCCTAGCA	ACATTGTGGTTTGGGGTGAT
AtPDR6	At2g36380	AGATGTTGACGTCACGAATCTTGCT	GTTGCCCTGCGTGAAAAGAATTG

	AGI Code	Forward	Reverse
AtPDR7	At1g15210	GGACATACACGCTTCCCACT	AAGCACACTTGTTCCCAACC
AtPDR8	At1g59870	AGAGCAGCGGCTATTGTGATGA	TGGCGTAGACGATGAGTGAGAT
AtPDR9	At3g53480	TGAGGAGAGGTATAACAGGAGGTC	GAGAGATTCAAAGAACGAGAGAGG
PAL1	At2g37040	GATTCTCGAAGCAATTACCAGTTT	GAGGAGAAGTACGAAGAGCGTAAC
PAL2	At3g53260	GTCAGAGTCAACACTCTTCTCCAA	TTACTTCAATTTGAGGACCTAGCC
PAL3	At5g04230	TACAACAACGGGTTACCCTCTAAT	GTTTTCCTTGAGATATAGCCCTGA
C4H	At2g30490	TCCTATCCTTAGACCATTCCTCAG	CTCAGACGAAGAGTCTCCTTAACC
4CL1	At1g51680	AGCTCGATAAGAGTGGTGAAATCT	CATCTTCTGATAACTCCGAATCCT
4CL2	At3g21240	TCGTTAGGGTTTGCTAAAGAGC	AAACACAACCTGTTTTGACACG
4CL3	At1g65060	CTGATCACTACCGATGAACCAA	CACCGGAAAGAACGAATCTAAC
FS1	At1g49390	GCCATCGATCTCAGTCTTCTCT	TGGACATGGAGGAAAGAAGTT
FS2	At5g63580	GGACCGAGAATCTTTTTCACAG	AGAGGAGGGAATGTAGTGGACA
FS3	At5g63590	CTGATACTGCGGTTGCTACAAG	CGTCGATCCATTGGTTATCTTT
F3H	At5g07990	TCCCTTAAAGGAACTGATCTTGAC	GAATCTCTCGGGTTTAAATGCTAA
CYP79B2	At4g39950	CAGAAGATCCTCTCTAACGGCTAC	CGATTTGAGTTCTCTTTCCTTCTC
CYP79B3	At2g22330	CTACACGACAATAGAGCTGAGGAA	ACCGTAACGGCTAAGTAAAACTTG
CYP79A2	At5g05260	GAGATTCTGAAGAAGCAAGACTCC	GATGGGTTAAAACCGTAAACAAAG
CYP71B15	At3g26830	GAGGAAGTGCTAAAGATCAACGAT	ATTCTTCAATCTCCTGTTCTGACC
β-Actin		CAACTGGGACGAYATGGAGA	GAGTCATCTTCTCTCTGTTGGCC

Table 2-2. Summary of the 454 pyrosequencing data for each sample.

	17d	**24d**	**31d**	**38d**	**Bulk Soil**
Total Number of Sequences uploaded to MG-RAST	14740	14740	14740	14740	14740
Mean Sequence Length	380	425	415	437	412
Predicted proteins with known function	7571	8632	7650	8699	8679
Predicted protein with unknown function	6194	5702	6114	5490	5959
Predicted Protein Features	14707	15904	15442	15829	15908
Identified Protein Features	7795	9093	8219	9245	9013
Identified Functional Categories to the M5NR protein datbase[a]	6858	8231	7165	8265	8154
Predicted KEGG Orthology[b]	4449	4255	4573	4750	4047
Unique Features in KEGG hierarchical classification	1694	1533	1489	1797	1366
Reads attributed to the unique feature of KEGG hierarchical classification[c]	2142	1611	1647	1916	1420
KEGG functions expressed in each sample after hierarchical classification	528	516	553	560	439
Unique KEGG functions for each time point	139	114	145	145	98
Reads attributed to unique KEGG functions for each time point	225	234	256	259	204

[a] M5NR protein database consists of NCBI's nr, KEGG database, EGGnogs, and SEED database
[b] Report abundances using the KEGG protein database that include all functional labels
[c] Report abundances using the KEGG protein database that supports hierarchical relationships between functions

Table 2-3. Table detailing the compounds released via root exudation by the plant as it develops. These were collected over a period of 3 days (7-10 days, 14-17 days, 21-24 days and 28-31 days). Compounds were detected using GC-MS. Numbers indicate the average area under the curve of three replicates, numbers in parenthesis indicate the standard deviation.

Compound	Category	7-10 days	14-17 days	21- 24 days	28-31 days
1-deoxyerythritol	Sugar Alcohol	1933.06 (260.45)	2735.5 (55.06)	11315.12 (1060.15)	3756.53 (90.38)
2-hydroxyvaleric acid	Phenolics	12057.96 (1049.56)	20930.5 (1357.67)	17989.5 (1910.57)	3681.93 (285.01)
3-hydroxybutanoic acid	Phenolics	26410.08 (7734.05)	10492.2 (2004.28)	2061.1 (162.69)	741.57 (5.28)
3-hydroxypropionic acid	Phenolics	9308.5 (905.44)	12501.3 (615.95)	22432.5 (476.55)	7714.9 (131.67)
4-hydroxybutyric acid	Phenolics	3118.83 (269.45)	3986.32 (115.35)	2332.9 (118.05)	905.57 (39.05)
alanine	Amino Acid	20380.66 (2154.44)	28883.25 (790.75)	125416 (11632.23)	257026.66 (8756.43)
arabinose	Sugar	4701.83 (683.13)	7430.8 (362.88)	3108.75 (312.58)	1560.53 (399.06)
arabitol	Sugar Alcohol	1252.54 (224.23)	1437.37 (114.38)	2638.05 (88.27)	1462.06 (48.26)
asparagine minor	Amino Acid	121.44 (26.6)	169.53 (38.16)	420.9 (34.18)	2687.23 (167.65)
benzoic acid	Phenolics	4851.96 (815.49)	6592.47 (301.82)	8785.3 (656.46)	6487.73 (708.56)
beta-alanine	Amino Acid	414.7 (75.56)	624.61 (44.88)	5441.97 (201.44)	5191 (352.57)
butyrolactam	Phenolics	9875.53 (1089.79)	10804.55 (615.98)	172080 (16103.72)	31995.33 (6279.88)
capric acid	Phenolics	8792.66 (1241.18)	9859.3 (457.12)	4583.37 (199.49)	1131.68 (195.91)
cyclohexylamine	Amino Acid	93297 (34137.08)	147046.75 (20157.1)	18729.75 (1323.14)	12392.66 (351.85)
erythritol	Sugar Alcohol	1000.66 (297.79)	1453.52 (55.92)	4323.15 (573.73)	2200.46 (107.52)
ethanolamine	Amino Acid	21175.93 (3919.15)	57607.25 (1869.85)	502022.5 (24338.64)	558026.66 (48001.73)
fructose	Sugar	266676.66 (13868.22)	214670 (2743.66)	67314.25 (9310.25)	12778.7 (1344.15)
fucose + rhamnose	Sugar	2614.9 (304.29)	9507.2 (794.75)	28415 (836.67)	7621.9 (331.57)
fumaric acid	Phenolics	282.23 (40.36)	319.67 (17.99)	4451.87 (820.7)	27146.33 (699.84)
γ-aminobutyric acid	Phenolics	68.8 (30.13)	754.51 (178.34)	63043.5 (6834.63)	287210 (29177.79)
galactose	Sugar	291386.66 (28268.19)	95830.25 (8682.93)	1939.85 (126.17)	745.21 (46.18)
glucose	Sugar	436110 (37741.68)	256895 (4910.03)	42315.75 (6521.47)	10486.56 (695.25)
glucose-1-phosphate	Sugar	784.83 (255.06)	857.51 (55.14)	2419.57 (55.19)	1199.7 (78.98)
glyceric acid	Phenolics	440.53 (20.64)	546.65 (142.85)	10077.07 (785.22)	9175.9 (637.01)
glycerol	Sugar Alcohol	274263.33 (62052.43)	389172.5 (39583.93)	103075.25 (7835.79)	38103.66 (620.74)
glycine	Amino Acid	60957.33 (15722.64)	64056.25 (2459.06)	127173 (8954.69)	120923.33 (5529.88)
glycolic acid	Phenolics	6675.4 (1728.38)	20901.25 (258.8)	2019.06 (737.81)	941.32 (266.5)
hydroxylamine	Amino Acid	3858.49 (983.25)	7684.4 (1160.47)	1927.55 (129.2)	1990.01 (970.9)
inositol myo-	Sugar Alcohol	2875.86 (619.56)	2402.42 (124.1)	17899.55 (2618.37)	8787.4 (725.12)
isoleucine	Amino Acid	342.89 (135.93)	1084.05 (210.45)	26118.75 (1584.97)	58209 (2707.19)
lactic acid	Phenolics	23300.73 (5238.51)	73611 (7104.4)	45983.25 (6544.83)	3574.86 (192.65)

Compound	Category	7-10 days	14-17 days	21- 24 days	28-31 days
lauric acid	Phenolics	12790.33 (460.52)	12362 (801.36)	6192.55 (328.03)	1550.8 (146.13)
levoglucosan	Phenolics	1869.33 (316.03)	2016.9 (79.29)	3851.27 (89.15)	521.8 (29.34)
N-acetyl-D-hexosamine	Phenolics	169.08 (90.71)	70.91 (29.34)	1896.42 (145.72)	1122.67 (123.55)
N-acetyl-D-mannosamine	Phenolics	220.73 (37.44)	489.11 (16.85)	5020 (357.59)	1980.56 (86.99)
oxoproline	Amino Acid	80.04 (17.55)	83.39 (34.88)	4896.02 (656.58)	331346.66 (67685.3)
palmitic acid	Phenolics	780.29 (31.94)	1068.28 (50.66)	641.72 (24.09)	199.69 (13.35)
pelargonic acid	Phenolics	19089.13 (3845.42)	21068.25 (1136.18)	8041.9 (610.54)	7801.7 (1545.58)
propane-1,3-diol	Phenolics	57358 (4163.92)	54362.75 (1825.69)	23968.75 (1078.21)	5671.4 (1075.56)
putrescine	Phenolics	9849.96 (1357.04)	8452.6 (2118.03)	3959.92 (156.26)	36865.66 (2999.73)
ribose	Sugar	287.19 (50.08)	2347.92 (272.29)	34444.25 (3057.63)	6991.33 (622.33)
serine	Amino Acid	7103.1 (1333.6)	22407.25 (1428.11)	99913.75 (3031.47)	61061.33 (3118.76)
shikimic acid	Phenolics	80701 (20108.85)	8955.52 (2721.09)	5485.27 (927.12)	1563.33 (468.42)
stearic acid	Phenolics	30803.33 (4158.47)	14150.6 (1359.91)	5694.55 (282.42)	1157.82 (86.15)
succinic acid	Phenolics	193.23 (28.36)	380.06 (46.95)	1215.59 (186.23)	5041.23 (395.21)
sucrose	Sugar	28497.63 (6433.51)	13086.02 (1755.13)	1982.77 (153.69)	205.75 (31.07)
tagatose	Sugar	515.14 (24.41)	868.72 (68.69)	54.51 (13.34)	15.98 (5.96)
threitol	Sugar Alcohol	217.08 (83.28)	1640.09 (130.4)	6245.45 (337.02)	1582.93 (45.25)
threonic acid	Phenolics	64.3 (12.65)	136.89 (17.58)	2410.9 (189.88)	3450 (188.38)
threonine	Amino Acid	1441.9 (361.03)	1719.07 (73.8)	2524.92 (229.55)	14676.66 (914.86)
tocopherol alpha	Phenolics	19102 (3951.27)	9558.47 (704.01)	2505.55 (159.38)	448.38 (87.43)
uracil	Phenolics	666.59 (181.86)	1212.99 (139.24)	1422.87 (84.62)	4530.8 (477.83)
urea	Phenolics	516.65 (101.66)	8072.19 (1675.14)	316592.5 (24815.98)	305800 (17260.82)
valine	Amino Acid	8361.83 (1076.37)	16381.5 (575.52)	90746.75 (3294.57)	122726.66 (3983.52)
xylitol	Sugar Alcohol	99.05 (39.74)	118.06 (23.84)	3719.12 (329.48)	2011.8 (78.4)

Table 2-4. Correlation of the compounds identified in the root exudates with the abundance or number of functional genes in each sample. * Correlation was statistically significant ($p<0.05$).

	Corresponding functional genes in each sample		**Abundance of corresponding functional genes in each sample**	
	Pearson	**Spearman**	**Pearson**	**Spearman**
Amino Acids	0.86	1.00*	0.63	0.80
Sugars	-0.55	-0.40	-0.24	0.20
Phenolics	0.95*	0.80	0.98*	0.80

Figures

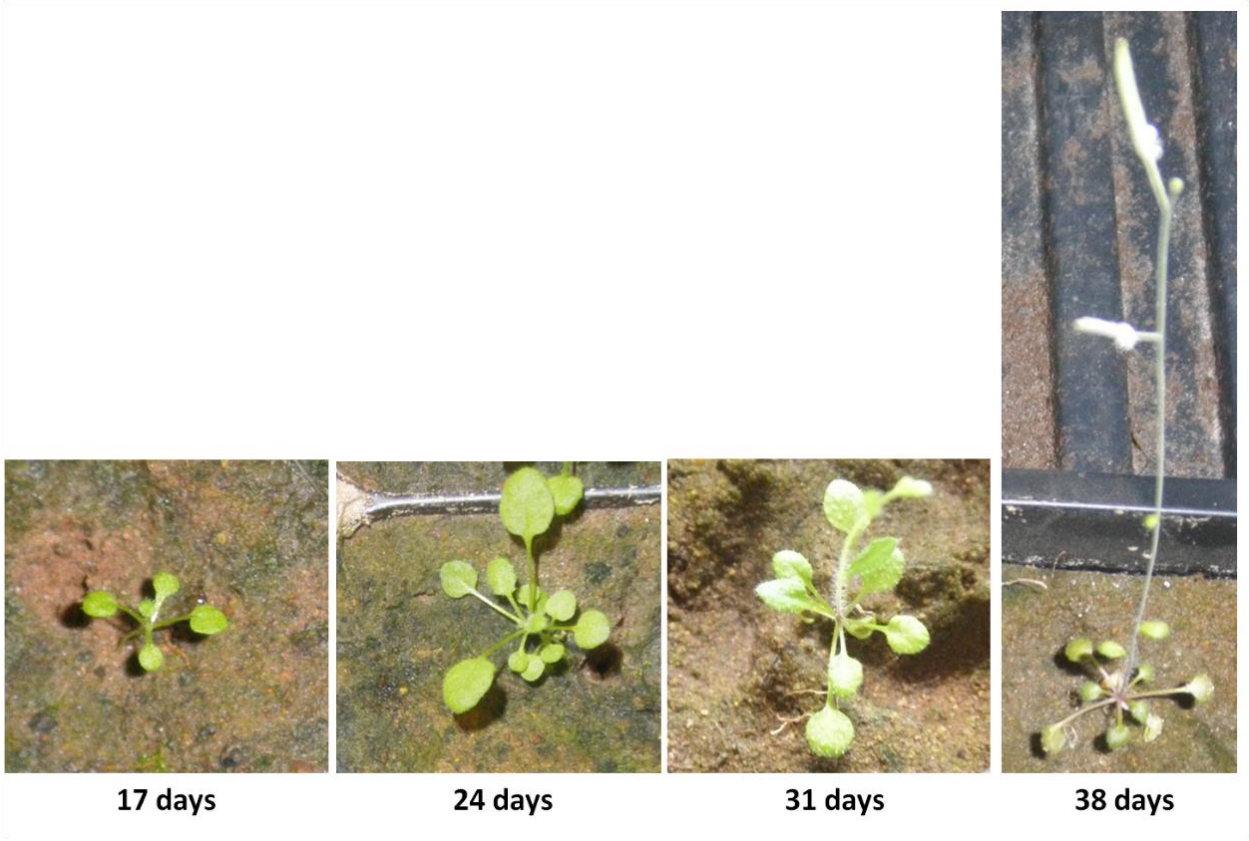

Figure 2-1. Soil grown Arabidopsis thaliana Col-0 at each plant developmental stage (17, 24, 31 and 38 days).

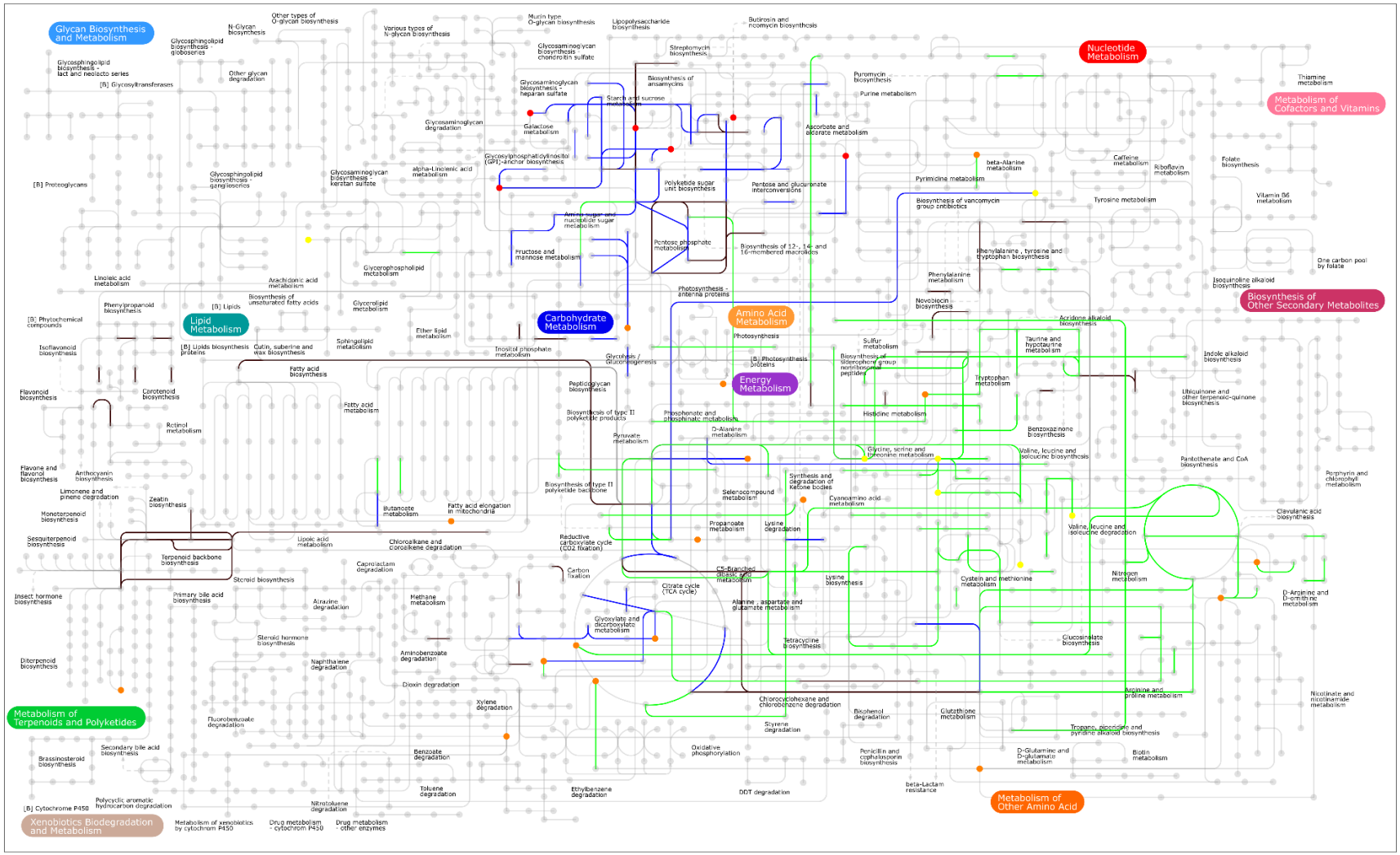

Figure 2-2. iPATH 2 KEGG Map exhibiting the functional genes involved in Metabolism with the identified root exudate compounds. Blue lines: functional genes involved in Carbohydrate Metabolism; Green lines: functional genes involved in Amino Acid Metabolism; Brown.

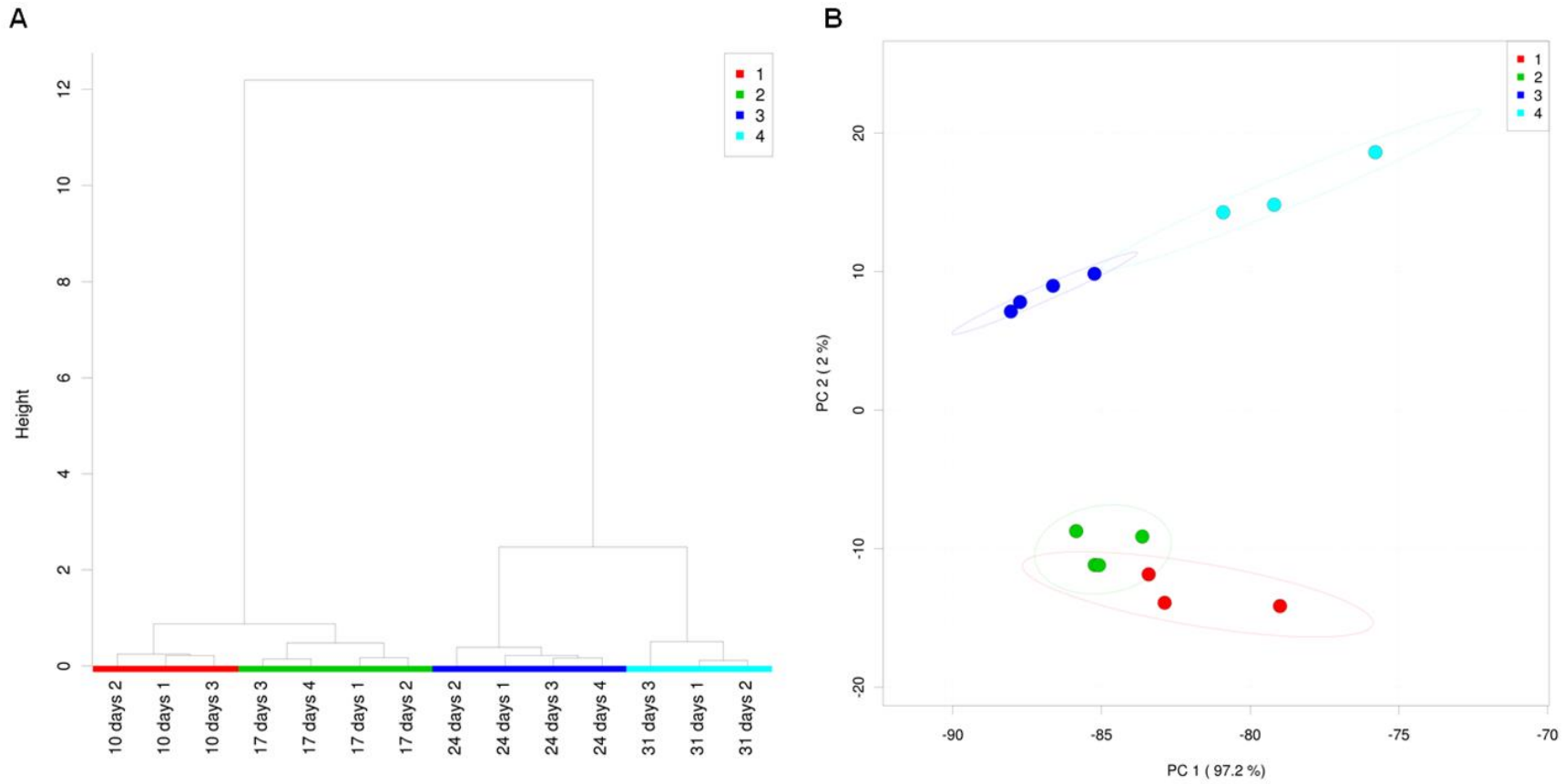

Figure 2-3. Multivariate analysis of the root exudates analyzed by GC-MS. (A) Cluster analysis (Ward method) based on the 107 compounds detected in root exudates collected at 7-10d, 14-17d, 21-24d, and 28-31d. (B) Principal Component Analysis (PCA) of the 107 root exudate compounds from samples collected at 7-10d, 14-17d, 21-24d, and 28-31d. Dashed ellipses indicate the 95% confidence region of each time point. Red: 7-10 days; Green: 14-17 days; Blue: 21-24 days; Aqua: 28-31 days. 1: 7-10d; 2: 14-17d; 3: 21-24d; 4: 28-31d.

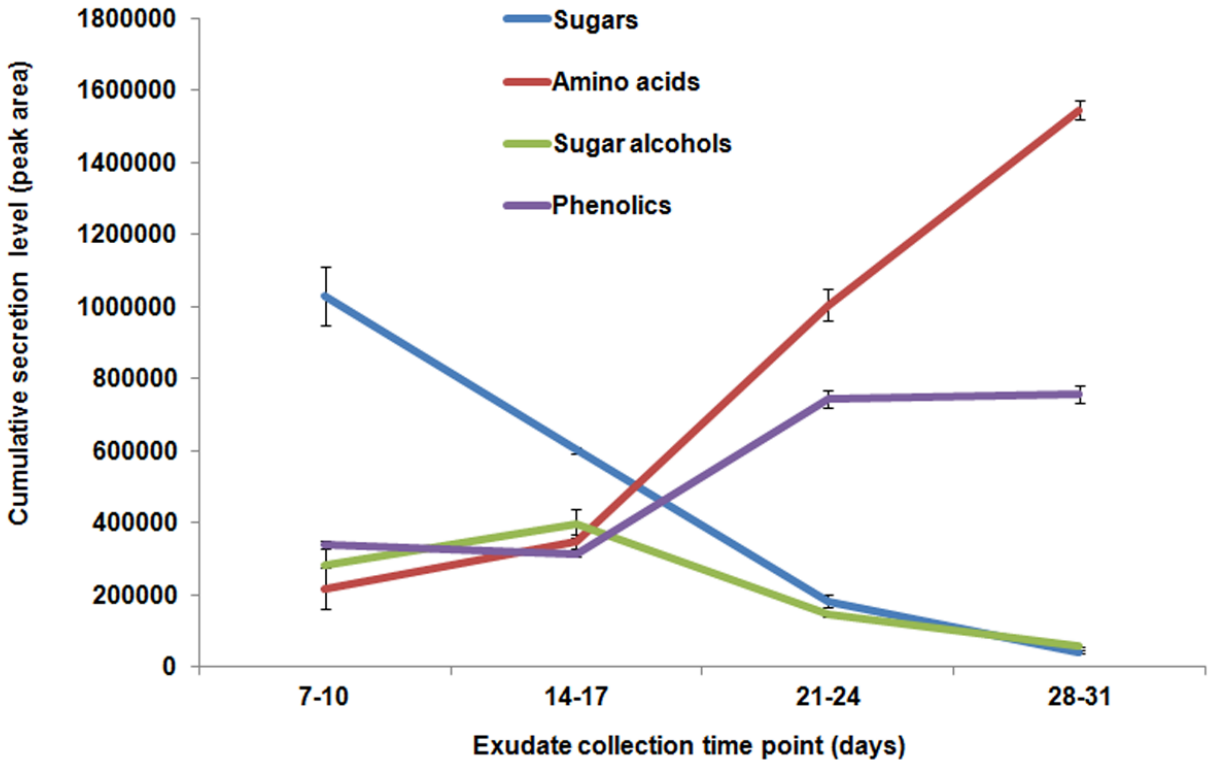

Figure 2-4. Arabidopsis root exudate composition across development. Identified compounds were grouped into chemical classes, and secretion levels were calculated based on cumulative peak area after normalization.

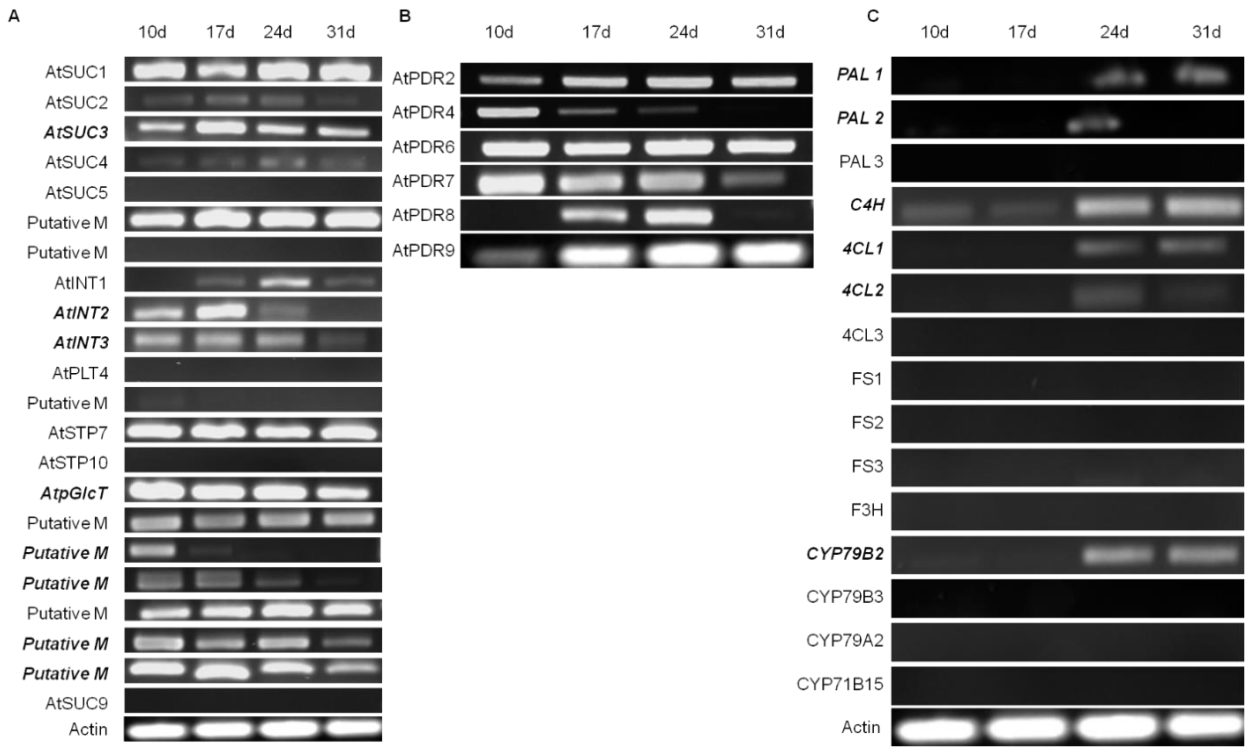

Figure 2-5. Arabidopsis root gene expression profiles. (A) sugar transporters, (B) ABC transporters, and (C) genes involved in secondary metabolism, as measured by semi-quantitative RT-PCR. Gene names are listed on the left side. The time points at which root tissues were collected are shown along the top. Bold and italicized text indicates genes whose expression profiles are most consistent with the GC-MS analysis of collected exudates.

REFERENCES

Abu-Nada Y, Kushalappa A, Marshall W, Al-Mughrabi K, Murphy A (2007). Temporal dynamics of pathogenesis-related metabolites and their plausible pathways of induction in potato leaves following inoculation with Phytophthora infestans. *Eur J Plant Pathol* **118:** 375.

Badri DV, Loyola-Vargas VM, Broeckling CD, De-la-Pena C, Jasinski M, Santelia D *et al* (2008). Altered profile of secondary metabolites in the root exudates of Arabidopsis ATP-binding cassette transporter mutants. *Plant Physiol* **146:** 762-771.

Badri DV, Quintana N, El Kassis EG, Kim HK, Choi YH, Sugiyama A *et al* (2009). An ABC transporter mutation alters root exudation of phytochemicals that provoke an overhaul of natural soil microbiota. *Plant Physiol* **151:** 2006-2017.

Badri DV, Vivanco JM (2009). Regulation and function of root exudates. *Plant Cell Environ* **32:** 666-681.

Badri DV, Loyola-Vargas VM, Broeckling CD, Vivanco JM (2010). Root secretion of phytochemicals in Arabidopsis is predominantly not influenced by diurnal rhythms. *Mol Plant* **3:** 491-498.

Badri DV, Chaparro JM, Manter DK, Martinoia E, Vivanco JM (2012). Influence of ATP-Binding Cassette Transporters in Root Exudation of Phytoalexins, Signals, and in Disease Resistance. *Front Plant Sci* **3:** 149.

Behera B, Wagner GH (1974). Microbial Growth Rate In Glucose-amended Soil. *Soil Sci Soc Am J* **38:** 591-594.

Bennett RN, Wallsgrove RM (1994). Tansley Review No. 72. Secondary Metabolites in Plant Defence Mechanisms. *New Phytologist* **127:** 617.

Berendsen RL, Pieterse CM, Bakker PA (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478-486.

Berg G (2009). Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Appl Microbiol Biotechnol* **84:** 11-18.

Biedrzycki ML, Jilany TA, Dudley SA, Bais HP (2010). Root exudates mediate kin recognition in plants. *Commun Integr Biol* **3:** 28-35.

Blencke HM, Homuth G, Ludwig H, Mader U, Hecker M, Stulke J (2003). Transcriptional profiling of gene expression in response to glucose in Bacillus subtilis: regulation of the central metabolic pathways. *Metab Eng* **5:** 133-149.

Bowra BJ, Dilworth MJ (1981). Motility and Chemotaxis Towards Sugars in Rhizobium-Leguminosarum. *J Gen Microbiol* **126:** 231-235.

Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR *et al* (2001). Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell* **13:** 1499-1510.

Broeckling CD, Broz AK, Bergelson J, Manter DK, Vivanco JM (2008). Root exudates regulate soil fungal community composition and diversity. *Appl Environ Microbiol* **74:** 738-744.

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

Chevrot R, Rosen R, Haudecoeur E, Cirou A, Shelp BJ, Ron E *et al* (2006). GABA controls the level of quorum-sensing signal in Agrobacterium tumefaciens. *Proc Natl Acad Sci U S A* **103:** 7460-7464.

De-la-Pena C, Lei Z, Watson BS, Sumner LW, Vivanco JM (2008). Root-microbe communication through protein secretion. *J Biol Chem* **283:** 25247-25255.

De-la-Pena C, Badri DV, Lei Z, Watson BS, Brandao MM, Silva-Filho MC *et al* (2010). Root secretion of defense-related proteins is development-dependent and correlated with flowering time. *J Biol Chem* **285:** 30654-30665.

de Weert S, Vermeiren H, Mulders IHM, Kuiper I, Hendrickx N, Bloemberg GV *et al* (2002). Flagella-Driven Chemotaxis Towards Exudate Components Is an Important Trait for Tomato Root Colonization by Pseudomonas fluorescens. *Mol Plant Microbe Interact* **15:** 1173-1180.

Edreva A, Velikova V, Tsonev T, Dagnon S, Gurel A, Aktas L *et al* (2008). Stress-protective role of secondary metabolites: diversity of functions and mechanisms. *Gen Appl Plant Physiology* **34:** 67-78.

Fiehn O, Wohlgemuth G, Scholz M (2005). Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. In: Ludascher B, Raschid L (eds). *Data Integration In The Life Sciences, Lecture Notes in Computer Science*. Springer-Verlag: Berlin/Heidelberg. pp 224-239.

Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee do Y, Lu Y *et al* (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J* **53:** 691-704.

Flores HE, Vivanco JM, Loyola-Vargas VM (1999). 'Radicle' biochemistry: the biology of root-specific metabolism. *Trends Plant Sci* **4:** 220-226.

Hamer U, Marschner B (2005). Priming effects in different soil types induced by fructose, alanine, oxalic acid and catechol additions. *Soil Biol and Biochem* **37:** 445.

Haudecoeur E, Planamente S, Cirou A, Tannieres M, Shelp BJ, Morera S *et al* (2009). Proline antagonizes GABA-induced quenching of quorum-sensing in Agrobacterium tumefaciens. *Proc Natl Acad Sci U S A* **106:** 14587-14592.

Horiuchi J, Prithiviraj B, Bais HP, Kimball BA, Vivanco JM (2005). Soil nematodes mediate positive interactions between legume plants and rhizobium bacteria. *Planta* **222:** 848-857.

Jimenez-Zurdo JI, Garcia-Rodriguez FM, Toro N (1997). The Rhizobium meliloti putA gene: its role in the establishment of the symbiotic interaction with alfalfa. *Mol Microbiol* **23:** 85-93.

Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res* **32:** D277-280.

Kuzyakov Y (2002). Review: Factors affecting rhizosphere priming effects. *J PLANT NUTR SOIL SC* **165:** 382-396.

Lambert JD, Chan XY, Spiecker B, Sweet HC (2010). Characterizing the embryonic transcriptome of the snail Ilyanassa. *Integr Comp Biol* **50:** 768-777.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25.

Li Y, Peng Q, Selimi D, Wang Q, Charkowski AO, Chen X *et al* (2009). The plant phenolic compound p-coumaric acid represses gene expression in the Dickeya dadantii type III secretion system. *Appl Environ Microbiol* **75:** 1223-1228.

Loake GJ, Ashby AM, Shaw CH (1988). Attraction of Agrobacterium-Tumefaciens C58C1 Towards Sugars Involves a Highly Sensitive Chemotaxis System. *J Gen Microbiol* **134:** 1427-1432.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332:** 1097-1100.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al* (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *Bmc Bioinformatics* **9**.

Micallef SA, Channer S, Shiaris MP, Colon-Carmona A (2009a). Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. *Plant Signal Behav* **4:** 777-780.

Micallef SA, Shiaris MP, Colon-Carmona A (2009b). Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. *J Exp Bot* **60:** 1729-1742.

Murashige T, Skoog F (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol Plant* **15:** 473.

Naher UA, Radziah O, Halimi MA, Shamsuddin ZH, Mohd Razi I (2008). Effect of Inoculation on Root Exudates Carbon Sugar and Amino Acids Production of Different Rice Varieties. *Res J Microbiol* **3:** 580-587.

Neal AL, Ahmad S, Gordon-Weeks R, Ton J (2012). Benzoxazinoids in Root Exudates of Maize Attract Pseudomonas putida to the Rhizosphere. *PLoS One* **7:** e35498.

Nicholson RL, Hammerschmidt R (1992). Phenolic Compounds and Their Role in Disease Resistance. *Annual Review of Phytopathology* **30:** 369-389.

Oku S, Komatsu A, Tajima T, Nakashimada Y, Kato J (2012). Identification of Chemotaxis Sensory Proteins for Amino Acids in Pseudomonas fluorescens Pf0-1 and Their Involvement in Chemotaxis to Tomato Root Exudate and Root Colonization. *Microbes Environ*.

Omrane S, Chiurazzi M (2009). A variety of regulatory mechanisms are involved in the nitrogen-dependent modulation of the nodule organogenesis program in legume roots. *Plant Signal Behav* **4:** 1066-1068.

Ordal GW, Villani DP, Rosendahl MS (1979). Chemotaxis Towards Sugars by Bacillus-Subtilis. *J Gen Microbiol* **115:** 167-172.

Ortiz-Castro R, Contreras-Cornejo HA, Macias-Rodriguez L, Lopez-Bucio J (2009). The role of microbial signals in plant growth and development. *Plant Signal Behav* **4:** 701-712.

Phillips DA, Fox TC, King MD, Bhuvaneswari TV, Teuber LR (2004). Microbial products trigger amino acid exudation from plant roots. *Plant Physiol* **136:** 2887-2894.

Pillai BV, Swarup S (2002). Elucidation of the flavonoid catabolism pathway in Pseudomonas putida PML2 by comparative metabolic profiling. *Appl Environ Microbiol* **68:** 143-151.

Poysti NJ, Loewen ED, Wang Z, Oresnik IJ (2007). Sinorhizobium meliloti pSymB carries genes necessary for arabinose transport and catabolism. *Microbiology* **153:** 727-736.

Ramos-Gonzalez MI, Campos MJ, Ramos JL (2005). Analysis of Pseudomonas putida KT2440 gene expression in the maize rhizosphere: in vivo expression technology capture and identification of root-activated promoters. *J Bacteriol* **187:** 4033-4041.

Rovira AD (1969). Plant Root Exudates. *Bot Rev* **35:** 35-57.

Rudrappa T, Czymmek KJ, Pare PW, Bais HP (2008). Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol* **148:** 1547-1556.

Ryan RP, Dow JM (2008). Diffusible signals and interspecies communication in bacteria. *Microbiology* **154:** 1845-1858.

Sana TR, Fischer S, Wohlgemuth G, Katrekar A, Jung KH, Ronald PC *et al* (2010). Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen Xanthomonas oryzae pv. oryzae. *Metabolomics* **6:** 451-465.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75:** 7537-7541.

Schmulling T, Schell J, Spena A (1988). Single genes from Agrobacterium rhizogenes influence plant development. *Embo J* **7:** 2621-2629.

Selvakumar G, Panneerselvam P, Ganeshamurthy AN, Maheshwari DK (2012). Bacterial Mediated Alleviation of Abiotic Stress in Crops. In: Maheshwari DK (ed). *Bacteria in Agrobiology: Stress Management*. Springer: New York, NY. pp 205-224.

Shaw CH (1991). Swimming against the tide: Chemotaxis in Agrobacterium. *BioEssays* **13:** 25-29.

Shcheglov AS, Zhulidov EA, Bogdanova DAS (2007). Normalization of cDNA Libraries. In: Buzdin A, Lukyanov S (eds). *Nucleic Acid Hybridization*. Springer: Dordrecht. pp 97-124.

Simons M, Permentier HP, deWeger LA, Wijffelman CA, Lugtenberg BJJ (1997). Amino acid synthesis is necessary for tomato root colonization by Pseudomonas fluorescens strain WCS365. *MPMI* **10:** 102-106.

Steinkellner S, Lendzemo V, Langer I, Schweiger P, Khaosaad T, Toussaint J-P *et al* (2007). Flavonoids and Strigolactones in Root Exudates as Signals in Symbiotic and Pathogenic Plant-Fungus Interactions. *Molecules* **12:** 1290-1306.

Tang CS, Cai WF, Kohl K, Nishimoto RK (1995). Plant Stress And Allelopathy. In: Inderjit, Dakshini KMM, Einhellig FA (eds). *Allelopathy: Organisms, Processes, And Applications*. Amer Chemical Soc: Washington. pp 142-157.

Volkers RJ, Ballerstedt H, Ruijssenaars H, de Bont JA, de Winde JH, Wery J (2009). TrgI, toluene repressed gene I, a novel gene involved in toluene-tolerance in Pseudomonas putida S12. *Extremophiles* **13:** 283-297.

Williams MA, Taylor EB, Mula HP (2010). Metaproteomic characterization of a soil microbial community following carbon amendment. *Soil Biol and Biochem* **42:** 1148.

Xia J, Psychogios N, Young N, Wishart DS (2009). MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res* **37:** W652-660.

Yamada T, Letunic I, Okuda S, Kanehisa M, Bork P (2011). iPath2.0: interactive pathway explorer. *Nucleic Acids Res* **39:** W412-415.

Zamioudis C, Pieterse CM (2012). Modulation of host immunity by beneficial microbes. *Mol Plant Microbe Interact* **25:** 139-150.

Zhang J, Subramanian S, Stacey G, Yu O (2009). Flavones and flavonols play distinct critical roles during nodulation of Medicago truncatula by Sinorhizobium meliloti. *Plant J* **57:** 171-183.

CHAPTER 3 RHIZOSPHERE MICROBIOME ASSEMBLAGE IS AFFECTED BY PLANT DEVELOPMENT[1]

Synopsis

There is a concerted understanding of the ability of root exudates to influence the structure of rhizosphere microbial communities. However, our knowledge of the connection between plant development, root exudation, and microbiome assemblage is limited. Here, we analyzed the structure of the rhizospheric bacterial community associated with Arabidopsis at four time points corresponding to distinct stages of plant development: seedling, vegetative, bolting, and flowering. Overall, there were no significant differences in bacterial community structure, but we observed that the microbial community at the seedling stage was distinct from the other developmental time points. At a closer level, phylum such as Acidobacteria, Actinobacteria, Bacteroidetes, Cyanobacteria, and specific genera within those phyla followed distinct patterns associated with plant development and root exudation. These results suggested that the plant can select a subset of microbes at different stages of development, presumably for specific functions. Accordingly, metatranscriptomics analysis of the rhizosphere microbiome revealed that 81 unique transcripts were significantly ($p<0.05$) expressed at different stages of plant development. For instance, genes involved in streptomycin synthesis were significantly induced at bolting and flowering stages, presumably for disease suppression. We surmise that plants secrete blends of compounds and specific phytochemicals in the root exudates that are

[1] Reprinted in part from Chaparro, JM; Badri, DV; Vivanco, JM, Rhizosphere microbiome assemblage is affected by plant development. The ISME J 2014, 8, 790-803, DOI: 10.1038/ismej.2013.196.

differentially produced at distinct stages of development to help orchestrate rhizosphere microbiome assemblage.

Introduction

Plants such as bean, maize, soybean, cowpea, cabbage, cotton, and *Arabidopsis* exhibit age-related resistance (ARR) (Develey-Riviere and Galiana 2007). For example, susceptibility to *Puccinia sorghi* (common rust) in maize is manifested in younger plants but as the plants mature, the level of disease resistance augments (Abedon and Tracy 1996). In *Arabidopsis*, transitions from the vegetative to the floral phase correlates with resistance to *Pseudomonas syringae* (Kus et al 2002, Rusterucci et al 2005). While ARR at the molecular level has been studied with respect to leaf pathogens, little discussion has focused on root defense strategies and their role in ARR. However, there are some indirect correlations; for example, cotton (Zaki et al 1998) or bean (Nicoli et al 2012) plants are more susceptible to root disease (*Rhizoctonia solani* or *Fusarium* root rot, respectively) at the seedling stage. Recent studies have shown that root secretion of defense-related proteins is enhanced during flowering time (De-la-Pena et al 2010) suggesting an involvement of plant roots in ARR. Similarly, Chaparro et al. (2013) have shown that *Arabidopsis* roots release more phenolic related compounds at later stages of life which might be correlated to defense strategies against pathogens as secondary metabolites are involved in plant immunity against bacterial and fungal pathogens (An and Mou 2011, Bednarek 2012, Clay et al 2009, Millet et al 2010, Rogers et al 1996). Thus, there is a need to understand the influence of plant development on microbial associations that might occur naturally in the rhizosphere related to defense but also to other vital plant necessities such as nutrient acquisition.

Under natural conditions, plants tend to require higher quantities of N at later stages of development (Kelly et al 1995, Malagoli et al 2004, Nazoa et al 2003, Rossato et al 2001) but

exactly how this N is obtained under natural conditions is unknown. The classically studied symbiotic relationship between rhizobia and legumes has shown that symbiosis occurs only when the plant is under N-limiting conditions (Davidson and Robson 1986, Eaglesham 1989, Zahran 1999). Similarly, the secretion of flavones and flavonols that initiate rhizobia-legume symbiosis is enhanced under N-limiting conditions (Coronado et al 1995, Zhang et al 2009). This combined information suggests that the plant may have some control over the identity and functionality of the rhizosphere microbiome.

Studies have shown that rhizospheric fungal and bacterial communities of a wide range of plants (i.e., *Arabidopsis*, *Medicago*, maize, pea, wheat, and sugar beet) change according to a plant developmental gradient (Baudoin et al 2002, Houlden et al 2008, Micallef et al 2009a, Mougel et al 2006). In these studies the microbial communities were assessed through automated ribosomal intergenic spacer analysis or denaturing gradient gel electrophoresis techniques that produce a fingerprint of the community structure but not of its members' identity. While these studies demonstrated that plant microbial communities change in response to plant development they were not able to distinguish how or which microbes contribute to the changes observed. For example, Micallef et al. (2009a) through denaturing gradient gel electrophoresis analysis observed that *Arabidopsis* rhizosphere microbial communities varied with plant development and observed that microbial communities in early plant development were more distinct to the bulk soil and that this difference decreased with plant age. Similarly, an assessment of the potato rhizosphere demonstrated that young potato plants showed cultivar-dependent rhizosphere microbial communities but these differences in the microbiomes disappeared as the plants aged (Inceoglu et al 2011). Soybean rhizosphere microbial communities were also influenced by plant development; early reproductive growth stages of the soybean plant produced more complex

microbial communities than late stage soybean plants (Xu et al 2009). An assessment of the microbial community structure through plant development focusing on the members that make up the community is warranted. Incidentally, the recent characterization of the *Arabidopsis thaliana* core microbiome provides a tool to decipher the influence of the plant on the rhizosphere microbiome at different stages of development (Bulgarelli et al 2012, Lundberg et al 2012).

Evidence demonstrating the close ties root exudates have on the microbial composition of the rhizosphere is mounting (Badri et al 2008, 2013a, Broeckling et al 2008, Chaparro et al 2012, 2013, Micallef et al 2009b), whereby many chemicals present in root exudates act as substrates, chemotactic, or signaling molecules to orchestrate changes in microbial composition (Badri and Vivanco 2009, Badri et al 2013a, Bais et al 2006, de Weert et al 2002, Horiuchi et al 2005, Jain and Nainawatee 2002, Neal et al 2012, Shaw 1991). Recently, it was reported that the composition of *Arabidopsis* root exudates change following a plant developmental gradient (Chaparro et al 2013). Cumulative secretion levels of sugars and sugar alcohols were higher in early time points and decreased through plant growth. In contrast, the cumulative secretion levels of amino acids and phenolics increased over time. Accordingly, it was hypothesized that seedlings of roots release sugars as substrates for a wide diversity of microbes at early stages of development but as the plant ages it releases specific substrates and potentially antimicrobial compounds in an effort to select for particular microbial inhabitants of the rhizosphere (Badri et al 2013a, Chaparro et al 2013). This potential selection of microbes in the rhizosphere as the plant ages might be associated with the ability of beneficial microbes to suppress pathogenic ones (Mendes et al 2011), trigger induced systemic tolerance to overcome abiotic stress (Selvakumar et al 2012), increase the plant's innate immunity (Zamioudis and Pieterse 2012),

help in mineral nutrition (Bolan 1991, van der Heijden et al 2008), and in overall plant health (Berendsen et al 2012, Chaparro et al 2012). Here, we tested this hypothesis by analyzing the rhizosphere microbial composition of *Arabidopsis* by 454 pyrosequencing at four distinct physiological stages of development: seedling (four-five leaf stage), vegetative (rosette), bolting, and flowering. We did not include samples past the flowering stage because previous studies have determined that rhizosphere microbial communities converge past the flowering stage (Lundberg et al 2012, Micallef et al 2009a). Further, a metatranscriptomics analysis of the rhizomicrobiome was also conducted to ascertain a relationship between plant growth and microbiome functioning.

Materials and methods

Soil experiment

Soil where *Arabidopsis thaliana* genotype (Pna-10) (Li et al 2010) has naturally grown for more than 8 years was collected in July 2011 from the Michigan Extension Station, Benton Harbor, MI (42° 05’ 34’’ N, 86° 21’ 19’’ W, elevation 630 feet). The soil is described in detail in Broeckling et al. (2008); soil from the same site was used in previous studies (Badri et al 2009, 2013a, Broeckling et al 2008). *Arabidopsis thaliana* seeds were surface-sterilized with Clorox for one minute and subsequently rinsed four times with sterile distilled water. Sterile seeds were placed on Murashige and Skoog (MS) media (Murashige and Skoog 1962) supplemented with 0.9% bactoagar and 3% sucrose Petri plates. Seeds were incubated in a growth chamber with photoperiod 16 h light/8 h night at 25°C for seven days. The *Arabidopsis* seven-day-old seedlings did not introduce any microbes to the system as they were surface sterilized with bleach and no bacterial growth was observed on the MS agar plates even after 7 days of growth. Six replicate pots were maintained for each of the four developmental time points, and one 7-

day-old seedling was transplanted to each pot. Individual plants were grown until the following stages were reached: seedling (17 days), vegetative (24 days), bolting (31 days), or flowering (38 days); see Chaparro et al. (2013).

Extraction of microbial DNA from soil

In our study we used the classical definition of the rhizosphere which consist of three zones: the endorhizosphere (root tissue area), the rhizoplane (root surface with epidermis), and the ectorhizosphere (soil directly surrounding the root), and we did not distinguish between these zones (Badri and Vivanco 2009, Brimecombe et al 2007, Lynch 1990, Morgan et al 2005). Rhizosphere soils (or 'soil' thereafter) for each of the time point's six biological replicates (24 samples; 4 time points) were collected by gently removing the plants from the pots and obtaining the soil attached to the roots (Figure 3-1) and stored at -80°C for future use. It is worth noting that the removal of the rhizosphere soil was done in such a manner to prevent mechanical root shearing. However, our rhizosphere soil (as per the classical definition) consists of the rhizosphere and the roots present in that soil (Figure 3-1). Once the total RNA was extracted from the soil using the PowerSoil total RNA isolation kit (see (Chaparro et al 2013)), total DNA was subsequently extracted using the RNA PowerSoil DNA Elution Accessory Kit (Mo Bio) according to the manufacturer's instructions. The isolated DNA was quantified using a Nanodrop ND-1000 spectrophotometer (Thermo, Waltham, MA, USA). The bacterial hypervariable regions V1-V3 of the 16S rRNA gene were PCR-amplified using individually bar-coded forward primers 27F, 5'-AGAGTTTGATYMTGGCTCAG-3' and reverse primer 533R, 5'-TTACCGCGGCTGCTGGC-3'. PCR was performed using Taq DNA polymerase (Takara, Mountain View, CA, USA) as previously described (Badri et al 2013a). Briefly, DNA samples were diluted to a concentration of 5 ng μl^{-1} and one microliter was used per PCR reaction. The

reaction mix (20 µl) contained 0.4 µmoles of each gene-specific primer, 200 µmoles of dNTPs, 1 X reaction buffer, and one unit of Taq DNA polymerase (Takara). PCR included 39 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 minute in thermal cycler (GeneAmp PCR system 2700; Applied Biosystems, Grand Island, NY, USA). After PCR amplification of the 24 soil DNA samples (6 reps per time point), repetitions were pooled in groups leaving three biological replicates per time point (12 samples). Amplicon products were gel purified using Wizard SV gel (Promega, Madison, WI, USA) and PCR clean-up system followed by Agencourt AMPure XP purification kit (Beckman Coulter, Brea, CA, USA). The concentration of DNA in each sample was determined using the Nanodrop ND-1000 spectrophotometer (Thermo). Approximately equal amounts of the 12 purified amplicon products (3 replicates per time point) were pooled and subjected to unidirectional pyrosequencing in 1/8 of a pico titer-plate at the W.M. Keck Center for Comparative and Functional Genomics, Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign on a Roche/454 Genome Sequencer GS-FLX+ instrument (Roche, Branford, CT, USA). Similarly, isolated total RNA (15µg per sample) was sent to the same facility. Briefly, mRNA isolation using the Ribozero rRNA removal Meta-bacterial kit (Epicentre Biotechnologies, Madison, WI, USA) was performed; using individually barcoded random hexamer primers the isolated mRNA was converted to cDNA. cDNA library normalization was performed using the Trimmer Direct kit (Evrogen, Moscow, Russia), the samples were pooled in equimolar concentrations and pyrosequencing was performed on 1/8 of a pico titer-plate (see (Chaparro et al 2013)).

16S rRNA sequencing analysis

Sequence reads were processed using Mothur v. 1.25.1 (Schloss et al 2009) as previously described (Badri et al 2013a). Sampling effort was equalized to the depth of the smallest sample

(2769 reads) and operational taxonomic units (OTUs) were defined at ≥ 97% sequence identity, using the average neighbor algorithm. Reads were classified within Mothur using the naïve Baysian classifier (Wang et al 2007). Final taxonomic assignment was based on the consensus identification for each OTU (see Supplementary Material 3-1). Sequences were also assigned to phylotypes using the phylotype command in Mothur. A multivariate data analysis of the OTUs was performed using METAGENassist (Arndt et al 2012), followed by normalization based on interquantile range (IQR) (Hackstadt and Hess 2009) and $\log_2$-transformation. IQR normalization allows one to increase statistical power by removing sequences that do not fall within the middle 50% of observations and thus reducing the number of statistical tests one has to perform. Principal component analysis (PCA) and significant features were identified for all treatments using METAGENassist (Arndt et al 2012). The Vegan package (Oksanen et al 2012) for R was used for community dissimilarity calculations (Bray-Curtis index) and principal coordinate analysis (PCoA).

Metatranscriptomics analysis

Sequence reads were processed using the Metagenomics-RAST (MG-RAST) server (Meyer et al 2008) and Mothur (Schloss et al 2009) (see (Chaparro et al 2013)). Briefly, 454 pyrosequencing of the isolated mRNA produced 166250 sequence reads for seedling, vegetative, bolting, flowering, and bulk soil samples (see (Chaparro et al 2013)). Host-specific species sequences (*Arabidopsis thaliana*) were removed using the DNA level matching Bowtie algorithm (Langmead et al 2009) within MG-RAST. Artificially replicated sequences were removed (Gomez-Alvarez et al 2009). Sequences were also removed if their length differed by more than two standard deviations from the mean length. Additionally, sequences identified via MG-RAST as ribosomal RNAs were removed using Mothur. Once the ribosomal RNAs were

removed (6473 rRNA sequences total) from each of the samples they were normalized to 14740 high-quality sequence reads per time point (Chaparro et al 2013). Sequence reads were assigned to the Kyoto Encyclopedia of Genes and Genomes (KEGG) protein database (Kanehisa et al 2004, Kanehisa et al 2008). Further quality control consisted of selecting sequences with a minimum percent identity cutoff of 70% and an E-value cutoff of 10^{-5}.

Root exudation

Root exudate data was obtained from Chaparro et al. (2013) from *in vitro* grown *Arabidopsis* plants that were grown to the designated plant developmental time points (seedling, vegetative, bolting, and flowering). Briefly, 7-day old seedlings (see above for seedling growth) were transferred to Magenta boxes containing 10 ml of MS media supplemented with 1% sucrose. Plants were grown until they were 7, 14, 21, or 28 days and were subsequently transferred to new Magenta boxes containing 10 ml of sterilized Millipore water. Root exudation was collected after 3 days of constant secretion (10, 17, 24, or 31 days) where solutions were filtered using nylon filters (0.45 μm; Millipore, MA, USA). Root exudates were freeze-dried, dissolved in 5 ml of 80% methanol, dried under N gas and sent to the Genome Center Core Services at the University of California, Davis for GC-MS analysis on an Agilent 6890 gas chromatograph (see (Chaparro et al 2013)).

Statistical analyses

All statistical analyses were done using SAS (ver. 9.3; SAS Institute). The PROC MIXED function was used to implement a two-way ANOVA analysis with a Tukey *post-hoc* adjustment to determine pairwise differences between the microbial communities at each plant developmental time point. To ensure that the data followed the assumptions of normality, sequences were $\log_2$ transformed. To identify if developmentally dependent root exudation

influenced the soil microbial community, Pearson correlation analysis was performed with the compounds released as root exudates (identified in Chaparro et al. (2013)) and the corresponding phylogenetic data. To ensure that the data abided by the assumptions of normality, the sequencing data was log_2 transformed and the metabolomics data was log transformed. To determine how the functional microbiome changes with plant growth pairwise comparisons between log_2 transformed and subsequently standardized early (seedling and vegetative) and late (bolting and flowering) metatranscriptomics data (see (Chaparro et al 2013)) were done using a two-sample t-test. Additionally, Pearson correlation analysis was performed with the log transformed root exudate compounds and the transformed and standardized functional genes that significantly changed (81 transcripts; 413 reads) with plant growth to determine if developmentally-dependent root exudation mediates the functions carried out by the rhizosphere microbiome. The interactive pathways explorer 2 (iPATH2) (Yamada et al 2011) was used to map the functional transcripts that were differentially expressed between early and late plant development (Figure 3-2).

Results

Plant development influences the soil microbial community

We analyzed the influence of plant development on the rhizosphere microbial community by 454 pyrosequencing and obtained 55921 high quality 16S rRNA sequence reads. After equalizing sampling effort, 33228 reads were retained for analysis. These reads clustered into 7452 OTUs at 3% distance sequence dissimilarity (Supplementary Material 3-1). We visualized the Bray-Curtis distances between samples using PCoA to determine how dissimilar the soil microbial communities were at each plant developmental time point. The rhizospheric microbial community of seedling was statistically ($p<0.05$) distinct from that of all other time points

(vegetative, bolting, and flowering; Figure 3-3A). On the other hand, the rhizosphere microbial community established at vegetative, bolting, and flowering were similar to each other (Figure 3-3A). We used PCA to identify the factors that influence the soil microbial community at each time point (Figure 3-3B). The first two principle components explain 90.6% of the variability in the data. Principle component 1 explains 86.4% of the data while principle component 2 explains 4.2% of the data. These data clearly show that the soil microbial community found at seedling was distinct from bolting and flowering. However, we observed that the soil microbial community corresponding to the vegetative stage overlapped with the microbial communities established at seedling and with those at bolting and flowering. This suggests that there is a transition state of the microbial community between seedling and bolting/flowering (Figure 3-3B).

We also determined the total OTU richness, evenness, and diversity of the sequencing data for each time point (seedling, vegetative, bolting, and flowering; Table 3-1). Although there were no statistically significant differences between the time points with respect to overall community characteristics, we observed that vegetative had the largest community richness, diversity, and evenness when compared to the other developmental stages, whereas seedling had the lowest diversity and evenness (Table 3-1). This suggests that while the structure of the overall soil microbial community does not change, specific microbes may be changing through development.

After aligning the OTUs with the SILVA database we classified the soil microbial community into phylotypes consisting of 21 phyla and unclassified (Figure 3-4). ANOVA analysis with a Tukey *post-hoc* test identified significant differences between the developmental time points and four phyla (Acidobacteria, Actinobacteria, Bacteroidetes, and Cyanobacteria).

All other phyla did not significantly change with development. For example, the abundance of Acidobacteria significantly increased ($p<0.05$) from seedling to vegetative where it reached its peak and then significantly decreased at bolting and flowering to levels similar to those at seedling (Figure 3-5A). Similarly, the abundance of Actinobacteria was at its highest at the early time points (seedling and vegetative) but then significantly decreased at the later time points bolting and flowering; Figure 3-5B). On the other hand, the abundance of Bacteroidetes increased with plant growth reaching its highest abundance at flowering (Figure 3-5C). Likewise, we observed an increase in the abundance of Cyanobacteria from early time points (seedling and vegetative) to the later time points (bolting and flowering; Figure 3-5D). These data indicates that while the soil microbial community as a whole did not dramatically change, specific-soil microbial phyla were influenced by plant development.

We further analyzed the soil microbial community at the genus level to determine which genera where influencing the changes observed through development within the phyla Acidobacteria, Actinobacteria, Bacteroidetes, and Cyanobacteria. We observed that four genera belonging to Acidobacteria increased in abundance from seedling to vegetative when they reached their peak abundance and then decreased at bolting reaching levels similar to that of seedling during flowering (Figure 3-6A-D). Eight genera belonging to Actinobacteria significantly changed according to plant development (Figure 3-7A-H). For example, *Streptomyces* increased in abundance from seedling to vegetative when the highest abundance was reached and then significantly decreased at bolting stage to levels below those observed at seedling and remaining at these levels throughout flowering (Figure 3-7E). On the other hand, six genera belonging to the Bacteroidetes phylum significantly increased with plant aging (Figure 3-8A-F). In general, bacteria classified as Cyanobacteria were significantly more

abundant at bolting than at seedling and vegetative and then decreased in abundance at flowering (Figure 3-9A-C).

Plant root exudation correlates with rhizosphere microbes through plant development

To determine how developmentally dependent root exudate changes may influence the soil microbiome, Pearson correlation analysis was performed between the compounds released as root exudates (data from (Chaparro et al 2013)) and the rhizosphere bacteria that significantly changed through the life span of the plant (Supplementary Material 3-2). We observed that Cyanobacteria significantly ($p<0.05$) correlated with the most root exudate compounds (373 correlations; Table 3-2) while Bacteroidetes correlated with the least (24 correlations). To determine whether a specific class of root exudate compounds were involved in soil microbial community dynamics root exudates from Chaparro et al. (2013) were categorized as amino acids, phenolics, sugars, or sugar alcohols. Phenolic compounds significantly correlated with the soil microbial community of Acidobacteria, Actinobacteria, Bacteroidetes, and Cyanobacteria (412 significant correlations; Table 3-3) followed by amino acids (151), sugars (137) and sugar alcohols (77).

Plant development influences the functional microbiome

To determine the functional genes that were differentially expressed at early (seedling and vegetative) compared to late plant development (bolting and flowering) two-sample t-tests on $\log_2$ transformed and standardized data were performed on the metatranscriptomics data from Chaparro et al. (2013). A total of 81 unique transcripts out of 1240 assigned to the KEGG database were significantly ($p<0.05$) differentially expressed with plant age (Table 3-4). iPATH2 (Yamada et al 2011) was used to map at what time point (early or late) these 81 functional transcripts were more abundant (Figure 3-2). Of those transcripts, 32 were more abundant during

seedling and vegetative stages, while 49 were more numerous during plant bolting and flowering. The majority of the transcripts that were differently expressed between early and late development are involved in metabolism and genetic information processing. Two transcripts are involved in toluene degradation (succinate dehydrogenase iron-sulfur protein) and a transcript involved in nitrotoluene degradation (hydrogenase large subunit) were significantly ($p<0.05$) more abundant early in plant development. On the other hand, a transcript involved in bacterial chemotaxis (two-component system, chemotaxis family, response regulator) and one involved in streptomycin biosynthesis (dTDP-glucose 4,6-dehydratase) were significantly more abundant in late developmental stages. Among five transcripts involved in nitrogen metabolism, three and two were significantly expressed at early or late plant developmental stages, respectively. For example, the transcript formamidase, carbonic anhydrase and nitric-oxide reductase NorQ protein were highly expressed during plant seedling and vegetative stages. Conversely, the transcripts nitrite reductase (NAD(P)H) large subunit and periplasmic nitrate reductase NapA were more abundant during plant bolting and flowering.

Beneficial microbes are more active during late plant development

The 81 transcripts that were significantly expressed through development were aligned to the BLAT database (Kent 2002) within MG-RAST (Meyer et al 2008) to determine which microbes were carrying out the expressed functions (Table 3-4). These transcripts were attributed to bacteria such as *Bradyrhizobia*, *Streptomyces*, *Azoarcus* or *Pseudomonas syringae*. Accordingly, we categorized these microbes into seven groups depending on their potential activity towards the plant (i.e. symbiotic N fixation, free N fixation, antagonists, PGPRs, pathogens, xenobiotic/metal detoxifiers or unclassified; Table 3-5). We observed that many of these transcripts aligned to symbiotic N fixers, free N fixers or pathogens (14, 17, or 12

transcripts respectively; Table 3-5), such as *Nitrobacter*, *Rhodospirillum*, *Nitrosospira*, *Mesorhizobium*, Cyanobacteria, *AzorhizobiumI*, or *Dickeya dadantii*. Interestingly, most of the transcripts that aligned to pathogens were significantly expressed early during development (8 genes) while transcripts that aligned to microbes involved in plant growth promotion or N fixation were more abundant late (23 genes; Table 3-5). For example, the transcripts involved in bacterial chemotaxis (two-component system, chemotaxis family, response regulator) were significantly more abundant at late time points and aligned to the nitrogen fixing PGPR *Methylobacterium extorquens* (Ivanova et al 2001, Lidstrom and Chistoserdova 2002, Sy et al 2001). On the other hand, the transcript involved in amoebiasis (Ras-related protein Rab-7A) was significantly more abundant at early time points, and it aligned with the pathogen *Moniliophthora perniciosa*. Additionally, microbes that are antagonistic and produce fungicides or bactericides such as *Streptosporangium, Streptomyces avermitilis*, and *Sorangium cellulosum* were transcriptionally active late (Table 3-4 and Table 3-5) (Burg et al 1979, Pradella et al 2002).

Plant root exudation correlates with the functional microbiome through plant development

To determine whether root exudation potentially mediates the functions carried out by the rhizosphere microbiome, we correlated the 81 unique and significant transcripts with the compounds released as root exudates (from (Chaparro et al 2013); see Supplementary Material 3-3). Similar to the 16S rRNA analysis, phenolic compounds appear to mediate the expression of the transcripts (449 significant correlations; Table 3-6). Additionally, dTDP-glucose 4,6-dehydratase positively correlated with the sugar alcohol *myo*-Inositol. The transcript two-component system, chemotaxis family, response regulator positively correlated with glycine and xylitol.

Discussion

Plant developmental changes affect the rhizosphere microbial community

Analysis of the overall bacterial rhizosphere community through plant development revealed that the community did not significantly change with respect to richness, diversity, and evenness (Table 3-1) but Bray-Curtis community dissimilarity analysis revealed that the microbial community at seedling was significantly different from the other developmental stages (vegetative, bolting, and flowering; Figure 3-3A). These results are in agreement with previous reports as the *Arabidopsis* rhizosphere microbial communities after the bolting stage were not distinct (Lundberg et al 2012, Micallef et al 2009a). A more detailed look at the assembled rhizosphere microbial communities through plant development revealed that a core microbiome established and this constituted bacteria comprising Actinobacteria, Bacteroidetes, and Proteobacteria, as was previously observed (Bulgarelli et al 2012, Lundberg et al 2012). In addition, the present study demonstrated that Chloroflexi, Firmicutes, Gemmatimonadetes, Nitrospirae, Planctomycetes, and Verrucomicrobia were also consistently present throughout plant development (Figure 3-4). Additionally, Acidobacteria (Figure 3-5A) and Cyanobacteria (Figure 3-5D) also comprised the core rhizosphere microbiome but these phyla along with Bacteroidetes (Figure 3-5C) and Actinobacteria (Figure 3-5B) significantly changed with plant development suggesting that the plant can select a subset of microbes at different stages of growth. Acidobacteria is one of the most abundant bacterial phyla found in terrestrial ecosystems (Barns et al 1999) and they have an important role in the carbon cycle due to their ability to degrade complex plant derived polysaccharides such as cellulose and lignin (Ward et al 2009). Unfortunately, the specific role they play in the soil ecosystem and their role in the rhizosphere (apart from being metabolically active (Lee et al 2008)) is relatively unknown. Actinobacteria,

on the other hand, has recently been associated with disease suppressive soils (Mendes et al 2011). *Streptomyces* species, which were significantly more abundant in the vegetative stage of *Arabidopsis* (Figure 3-7E), are able to increase root nodulation efficiency and promote plant growth of the legume *Pisum sativum* (Tokala et al 2002) while simultaneously triggering plant defense in *Arabidopsis* or apple trees (Cohen et al 2005, Conn et al 2008, Lin et al 2012). On the other hand, Bacteroidetes' role in the rhizosphere has not yet been elucidated but it has been reported that they are important contributors to nutrient turnover in the soil (Yousuf et al 2012). Bacterial species belonging to Bacteroidetes contain genes involved in denitrification indicating a possible involvement in N cycling (Van Spanning et al 2005). Cyanobacteria have been shown to colonize plant roots (Gantar et al 1991, Lundberg et al 2012), promote plant growth (Prasanna et al 2009), and are an important plant source for inorganic N due to their ability to fix N (Franche et al 2009).

Plant development influences the functional capacity of the rhizomicrobiome

Metatranscriptomics analyses of mRNA have only recently been used as a means to study microbial communities at a functional level in distinct environments such as the human gut (Gosalbes et al 2011), the mouse gut (Xiong et al 2012), and oceans (Baker et al 2013, Poretsky et al 2009). In soils, functional microarrays or GeoChip that target specific microbial functions have been successfully used to determine the metabolic potential of the microbial communities (Bai et al 2013, He et al 2007, He et al 2011, Zhang et al 2013). A drawback of the GeoChip is the fact that novel functions and transcripts cannot be identified, and one is biased to the probes present in the chip (Dugat-Bony et al 2012). Recently, metatranscriptomics in soil has been attempted on the rRNA to identify the active members of the microbial community in the rhizosphere of various crop species (Turner et al 2013); Urich et al. (2008) described the

isolation and sequencing of mRNA from a sandy lawn as a means of describing the microbial transcriptome in the soil. The rhizosphere microbiome plays an essential role in plant health and productivity and it is often referred as the plant's second genome (Berendsen et al 2012, Chaparro et al 2012). Accordingly, our metatranscriptomics data permitted a glimpse at the genes that the microbiome was expressing as a whole at each stage of plant development. More transcripts were significantly expressed at late plant developmental time points (Table 3-4) and this may be indicative of the soil microbial community selecting specific functions throughout plant development. For example, dTDP-glucose 4,6-dehydratase involved in streptomycin biosynthesis is more abundant during plant bolting and flowering. Streptomycin is an antibiotic mainly produced by *Streptomyces* and it is antagonistic against gram-positive and -negative bacteria and has been shown to enhance plant defenses and trigger systemic resistance (Conn et al 2008, Lin et al 2012, Schatz et al 2005). Additionally, we observe that dTDP-glucose-4,6-dehydratase positively correlates with the root exudate compound *myo*-Inositol which increases streptomycin biosynthesis (Heding 1964, Majumdar and Kutzner 1962). Interestingly, we did not find transcripts that are attributed to plant pathogens at late stages of plant development which is consistent with the increased production of streptomycin and in accordance with ARR. Furthermore, we determined that *Sorangium cellulosum* which produce bacteriocides and fungicides (Pradella et al 2002) was also present in late stages of plant development.

The rhizosphere microbiome can also supply the plant with essential nutrients such as nitrogen. Nitrogen is essential for plant growth (Burns 1996, Rossato et al 2001) and it is usually deficient in soils (Novoa and Loomis 1981). Thus, in natural environments (as compared to agricultural systems) the plant depends on N fixing and nitrifying bacteria for their N supply. It should be noted that in our study the plants grew in natural soil without external fertilization. In

accordance, we observed that throughout plant development functional genes involved in nitrogen metabolism were transcriptionally active (Table 3-5). Additionally, functions carried out by bacteria involved in N assimilation were prevalent (Table 3-4 and Table 3-5). This may be indicative of modulation of the microbiome to express specific functions throughout plant growth, that is, differentially express transcripts from beneficial bacteria. For example, transcripts more abundant in early time points align to bacteria involved in providing the plant with N such as *Nitrobacter*, *Rhodospirillum*, *Nitrosospira*, *Mesorhizobium*, or *Azorhizobium.* Similarly, during bolting and flowering, functional genes expressed in the rhizosphere align to PGPRs such as *Bacillus licheniformis* (Gutiérrez-Mañero et al 2001) or *Burkholderia ambifaria* (Chiarini et al 2006), free N-fixers such as *Cyanothece sp.* (Junier et al 2009), as well as symbiotic N-fixing bacteria (*Bradyrhizobium*) (Stacey et al 1995), or *Herbaspirillum* which is involved in endophytic N fixation (Elbeltagy et al 2001, Franche et al 2009). It should be noted that in addition to fixing N through legume symbiosis, *Bradyrhizobium* promotes plant growth of non-leguminous plants (Antoun et al 1998). The functional microbiome also expressed genes that aligned to the free N-fixing Cyanobacteria (Table 3-4). These wide changes in the soil bacterial community hint at the soil microbiome shifting and changing according to the needs of the plant and that the rhizosphere functional microbiome can express specific functions at precise stages of plant growth.

Interestingly, we determined that the expression of a functional gene is altered through plant development even when the bacterium performing the function does not change in abundance. For example, *Bradyrhizobia* classified as Proteobacteria does not significantly change with plant growth (Figure 3-4) yet transcripts expressed by *Bradyrhizobia* such as imidazolonepropionase, phenol 2-monooxygenase, or glutaminase are more abundant in later

stages of plant development (Table 3-4). Alternatively, while Actinobacteria, and more specifically *Streptomyces*, were more abundant during the vegetative stage (Figure 3-6B, Figure 3-7E) we observed that genes involved in synthesis of the antibiotic streptomycin were higher during plant bolting and flowering (Table 3-4). This indicates that root exudates are presumably able to modulate the expression of specific functional genes without altering the bacterial taxonomic composition of the rhizosphere.

Root exudates act as potential stimulants for rhizomicrobiome functions

While microbial colonization of the rhizosphere has shown an array of benefits to the plant, the exact mechanism by which the plant is able to attract these microbes varies. One such mechanism used is bacterial chemotaxis towards root exudate compounds. Interestingly, we see that the two-component system, chemotaxis family, response regulator involved in bacterial chemotaxis is significantly expressed late in plant development and its expression positively correlates with the root exudates glycine and xylitol (Supplementary Material 3-3). Interestingly, PGPRs and endophytic bacteria show chemotaxis towards glycine (de Weert et al 2002, Gaworzewska and Carlile 1982, Gupta Sood 2003), while the non-symbiotic nitrogen-fixing bacteria *Azotobacter vinelandii* shows chemotactic behavior towards xylitol (Haneline et al 1991). This provides an additional mechanism for the plants ability to manipulate and orchestrate the rhizosphere microbiome. While the results presented are only correlative they do highlight the importance of root exudation in rhizosphere plant-microbiome interactions and provide strong evidence to warrant further investigation that would conclusively determine how specific components of the root exudates in the absence of the plant could influence the rhizosphere metatranscriptome. Incidentally, a recent study depicting the microbial communities established from the addition of root exudate blends to the soil demonstrated that different blends produced

changes in the microbial community. For example, soil supplemented with ethyl acetate- or water-extracted root exudates generated microbial communities that presumably had the ability to metabolize the pesticide atrazine (Badri et al 2013a). Similar experiments that add specific compounds (sugars or oxalic acids) to the soil have also shown dramatic shifts in the composition of the microbial community (Eilers et al 2010, Shi et al 2011).

The general root exudation patterns reported in Chaparro et al. (2013) do not seem to correlate with general microbiome characteristics as initially hypothesized (i.e. increased root exudation of sugars at early plant development would result in enhanced richness of the rhizosphere microbial community at these early time points). This could be due to the fact that root exudates were collected *in vitro* while rhizosphere microbial communities were analyzed *in vivo*; in other words, the root secretion patterns could be different under soil conditions. Accordingly, it has been reported that under *in vitro* conditions plants grown alone or co-cultured with a microorganism show different patterns of proteins present in the root exudates (De-la-Pena et al 2008). However, Chaparro et al. (2013) showed evidence that root exudation through development is genetically programmed since there were strong correlations between *in vitro* root exudation patterns and the ability of the soil microbiome to utilize these compounds *in vivo*. During plant vegetative stage (transition state; Figure 3-3B) we identified the largest OTU richness, diversity, and evenness in the rhizosphere (Table 3-1). This rhizospheric microbial transition state corresponds with a root exudation transition state where the highest diversity of phytochemicals (sugars, sugar alcohols, phenolics, and amino acids) was observed (Chaparro et al 2013). Therefore, it appears that sugars present in the root exudates do not necessarily function as general substrates for soil microbes as is widely reported in the literature (Behera and Wagner 1974, Eilers et al 2010, Fierer et al 2007, Jaeger et al 1999), whereby compounds such as

phenolics or amino acids more readily influence soil rhizosphere microbial communities as well as modulate their transcription (Table 3-3 and Table 3-6). Recently, Badri et al. (2013a) reported that fractions of root exudates containing phenolic and phenolic-related compounds when applied to the soil (in the absence of the plant) significantly modulated the soil microbiome. In addition, analysis of the rhizosphere microbial community of an *Arabidopsis* mutant that secretes more phenolic compounds than sugars showed that it cultured more beneficial microbes such as PGPRs and those involved in N-fixation such as *Bradyrhizobia* and Cyanobacteria when compared to wild type (Badri et al 2009). Interestingly, the reduction of phenolic exudates by inhibition of *phenylalanine ammonia-lyase* gene expression in transgenic rice influenced the rhizospheric microbial community, with eight phyla decreasing in abundance in transgenic rice when compared to wild type plants (Fang et al 2013). Further studies pinpointing which phenolic compounds are involved in coordinating these microbial interactions are needed.

The data presented here implies and alludes to the fact that the plant through root exudation may be selecting microbes for different functions; however, we would be remiss not to point out other potential mechanisms that may be playing a role. For example, the changes observed in the rhizosphere microbial community could be due to microbial community succession with respect to microbial competition. Unfortunately, presently there is little research regarding microbial succession and stability in soils (Fierer et al 2010). Recently, it was shown that microbial succession is similar to that of previously described plant and animal succession with respect to species-time relationships (Shade et al 2013) but what this could mean in the rhizosphere is unclear. There is still much to be explored with respect to plant-microbiome interactions to better understand and decipher the complex patterns and associations that arise in this unique ecological niche.

Much like the plant can influence the rhizosphere microbiome the rhizosphere microbiome can also influence the plant. For example, fungal communities have been shown to influence root exudation rates which can in turn influence the rhizosphere microbiome (Meier et al 2013). Similarly, Lau and Lennon (2011) demonstrated that microbial community structure affects natural plant trait selections. Additionally, it has been established that distinct microbial communities influence plants' ability to tolerate abiotic stress such as drought (Zolla et al 2013) and even affect leaf metabolome and subsequent insect feeding (Badri et al 2013b). These examples highlight the multifaceted nature of the interactions in the rhizosphere microbiome.

Conclusions

The conclusions of this study could be summarized as follows: (1) the plant maintains a core rhizosphere microbiome; (2) this core microbiome is likely to express different functions at different stages of plant development; (3) the plant can enhance the expression of a subset of microbial functions at specific times that the core microbiome is not currently expressing; (4) the plant can select a subset of microbes at different stages of development presumably for specific functions that the core microbiome can't express (i.e. N fixation, antibiosis against pathogens, etc.); and (5) the plant secretes blends of compounds and specific phytochemicals in the root exudates that are differentially produced at distinct stages of plant growth to help orchestrate the activities described in 1, 2, 3, and 4. Overall, these concepts suggest that plants and the rhizomicrobiome are in constant communication through the exchange of signals. Experiments targeting some of these ideas are essential to conclusively determine the interactive functionalities that occur in the rhizosphere between crops and their microbiome.

Tables

Table 3-1. Observed species richness (Sobs), Shannon diversity and evenness of the OTU soil microbial community for each plant developmental time point.

	Richness	Shannon	
	Sobs	Evenness	Diversity
Seedling	1107.33	0.8009	5.6136
Vegetative	1262.67	0.8638	6.1675
Bolting	1114.00	0.8192	5.7484
Flowering	1040.33	0.8245	5.7278

Table 3-2. Pearson correlation analysis of the OTUs classified as Acidobacteria, Actinobacteria, Bacteroidetes, or Cyanobacteria with the compounds released as root exudates. The values indicate the number of significant ($p<0.05$) Pearson correlations for each phyla.

Phylum	Number of positive correlations	Number of negative correlations	Total number of correlations
Acidobacteria	156	75	231
Actinobacteria	103	46	149
Bacteroidetes	8	16	24
Cyanobacteria	130	243	373

Table 3-3. Pearson correlations analysis of the OTUs classified as Acidobacteria, Actinobacteria, Bacteroidetes or Cyanobacteria with the group of compounds released as root exudates. The values indicate the number of significant ($p<0.05$) Pearson correlations.

Phylum	Amino Acids		Phenolics		Sugars		Sugar Alcohols	
	Positively	Negatively	Positively	Negatively	Positively	Negatively	Positively	Negatively
Acidobacteria	35	12	60	38	24	18	37	7
Actinobacteria	3	21	58	25	24	0	18	0
Bacteroidetes	2	0	4	10	1	5	1	1
Cyanobacteria	67	11	60	157	0	65	3	10
Total Correlations	151		412		137		77	

Table 3-4. Statistically significant (t-test $p<0.05$) transcripts (81 total) classified under hierarchical KEGG orthology expressed by the rhizomicrobiome at early (seedling (seed) and vegetative (veg)) vs. late (bolting (bolt) and flowering)

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Cellular Processes	Cell Growth and Death	04112 Cell cycle - Caulobacter [PATH:ko04112]	modification methylase [EC:2.1.1.72]	K13581	2	2	0	0	Early	Anaeromyxobacter dehalogenans (strain 2CP-C)
Cellular Processes	Cell Motility	02030 Bacterial chemotaxis [PATH:ko02030]	two-component system, chemotaxis family, response regulator	K03413	0	0	3	2	Late	Methylobacterium extorquens DSM 13060
Environmental Information Processing	Membrane Transport	02010 ABC transporters [PATH:ko02010]	D-methionine transport system ATP-binding protein	K02071	0	0	1	1	Late	Anabaena variabilis ATCC 29413
Environmental Information Processing	Membrane Transport	02010 ABC transporters [PATH:ko02010]	D-xylose transport system permease protein	K10544	0	0	2	2	Late	Rhizobium loti (strain MAFF303099) (Mesorhizobium loti)

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Environmental Information Processing	Membrane Transport	02010 ABC transporters [PATH:ko02010]	iron(III) transport system ATP-binding protein [EC:3.6.3.30]	K02010	0	0	2	4	Late	Azorhizobium caulinodans (strain ATCC 43989 / DSM 5975 / ORS 571)
Environmental Information Processing	Membrane Transport	02010 ABC transporters [PATH:ko02010]	iron(III) transport system substrate-binding protein	K02012	2	2	0	0	Early	Nitrosospira multiformis (strain ATCC 25196 / NCIMB 11849)
Environmental Information Processing	Membrane Transport	03070 Bacterial secretion system [PATH:ko03070]	general secretion pathway protein E	K02454	0	0	4	3	Late	Saccharophagus degradans (strain 2-40 / ATCC 43961 / DSM 17024)
Environmental Information Processing	Signal Transduction	02020 Two-component system [PATH:ko02020]	beta-lactamase [EC:3.5.2.6]	K01467	2	2	0	0	Early	Azoarcus sp. (strain BH72)
Environmental Information Processing	Signal Transduction	04070 Phosphatidylinositol signaling system	myo-inositol-1(or 4)-monophosphatase [EC:3.1.3.25]	K01092	2	2	0	0	Early	Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155)

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
		[PATH:ko04070]								
Genetic Information Processing	Folding, Sorting and Degradation	03018 RNA degradation [PATH:ko03018]	ribonuclease R [EC:3.1.-.-]	K12573	1	1	6	7	Late	Chelativorans sp. BNC1
Genetic Information Processing	Replication and Repair	03410 Base excision repair [PATH:ko03410]	endonuclease III [EC:4.2.99.18]	K10773	1	2	0	0	Early	Synechococcus sp. (strain JA-3-3Ab) (Cyanobacteria bacterium Yellowstone A-Prime)
Genetic Information Processing	Replication and Repair	03440 Homologous recombination [PATH:ko03440]	exodeoxyribonuclease V alpha subunit [EC:3.1.11.5]	K03581	1	2	0	0	Early	Gluconacetobacter diazotrophicus (strain ATCC 49037 / DSM 5601 / PAl5)
Genetic Information Processing	Replication and Repair	03440 Homologous recombination [PATH:ko03440]	recombination protein RecA	K03553	5	5	3	3	Early	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NCIMB 10081)

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Genetic Information Processing	Translation	03010 Ribosome [PATH:ko03010]	large subunit ribosomal protein L10e	K02866	0	0	1	1	Late	Candida albicans SC5314
Genetic Information Processing	Translation	03010 Ribosome [PATH:ko03010]	large subunit ribosomal protein L22	K02890	0	0	2	1	Late	Gemmatimonas aurantiaca (strain T-27 / DSM 14586 / JCM 11422 / NBRC 100505)
Genetic Information Processing	Translation	03010 Ribosome [PATH:ko03010]	large subunit ribosomal protein L24	K02895	0	0	4	3	Late	Polaromonas naphthalenivorans (strain CJ2)
Genetic Information Processing	Translation	03010 Ribosome [PATH:ko03010]	large subunit ribosomal protein L26e	K02898	0	0	1	1	Late	Schizosaccharomyces pombe 972h-
Genetic Information Processing	Translation	03010 Ribosome [PATH:ko03010]	large subunit ribosomal protein L30e	K02908	1	1	0	0	Early	Trypanosoma brucei TREU927
Genetic Information Processing	Translation	03010 Ribosome [PATH:ko03010]	small subunit ribosomal protein S15Ae	K02957	1	1	0	0	Early	Paramecium tetraurelia d4-2

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Human Diseases	Infectious Diseases	05111 Vibrio cholerae pathogenic cycle [PATH:ko05111]	RNA polymerase sigma factor for flagellar operon FliA	K02405	0	0	1	2	Late	Herbaspirillum seropedicae SmR1
Human Diseases	Infectious Diseases	05146 Amoebiasis [PATH:ko05146]	Ras-related protein Rab-7A	K07897	1	1	0	0	Early	Moniliophthora perniciosa FA553
Metabolism	Amino Acid Metabolism	00250 Alanine, aspartate and glutamate metabolism [PATH:ko00250]	glutamate synthase (NADPH/NADH) large chain [EC:1.4.1.13	K00265	6	6	11	10	Late	Streptosporangium roseum (strain ATCC 12428 / DSM 43021 / JCM 3005 / NI 9100)
Metabolism	Amino Acid Metabolism	00260 Glycine, serine and threonine metabolism [PATH:ko00260]	5-aminolevulinate synthase [EC:2.3.1.37]	K00643	0	0	1	1	Late	Nitrobacter winogradskyi Nb-255

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Metabolism	Amino Acid Metabolism	00260 Glycine, serine and threonine metabolism [PATH:ko00260]	choline dehydrogenase [EC:1.1.99.1]	K00108	0	0	2	3	Late	Burkholderia ambifaria (strain ATCC BAA-244 / AMMD) (Burkholderia cepacia (strain AMMD))
Metabolism	Amino Acid Metabolism	00260 Glycine, serine and threonine metabolism [PATH:ko00260]	threonine 3-dehydrogenase [EC:1.1.1.103]	K00060	0	0	1	1	Late	Arthrobacter aurescens TC1
Metabolism	Amino Acid Metabolism	00280 Valine, leucine and isoleucine degradation [PATH:ko00280]	dihydrolipoamide dehydrogenase [EC:1.8.1.4]	K00382	12	5	0	0	Early	Solibacter usitatus (strain Ellin6076)
Metabolism	Amino Acid Metabolism	00290 Valine, leucine and isoleucine biosynthesis	leucine dehydrogenase [EC:1.4.1.9]	K00263	0	0	1	1	Late	Rhodocista centenaria SW

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
		[PATH:ko00290]								
Metabolism	Amino Acid Metabolism	00300 Lysine biosynthesis [PATH:ko00300]	aspartate-semialdehyde dehydrogenase [EC:1.2.1.11]	K00133	0	0	2	1	Late	Thiomonas intermedia (strain K12) (Thiobacillus intermedius)
Metabolism	Amino Acid Metabolism	00330 Arginine and proline metabolism [PATH:ko00330]	acetylglutamate/acetylaminoadipate kinase [EC:2.7.2.8	K00930	1	2	0	0	Early	Nitrobacter winogradskyi (strain Nb-255 / ATCC 25391)
Metabolism	Amino Acid Metabolism	00340 Histidine metabolism [PATH:ko00340]	cyclase [EC:4.1.3.-]	K02500	0	0	1	1	Late	Magnetospirillum magneticum AMB-1
Metabolism	Amino Acid Metabolism	00340 Histidine metabolism [PATH:ko00340]	imidazolonepropionase [EC:3.5.2.7]	K01468	0	0	1	1	Late	Bradyrhizobium sp. ORS278
Metabolism	Amino Acid Metabolism	00360 Phenylalanine metabolism	ferredoxin subunit of phenylpropionate dioxygenase	K05710	0	0	1	1	Late	Pseudomonas syringae pv. phaseolicola 1448A

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
		[PATH:ko00360]								
Metabolism	Biosynthesis of Other Secondary Metabolites	00521 Streptomycin biosynthesis [PATH:ko00521]	dTDP-glucose 4,6-dehydratase [EC:4.2.1.46]	K01710	0	0	7	8	Late	Cyanothece sp. (strain PCC 7425 / ATCC 29141)
Metabolism	Carbohydrate Metabolism	00040 Pentose and glucuronate interconversions [PATH:ko00040]	xylulokinase [EC:2.7.1.17]	K00854	0	0	1	1	Late	Acidobacterium capsulatum ATCC 51196
Metabolism	Carbohydrate Metabolism	00500 Starch and sucrose metabolism [PATH:ko00500]	maltose alpha-D-glucosyltransferase [EC:5.4.99.16]	K05343	0	0	2	1	Late	Acidobacterium capsulatum (strain ATCC 51196 / DSM 11244 / JCM 7670)
Metabolism	Carbohydrate Metabolism	00500 Starch and sucrose metabolism [PATH:ko00500]	sucrose phosphorylase [EC:2.4.1.7]	K00690	0	0	1	1	Late	Polaromonas naphthalenivorans CJ2

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Metabolism	Carbohydrate Metabolism	00520 Amino sugar and nucleotide sugar metabolism [PATH:ko00520]	CDP-paratose 2-epimerase [EC:5.1.3.10]	K12454	1	1	0	0	Early	Mesorhizobium ciceri bv biserrulae WSM1271
Metabolism	Carbohydrate Metabolism	00620 Pyruvate metabolism [PATH:ko00620]	L-lactate dehydrogenase (cytochrome) [EC:1.1.2.3]	K00101	0	0	2	1	Late	Laccaria bicolor (strain S238N-H82 / ATCC MYA-4686)
Metabolism	Carbohydrate Metabolism	00620 Pyruvate metabolism [PATH:ko00620]	pyruvate carboxylase [EC:6.4.1.1]	K01958	0	0	1	2	Late	Bacillus licheniformis DSM 13 Goettingen
Metabolism	Carbohydrate Metabolism	00630 Glyoxylate and dicarboxylate metabolism [PATH:ko00630]	oxalyl-CoA decarboxylase [EC:4.1.1.8]	K01577	1	1	0	0	Early	Methylobacterium extorquens PA1

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Metabolism	Energy Metabolism	00190 Oxidative phosphorylation [PATH:ko00190]	NADH dehydrogenase I subunit G [EC:1.6.5.3]	K00336	3	3	0	0	Early	Legionella pneumophila subsp. pneumophila (strain Philadelphia 1 / ATCC 33152 / DSM 7513)
Metabolism	Energy Metabolism	00190 Oxidative phosphorylation [PATH:ko00190]	NADH dehydrogenase I subunit J [EC:1.6.5.3]	K00339	4	2	0	0	Early	Roseiflexus castenholzii (strain DSM 13941 / HLO8)
Metabolism	Energy Metabolism	00195 Photosynthesis [PATH:ko00195]	apocytochrome f	K02634	0	0	2	1	Late	cyanobacterium UCYN-A
Metabolism	Energy Metabolism	00195 Photosynthesis [PATH:ko00195]	photosystem I subunit IX	K02697	0	0	1	1	Late	Trichodesmium erythraeum IMS101
Metabolism	Energy Metabolism	00195 Photosynthesis [PATH:ko00195]	photosystem I subunit XI	K02699	0	0	2	1	Late	Synechococcus sp. (strain ATCC 27264 / PCC 7002 / PR-6)

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
										(Agmenellum quadruplicatum)
Metabolism	Energy Metabolism	00195 Photosynthesis [PATH:ko00195]	photosystem II cytochrome b559 subunit beta	K02708	0	0	1	1	Late	Anabaena variabilis ATCC 29413
Metabolism	Energy Metabolism	00195 Photosynthesis [PATH:ko00195]	photosystem II PsbL protein	K02713	0	0	1	1	Late	Prochlorococcus sp. WH8102
Metabolism	Energy Metabolism	00196 Photosynthesis - antenna proteins [PATH:ko00196]	phycocyanobilin lyase alpha subunit [EC:4.-.-.-]	K02288	0	0	2	1	Late	Cyanothece sp. (strain PCC 7425 / ATCC 29141)
Metabolism	Energy Metabolism	00680 Methane metabolism [PATH:ko00680]	6-phosphofructokinase [EC:2.7.1.11]	K00850	2	3	0	0	Early	Gallionella capsiferriformans (strain ES-2) (Gallionella ferruginea capsiferriformans (strain ES-2))

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Metabolism	Energy Metabolism	00710 Carbon fixation in photosynthetic organisms [PATH:ko00710]	triosephosphate isomerase (TIM) [EC:5.3.1.1]	K01803	7	3	0	0	Early	Clostridium botulinum (strain Alaska E43 / Type E3)
Metabolism	Energy Metabolism	00720 Carbon fixation in autotrophic prokaryotes [PATH:ko00720]	2-oxoglutarate ferredoxin oxidoreductase subunit alpha	K00174	0	0	3	2	Late	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NCIMB 10081)
Metabolism	Energy Metabolism	00910 Nitrogen metabolism [PATH:ko00910]	carbonic anhydrase [EC:4.2.1.1]	K01673	4	3	0	0	Early	Rhodococcus sp. (strain RHA1)
Metabolism	Energy Metabolism	00910 Nitrogen metabolism [PATH:ko00910]	formamidase [EC:3.5.1.49]	K01455	1	1	0	0	Early	Brachybacterium faecium 6-10, DSM 4810
Metabolism	Energy Metabolism	00910 Nitrogen metabolism	nitric-oxide reductase	K04748	2	1	0	0	Early	Nitrosomonas europaea ATCC 19718

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
		[PATH:ko00910]	NorQ protein [EC:1.7.99.7]							
Metabolism	Energy Metabolism	00910 Nitrogen metabolism [PATH:ko00910]	nitrite reductase (NAD(P)H) large subunit [EC:1.7.1.4]	K00362	0	0	3	3	Late	Dickeya dadantii (strain Ech586)
Metabolism	Energy Metabolism	00910 Nitrogen metabolism [PATH:ko00910]	periplasmic nitrate reductase NapA [EC:1.7.99.4]	K02567	0	0	2	1	Late	Rhodobacter sphaeroides (strain ATCC 17025 / ATH 2.4.3)
Metabolism	Lipid Metabolism	00564 Glycerophospholipid metabolism [PATH:ko00564]	CDP-diacylglycerol--glycerol-3-phosphate	K00995	1	1	0	0	Early	Parvibaculum lavamentivorans DS-1
Metabolism	Lipid Metabolism	00564 Glycerophospholipid metabolism [PATH:ko00564]	glycerol-3-phosphate dehydrogenase (NAD(P)+) [EC:1.1.1.94]	K00057	2	3	0	0	Early	Nitrobacter winogradskyi (strain Nb-255 / ATCC 25391)
Metabolism	Lipid Metabolism	00564 Glycerophospholipid metabolism	glycerophosphoryl diester phosphodiesterase [EC:3.1.4.46]	K01126	1	2	0	0	Early	Parabacteroides distasonis (strain ATCC 8503 / DSM

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
		[PATH:ko00564]								20701 / NCTC 11152)
Metabolism	Lipid Metabolism	00600 Sphingolipid metabolism [PATH:ko00600]	arylsulfatase [EC:3.1.6.1]	K01130	1	0	6	8	Late	Thalassiosira pseudonana (Marine diatom) (Cyclotella nana)
Metabolism	Metabolism of Cofactors and Vitamins	00130 Ubiquinone and other terpenoid-quinone biosynthesis [PATH:ko00130]	4-hydroxyphenylpyruvate dioxygenase [EC:1.13.11.27]	K00457	2	2	0	0	Early	Salinispora arenicola (strain CNS-205)
Metabolism	Metabolism of Cofactors and Vitamins	00670 One carbon pool by folate [PATH:ko00670]	5-methyltetrahydrofolate--homocysteine methyltransferase	K00548	6	13	0	0	Early	Chloroflexus aggregans (strain MD-66 / DSM 9485)
Metabolism	Metabolism of Cofactors and Vitamins	00670 One carbon pool by folate	formyltetrahydrofolate deformylase [EC:3.5.1.10]	K01433	0	0	2	1	Late	Arthrobacter aurescens (strain TC1)

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
		[PATH:ko00670]								
Metabolism	Metabolism of Cofactors and Vitamins	00670 One carbon pool by folate [PATH:ko00670]	glycine hydroxymethyltransferase [EC:2.1.2.1]	K00600	5	3	0	0	Early	Azorhizobium caulinodans (strain ATCC 43989 / DSM 5975 / ORS 571)
Metabolism	Metabolism of Cofactors and Vitamins	00770 Pantothenate and CoA biosynthesis [PATH:ko00770]	holo-[acyl-carrier protein] synthase [EC:2.7.8.7]	K00997	0	0	2	1	Late	Paracoccus denitrificans (strain Pd 1222)
Metabolism	Metabolism of Cofactors and Vitamins	00860 Porphyrin and chlorophyll metabolism [PATH:ko00860]	light-independent protochlorophyllide reductase subunit N	K04038	0	0	1	1	Late	Cyanothece sp. PCC 8801
Metabolism	Metabolism of Cofactors and Vitamins	00860 Porphyrin and chlorophyll metabolism [PATH:ko00860]	magnesium chelatase subunit H [EC:6.6.1.1]	K03403	0	0	1	1	Late	Thalassiosira pseudonana CCMP1335

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Metabolism	Metabolism of Terpenoids and Polyketides	00900 Terpenoid backbone biosynthesis [PATH:ko00900]	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase /	K12506	0	0	2	1	Late	Hyphomonas neptunium (strain ATCC 15444)
Metabolism	Metabolism of Terpenoids and Polyketides	00900 Terpenoid backbone biosynthesis [PATH:ko00900]	farnesyl diphosphate synthase [EC:2.5.1.12.5.1.10]	K00795	2	1	0	0	Early	Rhodospirillum rubrum S1, ATCC 11170
Metabolism	Nucleotide Metabolism	00230 Purine metabolism [PATH:ko00230]	GTP pyrophosphokinase [EC:2.7.6.5]	K00951	4	4	0	0	Early	Acidovorax citrulli (strain AAC00-1) (Acidovorax avenae subsp. citrulli)
Metabolism	Nucleotide Metabolism	00230 Purine metabolism [PATH:ko00230]	phosphoribosylaminoimidazole-succinocarboxamide synthase	K01923	0	0	7	5	Late	Azoarcus sp. (strain BH72)
Metabolism	Nucleotide Metabolism	00240 Pyrimidine metabolism	nucleoside-triphosphate pyrophosphat	K02428	0	0	2	2	Late	Streptomyces avermitilis

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
		[PATH:ko00240]	ase [EC:3.6.1.19]							
Metabolism	Xenobiotics Biodegradation and Metabolism	00362 Benzoate degradation [PATH:ko00362]	p-hydroxybenzoate 3-monooxygenase [EC:1.14.13.2]	K00481	0	0	2	2	Late	Rhizobium meliloti (strain 1021) (Ensifer meliloti) (Sinorhizobium meliloti)
Metabolism	Xenobiotics Biodegradation and Metabolism	00623 Toluene degradation [PATH:ko00623]	phenol 2-monooxygenase [EC:1.14.13.7]	K03380	0	0	1	1	Late	Bradyrhizobium japonicum USDA 110
Metabolism	Xenobiotics Biodegradation and Metabolism	00623 Toluene degradation [PATH:ko00623]	succinate dehydrogenase iron-sulfur protein [EC:1.3.99.1]	K00240	4	2	0	0	Early	Sanguibacter keddieii (strain ATCC 51767 / DSM 10542 / NCFB 3025 / ST-74)
Metabolism	Xenobiotics Biodegradation and Metabolism	00627 Aminobenzoate degradation [PATH:ko00627]	anthraniloyl-CoA monooxygenase [EC:1.14.13.40]	K09461	0	0	5	6	Late	Sorangium cellulosum (strain So ce56) (Polyangium cellulosum

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
										(strain Source56))
Metabolism	Xenobiotics Biodegradation and Metabolism	00633 Nitrotoluene degradation [PATH:ko00633]	hydrogenase large subunit [EC:1.12.99.6]	K06281	1	1	0	0	Early	Thermobispora bispora R51, DSM 43833
Metabolism	Xenobiotics Biodegradation and Metabolism	00930 Caprolactam degradation [PATH:ko00930]	acyl-CoA dehydrogenase [EC:1.3.99.-]	K06446	0	0	1	2	Late	Cupriavidus metallidurans CH34
Organismal Systems	Endocrine System	03320 PPAR signaling pathway [PATH:ko03320]	ubiquitin C	K08770	0	0	2	1	Late	Volvox carteri (Green alga)
Organismal Systems	Excretory System	04964 Proximal tubule bicarbonate reclamation [PATH:ko04964]	glutaminase [EC:3.5.1.2]	K01425	0	0	2	2	Late	Bradyrhizobium japonicum

Level 1	Level 2	Level 3	Function	KEGG ID	Seed	Veg	Bolt	Flowering	Significantly more abundant	BLAT taxonomy sequence alignment
Organismal Systems	Excretory System	04964 Proximal tubule bicarbonate reclamation [PATH:ko04964]	phosphoenolpyruvate carboxykinase (GTP) [EC:4.1.1.32]	K01596	6	3	0	0	Early	Mycobacterium bovis BCG str. Mexico

Table 3-5. Taxonomic assignments of the differentially expressed (81) transcripts were categorized based on their activity and whether the corresponding transcript were significantly (t-test $p<0.05$) expressed early or late in plant development.

	Early	**Late**	**Total**
Symbiotic N fixers	5	9	14
Free N fixers	6	11	17
Antagonistic	5	6	11
Plant growth promoting	0	3	3
Pathogen	8	4	12
Xenobiotic/ metal detoxification	4	5	9
Unclassified function	4	11	15

Table 3-6. Pearson correlation analysis of the significantly expressed transcripts at early or late plant development correlated with the group of compounds released as root exudate. The values indicate the number of significant ($p<0.05$) correlations.

	Early			Late			Overall
	Positive	**Negative**	**Total**	**Positive**	**Negative**	**Total**	**Total**
Amino Acids	7	160	167	198	1	199	366
Phenolics	31	139	170	222	57	279	449
Sugars	51	0	51	0	82	82	133
Sugar Alcohols	0	47	47	85	6	91	138

Figures

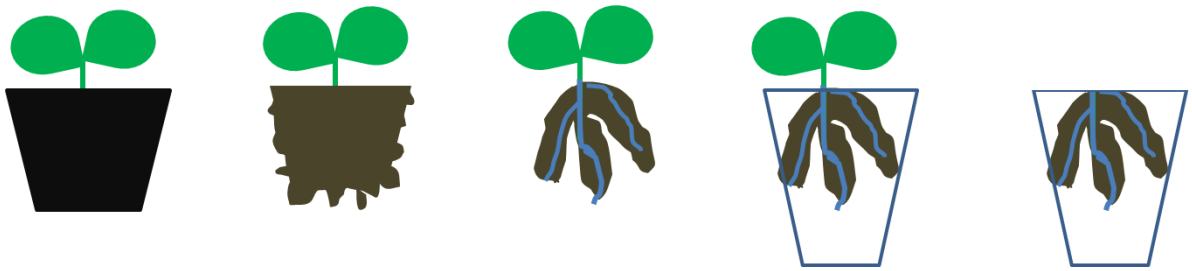

Figure 3-1. Schematic representation of Arabidopsis rhizosphere soil collection.

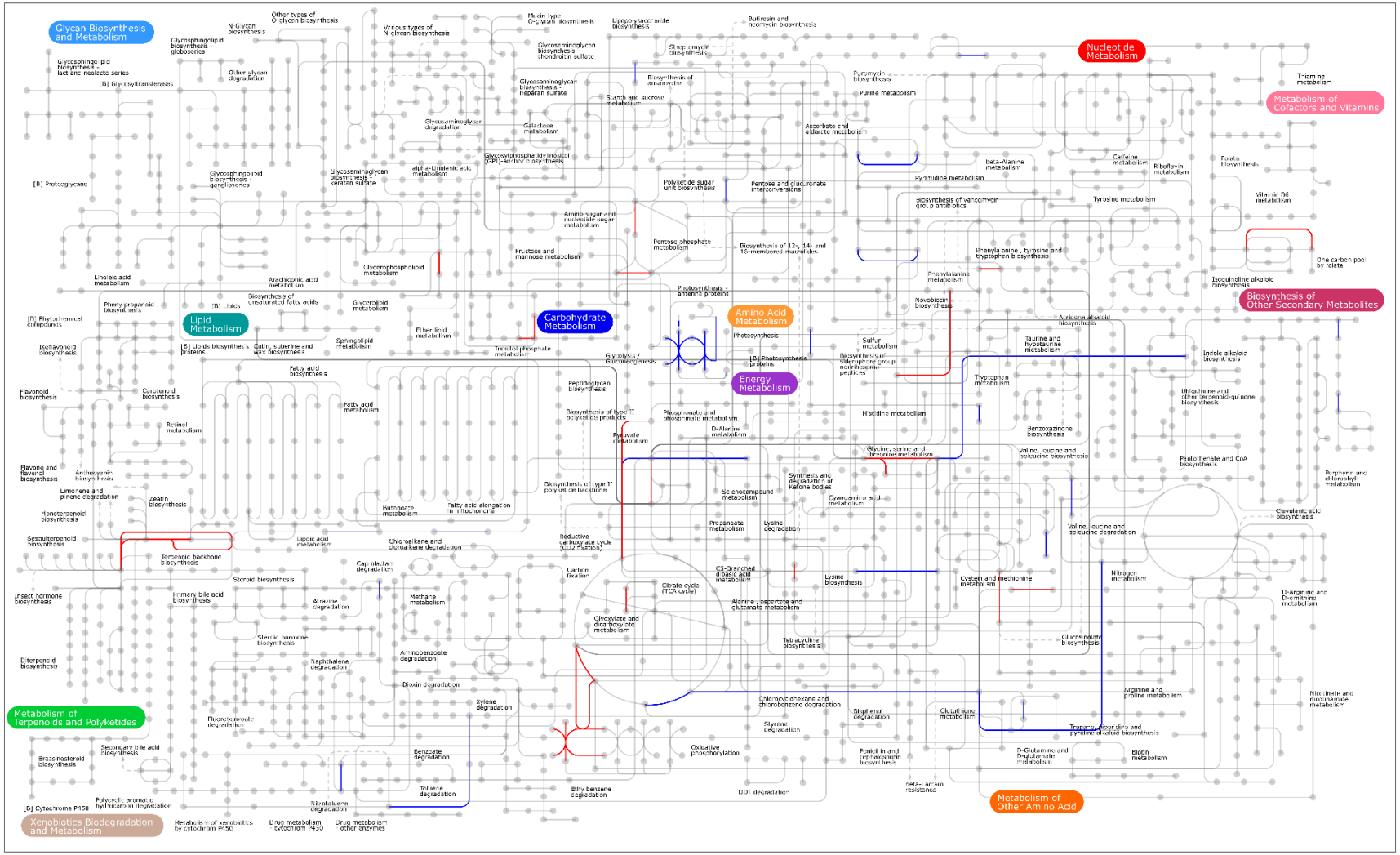

Figure 3-2. iPATH2 image of the soil microbial transcripts significantly expressed during early (seedling and vegetative-red) or late (bolting and flowering-blue) developmental stages.

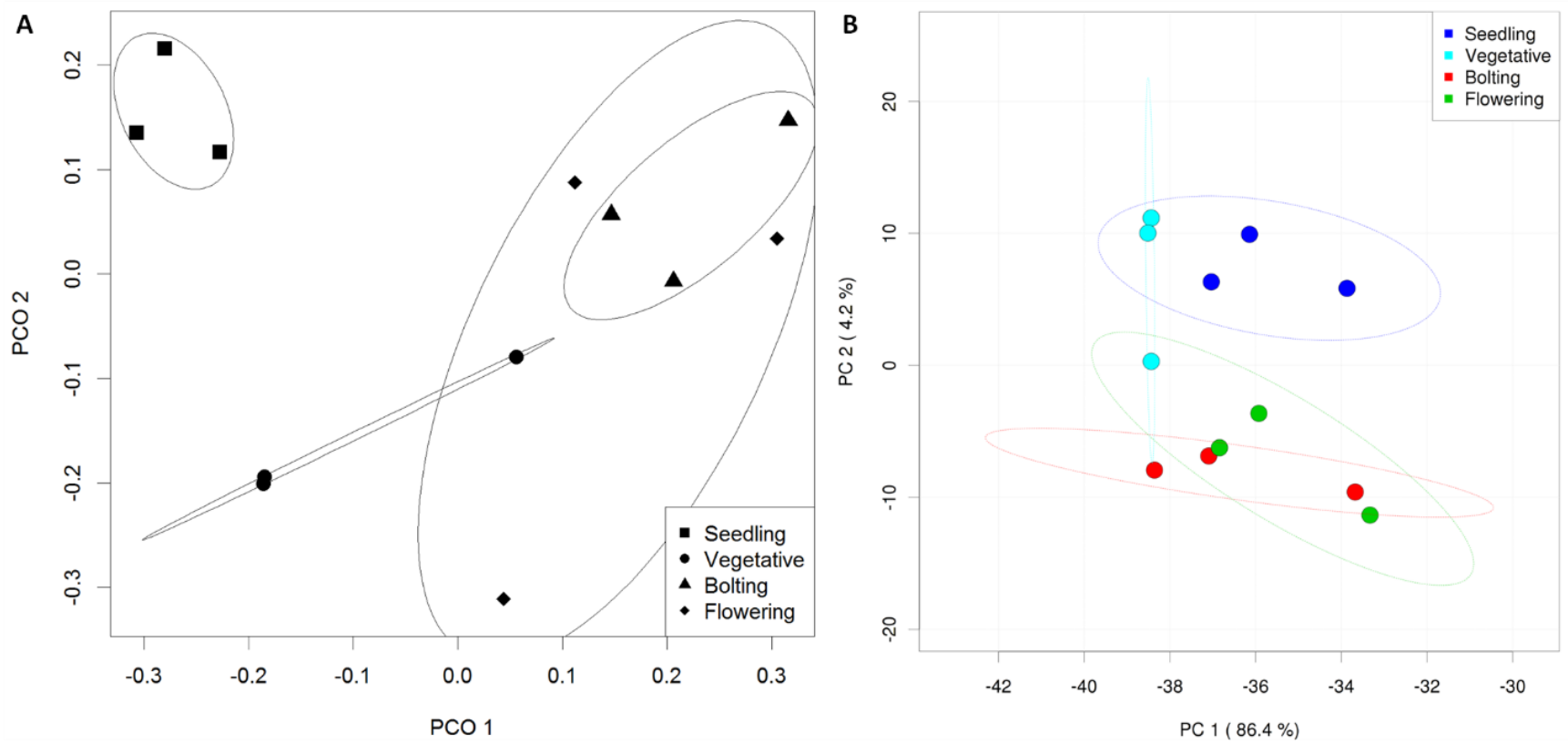

Figure 3-3. Multivariate analyses of the rhizosphere microbial community through plant development; analyzed by 454 pyrosequencing. (A) Principal Coordinate Analysis (PCoA) for the visualization of pairwise community dissimilarity (Bray-Curtis index) of the rhizosphere microbial community at each plant developmental stage (seedling, vegetative, bolting and flowering). 95% confidence ellipses are shown around each developmental stage. (B) Principal Component Analysis (PCA) of the rhizosphere microbial community at each plant developmental stage. 95% confidence ellipses are shown around each developmental stage.

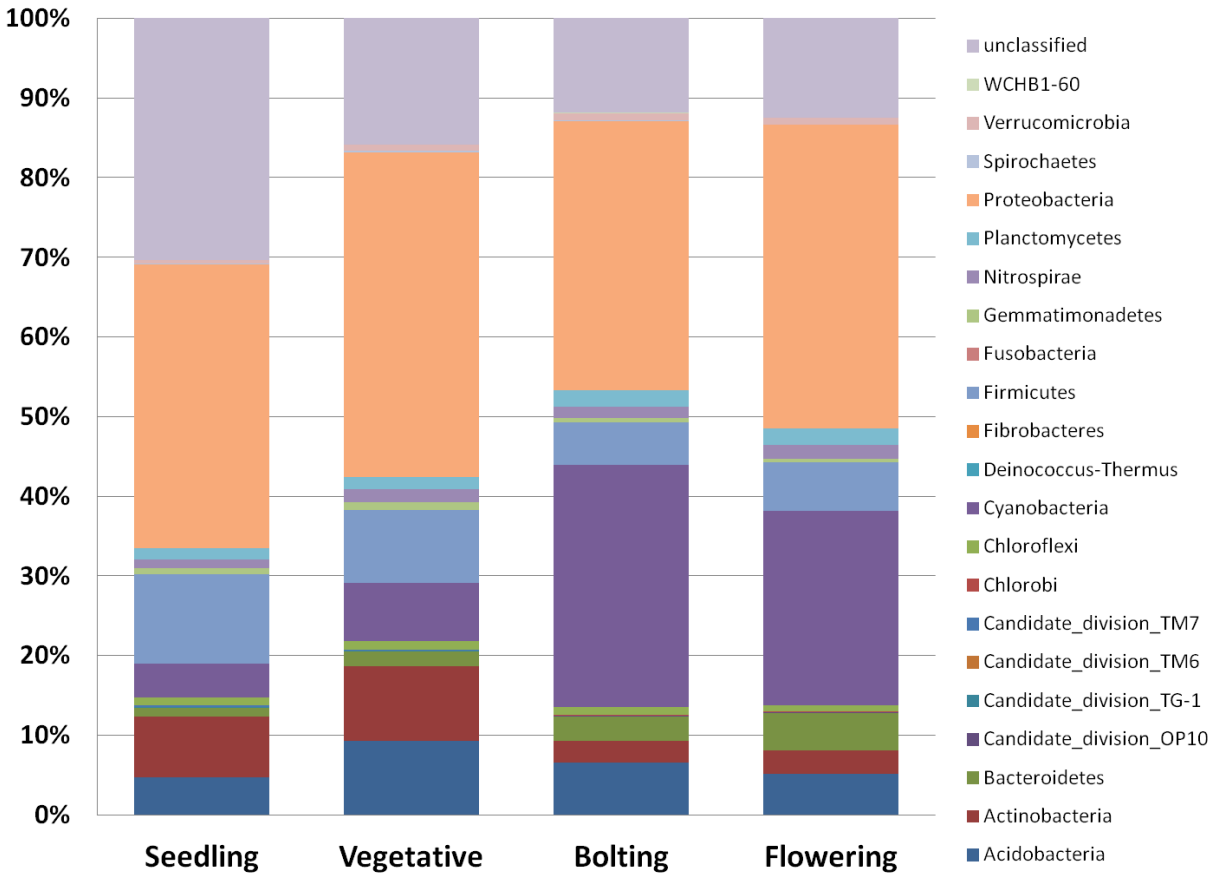

Figure 3-4. Relative abundance (%) of the major bacterial phyla present in the rhizosphere microbial community at each plant developmental stage.

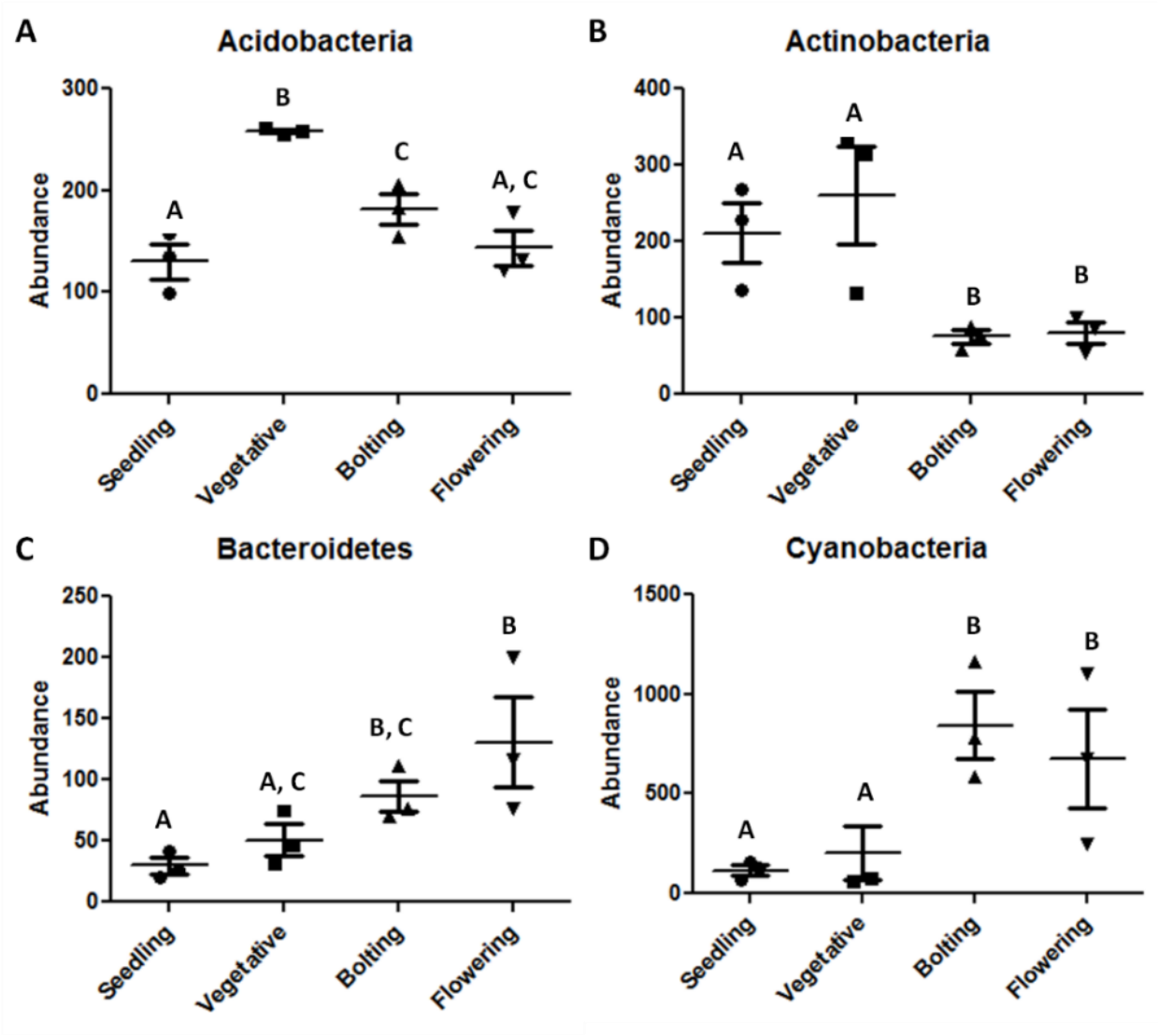

Figure 3-5. Bacterial phyla that significantly ($p<0.05$) change with plant development. (A) Acidobacteria, (B) Actinobacteria, (C) Bacteroidetes and (D) Cyanobacteria. The bars with different letters are significantly different (ANOVA Tukey post-hoc $p<0.05$) from one another. Each point represents one repetition and graphs show mean ± SE.

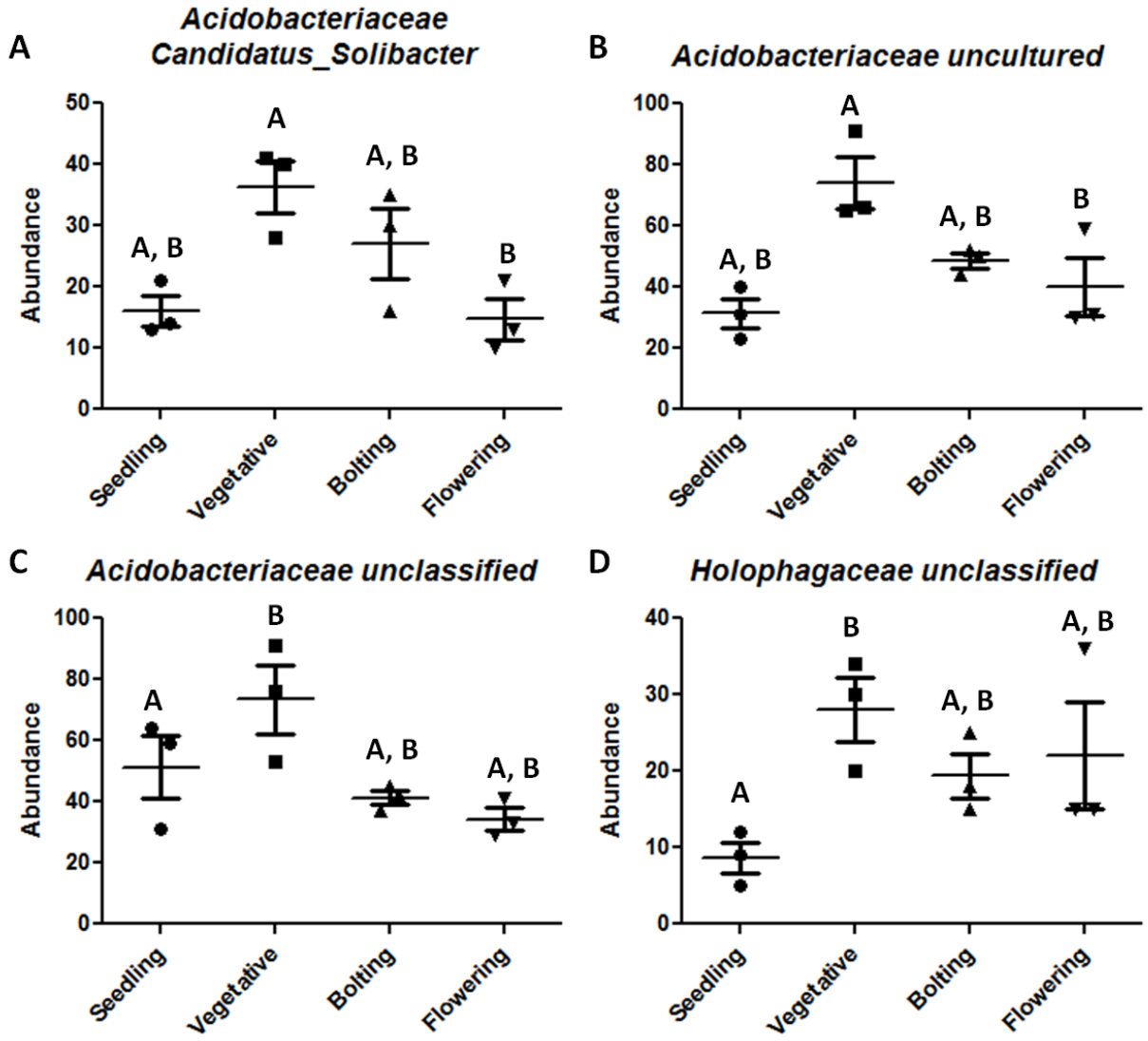

Figure 3-6. Bacteria classified as Acidobacteria that significantly ($p<0.05$) change with plant development. (A) Acidobacteriaceae Candidatus_Solibacter, (B) Acidobacteriaceae uncultured, (C) Acidobacteriaceae unclassified and (D) Holophagaceae unclassified. The bars with different letters are significantly different (ANOVA Tukey post-hoc $p<0.05$) from one another. Each point represents one repetition and graphs show mean ± SE.

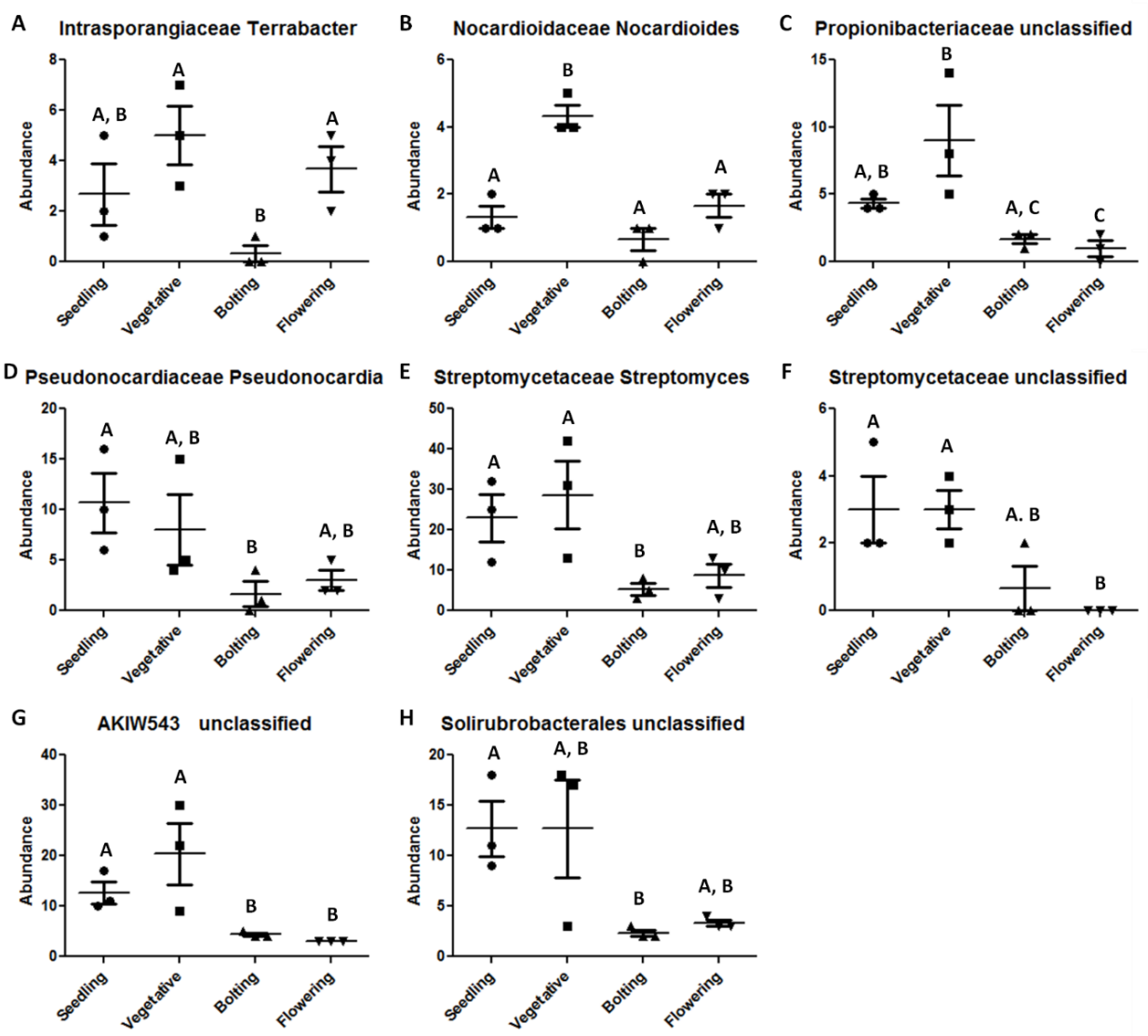

Figure 3-7. Bacteria classified as Actinobacteria that significantly (ANOVA Tukey post-hoc $p<0.05$) change with plant development. (A) *Intrasporangiaceae Terrabacter*, (B) *Nocardioidaceae Nocardioides*, (C) *Propionibacteriaceae* unclassified, (D) *Pseudonocardiaceae Psuedonocardia*, (E) *Streptomycetaceae Streptomyces*, (F) *Streptomycetaceae* unclassified, (G) AKIW543 unclassified and (H) *Solirubrobacterales* unclassified. The bars with different letters are significantly different (ANOVA Tukey post-hoc $p<0.05$) from one another. Each point represents one repetition and graphs show mean ± SE.

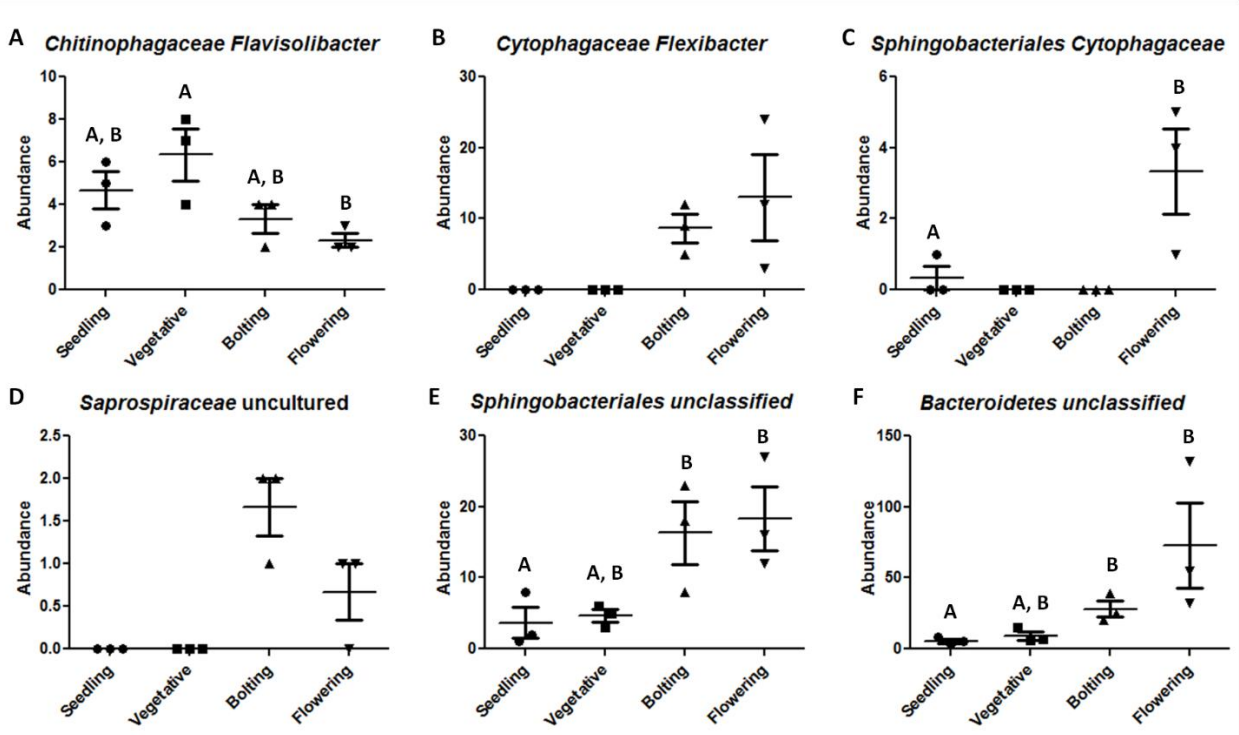

Figure 3-8. Bacteria classified as Bacteroidetes that significantly (ANOVA Tukey post-hoc $p<0.05$) change with plant development. (A) *Chitinophagaceae Flavisolibacter*, (B) *Cytophagaceae Flexibacter*, (C) *Sphingobacteriales Cytophagaceae*, (D) *Saprospiraceae* uncultured, (E) *Sphingobacteriales* unclassified and (F) Bacteroidetes unclassified. The bars with different letters are significantly different (ANOVA Tukey post-hoc $p<0.05$) from one another. Each point represents one repetition and graphs show mean ± SE.

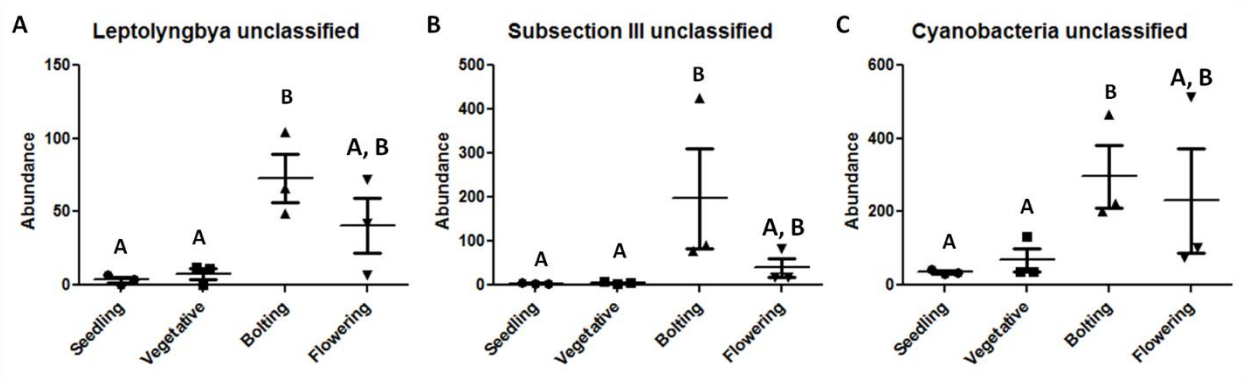

Figure 3-9. Bacteria classified as Cyanobacteria that significantly (ANOVA Tukey *post-hoc* $p<0.05$) change with plant development. (A) *Leptolyngbya* unclassified, (B) Subsection III unclassified and (C) Cyanobacteria unclassified. The bars with different letters are significantly different (ANOVA Tukey post-hoc $p<0.05$) from one another. Each point represents one repetition and graphs show mean ± SE.

REFERENCES

Abedon BG, Tracy WF (1996). Corngrass 1 of Maize (Zea mays L.) Delays Development of Adult Plant Resistance to Common Rust (Puccinia sorghi Schw.) and European Corn Borer (Ostrinia nubilalis Hubner). *Journal of Heredity* **87:** 219-223.

An C, Mou Z (2011). Salicylic acid and its function in plant immunity. *J Integr Plant Biol* **53:** 412-428.

Antoun H, Beauchamp C, Goussard N, Chabot R, Lalande R (1998). Potential of Rhizobium and Bradyrhizobium species as plant growth promoting rhizobacteria on non-legumes: Effect on radishes (Raphanus sativus L.). *Plant and Soil* **204:** 57-67.

Arndt D, Xia J, Liu Y, Zhou Y, Guo AC, Cruz JA *et al* (2012). METAGENassist: a comprehensive web server for comparative metagenomics. *Nucleic Acids Res* **40:** W88-95.

Badri DV, Loyola-Vargas VM, Broeckling CD, De-la-Pena C, Jasinski M, Santelia D *et al* (2008). Altered profile of secondary metabolites in the root exudates of Arabidopsis ATP-binding cassette transporter mutants. *Plant Physiol* **146:** 762-771.

Badri DV, Quintana N, El Kassis EG, Kim HK, Choi YH, Sugiyama A *et al* (2009). An ABC transporter mutation alters root exudation of phytochemicals that provoke an overhaul of natural soil microbiota. *Plant Physiol* **151:** 2006-2017.

Badri DV, Vivanco JM (2009). Regulation and function of root exudates. *Plant Cell Environ* **32:** 666-681.

Badri DV, Chaparro JM, Zhang R, Shen Q, Vivanco JM (2013a). Application of natural blends of phytochemicals derived from the root exudates of Arabidopsis to the soil reveal that phenolic related compounds predominantly modulate the soil microbiome. *J Biol Chem* **288:** 4502-4512.

Badri DV, Zolla G, Bakker MG, Manter DK, Vivanco JM (2013b). Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. *New Phytol*.

Bai S, Li J, He Z, Van Nostrand JD, Tian Y, Lin G *et al* (2013). GeoChip-based analysis of the functional gene diversity and metabolic potential of soil microbial communities of mangroves. *Appl Microbiol Biotechnol* **97:** 7035-7048.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* **57:** 233-266.

Baker BJ, Sheik CS, Taylor CA, Jain S, Bhasi A, Cavalcoli JD *et al* (2013). Community transcriptomic assembly reveals microbes that contribute to deep-sea carbon and nitrogen cycling. *Isme J* **7:** 1962-1973.

Barns SM, Takala SL, Kuske CR (1999). Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. *Appl Environ Microbiol* **65:** 1731-1737.

Baudoin E, Benizri E, Guckert A (2002). Impact of growth stage on the bacterial community structure along maize roots, as determined by metabolic and genetic fingerprinting. *Appl Soil Ecol* **19:** 135-145.

Bednarek P (2012). Chemical warfare or modulators of defence responses - the function of secondary metabolites in plant immunity. *Curr Opin Plant Biol* **15:** 407-414.

Behera B, Wagner GH (1974). Microbial Growth Rate In Glucose-amended Soil. *Soil Sci Soc Am J* **38:** 591-594.

Berendsen RL, Pieterse CM, Bakker PA (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478-486.

Bolan NS (1991). A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants. *Plant and Soil* **134:** 189-207.

Brimecombe MJ, de Leij FA, Lynch JM (2007). Rhizodeposition and Microbial Populations. In: Pinton R, Varanini Z, Nannapaneni P (eds). *The Rhizosphere Biochemistry and Organic Substances at the Soil-Plant Interface, Second Edition*. CRC Press: Boca Raton, Fl. pp 73-109.

Broeckling CD, Broz AK, Bergelson J, Manter DK, Vivanco JM (2008). Root exudates regulate soil fungal community composition and diversity. *Appl Environ Microbiol* **74:** 738-744.

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.

Burg RW, Miller BM, Baker EE, Birnbaum J, Currie SA, Hartman R *et al* (1979). Avermectins, new family of potent anthelmintic agents: producing organism and fermentation. *Antimicrob Agents Chemother* **15:** 361-367.

Burns IG (1996). Nitrogen supply, growth and development. *Acta Hort* **428:** 21-30.

Chaparro JM, Sheflin AM, Manter DK, Vivanco JM (2012). Manipulating the soil microbiome to increase soil health and plant fertility. *Biol Fertil Soils* **48:** 489-499.

Chaparro JM, Badri DV, Bakker MG, Sugiyama A, Manter DK, Vivanco JM (2013). Root exudation of phytochemicals in Arabidopsis follows specific patterns that are developmentally programmed and correlate with soil microbial functions. *PLOS ONE* **8:** e55731.

Chiarini L, Bevivino A, Dalmastri C, Tabacchioni S, Visca P (2006). Burkholderia cepacia complex species: health hazards and biotechnological potential. *Trends Microbiol* **14:** 277-286.

Clay NK, Adio AM, Denoux C, Jander G, Ausubel FM (2009). Glucosinolate metabolites required for an Arabidopsis innate immune response. *Science* **323:** 95-101.

Cohen MF, Yamasaki H, Mazzola M (2005). Brassica napus seed meal soil amendment modifies microbial community structure, nitric oxide production and incidence of Rhizoctonia root rot. *Soil Biol and Biochem* **37:** 1215.

Conn VM, Walker AR, Franco CM (2008). Endophytic actinobacteria induce defense pathways in Arabidopsis thaliana. *Mol Plant Microbe Interact* **21:** 208-218.

Coronado C, Zuanazzi J, Sallaud C, Quirion JC, Esnault R, Husson HP *et al* (1995). Alfalfa Root Flavonoid Production Is Nitrogen Regulated. *Plant Physiol* **108:** 533-542.

Davidson IA, Robson MJ (1986). Effect of Contrasting Patterns of Nitrate Application on the Nitrate Uptake, N2-Fixation, Nodulation and Growth of White Clover. *Annals of Botany* **57:** 331-338.

De-la-Pena C, Lei Z, Watson BS, Sumner LW, Vivanco JM (2008). Root-microbe communication through protein secretion. *J Biol Chem* **283:** 25247-25255.

De-la-Pena C, Badri DV, Lei Z, Watson BS, Brandao MM, Silva-Filho MC *et al* (2010). Root secretion of defense-related proteins is development-dependent and correlated with flowering time. *J Biol Chem* **285:** 30654-30665.

de Weert S, Vermeiren H, Mulders IHM, Kuiper I, Hendrickx N, Bloemberg GV *et al* (2002). Flagella-Driven Chemotaxis Towards Exudate Components Is an Important Trait for Tomato Root Colonization by Pseudomonas fluorescens. *Mol Plant Microbe Interact* **15:** 1173-1180.

Develey-Riviere MP, Galiana E (2007). Resistance to pathogens and host developmental stage: a multifaceted relationship within the plant kingdom. *New Phytol* **175:** 405-416.

Dugat-Bony E, Peyretaillade E, Parisot N, Biderre-Petit C, Jaziri F, Hill D *et al* (2012). Detecting unknown sequences with DNA microarrays: explorative probe design strategies. *Environ Microbiol* **14:** 356-371.

Eaglesham ARJ (1989). Nitrate inhibition of root-nodule symbiosis in doubly rooted soybean plants. *Crop science* **29:** 115-119.

Eilers KG, Lauber CL, Knight R, Fierer N (2010). Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. *Soil Biol and Biochem* **42:** 896.

Elbeltagy A, Nishioka K, Sato T, Suzuki H, Ye B, Hamada T *et al* (2001). Endophytic colonization and in planta nitrogen fixation by a Herbaspirillum sp. isolated from wild rice species. *Appl Environ Microbiol* **67:** 5285-5293.

Fang C, Zhuang Y, Xu T, Li Y, Li Y, Lin W (2013). Changes in rice allelopathy and rhizosphere microflora by inhibiting rice phenylalanine ammonia-lyase gene expression. *J Chem Ecol* **39:** 204-212.

Fierer N, Bradford MA, Jackson RB (2007). Toward an ecological classification of soil bacteria. *Ecology* **88:** 1354.

Fierer N, Nemergut D, Knight R, Craine JM (2010). Changes through time: integrating microorganisms into the study of succession. *Res Microbiol* **161:** 635-642.

Franche C, Lindstrom K, Elmerich C (2009). Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant and Soil* **321:** 35-59.

Gantar M, Kerby NW, Rowell P, Obreht Z (1991). Colonization of wheat Triticum vulgare L.) by N2-fixing cyanobacteria: I. A survey of soil cyanobacterial isolates forming associations with roots. *New Phytologist* **118:** 477.

Gaworzewska ET, Carlile MJ (1982). Positive Chemotaxis Of Rhizobium-Leguminosarum And Other Bacteria Towards Root Exudates From Legumes And Other Plants. *J Gen Microbiol* **128:** 1179-1188.

Gomez-Alvarez V, Teal TK, Schmidt TM (2009). Systematic artifacts in metagenomes from complex microbial communities. *Isme J* **3:** 1314-1317.

Gosalbes MJ, Durban A, Pignatelli M, Abellan JJ, Jimenez-Hernandez N, Perez-Cobas AE *et al* (2011). Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS One* **6:** e17447.

Gupta Sood S (2003). Chemotactic response of plant-growth-promoting bacteria towards roots of vesicular-arbuscular mycorrhizal tomato plants. *FEMS Microbiol Ecol* **45:** 219-227.

Gutiérrez-Mañero FJ, Ramos-Solano B, Probanza An, Mehouachi J, R. Tadeo F, Talon M (2001). The plant-growth-promoting rhizobacteria Bacillus pumilus and Bacillus licheniformis produce high amounts of physiologically active gibberellins. *Physiologia Plantarum* **111:** 206.

Hackstadt AJ, Hess AM (2009). Filtering for increased power for microarray data analysis. *BMC Bioinformatics* **10:** 11.

Haneline S, Connelly CJ, Melton T (1991). Chemotactic Behavior of Azotobacter vinelandii. *Appl Environ Microbiol* **57:** 825-829.

He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC *et al* (2007). GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *Isme J* **1:** 67-77.

He ZL, Van Nostrand JD, Deng Y, Zhou JZ (2011). Development and applications of functional gene microarrays in the analysis of the functional diversity, composition, and structure of microbial communities. *Front Environ Sci Engin China* **5:** 1-20.

Heding H (1964). Radioactive Myoinositol: Incorporation into Streptomycin. *Science* **143:** 953-954.

Horiuchi J, Prithiviraj B, Bais HP, Kimball BA, Vivanco JM (2005). Soil nematodes mediate positive interactions between legume plants and rhizobium bacteria. *Planta* **222:** 848-857.

Houlden A, Timms-Wilson TM, Day MJ, Bailey MJ (2008). Influence of plant developmental stage on microbial community structure and activity in the rhizosphere of three field crops. *FEMS Microbiol Ecol* **65:** 193-201.

Inceoglu O, Al-Soud WA, Salles JF, Semenov AV, van Elsas JD (2011). Comparative analysis of bacterial communities in a potato field as determined by pyrosequencing. *PLoS One* **6:** e23321.

Ivanova EG, Doronina NV, Trotsenko Iu A (2001). Aerobic methylobacteria are capable of synthesizing auxins. *Mikrobiologiia* **70:** 452-458.

Jaeger CH, 3rd, Lindow SE, Miller W, Clark E, Firestone MK (1999). Mapping of sugar and amino acid availability in soil around roots with bacterial sensors of sucrose and tryptophan. *Appl Environ Microbiol* **65:** 2685-2690.

Jain V, Nainawatee HS (2002). Plant flavonoids: Signals to legume nodulation and soil microorganisms. *J Plant Biochem Biot* **11:** 1-10.

Junier P, Junier T, Witzel KP, Caru M (2009). Composition of diazotrophic bacterial assemblages in bean-planted soil compared to unplanted soil. *Eur J Soil Biol* **45:** 153-162.

Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res* **32:** D277-280.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M *et al* (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res* **36:** D480-484.

Kelly JT, Bacon RK, Wells BR (1995). Genetic variability in nitrogen utilization at four growth stages in soft red winter wheat. *Journal of Plant Nutrition* **18:** 969.

Kent WJ (2002). BLAT--the BLAST-like alignment tool. *Genome Res* **12:** 656-664.

Kus JV, Zaton K, Sarkar R, Cameron RK (2002). Age-related resistance in Arabidopsis is a developmentally regulated defense response to Pseudomonas syringae. *Plant Cell* **14:** 479-490.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25.

Lau JA, Lennon JT (2011). Evolutionary ecology of plant-microbe interactions: soil microbial structure alters selection on plant traits. *New Phytologist* **192:** 215-224.

Lee SH, Ka JO, Cho JC (2008). Members of the phylum Acidobacteria are dominant and metabolically active in rhizosphere soil. *FEMS Microbiol Lett* **285:** 263-269.

Li X, Bergelson J, Chapple C (2010). The ARABIDOPSIS Accession Pna-10 Is a Naturally Occurring sng1 Deletion Mutant. *Mol Plant* **3:** 91-100.

Lidstrom ME, Chistoserdova L (2002). Plants in the pink: cytokinin production by methylobacterium. *J Bacteriol* **184:** 1818.

Lin L, Ge HM, Yan T, Qin YH, Tan RX (2012). Thaxtomin A-deficient endophytic Streptomyces sp. enhances plant disease resistance to pathogenic Streptomyces scabies. *Planta* **236:** 1849-1861.

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.

Lynch JM (ed) (1990) *The Rhizosphere*. Wiley Interscience: Chinchester, UK, 458pp.

Majumdar SK, Kutzner HJ (1962). Myo-inositol in the biosynthesis of streptomycin by Streptomyces griseus. *Science* **135:** 734.

Malagoli P, Laine P, Le Deunff E, Rossato L, Ney B, Ourry A (2004). Modeling nitrogen uptake in oilseed rape cv Capitol during a growth cycle using influx kinetics of root nitrate transport systems and field experimental data. *Plant Physiol* **134:** 388-400.

Meier IC, Avis PG, Phillips RP (2013). Fungal communities influence root exudation rates in pine seedlings. *FEMS Microbiol Ecol* **83:** 585-595.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332:** 1097-1100.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al* (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *Bmc Bioinformatics* **9**.

Micallef SA, Channer S, Shiaris MP, Colon-Carmona A (2009a). Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. *Plant Signal Behav* **4:** 777-780.

Micallef SA, Shiaris MP, Colon-Carmona A (2009b). Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. *J Exp Bot* **60:** 1729-1742.

Millet YA, Danna CH, Clay NK, Songnuan W, Simon MD, Werck-Reichhart D *et al* (2010). Innate Immune Responses Activated in Arabidopsis Roots by Microbe-Associated Molecular Patterns. *Plant Cell* **22:** 973-990.

Morgan JA, Bending GD, White PJ (2005). Biological costs and benefits to plant-microbe interactions in the rhizosphere. *J Exp Bot* **56:** 1729-1739.

Mougel C, Offre P, Ranjard L, Corberand T, Gamalero E, Robin C *et al* (2006). Dynamic of the genetic structure of bacterial and fungal communities at different developmental stages of Medicago truncatula Gaertn. cv. Jemalong line J5. *New Phytol* **170:** 165-175.

Murashige T, Skoog F (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol Plant* **15:** 473.

Nazoa P, Vidmar JJ, Tranbarger TJ, Mouline K, Damiani I, Tillard P *et al* (2003). Regulation of the nitrate transporter gene AtNRT2.1 in Arabidopsis thaliana: responses to nitrate, amino acids and developmental stage. *Plant Mol Biol* **52:** 689-703.

Neal AL, Ahmad S, Gordon-Weeks R, Ton J (2012). Benzoxazinoids in Root Exudates of Maize Attract Pseudomonas putida to the Rhizosphere. *PLoS One* **7:** e35498.

Nicoli A, Zambolim L, Junior TJP, Vieira RF, Teixeira H, Carneiro JES (2012). Resistance of advanced common bean lines to Fusarium root rot. *Trop plant pathol* **37:** 393-398.

Novoa R, Loomis RS (1981). Nitrogen and plant production. *Plant and Soil* **58:** 177-204.

Oksanen J, Blanchet G, Kindt R, Legendre P, Minchin PR, O'Hara RB *et al* (2012). vegan: Community Ecology Package. R package.

Poretsky RS, Gifford S, Rinta-Kanto J, Vila-Costa M, Moran MA (2009). Analyzing gene expression from marine microbial communities using environmental transcriptomics. *J Vis Exp* **24:** e1086.

Pradella S, Hans A, Sproer C, Reichenbach H, Gerth K, Beyer S (2002). Characterisation, genome size and genetic manipulation of the myxobacterium Sorangium cellulosum So ce56. *Arch Microbiol* **178:** 484-492.

Prasanna R, Jaiswal P, Nayak S, Sood A, Kaushik BD (2009). Cyanobacterial diversity in the rhizosphere of rice and its ecological significance. *Indian J Microbiol* **49:** 89-97.

Rogers EE, Glazebrook J, Ausubel FN (1996). Mode of action of the Arabidopsis thaliana phytoalexin camalexin and its role in Arabidopsis-pathogen interactions. *Molecular Plant-Microbe Interactions* **9:** 748-757.

Rossato L, Laine P, Ourry A (2001). Nitrogen storage and remobilization in Brassica napus L. during the growth cycle: nitrogen fluxes within the plant and changes in soluble protein patterns. *J Exp Bot* **52:** 1655-1663.

Rusterucci C, Zhao Z, Haines K, Mellersh D, Neumann M, Cameron RK (2005). Age-related resistance to Pseudomonas syringae pv. tomato is associated with the transition to flowering in Arabidopsis and is effective against Peronospora parasitica. *Physiological and Molecular Plant Pathology* **66:** 222.

Schatz A, Bugie E, Waksman SA (2005). Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944. *Clin Orthop Relat Res***:** 3-6.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75:** 7537-7541.

Selvakumar G, Panneerselvam P, Ganeshamurthy AN, Maheshwari DK (2012). Bacterial Mediated Alleviation of Abiotic Stress in Crops. In: Maheshwari DK (ed). *Bacteria in Agrobiology: Stress Management*. Springer: New York, NY. pp 205-224.

Shade A, Gregory Caporaso J, Handelsman J, Knight R, Fierer N (2013). A meta-analysis of changes in bacterial and archaeal communities with time. *Isme J* **7:** 1493-1506.

Shaw CH (1991). Swimming against the tide: Chemotaxis in Agrobacterium. *BioEssays* **13:** 25-29.

Shi S, Richardson AE, O'Callaghan M, DeAngelis KM, Jones EE, Stewart A *et al* (2011). Effects of selected root exudate components on soil bacterial communities. *FEMS Microbiol Ecol* **77:** 600-610.

Stacey G, Sanjuan J, Luka S, Dockendorff T, Carlson RW (1995). Signal exchange in the Bradyrhizobium-soybean symbiosis. *Soil Biol and Biochem* **27:** 473.

Sy A, Giraud E, Jourand P, Garcia N, Willems A, de Lajudie P *et al* (2001). Methylotrophic Methylobacterium bacteria nodulate and fix nitrogen in symbiosis with legumes. *J Bacteriol* **183:** 214-220.

Tokala RK, Strap JL, Jung CM, Crawford DL, Salove MH, Deobald LA *et al* (2002). Novel plant-microbe rhizosphere interaction involving Streptomyces lydicus WYEC108 and the pea plant (Pisum sativum). *Appl Environ Microbiol* **68:** 2161-2171.

Turner TR, Ramakrishnan K, Walshaw J, Heavens D, Alston M, Swarbreck D *et al* (2013). Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *Isme J*.

Urich T, Lanzen A, Qi J, Huson DH, Schleper C, Schuster SC (2008). Simultaneous Assessment of Soil Microbial Community Structure and Function through Analysis of the Meta-Transcriptome. *PLoS ONE* **3:** e2527.

van der Heijden MGA, Bardgett RD, Van Straalen NM (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* **11:** 296-310.

Van Spanning RJM, Delgado MJ, Richardson DJ (2005). The Nitrogen Cycle: Denitrification and its Relationship to N2 Fixation. In: Werner D, Newton WE (eds). *Nitrogen Fixation in Agriculture, Forestry, Ecology, and the Environment*. Springer Netherlands. pp 277-342.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *App Environ Microb* **73:** 5261-5267.

Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinho PM, Wu M *et al* (2009). Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microbiol* **75:** 2046-2056.

Xiong X, Frank DN, Robertson CE, Hung SS, Markle J, Canty AJ *et al* (2012). Generation and analysis of a mouse intestinal metatranscriptome through Illumina based RNA-sequencing. *PLoS One* **7:** e36009.

Xu Y, Wang G, Jin J, Liu J, Zhang Q, Liu X (2009). Bacterial communities in soybean rhizosphere in response to soil type, soybean genotype, and their growth stage. *Soil Biol Biochem* **41:** 919.

Yamada T, Letunic I, Okuda S, Kanehisa M, Bork P (2011). iPath2.0: interactive pathway explorer. *Nucleic Acids Res* **39:** W412-415.

Yousuf B, Keshri J, Mishra A, Jha B (2012). Application of targeted metagenomics to explore abundance and diversity of CO(2)-fixing bacterial community using cbbL gene from the rhizosphere of Arachis hypogaea. *Gene* **506:** 18-24.

Zahran HH (1999). Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol Mol Biol Rev* **63:** 968-989, table of contents.

Zaki K, Misaghi IJ, Heydari A (1998). Control of Cotton Seedling Damping-off in the Field by Burkholderia (Pseudomonas) cepacia. *Plant Disease* **82:** 291-293.

Zamioudis C, Pieterse CM (2012). Modulation of host immunity by beneficial microbes. *Mol Plant Microbe Interact* **25:** 139-150.

Zhang J, Subramanian S, Stacey G, Yu O (2009). Flavones and flavonols play distinct critical roles during nodulation of Medicago truncatula by Sinorhizobium meliloti. *Plant J* **57:** 171-183.

Zhang Y, Lu Z, Liu S, Yang Y, He Z, Ren Z *et al* (2013). Geochip-based analysis of microbial communities in alpine meadow soils in the Qinghai-Tibetan plateau. *BMC Microbiol* **13:** 72.

Zolla G, Badri DV, Bakker MG, Manter DK, Vivanco JM (2013). Soil microbiomes vary in their ability to confer drought tolerance to Arabidopsis. *App Soil Ecol* **68:** 1.

CHAPTER 4 APPLICATION OF NATURAL BLENDS OF PHYTOCHEMICALS DERIVED FROM THE ROOT EXUDATES OF *ARABIDOPSIS* TO THE SOIL REVEAL THAT PHENOLIC-RELATED COMPOUNDS PREDOMINANTLY MODULATE THE SOIL MICROBIOME[1]

Synopsis

The roots of plants have the ability to influence its surrounding microbiology, the so-called rhizosphere microbiome, through the creation of specific chemical niches in the soil mediated by the release of phytochemicals. Here we report how these phytochemicals could modulate the microbial composition of a soil in the absence of the plant. For this purpose, root exudates of *Arabidopsis* were collected and fractionated to obtain natural blends of phytochemicals at various relative concentrations that were characterized by GC-MS and applied repeatedly to a soil. Soil bacterial changes were monitored by amplifying and pyrosequencing the 16S ribosomal small subunit region. Our analysis reveal that one phytochemical can culture different operational taxonomic units (OTUs), mixtures of phytochemicals synergistically culture groups of OTUs, and the same phytochemical can act as a stimulator or deterrent to different groups of OTUs. Furthermore, phenolic-related compounds showed positive correlation with a higher number of unique OTUs compared with other groups of compounds (i.e. sugars, sugar alcohols, and amino acids). For instance, salicylic acid showed positive correlations with species of Corynebacterineae, Pseudonocardineae, and Streptomycineae, and GABA correlated with

[1] This research was originally published in Journal of Biological Chemistry. Badri, DV; Chaparro, JM; Zhang, R; Shen, Q; Vivanco, JM. Application of natural blends of phytochemicals derived from root exudates of *Arabidopsis* to the soil reveal that phenolic-related compounds predominantly modulate the soil microbiome. *JBC*. 2013; 288:4502-4512.

species of *Sphingomonas*, *Methylobacterium*, Frankineae, *Variovorax*, Micromonosporineae, and *Skermanella*. These results imply that phenolic compounds act as specific substrates or signaling molecules for a large group of microbial species in the soil.

Introduction

It is assumed that the roots of plants create physical and chemical niches that allow the colonization of microbes in the rhizosphere, the soil immediately surrounding the roots of a plant. Increasing evidences demonstrate that plants predominantly drive and shape the selection of microbes (Badri and Vivanco 2009, Berendsen et al 2012, Broeckling et al 2008, Chaparro et al 2012, Hartmann et al 2009) by the active secretion of compounds that specifically stimulates or represses distinct microbial members of the soil community (Abdel-Lateif et al 2012, Doornbos et al 2012, Neal et al 2012, Somers et al 2004) in order to shape the rhizosphere microbiome (Bakker et al 2012, Chaparro et al 2012). For example, legumes exude specific flavonoids that act as signaling molecules to attract nitrogen-fixing bacteria (Broughton et al 2003). Yet when soil nitrogen is not limiting to the plant, this symbiosis does not occur (Omrane and Chiurazzi 2009). On the other hand, certain legumes release canavanine, an antimicrobial that acts against a broad range of microbes without affecting *Rhizobia*, which further cultures this beneficial microbe in the rhizosphere of legumes (Cai et al 2009). Other plant roots release strigolactones (sesquiterpenes) to attract mycorrhizae, and parasitic plants have benefited from these chemical cues to recognize their host plants (Akiyama et al 2005, Yoneyama et al 2008). Plants can also culture beneficial microbes such as plant-growth promoting rhizobacteria by the release of organic acids. For example, citric and fumaric acids released from tomato roots have been shown to attract *Pseudomonas fluorescens* strains (de Weert et al 2002, Gupta Sood 2003). Besides these specific signaling molecules, sugars serve as sources of energy for a broad range of

microbes (Behera and Wagner 1974), and secondary metabolites such as phenolics may function as general antimicrobials (de Weert et al 2002, Rudrappa et al 2008, Steinkellner et al 2007, Zhang et al 2009).

Two recent studies have characterized the core rhizosphere microbiome of *Arabidopsis*, which is mainly comprised of the groups: Actinobacteria, Proteobacteria, and Bacteroidetes, which are common inhabitants of diverse soils (Bulgarelli et al 2012, Janssen 2006, Lin et al 2012, Lundberg et al 2012, Philippot et al 2010, Singh et al 2007). Furthermore, the root exudates of a loss-of-function *Arabidopsis* mutant in an ABC transporter was found to have an increase in phenolic compounds and a decline in sugars compared with the wild type (Badri et al 2009a). When this mutant was grown in soil, it elicited dramatic quantitative and qualitative changes in the *Arabidopsis* native soil microbial community; it cultivated a microbial community with a relatively greater abundance of beneficial bacteria (i.e. plant growth-promoting rhizobacteria and nitrogen fixers). These studies suggest a correlation between components in the root exudates and the soil microbiome as a whole.

The goal of this study was to get a finer and initial understanding of how the chemical diversity present in the root exudates of plants can promote or inhibit the growth of specific groups within a natural soil microbiome in the absence of the plant. For this purpose we collected, fractionated, and chemically characterized root exudates of *in vitro* grown *Arabidopsis thaliana* plants, and repeatedly supplemented those fractions to *Arabidopsis* co-adapted soil (defined as a natural soil that has supported growth of *Arabidopsis* for a long period of time) without the presence of *Arabidopsis*. The soil microbial communities related to each fraction were characterized by 454 sequencing, and correlation analysis between microbial groups and phytochemicals was conducted.

Materials and methods

Plant growth conditions and collection of root exudates

A. thaliana wild type (Col-0) seeds were surface-sterilized with Clorox® (laundry bleach) for 1 min followed by four rinses in sterile distilled water and plated on full-strength Murashige and Skoog agar media supplemented with 3% sucrose. Plates were incubated in a growth chamber (Percival Scientific) at 25 °C, with a photoperiod of 16h light/8h dark for germination. We followed the methodology for collecting root exudates described by Badri et al. (2008a, 2009a, 2012a, 2012b). Briefly, 7-day-old seedlings were transferred to 6-well culture plates, each well containing 5 ml of liquid Murashige and Skoog (full strength Murashige and Skoog salts supplemented with 1% sucrose); these were incubated on an orbital shaker at 90 rpm and illuminated under cool white fluorescent light (45 µmol m^{-2} s^{-1}) with a photoperiod of 16-h light/8-h dark at 25 ±2 °C. When plants were 18 days old they were gently washed with sterile distilled water to remove the surface-adhering exudates and transferred to new 6-well plates containing 5 ml sterile distilled water and incubated for 3 days on an orbital shaker under the same conditions described above. We used sterile distilled water to prevent the interference of exogenously supplemented salts and sucrose present in the Murashige and Skoog liquid media in subsequent GC-MS analyses of the root exudates. The 3-day period of incubation did not create any visible toxicity symptoms on the plants (Figure 0-1) as it has been previously reported (Behera and Wagner 1974, Rudrappa et al 2008, Steinkellner et al 2007, Zhang et al 2009). The particular window frame of collection of exudates (18-21 days) was selected because during this period *Arabidopsis* roots were reported to secrete a high diversity of phytochemicals (Behera and Wagner 1974, Rudrappa et al 2008, Steinkellner et al 2007, Zhang et al 2009). The exudates contained in the media were collected and filtered using nylon filters of pore size 0.45 µm

(Millipore) to remove root sheathing and root border-like cells. By following this method, we collected a total of three liters of root exudates by growing 600 individual *Arabidopsis* plants in 6-well plates. We pooled all the root exudates collected from 600 individual *Arabidopsis* plants for further fractionation analyses. In other words, we considered that these 3 liters of pooled root exudates represented 600 individual biological replicates. It should be noted that the profiles of root exudates of *Arabidopsis* 18-21-day-old plants have been previously reported and found to be consistent and reproducible (Badri et al 2008a, 2009a, 2012a, 2012b).

Fractionation of root exudates

Filtered root exudates were freeze-dried and dissolved in 25 ml of sterile distilled water, and the pH was adjusted to 2 with 1N HCl and partitioned with 25 ml of ethyl acetate; the organic phase (ethyl acetate) was separated and dried under nitrogen gas. The aqueous phase was subsequently fractionated with 25 ml of chloroform to separate the organic (chloroform) and water phases. For each solvent type, we performed the extraction two times and then pooled the extractions together. Both fractions were collected independently and dried under nitrogen gas. At the end we had three types of fractions: an ethyl acetate (EtoAc) fraction, a chloroform ($CHCl_3$) fraction, and a water fraction. Exudates collected from the plants without fractionation served as whole exudates.

Gas chromatography-mass spectrometry (GC-MS) analyses of exudate fractions and data analyses

The fractions and whole exudates were subjected to GC-MS analyses at the Genome Center Core Services, University of California Davis to identify the compounds present in each fraction. Briefly, the whole exudates and fractions were dried under nitrogen gas followed by methoximation and trimethylsilylation derivatization as described by Sana et al. (2010). An

Agilent 6890 gas chromatograph (Santa Clara, CA) containing a 30-m-long, 0.25-mm inner diameter rtx5Sil-MS column with an additional 10 m integrated guard column was used to run the samples controlled by Leco ChromaTOF software Version 2.32 (St. Joseph, MI). The resulting text files were exported to the data server with absolute spectra intensities and further processed by a filtering algorithm implemented in the metabolomics BinBase database (Fiehn et al 2005). Quantification was reported as peak height using the unique ion as default. Metabolites were unambiguously assigned by the BinBase identifier numbers using retention index and mass spectrum as the two most important identification criteria. Additional confidence criteria were used by giving mass spectral metadata using the combination of unique ions, apex ions, peak purity, and signal/noise ratios. All entries in BinBase were matched against the Fiehn mass spectral library. Data normalization was performed as described in Fiehn et al. (2008) by using "total metabolite content." Furthermore, we determined the concentration of four target analytes (D-(+) glucose, serine, and valine) producing multiple derivatization products were summed before calculating the linear calibration curves. Final analyte masses were adjusted to sample preparation volume and were reported in nanomoles.

Supplementing exudate fractions to Arabidopsis co-adapted soil

Top soil (0-10 cm) was collected in July, 2011 from long standing fallow soil where *Arabidopsis* genotypes grow naturally. This field has been fallow for approximately more than eight years and is located at the Michigan Extension Station, Benton Harbor, MI (N42° 05' 34'', W86° 21' 19'' W, elevation 630 feet). Recently, it has been reported that the *Arabidopsis* genotype Pna-10 that harbors a *sng1* mutation grows naturally on this site (Li et al 2010) along with some natural grasses. Soil was collected from three spots at this location, transported to the laboratory in air tight coolers and stored in a cold room (4^{o}C) until use. Before the start of the

experiment, all soils collected from the three spots were dried under room temperature, pooled, and thoroughly homogenized by hand. Cubical pots (length 2.0 X width 1.0 X height 2.0 inches) were lined with Whatmann No. 3MM filter paper to prevent soil loss. The pots were then filled with soil and incubated in a growth chamber under the photoperiod of 16h light/8h dark at 25 ± 2°C for 2 weeks and sufficiently watered before supplementing them with the exudate fractions. During this 2-week period, the *Arabidopsis* seedlings (that were present in the natural soil) were continuously removed from the existing seed bank present in the soil. After complete removal of the existing seed bank seedlings, the exudate fractions were independently supplemented to each of the pots, of which none contained a plant. In total, there were eight treatments including controls: pots supplemented with whole exudates, EtoAc fraction, $CHCl_3$ fraction, water fraction, and four controls (only the solvents: methanol (2%), chloroform (2%), and sterile water; the negative control pots receiving no solvents (nothing)). For each treatment nine pots (considered as nine biological replicates) were maintained, and each pot received 2 ml of solution, which is equivalent to the root exudates of two plants. Additionally, we determined the absolute concentrations of four marker analytes present in each fraction and whole exudates that were supplemented to the soil (Table 4-1 and Table 4-2). These fractions were added twice a week with an interval of 3 days between fractions for 4 weeks. None of the pots received additional supplementation of water during the experimental period. The soil samples were collected at the end of the 4-week period. The collected soils were stored at -80°C before pyrosequencing. Further, we made three soil replicates by pooling three sub replicates into one for subsequent pyrosequencing analysis to improve the coverage of the microbiome and to reduce the variability between the replicates within a given treatment (see Figure 0-2).

Soil DNA extraction and pyrosequencing

To characterize the soil microbial community, total DNA was extracted from soil by using a MoBio ultraclean soil DNA kit (Mo Bio, Carlsbad, CA) according to the manufacturer's instructions. The DNA was quantified using a Nanodrop (Nanodrop Technologies, Wilmington, DE) spectrophotometer, and all DNA had an absorbance ratio (A_{260}/A_{280}) between 1.7 and 1.9. PCR amplification was performed by using the primer pairs 27F (AGAGTTTGATYMTGGCTCAG) and 533R (TTACCGCGGCTGCTGGC) with 454 adaptor (454 Life Sciences, Branford, CT, USA) and 10-base long barcode sequences, which are not shown here, to amplify the variable regions (V1-V3) of 16S ribosomal small subunit region with the following PCR conditions: the reaction mix (50 µl) contained 0.4 µmoles of each primer, 200 µmoles of dNTPs, 1 X reaction buffer, and one unit of TaqDNA polymerase (Takara). PCR included 39 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 1 min in an Applied Biosystems thermal cycler (GeneAmp PCR system 2700). For each exudate fraction applied to the soil, nine biological replicates were subjected to PCR amplification. At the end of the PCR procedure, three biological replicates were derived from the nine biological replicates by pooling three samples together. After pooling, the PCR products were purified using AMPure XP beads (Agencourt) before running the agarose gel electrophoresis. The specific amplicon product (~600 bp) was excised from the gel using a QIAquick gel extraction kit (Qiagen) according to the manufacturer's instructions. The purified amplicon products were subjected to unidirectional pyrosequencing by using a 454 GS FLX Titanium sequencing platform. Pyrosequencing was performed under contract with Roy J. Carver Biotechnology Center, University of Illinois, Urbana-Champaign.

Sequencing analysis

Sequence reads were processed using Mothur v. 1.25.1 (Schloss et al 2009) following the Schloss SOP. Reads having a minimum flow length of 360 flows were de-noised. This was done by processing through the default parameters of the Mothur-based re-implementation of the PyroNoise algorithm (Quince et al 2011). De-noised reads were screened by the following quality criteria: no more than 2 mismatches to the forward primer sequence, no more than 1 mismatch in the barcode sequence (for sample multiplexing), no homopolymeric runs of more than 8 nucleotides, and read length of greater than 200 bases. Reads passing these quality criteria were aligned to the SILVA bacterial reference database (Pruesse et al 2007). Aligned reads were screened to begin at the same position. Reads were removed if they did not reach the position at which 95% of reads ended. Chimeras were detected using the UChime method (Edgar et al 2011) and excluded from further analysis. Sampling effort was equalized to the depth of the smallest sample (2185 reads) and operational taxonomic units (OTU) were defined at 3% sequence dissimilarity, with the average neighbor algorithm. Reads were classified using the naïve Baysian classifier embedded in Mothur (Wang et al 2007). Final taxonomic assignment was based on the consensus identification for each OTU. A multivariate data analysis was performed by using METAGENassist a web server tool (Arndt et al 2012) that assigns probable microbial functions based on taxonomy (16S ribosomal subunit). Data-filtering was performed by the interquantile range method followed by quantile normalization within replicates after log transformation. Principal component analysis and identification of significant features were performed for all treatments together. Cluster analysis was performed using the Ward method.

Statistical analysis

We computed the Pearson correlation of the OTU abundances to the exudate fractions using SAS Version 9.3 (SAS Institute, Cary, NC). For the data to follow a normal distribution, it was transformed. For the GC-MS data of the root exudate fractions, the peak area of each identified compound underwent a log transformation. On the other hand, the abundance of each OTU in each sample was normalized using a $\log_2$ transformation procedure and subsequently standardized to a mean of 0 and S.D. of 1 as is usually implemented for microarray data (Speed 2003). The average of the transformed OTU abundances for each treatment was correlated to the transformed GC-MS-identified compounds found in each treatment. A p-value of <0.05 indicated a significant correlation.

Results

Composition of compounds in each fraction by GC-MS analyses

In total, we detected 415 compounds in the fractions and whole exudates. We observed that all identified compounds were present in all the fractions, but the abundance of a particular compound in each fraction varied. Furthermore, we identified 130 compounds by broadly categorizing them based on its chemical nature (i.e. sugars, sugar alcohols, amino acids, and phenolics). The phenolics category includes organic acids, fatty acids, and aliphatic and aromatic amino acids. In total we identified 12 sugars, 11 sugar alcohols, 29 amino acids and 59 phenolic compounds (Table 4-3). The fractionation of exudates modified the composition of the major types of compounds present in a particular fraction compared to the whole exudates (Table 4-4). For instance, the water fraction had a higher abundance of amino acids (29.55 *versus* 9.85%) and phenolics (32.01 *versus* 18.55%) compared with the whole exudates. Similarly, the $CHCl_3$ fraction had a higher concentration of sugars (27.25 *versus* 3.10%) compared with the whole

exudates. The EtoAc fraction showed an increase in amino acids (21.47 *versus* 9.85%) and phenolics (30.24 *versus* 18.55%) compared with the whole exudates. It should be noted that in all fractions as well as in the whole exudates the unknowns accounted for the highest percentage of compounds.

Influence of whole exudates and fractions on soil microbial composition

We analyzed the influence of the water/solvent-extracted fractions, whole exudates, and their respective controls on the soil microbial community structure by 454 pyrosequencing analyses. We used principal component analyses on pair-wise and normalized OTUs between all treatments to identify the main factors driving community composition differences. Based on our principle component analyses, we did not observe significant differences between the treatments and controls at phylum level (Figure 0-3A). However, we observed that the controls (nothing added, water, $CHCl_3$, and EtoAc) and treatments (water, $CHCl_3$, EtoAc fractions, and whole exudates) formed two different clusters at the genus level (Figure 0-3B). The second principal component (2.7%) revealed that the controls separated from their respective treatments. This pattern was recapitulated by hierarchical clustering using Ward method where controls clustered separately from the treatments (Figure 0-4). These data clearly indicate that the compounds present in the whole root exudates and extracted fractions have a significant impact on the soil microbial composition.

We also determined the total estimated species richness, evenness, and diversity of the sequencing data of all controls and treatments, and only the $CHCl_3$ control showed significant differences in the evenness compared with other control and treatments (Table 4-5). Furthermore, we performed the analysis of variance and Tukey post-hoc significance test to determine significant differences at the phylum level in pair-wise combinations of all controls

and treatments. Overall, we classified all the OTUs into 21 phyla; among those Proteobacteria showed a higher abundance in all the controls and treatments followed by Actinobacteria (Figure 0-5). However, the EtoAc fraction showed a significantly higher abundance of Proteobacteria, and the water fraction showed a significantly higher abundance of Actinobacteria compared with the other controls and treatments. Interestingly, the $CHCl_3$ control treatment showed an increased abundance of Firmicutes than the other treatments and controls.

OTUs present uniquely for a given treatment

We performed qualitative analyses to identify the OTUs shared by all treatments and controls and also uniquely present in a given treatment or control (Figure 0-6). We identified 138 OTUs shared by all four controls (nothing added, $CHCl_3$, EtoAc, and water). Similarly**,** only 11 OTUs were shared by all four treatments (whole exudates, water, EtoAc, and $CHCl_3$ fractions). We also identified OTUs that were specific to a given treatment or control. The controls (nothing added, $CHCl_3$, water, and EtoAc) cultured seven, five, four, and two unique OTUs (Figure 0-6A). Similarly, the treatments revealed that 3, 6, 25, and 32 OTUs are unique to whole exudates, $CHCl_3$, water, and EtoAc fractions, respectively (Figure 0-6B). Among the four treatments, the EtoAc fraction cultured a higher number (32) of specific OTUs including *Methylobacterium*, *Sphingomonas*, Pseudocardineae, and *Bradyrhizobium*. The water fraction cultured the second highest number of OTUs (25) including Micromonosporineae, *Skermanella*, *Burkholderia*, *Varivorax*, and Frankineae (Supplementary Material 4-1). Similarly, the $CHCl_3$ fraction cultured only six OTUs including species of Propionibacterineae, *Bacillus*, Streptomycineae, *Duganella*, and unclassified.

Taxonomic to phenotype mapping

We assigned the OTUs from taxonomic to phenotype mapping by employing METAGENassist webserver tool (Arndt et al 2012) for nearly 20 phenotype categories classified based on oxygen requirement, metabolism, energy source, habitat, etc. Based on our analyses we observed that only two categories, habitat and metabolism, showed significant differences in the abundance of sequence reads among the 20 phenotype categories analyzed in this study. For instance, the EtoAc fraction significantly enriched the number of sequence reads assigned to symbiotic bacteria (Figure 0-7A) whereas the $CHCl_3$ control enriched the number of sequence reads assigned to free-living bacteria (Figure 0-7B). Interestingly, whole exudates showed the least number of free-living bacteria compared with other controls and treatments. In the metabolism category all the controls significantly reduced the number of sequence reads assigned to carbon fixation compared with the treatments (Figure 0-8A). The $CHCl_3$ control significantly enriched the number of sequence reads assigned to nitrite-reducing bacteria compared with the controls and treatments (Figure 0-8B). On the contrary, $CHCl_3$ control and nothing added significantly reduced the number of sequence reads assigned to atrazine metabolism, while EtoAc and water fractions significantly increased the number of sequence reads compared with other controls and treatments (Figure 0-8C).

Correlation analyses between compounds and soil microbes

We performed correlation analyses to determine the relationship between the abundances of sequences at phyla and OTU levels with each category of compounds detected in the root exudate fractions. Initially, the correlation analyses of these broad groups of compounds (sugars, sugar alcohols, amino acids, and phenolics) at phyla level revealed that the majority of phenolics and related compounds showed positive correlation with the phyla Cyanobacteria and were

negatively correlated with the phyla Actinobacteria, Chlorobi, Fibrobacteres, and Candidate division TM6 (Table 4-6; Supplementary Material 4-2). Other groups of compounds (sugars, sugar alcohols, and amino acids) showed correlation with all 21 phyla at various levels. Interestingly, the majority of compounds related to phenolics, sugars, sugar alcohols, and amino acids did not show correlation with the phyla Proteobacteria and Bacteroidetes. The majority of compounds related to amino acids and phenolics showed negative correlation with the phyla Actinobacteria.

We also determined the number of OTUs that were significantly (positively and negatively) correlated with each group of compounds. We found that phenolics showed the highest number of OTUs (966) followed by amino acids (389), sugars (206), and sugar alcohols (205) (Table 4-7). Furthermore, we calculated the number of OTUs positively correlated with each group of compounds and found that phenolics (742) showed the highest number of OTUs positively correlated followed by amino acids (319), sugar alcohols (166), and sugars (161). In addition, we determined the number of OTUs that did not significantly correlate with any group of compounds and the OTUs that uniquely showed significant positive correlation with each group of compounds. For instance, 20 OTUs significantly positively correlated with phenolics, followed by nine OTUs with amino acids, two OTUs with sugars, and one OTU with sugar alcohols. On the other hand, there are 20 OTUs which are not correlated with any group of compounds analyzed in this study.

Furthermore, we analyzed the correlations at the individual compound level and observed that four compounds (cellobiotol, urea, citramalic acid, and 4-hydroxybenzoate) significantly correlated with a higher number of OTUs (50) including species of *Burkholderia, Variovorax, Pesudomonas, Pseudocardineae,* Frankineae, *Skermanella*, and some unclassified.

Similarly, two sugars (glucose and fructose) were significantly positively correlated with 28 OTUs including species of *Bacillus, Methylobacterium, Sphingomonas,* Propionibacterineae, *Variovorax, Nitrobacter, Pseudomonas* and some unclassified (Supplementary Material 4-3). Based on these analyses, it is clear that one compound can culture many different OTUs and that a mixture of compounds can synergistically culture groups of OTUs. Besides these two scenarios, we also observed that a single compound showed both positive and negative correlation with some OTUs. For example, isoleucine shows a significant positive correlation with one OTU (Propionibacterineae) and significantly negatively correlates with a different OTU (unclassified). Similarly, azelaic acid shows a significant positive correlation with one OTU (unclassified), while it significantly negatively correlates with a different OTU (*Kineosporiineae*). On the other hand, some compounds such as salicylic acid, ferulic acid, and GABA showed significant positive and negative correlation with many OTUs. For instance, salicylic acid positively correlated with three OTUs (Corynebacterineae, Streptomycineae and Pseudonocardineae) and negatively correlated with four OTUs (*Nitrobacter* and three unclassified). Similarly, GABA showed significant positive correlation with 18 OTUs (species of Propionibacterineae, Micromonosporineae, *Methylobacterium, Sphingomonas,* Frankineae, *Variovorax*, and unclassified) and negatively correlated with three OTUs (species of *Bacillus*, Streptomycineae, and unclassified). This suggests that the same compound could act as a positive regulator for some OTUs while also acting as a negative regulator for other OTUs.

Discussion

Plant root-secreted phytochemicals mediate a number of rhizospheric interactions, and these can vary from neutral to beneficial to deleterious interactions (Badri et al 2009b, Bais et al 2004, Doornbos et al 2012, Mercado-Blanco and Bakker 2007, Raaijmakers et al 2009, Walker

et al 2003). The majority of these findings were demonstrated under highly controlled conditions in such a way that a specific compound could attract or deter a specific microbe. However, in nature, plants tend to release an array of compounds that interact with a community of rhizospheric microbes. It has been proposed that root exudates among other factors such as the creation of microenvironments by roots (i.e. by soil fertility or modification of soil structure and chemistry) and plant gentotype characteristics allow the culturing of rhizosphere-specific microbiomes (Berg and Smalla 2009, Dennis et al 2010, Micallef et al 2009a, 2009b). Therefore, in this study we removed all plant/root-associated characteristics to study the effect of different components of root exudates on the soil microbial composition.

We collected root exudates from plants that were 18-21-days old because it has been previously shown that at this time point (that corresponds to vegetative) *Arabidopsis* plants secrete the largest number of phytochemicals (Behera and Wagner 1974, Rudrappa et al 2008, Steinkellner et al 2007, Zhang et al 2009). Furthermore, our previous studies clearly demonstrated that the root exudate profiles collected from individual grown plants at this time point were very consistent and reproducible (Badri et al 2008a, 2009a, 2012a, 2012b). Based on that information, in the present study we pooled the root exudates collected from 600 individual plants for further extraction studies. It is also worth mentioning that *Arabidopsis* secretes different blends of phytochemicals at distinct developmental stages (Chaparro et al 2013). Therefore, it is likely that root exudates collected at other time points when applied to the soil might create different microbial scenarios. Keeping in mind this situation, we artificially separated and extracted the root exudates of 18-21-day-old plants and when applied to the soil observed significant differences between the controls and treatments with regards to carbon fixation, where carbon fixation significantly increased in the treatments (Figure 0-8A). This

observation suggests that the root-secreted phytochemicals impact soil microbial activity (Chen et al 2003) by utilizing the compounds present in the treatments. In addition, we also observed specific OTUs uniquely present in each fraction. Overall, we observed that the controls shared more OTUs (138) than the treatments (11), suggesting that the treatments culture more specific microbes upon the quantitative distribution of compounds present in a given fraction (Figure 0-6). A higher number of unique OTUs were observed in the EtoAc fraction (32) and water fraction (25) compared with the $CHCl_3$ fraction, whole exudates, and controls. This is probably due to the higher percentage of phenolics present in the EtoAc fraction. EtoAc fraction has higher percentage of phenolic-related compounds than sugars, amino acids, and sugar alcohols. However, when compared with the water fraction, the EtoAc fraction had less percentage of amino acids that act as substrates (carbon and nitrogen source) for the majority of microbes. Conversely, the EtoAc fraction (having more phenolics and less sugars) had a significant increase in symbiotic bacteria than other fractions (Figure 0-7A). Symbiotic bacteria usually interact with plants through specific signals such as legume flavonoids that induce *nod* factors in *Rhizobia* (Zhang et al 2009), although we did not identify flavonoids due to the limitation of the GC-MS analyses.

Plant roots attract microbes through the release of cues in which carbohydrates and amino acids predominantly act as chemo-attractants (Somers et al 2004). In contrast, secondary metabolites such as flavonoids act as chemo-attractants to draw *Rhizobia* to the root surface regulating *nod* gene expression (Abdel-Lateif et al 2012), and root-secreted malic acid is involved in recruiting the plant growth-promoting rhizobacteria *Bacillus subtilis* to the rhizosphere upon infection with foliar pathogens (Rudrappa et al 2008). However, for the most part secondary metabolites are considered antimicrobial compounds (Dixon 2001, Wallace

2004). We performed correlation analyses to determine the relationship between the compounds added to the soil and the subsequent microbes being influenced by these compounds. The correlation analyses revealed that the phenolic compounds showed significant correlation with most OTUs (966), among those, 742 OTUs were positively correlated and 224 OTUs were negatively correlated compared with the other groups of compounds (sugars, sugar alcohols, and amino acids). These data clearly suggest that phenolic compounds play a major role in attracting microbes. However, it is unclear if these phenolics accomplish this function by acting as attractants, signaling molecules, or specific substrates.

Furthermore, we analyzed the number of OTUs showing a positive correlation to a particular group of compounds, and this revealed that phenolics significantly correlated with 31 OTUs (species of *Rhizobium, Bacillus, Sphingomonas,* Streptoycineae, Pseudonocardineae, etc.) followed by amino acids with nine OTUs, sugars with two OTUs, and sugar alcohols with one OTU. These data suggest that sugars, amino acids, and sugar alcohols act as general attractants to a broad range of microbes, but the phenolic compounds act as specific substrates or signaling molecules for specific microbe(s). Two sugars (glucose and fructose) positively correlated with 28 OTUs, reinforcing the statement that sugars act as general chemo-attractants. It should be noted that the values of glucose supplied to the soil at a given application ranged from 10 to 95 nmol depending on the fraction. These quantities are comparable (or even lower) to previous studies that applied glucose to the soil to measure its effects on soil microbes (Baudoin et al 2003, Eilers et al 2010).

Root-secreting compounds could act as stimulators for certain microbes while also acting as deterrents for other microbes. For instance, the compound canavanine secreted from the seed coat or roots of leguminous plants acts as an antimicrobial for many rhizosphere bacteria but not

for *Rhizobia*, suggesting that the host plant secretes this compound for selection of the beneficial *Rhizobia* (Cai et al 2009). In the present study we observed a few compounds such as GABA, ferulic acid, salicylic acid, idonic acid, and isoleucine that show a positive correlation with some OTUs while also being negatively correlated with other OTUs. For example, GABA showed a positive correlation with certain OTUs (species of *Sphingomonas, Methylobacterium,* Frankineae, *Variovorax,* Micromonosporineae, *Skermanella,* and unclassified) but negatively correlated with other OTUs (species of *Bacillus* and Streptomycineae). Accordingly, GABA has been reported to have multiple functions. For example, it acts as a carbon and nitrogen source for a wide variety of microbes (Hosie et al 2002), and it specifically reduces *Agrobacterium* quorum sensing ability (Chevrot et al 2006). Interestingly, we did not observe any OTUs assigned to *Agrobacterium* in our studies.

Similarly, salicylic acid showed a positive correlation with certain OTUs (species of Corynebacterineae, Pseudonocardineae and Streptomycineae). Both salicylic acid and GABA showed positive correlation with OTUs belonging to the group of endophytic Actinobacteria (Pseudonocardineae, Corynebacterineae, Streptomycineae, Frankineae and Micromonosporineae), which are known to promote plant growth and reduce the disease symptoms caused by plant pathogens through various mechanisms (Castillo et al 2007, Conn et al 2008, Qin et al 2009). Therefore, the ability of salicylic acid and GABA to attract beneficial microbes should be explored in detail. Similarly, the species of *Variovorax* and *Methylobacterium,* whose presence was correlated with GABA, are reported to produce ACC deaminase, which helps to alleviate the impact of both biotic and abiotic stresses (Saleem et al 2007). It should be noted that the amount of salicylic acid applied in the different treatments (15-62 nmol) in the present study is lower than the usual concentrations of salicylic acid used in

bioactivity studies (Badri et al 2008b, Prithiviraj et al 2005). This suggests that the soil microbiome might be very sensitive to this signaling molecule.

Here, we demonstrated the effect of different groups of natural compounds derived from plant root exudates on soil microbial composition. Based on our analyses, we formulated three scenarios: 1) one compound can culture different OTUs, 2) mixtures of compounds synergistically culture groups of OTUs, and 3) the same compound can act as a stimulator or deterrent to a group or groups of OTUs. Furthermore, our correlation analyses revealed that the phyla Proteobacteria and Bacteroidetes recently determined to be part of the *Arabidopsis* core microbiome (Bulgarelli et al 2012, Lundberg et al 2012) did not show any correlation with the groups of compounds (i.e. sugars, sugar alcohols, amino acids, and phenolics) present in the root exudates. This suggests that the components present in the root exudates are not necessary for culturing these two core microbiome groups. This finding is supported by the fact that these groups of microbes are widely present in a variety of dissimilar soil types, some of them supporting plants while others do not support vegetation (Barberan et al 2011, Janssen 2006, Lin et al 2012, Nemergut et al 2011, Philippot et al 2010, Singh et al 2007).

We have provided a glimpse of how the natural chemical diversity of compounds present in the root exudates (excluding the plant) could affect the soil microbial composition. Further studies are warranted to include additional natural mixes of compounds present in the root exudates at different stages of development so extensive correlations with OTUs could be accomplished. Additionally, this study provides some level of mechanistic understanding of the microbiome by employing a webserver tool, METAGENassist, where microbial functions are assigned based on taxonomic (16S pyrosequencing) data. However, we believe that the only way to provide a functional understanding of the microbiome is by performing metatranscriptomics

analysis. Yet, our taxonomic to phenotypic mapping analysis via the use of METAGENassist does provide a starting point and hints at the potential roles of the soil microbiome. In summary, these studies hold the promise to develop natural mixes of compounds that could influence plant-microbiome interactions to increase crop yield and sustainability.

Tables

Table 4-1. Calibration curve properties of four analytes measured in this study

	Glucose	Serine	Valine	Salicylic acid
slope	7.71E+09	1.19E+10	2.27E+10	8.77E+09
intercept	-3268.71	-359.882	11338.4	-13264.1
R^2	0.97	0.99	0.99	0.98
Lowest standard (M)	5.55E-07	9.52E-07	8.54E-07	7.24E-07
Highest standard (M)	2.78E-05	4.76E-05	4.27E-05	3.62E-05

Table 4-2. Absolute concentrations (nanomoles) of four compounds present in the whole exudates and fractions added to the soil. The values represented are the equivalent of two plant exudates added to the soil per supplemental time. LOD: below the limit of detect

Sample name	Glucose	Serine	Valine	Salicylic acid
Whole exudates	95.9931	253.116	17.3453	40.1557164
Water fraction	236.336	2809.02	1147.67	62.5506512
$CHCl_3$ fraction	53.0792	2.83566	LOD	15.6158104
EtoAc fraction	10.8169	2.84518	LOD	15.6158104

Table 4-3. List of compounds and their relative concentrations (peak areas in %) present in the whole exudates, water fraction, ethylacetate fraction and chloroform fraction analyzed by GC-MS.

	Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
Sugars					
	xylose	0.041284673	0.174157072	0.164066156	0.340677431
	trehalose	0.030752536	0.002101556	0.224478717	0.159986845
	maltose	0.589739639	0.002461411	0.19618472	0.050189753
	levanbiose	0.35271553	0.010634448	0.250079906	0.052170436
	glucose	0.016255733	0.095240777	0.885660641	0.055982316
	glucoheptulose	0.158000064	0.051977518	0.024391377	0.014574835
	fucose	1.311145997	0.351427613	0.251484849	0.253340534
	fructose	0.226582019	1.245979263	24.92349889	0.527833264
	cellobiose	0.14137385	0.022440584	0.098424083	0.043163935
	beta-gentiobiose	0.148707153	0.012325768	0.054909867	0.03703503
	arabinose	0.081254652	0.238243005	0.132767141	0.178560417
	1-kestose	0.004318854	0.000844941	0.045738709	0.03049504
Amino acids					
	valine	0.304772481	2.396114632	1.182298807	0.674142187
	tryptophan	0.098429728	0.925225727	0.072393606	0.079003081
	thymine	0.148467026	0.010955439	0.025601189	0.061289051
	threonine	0.116010447	0.504432382	0.367392671	0.22273338
	serine	0.088652275	2.194165206	0.399316105	0.26059058
	saccharopine	0.010917198	0.013945117	0.041484853	0.029037556
	proline	0.103868602	0.709234742	0.146192155	0.149634975
	phenylalanine	0.068707164	0.093363771	0.283837571	0.136405509
	oxyproline	6.247774967	4.394803492	3.22360336	15.09127023
	ornithine	0.009647955	0.028943892	0.050812116	0.044172962
	O-acetylserine	0.011751639	0.003605751	0.029269652	0.021264311
	N-methylalanine	0.011956604	0.012641001	0.222839617	0.133527914
	methionine	0.023693663	0.128838319	0.117312765	0.048844384
	lysine	0.094285824	0.835359751	0.11325404	0.049404954
	leucine	0.042329225	0.00293642	0.026850027	0.022572309
	isoleucine	0.06129925	0.464119938	1.34351605	0.61595496
	homoserine	0.030192526	0.009337529	0.025210927	0.022460195
	homocystine	0.057805403	0.023245221	0.080901318	0.040959024
	glycine	1.292046476	3.837512665	0.928550438	0.342247029
	glutamine	0.30554689	4.520863727	1.273776226	2.041859674
	cytidine	0.001099438	0.005544652	0.054480579	0.037707715
	cysteine-glycine	0.124563252	0.036487901	0.055104998	0.056393401
	cysteine	0.003429527	0.02382099	0.023805984	0.016480775
	cyano-L-alanine	0.014191499	0.212347809	0.062637055	0.118504622
	citrulline	0.072391397	0.519601008	0.050421854	0.07788194

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
beta-alanine	0.002324085	0.000335385	0.031533172	0.014911178
asparagine	0.381618232	4.679773004	0.419024337	0.348338562
arginine + ornithine	0.105692709	1.266991941	0.108570896	0.060205281
alanine	0.022509323	1.697154482	0.117117634	0.652653648
Sugar alcohols				
xylitol	0.091536371	0.072721024	0.042889797	0.03539069
threitol	0.628500407	0.173620167	0.048392491	0.074668002
ribitol	0.007877877	0.059219248	0.021542464	0.015920205
pentitol	0.362818011	0.136386646	0.031299014	0.033036293
inositol-4-monophosphate	0.002829209	0.001312753	0.024391377	0.014350607
inositol myo-	0.078385136	0.878563994	0.254567919	0.344339825
galactinol	0.846104253	0.057155118	0.19618472	0.294523786
erythritol	0.048332398	0.196224846	0.050812116	0.026047847
cellobiotol	0.06730328	0.006642931	0.147401967	1.347350089
arabitol	0.009102524	0.06372032	0.025835346	0.014686949
2-deoxyerythritol	1.659617284	0.439366203	0.277320196	0.491097205
Phenolic-related				
xylonolactone NIST	0.209576748	0.03915371	0.083359969	0.065100931
xylonic acid	0.008188327	0.028240014	0.030713621	0.034605891
urea	0.60209674	1.891310879	2.103317188	9.832557566
threonic acid	0.402907196	0.403415206	0.042304404	0.109273893
succinic acid	0.0721264	0.613893206	0.596125244	0.631613565
salicylic acid	0.153840723	0.038100053	0.088862663	0.084795644
saccharic acid	0.011597271	0.012491302	0.024196246	0.016144433
pyrazine 2,5-dihydroxy NIST	0.024433769	0.013696098	0.052060954	0.055870202
phthalic acid	5.365674835	1.202481379	0.505857638	0.742120713
phosphoric acid	3.046069648	0.594553136	0.170778663	1.258929422
parabanic acid NIST	0.248376965	0.07183434	0.152904662	0.229871312
oxalic acid	0.698702361	0.09386469	1.229676617	1.19894837
N-acetyl-glutamic acid	0.020749536	4.36528815	0.040665303	0.045069875
myristic acid	0.015326099	0.011154079	0.291955022	0.154082168
methylmaleic acid	0.001749496	0.002434062	0.024391377	0.018349344
methylhexadecanoic acid	0.034951327	0.00806364	0.870206265	0.391876211
methionine sulfoxide	0.046518582	0.27377513	0.090892026	0.0571782
mannonic acid NIST	0.077434919	0.022292324	0.02805984	0.023581336
malic acid	0.003429527	1.553825505	0.230137517	0.459182053
maleimide	0.172861346	0.058678026	0.059788142	0.573089996
lauric acid	0.201828367	0.065117999	1.05421481	0.512922086
lactic acid	0.029992706	0.202195567	1.450682002	1.846594253
kynurenine	0.014706057	0.010458838	0.026850027	0.028140643
itaconic acid	0.100774396	0.082262951	0.455435784	0.946168407
isonicotinic acid	0.574084223	0.09107653	0.090892026	0.071416693

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
isocitric acid	0.002574503	0.066495525	0.021737595	0.01823723
idonic acid NIST	0.000204965	0.012491302	0.036411447	0.022348081
glycolic acid	0.028672866	0.108803008	0.453601553	0.721641201
glycerol-alpha-phosphate	0.002494747	0.010371033	0.031103883	0.067978526
glycerol-3-galactoside	0.178074672	2.738423508	0.261280426	0.108825436
glycerol	0.268511605	1.398193791	1.459228741	1.855600754
glyceric acid	0.051746487	0.217133886	0.148416648	0.297849838
glutamic acid	0.887514136	2.57578181	0.315761005	0.491657775
glucuronic acid	0.031512367	0.015089458	0.044333766	0.016929232
galactonic acid	0.265287902	0.073935896	0.030908752	0.026720531
galactinol	0.846104253	0.057155118	0.19618472	0.294523786
GABA	0.011982332	4.270156769	0.549176723	0.216641846
fumaric acid	0.634098794	2.207193412	0.557684435	0.669358651
ferulic acid	0.005573517	0.001128507	0.023181564	0.023917678
erythronic acid lactone	0.159319047	0.365437504	0.788875659	0.672908932
citric acid	0.03129711	1.54293484	0.149236199	0.194929078
citramalic acid	0.01480125	0.016307209	0.035396766	0.086215756
capric acid	0.031171901	0.011659316	0.669728663	0.297289268
beta-sitosterol	0.015306374	0.00116881	0.259641326	0.04929284
azelaic acid	0.081089136	0.009170556	0.051436535	0.163985582
aspartic acid	1.313655323	3.46528397	0.512140856	0.673805845
aminomalonic acid	0.008077697	0.000695241	0.035396766	0.018349344
alpha ketoglutaric acid	0.037500102	0.021382609	0.068100724	0.033148407
adipic acid	0.034776377	0.038873023	0.260851138	0.336006009
aconitic acid	0.003848891	0.056989584	0.030479464	0.021712767
4-hydroxybutyric acid	0.031382012	0.123428972	0.037191971	0.041071138
4-hydroxybenzoate	0.026218426	0.039221363	0.128513285	0.88962552
3-hydroxypropionic acid	0.021309546	0.193692903	0.365987728	0.312312559
3-hydroxy-3-methylglutaric acid	0.020844729	0.155515121	0.081955025	0.107143725
3-aminoisobutyric acid	0.012936836	0.006343531	0.038011521	0.017564545
2-hydroxyvaleric acid	0.308466147	0.081874307	0.894792773	1.858814692
2-hydroxyglutaric acid	1.026663381	0.381023562	0.181354763	0.211933053
2-5-diketopiperazine NIST	0.003624201	0.001925946	0.027045158	0.018349344
2,3-dihydroxybutanoic acid NIST	0.021684316	0.003854771	0.022166883	0.019470485
Unclassified				
214152	0.311505467	0.369876681	0.052256085	0.024814591
268610	0.712019112	0.034510135	0.150875299	0.267579026
357788	0.918795809	0.082827204	0.023571826	0.01334158
288866	0.019109813	0.752231708	0.022557145	0.02313288
214165	0.311140989	0.660498802	0.055924548	0.046489987
226848	0.844585451	0.075978435	0.025015796	0.018125116

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
213160	0.016760857	0.163748613	0.026030477	0.014686949
213148	0.024013547	0.124425052	0.041055565	0.022908651
216832	1.461218444	0.091604797	0.054480579	0.084347188
373725	0.016600487	0.329521054	0.091906707	0.039837883
352812	1.223668628	0.105525445	0.022986433	0.026608417
201005	0.001924446	1.962041183	0.024586508	0.050413981
385048	0.002749453	0.322775924	0.034967478	0.026384189
223830	0.003958663	0.15494511	0.031533172	0.019806827
371568	0.02537884	0.043233031	0.029074521	0.015471748
362008	0.369340887	0.080666632	0.072978999	0.053291577
213179	0.746090546	0.263956834	0.031923434	0.094699058
367939	4.829931158	0.992692862	0.180340082	0.193807937
213188	0.007328158	0.408648944	0.024820665	0.025375162
217809	0.669859691	0.213260402	0.033952796	0.038716742
267904	0.105913111	0.13909132	0.051826797	0.034830119
199203	19.84754412	3.691949706	0.427571075	3.570685148
228018	0.004014407	0.143811183	0.028450102	0.022908651
238134	0.034236092	0.186138818	0.090267607	0.034269548
327468	0.002384117	3.169514476	0.025601189	0.027131616
385107	0.002244329	0.264975945	0.030713621	0.032812065
357841	0.007517687	0.121197868	0.055924548	0.029598127
362010	0.010507267	0.158997082	0.113839433	0.021712767
269256	0.006793018	0.060419726	0.02743542	0.015359634
362028	0.020740103	0.155672018	0.051436535	0.028140643
321749	0.146263005	0.054556962	0.021347333	0.020591626
367944	0.628500407	0.173620167	0.048392491	0.074668002
267692	0.065478315	0.101269075	0.033133246	0.021152197
362023	0.016660518	0.114334706	0.038831072	0.027804301
307912	0.002209167	0.082395377	0.021737595	0.034269548
219507	1.593873972	3.874756263	0.3374986	0.532504685
207188	0.042004196	0.067511757	0.025015796	0.026047847
200490	0.040184378	0.027685837	0.019513101	0.017041346
328420	0.00459843	0.164333019	0.025015796	0.018349344
226841	0.030641907	0.047608873	0.028684259	0.021824881
200906	0.004014407	0.143811183	0.028450102	0.022908651
218492	0.02770807	0.101269075	0.033133246	0.019246257
224843	0.313825265	0.049062689	0.031728303	0.028252758
330992	3.689461352	0.274942501	0.128279128	0.083113932
299185	0.048622265	0.000106517	0.054909867	0.051385637
362130	0.048622265	1.581630814	0.054909867	0.051385637
268345	0.094630577	0.032427292	0.034343058	0.054300604
352980	0.355344919	0.080647919	0.032118565	0.021488539

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
269249	0.004043565	0.046763932	0.031103883	0.017564545
213141	0.011671882	0.025138061	0.037191971	0.027804301
224632	0.005598387	0.030331494	0.021737595	0.013677922
385023	0.090641898	0.015845154	0.021542464	0.013453694
359697	0.023474119	0.070354615	0.040665303	0.022796537
269294	0.053081765	0.06302364	0.039845753	0.030607154
385021	0.004318854	0.008502664	0.024391377	0.014686949
200610	0.319403927	0.05204805	0.044138635	0.037819829
310063	0.319403927	0.05204805	0.044138635	0.037819829
310897	0.009817759	0.051194472	0.056119679	0.026907388
199177	0.141863537	0.034466953	0.036411447	0.029037556
268353	0.002269199	0.015575982	0.029074521	0.013790036
228911	0.056356066	0.018107925	0.024001115	0.01823723
213154	0.043414084	0.189602067	0.030089202	0.064876702
362086	0.324782769	0.084593374	0.112044228	0.056953972
381469	0.253225813	1.285595028	0.146192155	0.064204018
268565	0.014506237	0.060517607	0.032547853	0.024814591
225327	0.079084934	0.046061495	0.01767887	0.019246257
385058	0.015156295	0.016734717	0.030089202	0.01334158
385112	0.00417392	0.001055096	0.027045158	0.014686949
303838	0.896607227	0.327623896	0.090072476	0.092232547
385028	0.033201831	0.031072796	0.03781639	0.032363609
280930	0.086003161	0.01754799	0.024001115	0.018909914
285340	0.140364459	0.004843654	0.031103883	0.022235967
224589	0.014146047	0.043002724	0.053270767	0.039389427
280945	0.038860249	0.0103552	0.024391377	0.014911178
352849	0.003159384	0.001769049	0.023181564	0.020031056
385024	0.003634492	0.001088203	0.025406058	0.014014265
307669	0.035575657	0.020821235	0.024391377	0.015808091
218748	0.308481584	0.03074029	0.032547853	0.133415799
268583	0.021464771	0.031818417	0.026420739	0.020591626
303956	0.370275666	0.181385848	0.039455491	0.057962999
208897	0.020514555	0.019617878	0.027864709	0.013677922
250380	0.269056179	0.091963213	0.115673664	0.026496303
385034	0.019185281	0.064785492	0.076842593	0.043948734
223191	0.012207022	0.026819305	0.041055565	0.033708978
310871	0.007108614	0.010326411	0.022166883	0.014126379
199205	0.375624493	0.076606023	0.030713621	0.039613655
202899	0.008957591	0.058309534	0.055104998	0.045854674
208686	0.168542492	0.088708681	0.105526852	0.075004344
352777	0.136100492	0.035366591	0.039845753	0.032475723
362109	0.006038334	0.106397734	0.030479464	0.031616181

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
367914	0.0813447	1.667532623	0.150641142	0.329428648
211945	0.033091201	0.011593102	0.061622374	0.074555888
245705	0.228016778	0.264626165	0.041875115	0.120373191
385065	0.040129492	0.004683878	0.034343058	0.027580073
200509	0.291086106	0.009000704	0.857171514	0.034830119
213227	0.01975987	0.029156926	0.021542464	0.038492514
310888	0.006103511	0.051033257	0.102482808	0.039837883
231350	0.097770237	0.02655733	0.024820665	0.018909914
237415	0.024433769	0.013696098	0.052060954	0.055870202
359567	0.024338576	0.032057361	0.03699684	0.031391953
224818	0.024743361	0.005120023	0.02887939	0.012444667
303992	0.038015517	0.005642533	0.027669578	0.014686949
281187	0.12617296	0.038524683	0.041875115	0.055309631
213253	0.066683238	0.306867435	0.316775686	0.121045875
208538	0.043018733	0.114042503	0.302531122	0.198479359
267751	0.049687399	0.022535586	0.065681099	0.041183252
310757	0.26849188	0.058408854	0.026030477	0.037819829
240551	0.043748547	0.03054165	0.059553985	0.027692187
267691	0.022954416	0.024726386	0.038831072	0.021824881
267649	0.056065341	0.009966556	0.042694666	0.171459857
310162	0.22674239	0.067226751	0.046363129	0.07108035
385117	0.006898503	0.032502142	0.145177474	0.096007056
236810	0.008957591	0.048384721	0.062637055	0.045854674
202599	0.026333344	0.025926864	0.10287307	0.045294103
228619	0.060429647	0.016226601	0.027669578	0.015471748
202573	0.030432653	0.848688796	0.155324286	0.022572309
227816	0.007227819	0.005311466	0.020137521	0.017041346
225540	0.001079713	0.009883069	0.024391377	0.013902151
310875	0.003993825	0.021461778	0.025406058	0.032363609
324275	0.00301445	0.102829408	0.03617729	0.027467959
362073	0.003993825	0.021461778	0.025406058	0.032363609
199242	0.017850862	0.003481961	0.026225608	0.017340317
236605	0.254935002	0.170290785	0.027045158	0.038268285
218734	0.510195033	0.131305488	0.189706371	0.173440539
231254	0.037500102	0.021382609	0.068100724	0.033148407
281132	0.00606835	0.006815662	0.023571826	0.017041346
284607	0.003023884	0.005995191	0.025015796	0.020927969
375029	0.073386208	0.019104005	0.031923434	0.183568181
212022	1.437288941	0.116185802	0.163246605	0.258236183
227598	0.066022889	0.008748805	0.041055565	0.022348081
307924	0.003019596	0.005829658	0.024391377	0.012332553
215445	0.955956303	0.268555787	0.046167998	0.105798355

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
267696	0.024918311	0.006273	0.020332652	0.017676659
202572	0.355209419	0.042200966	0.044333766	0.024590363
208397	0.085498037	0.002651415	0.302531122	0.141525387
213143	0.009992709	0.030579075	0.033328377	0.012668895
374786	0.275249737	0.083142437	0.124220403	0.135844939
323686	0.298953692	0.213516619	0.155324286	0.192275711
200710	0.029172845	0.008253644	0.040860434	0.0285891
212761	0.078765051	0.014493537	0.045543578	0.071865149
212016	0.193140919	0.066325673	0.06443226	0.109610235
200557	0.00515844	0.000903957	0.039845753	0.016817118
233408	0.019019765	0.012055157	0.022166883	0.017340317
353047	0.038644993	0.004575922	0.087613825	0.340341089
280546	0.024233949	0.005982237	0.028684259	0.014350607
228680	0.004763946	0.00131995	0.023571826	0.014686949
268365	0.006358217	0.091669571	0.060997955	0.132518887
238267 trisaccharide	0.011037261	0.001224948	0.034343058	0.015920205
269250	0.020224687	0.029578677	0.025835346	0.103817673
235449	0.005963723	0.022266414	0.036801709	0.019246257
215397	1.051327843	0.261394664	1.530997927	1.749690285
296119	0.014621155	0.004393115	0.022557145	0.015135406
237606	0.014111743	0.012698578	0.030284333	0.026159961
224811	0.024258819	0.003041498	0.027240289	0.016705003
227652	0.095365537	0.037391858	0.043084928	0.039613655
214416	0.021284676	0.003255972	0.023805984	0.018461458
237520	0.013216413	0.004060609	0.018927708	0.017564545
224635	0.008227776	0.015191657	0.029464783	0.016032319
385085	0.503076986	0.025653374	0.179520532	0.269260738
310053	0.00913254	0.004935777	0.031923434	0.025711504
200624	0.018675011	0.003181122	0.022166883	0.016705003
212177	0.088017654	0.014772785	0.044919159	0.047835357
205672	0.090537272	0.031235451	0.413131381	0.244782489
205674	0.26297325	0.055512738	0.803510485	2.902559752
237605	0.013606619	0.006950967	0.032547853	0.022908651
232755	0.07668538	0.038089977	0.021152202	0.030382926
240436	0.574084223	0.09107653	0.090892026	0.071416693
310380	0.097045568	0.019245068	0.059358854	0.029037556
227728	0.012507181	0.011601739	0.121995909	0.065549387
213243	0.406216659	0.098347049	0.031923434	0.241306952
296071	0.005153295	0.001283964	0.02969894	0.020479512
385030	0.122633661	0.016667064	0.095965432	0.033148407
374402	0.054820969	0.00988163	0.02661587	0.018573572
211896	0.830468562	0.182429428	0.131947591	0.114169543

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
200615	0.008977315	0.002356333	0.024391377	0.0187978
231248	0.215190571	0.006081557	0.039650622	0.026496303
238938	0.006913082	0.001288282	0.026850027	0.019358371
233790	0.021419319	0.00440463	0.112044228	0.345909423
212261	0.07530122	0.08726638	0.075827912	0.457724569
310581	0.002673984	0.009943525	0.025835346	0.018685686
331031	0.07094206	0.016010688	0.03844081	0.07978788
267654	0.003514429	0.021487687	0.063651736	0.063194991
227822	0.008662578	0.007404385	0.030908752	0.021712767
268671	0.029237164	0.005085477	0.028450102	0.011585125
231098	0.011367436	0.004518345	0.044724028	0.031690924
322204	0.084228795	0.062414765	0.124220403	0.070445037
200702	0.234975312	0.246466421	0.046167998	0.22329395
224322	0.078560085	0.026827942	0.360680164	0.176355506
208714	0.005653273	0.001728745	0.02805984	0.023020765
232946	0.021234936	0.000808955	0.026030477	0.019022029
228164	0.118314807	0.033243444	0.106541533	0.184689322
304391	0.006748423	0.00039872	0.033562534	0.022684423
384918	0.726255207	0.196201815	0.552611029	0.702058602
268506	0.013376783	0.00627156	0.029269652	0.038380399
318770	0.025033228	0.091315474	0.118951865	0.038156171
211921	0.010032158	0.002621187	0.034577215	0.040398454
202083	0.256774546	0.009144646	0.043514216	0.0473869
241312	0.00605377	0.00458024	0.022557145	0.036063374
385104	0.008337549	0.022886805	0.034577215	0.017340317
310006	0.153810707	0.01503332	0.033757665	0.033372636
310367	0.043983528	0.010552401	0.07543765	0.049180726
237333	0.028312676	0.002126026	0.025406058	0.019358371
310448	0.098254778	0.016010688	0.040040884	0.07978788
353091	0.026078638	0.055230611	0.063846867	0.023020765
231576	0.015170874	0.020455622	0.024820665	0.022235967
241168	0.004783671	0.00176617	0.04901691	0.042042794
220122	0.079934812	0.01680093	0.053895186	0.040959024
231260	0.004783671	0.001984963	0.04901691	0.042042794
237392	0.022799191	0.003140818	0.028450102	0.022348081
211916	0.018915138	0.011191504	0.096784982	0.049965525
237652	0.018235065	0.002929223	0.037191971	0.017116089
385042	0.011807382	0.002036782	0.022752276	0.016032319
268712	0.070691642	0.015563027	0.034577215	0.027467959
214201	0.007082886	0.003057332	0.050226723	0.022796537
310193	0.017485526	0.002504594	0.032742984	0.024814591
269776	0.391420554	0.038995373	1.088557869	0.662108605

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
199562	0.0730466	0.006093072	0.128513285	0.027019502
206309	0.075255768	0.01630433	1.11927149	0.068539097
229277	0.005798207	0.000960094	0.151460692	0.071528807
385075	0.078719598	0.009636928	0.024820665	0.019246257
206965	0.018434884	0.021458899	0.667504169	0.033036293
213960	0.001204922	0.001948977	0.029074521	0.028701214
231210	0.017590152	0.003417187	0.031728303	0.016256547
241387	0.082843777	0.020226754	0.927340626	0.068539097
303966	0.054691472	0.004958808	0.009132131	0.142982871
233471	0.003479267	0.00054842	0.025015796	0.016144433
267880	0.003119077	0.000279248	0.021152202	0.014799063
233289	0.015356115	0.00169276	0.037621259	0.022684423
211910	0.201828367	0.057031327	1.219100516	0.606798974
319168	0.276649334	0.115939661	1.346169832	1.345817863
311041	0.016050767	0.002292999	0.027669578	0.020815854
305637	0.013431669	0.001053657	0.048587622	0.024141907
225867	0.023719391	0.002973845	0.022362014	0.011809354
238549	0.02007032	0.009255482	0.027669578	0.016032319
239332	0.287766351	0.098098029	0.546913203	0.127286894
374356	0.020449378	0.011467873	0.03844081	0.146645265
211636	0.435849175	0.04515178	0.878753004	1.705405208
211590	0.606365853	0.053369439	1.810737747	2.863618782
322260	0.009238025	0.006929376	0.032938115	0.031616181
199942	0.017315722	0.016654109	0.549371854	0.295046985
385006	0.745320424	0.003231502	0.057524623	0.236747644
227658	0.001379872	0.000394402	0.035787028	0.050974552
227923	0.002704	0.001853975	0.030089202	0.025263048
238550	0.070087036	0.01040414	0.088862663	0.079899994
226256	0.003184254	0.000303718	0.027045158	0.014014265
271049	0.112941111	0.062810606	1.095075244	0.470057122
200486	0.070087036	0.01040414	0.266353833	0.079899994
236965	0.021739202	0.004252052	0.165900387	0.112263603
299487	0.000825007	0.000175609	0.024196246	0.019582599
385055	0.122843772	0.02282491	0.039650622	0.068426983
218694	0.225817901	0.020382211	0.17913027	0.122727587
203592	0.034456494	0.008027655	0.556084361	0.330325561
199246	0.011606705	0.002673006	0.115673664	0.079563652
200905	0.28442773	0.050153771	0.152904662	0.658894667
373752	0.086897634	0.011777348	0.090462738	0.059308368
309617	0.02906736	0.005000551	0.108961158	0.326887394
200567	0.028632559	0.006628537	0.250079906	0.04682633
308219	0.01046696	0.003939697	0.117507896	0.061849621

Compounds	Whole exudates	Water fraction	$CHCl_3$ fraction	EtoAc fraction
268438	0.009572487	2.87884E-05	0.042694666	0.029934469
219512	0.628700227	0.011007258	0.146582417	5.904265755
211919	0.023564166	0.002854373	0.037621259	0.039277312
299441	0.008158311	0.001223508	0.042889797	0.005904677
385120	0.056890349	0.015748713	0.021971752	0.024478249
224627	0.006753569	0.006319061	0.068920274	0.197059247
206022	0.022629387	0.011505298	0.60994052	0.359923688
224574	0.008918141	0.001458134	0.103497489	0.055085403
303839	0.090081888	0.054220137	0.544493578	0.155875994
238506	0.106082915	0.019612121	0.027864709	0.024478249
234717	0.035810638	0.02434206	1.347379644	0.592112024
201042	0.061903855	0.012459634	1.377273715	0.43268575
294511	0.003408944	0.000601678	0.074618099	0.021264311
267987	0.081939014	0.016904569	0.064237129	0.097016083
268437	0.002509326	0.001669729	0.071144767	0.031279839
235327	0.033476261	0.007780074	0.978972292	0.358914661
384992	0.072946262	0.020090009	1.658886793	1.063215545
213191	0.006718407	0.001068051	0.035591897	0.050301867
357685	0.032201874	0.006352168	0.236459761	0.251770936
241881	0.00073496	0.000152579	0.021347333	0.010463984
288810	0.016501005	0.002875965	0.223034748	0.137078194
368156	0.006808455	0.001653896	0.210624415	0.101911733
357502	0.045263061	0.008656682	0.03844081	0.020367398
200540	0.257934015	0.047136743	0.132376879	0.07108035
224849	0.005663564	0.001707154	0.211053704	0.062970763
216428	0.085528053	0.017316243	3.323237255	1.537047176
327143	0.002099395	0.000418872	0.055690391	0.035502804

Table 4-4. Abundance (%) of different categories of compounds present in the whole exudates and fractions of exudates analyzed by GC-MS. The percentage of the compounds in each category was calculated by dividing the sum of compounds in each category with the sum of compounds in all categories.

Categories	**Water fraction**	**$CHCl_3$ fraction**	**EtoAc fraction**	**Whole exudates**
Sugars	2.20	27.25	1.74	3.10
Amino acids	29.55	10.87	21.47	9.85
Sugar alcohols	2.08	1.12	2.69	3.80
Phenolics	32.01	18.56	30.24	18.55
Unknown	34.13	42.18	43.85	64.68

Table 4-5. Total observed (Sobs) and estimated (Chao and ACE) species richness, evenness, and diversity (Shannon) of the soil samples supplemented with whole exudates and their fractions with respective controls. * indicates significantly different at $p<0.05$.

	Richness			Shannon	
Treatments	**Sobs**	**Chao**	**Ace**	**Evenness**	**Diversity**
Nothing	1227.67	3790.93	7419.23	0.92	6.57
Water control	1140.00	3415.79	6643.4	0.93	6.53
$CHCl_3$ control	987.67	3276.23	6216.58	0.83*	5.70
EtoAc control	1033.33	3082.48	5997.75	0.92	6.39
Water fraction	1288.00	3411.12	4993.73	0.96	6.84
$CHCl_3$ fraction	957.67	2687.32	4466.56	0.92	6.29
EtoAc fraction	849.67	1849.17	2623.46	0.92	6.17
Whole exudates	1009.67	2998.14	5259.03	0.95	6.44

Table 4-6. Pearson correlation analyses of compounds identified by GC-MS with the pyrosequencing data classified at the phyla level. The numbers represented are significant at p value 0.05. + indicates positive correlation. – indicates negative correlation.

	Sugars		**Amino acids**		**Sugar alcohols**		**Phenolics**	
	+	-	+	-	+	-	+	-
Actinobacteria	0	0	2	8	0	3	1	9
Bacteriodetes	0	0	1	0	0	0	1	2
Proteobacteria	0	1	0	0	0	0	0	1
Acidobacteria	0	0	0	0	0	0	1	0
BD1–5	0	0	0	0	0	0	0	0
Candidate division BRC1	0	2	0	1	0	0	0	1
Candidate division OP10	0	0	0	1	0	0	1	0
Candidate division TM6	0	0	0	7	0	3	1	10
Candidate division TM7	2	1	1	2	0	0	1	1
Candidate division WSB	0	0	0	0	0	0	0	0
Chlorobi	0	0	0	7	0	3	1	10
Chloroflexi	0	1	0	0	0	0	0	1
Cyanobacteria	2	0	5	0	0	0	8	0
Fibrobacteres	0	4	0	1	1	3	3	7
Firmicutes	0	2	0	0	0	0	0	0
Gemmatimonadetes	0	0	0	0	0	0	0	0
Nitrospirae	0	0	0	0	0	0	0	0
Planctomycetes	0	1	0	0	0	0	0	1
Verrucomicrobia	0	0	0	2	0	0	1	0
WCHB1-60	0	0	1	0	0	0	0	0
Unclassified	1	0	0	0	0	0	0	0

Table 4-7. Pearson correlation analyses of the groups of compounds with OTUs at the genus level. The numbers represented are significant at p value 0.05

Compounds group	Total number of OTUs correlated	Number of OTUs positively correlated	Number of OTUs negatively correlated
Sugars	206	161	45
Amino acids	389	319	70
Sugar alcohols	205	166	39
Phenolics	966	742	224

Figures

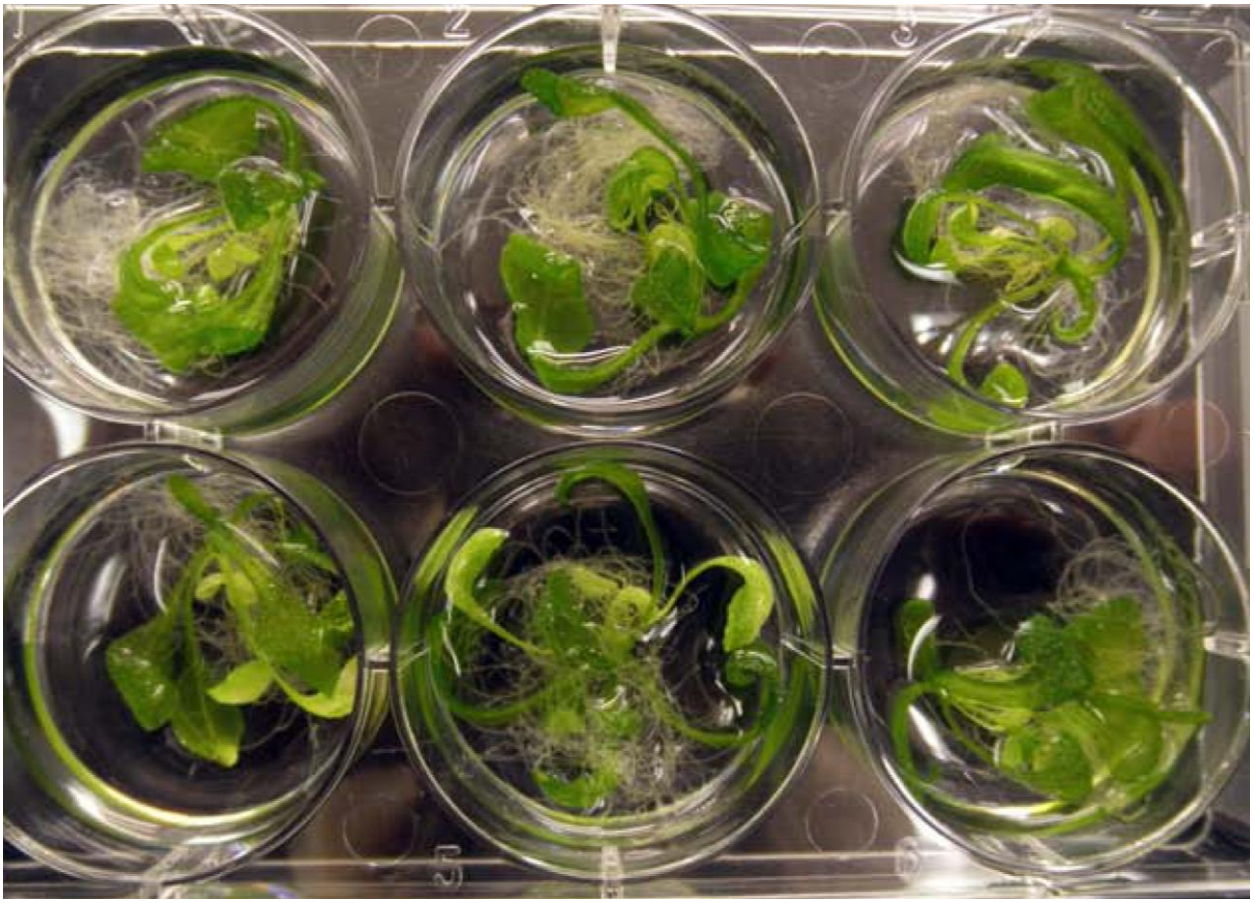

Figure 0-1. Picture showing the *Arabidopsis* plants were transferred into sterile distilled water for three days prior to exudate collection

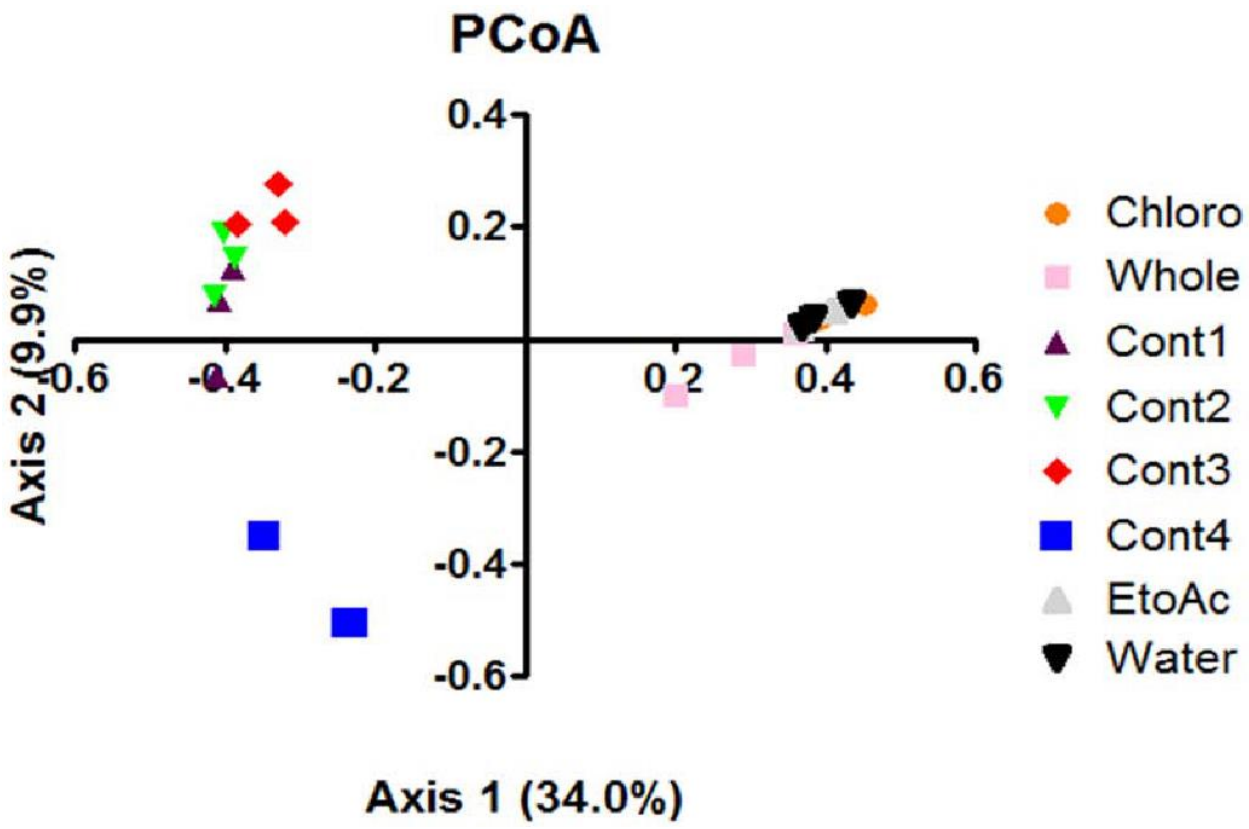

Figure 0-2. Principal Coordinate Analyses (PCoA) of soil microbiome sequence data at genus level. Whole: whole exudates; Chloro: CHCl3 fraction; EtoAc: ethylacetate fraction, Water: water fraction; Cont 1: nothing added in the soil; Cont 2: water control; Cont 3: ethylacetate control; Cont 4: chloroform control.

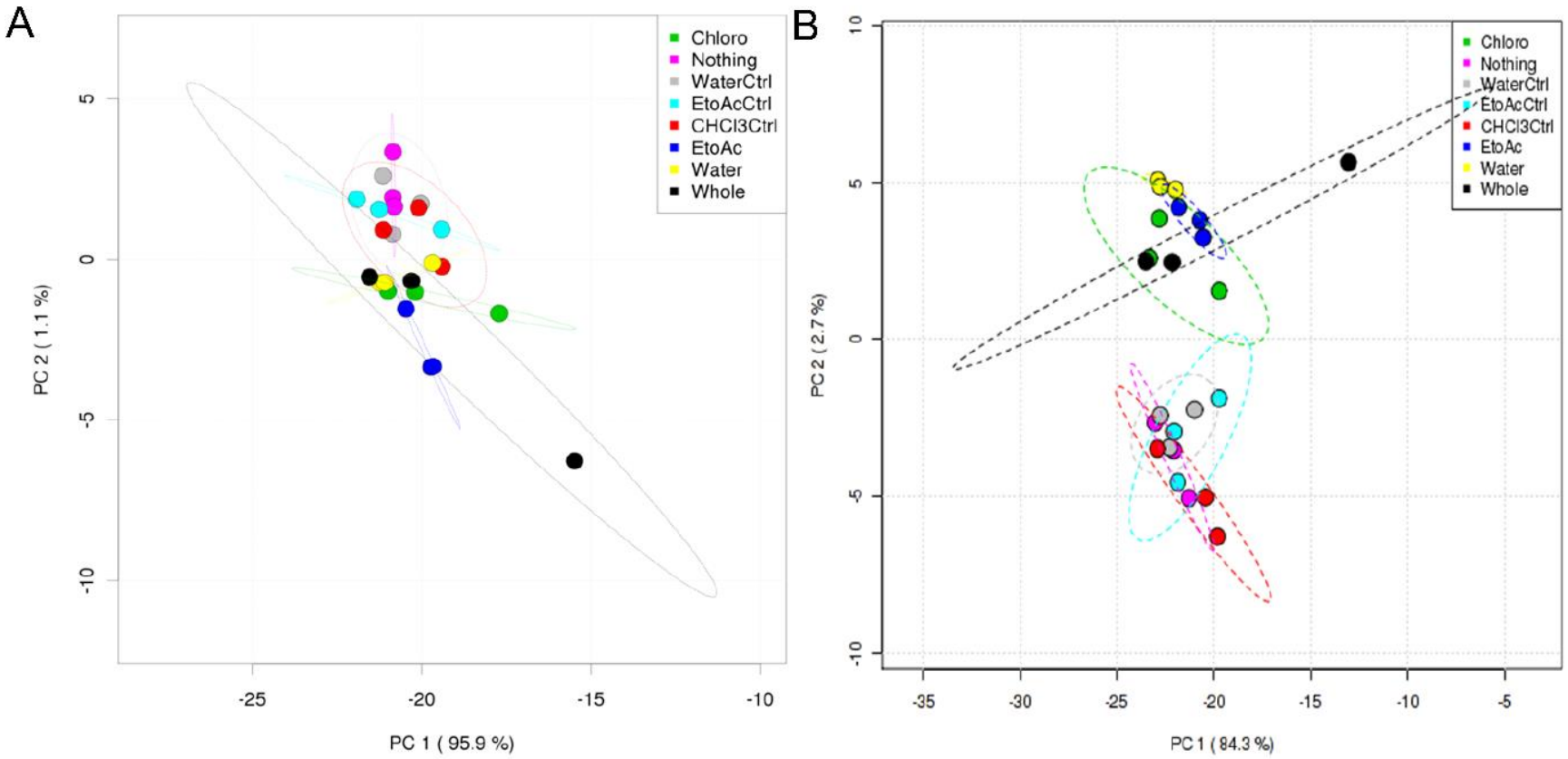

Figure 0-3. Soil microbiome sequencing data of treatments and controls analyzed by Principal Component Analyses (PCA) at phyla level (A) and genus level (B). Whole: whole exudates; Chloro: $CHCl_3$ fraction; EtoAc: ethylacetate fraction, Water: water fraction; Nothing: nothing added in the soil; Water ctrl: water control; EtoAc ctrl: ethylacetate control; $CHCl_3$ Ctrl: chloroform control.

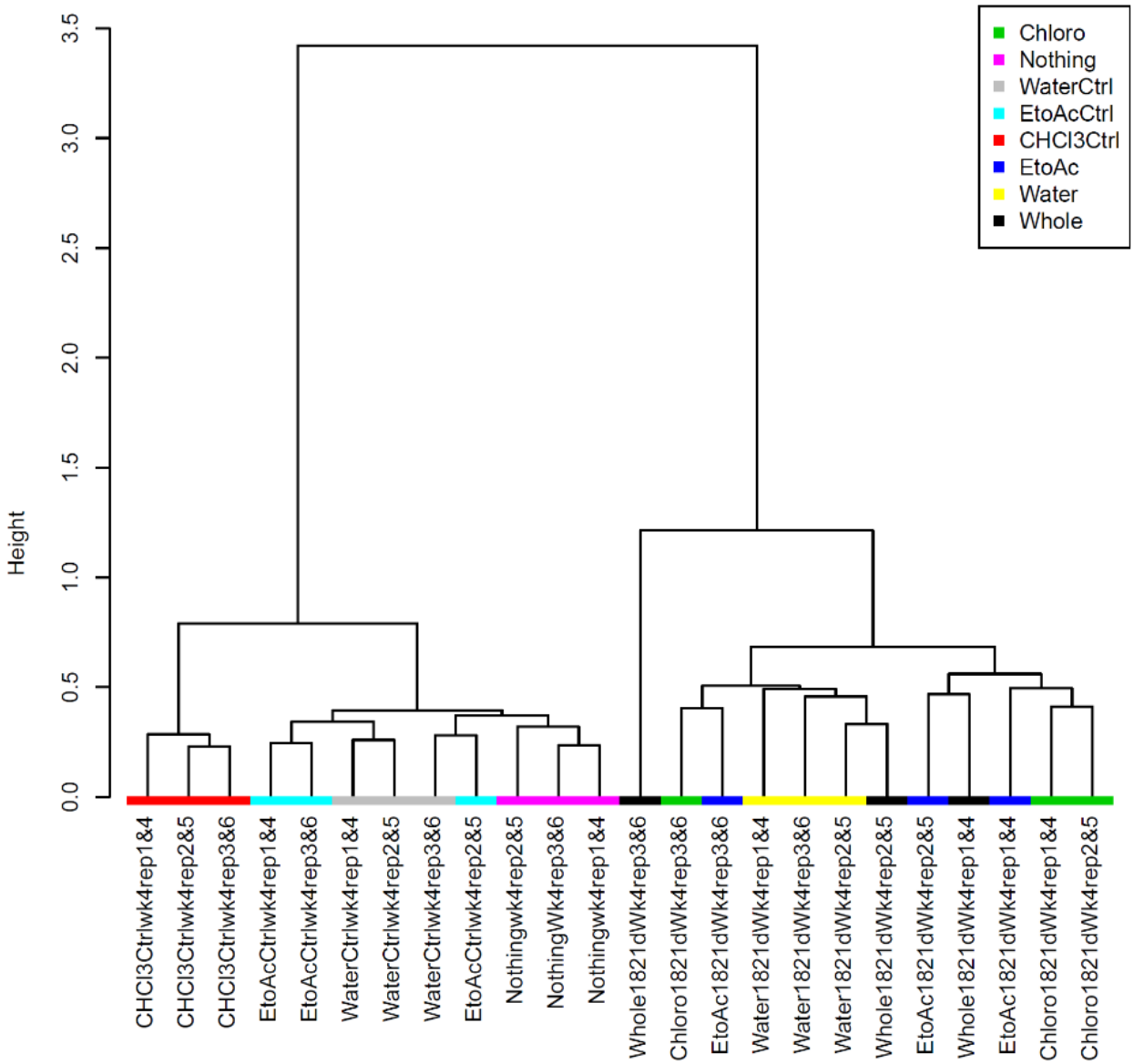

Figure 0-4. Cluster analysis of the soil microbiome sequencing data of controls and treatments by Ward method. Whole: whole exudates; Chloro: $CHCl_3$ fraction; EtoAc: ethylacetate fraction, Water: water fraction; Nothing: nothing added in the soil; Water ctrl: water control; EtoAc ctrl: ethylacetate control; $CHCl_3$ Ctrl: chloroform control

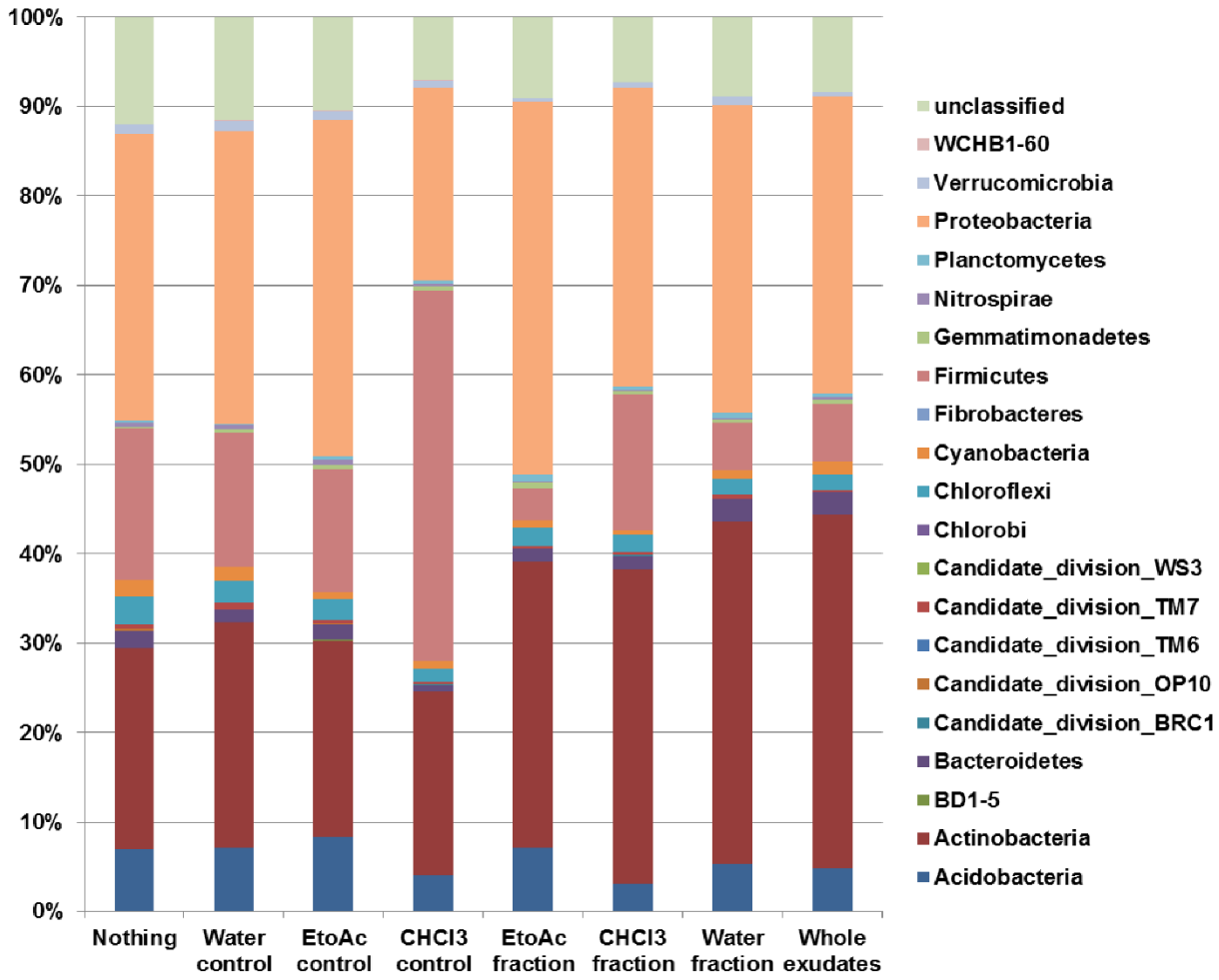

Figure 0-5. Relative abundance (%) of the major bacterial phyla present in the treatments and controls revealed by pyrosequencing.

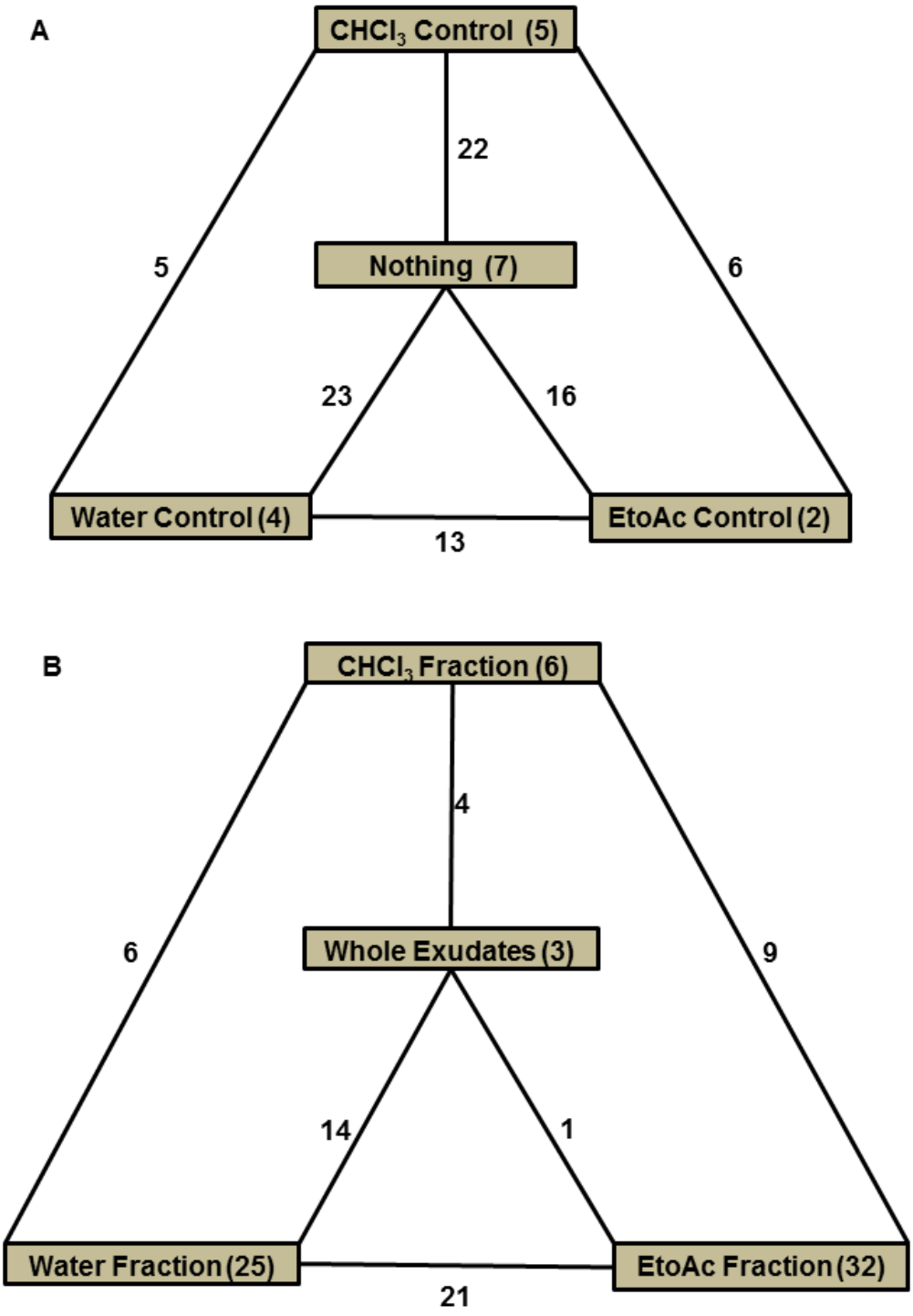

Figure 0-6. Flow diagram indicating the shared and unique OTUs present in controls and treatments. (A) Controls and (B) Treatments. Overall, 138 OTUs were shared by controls and 11 OTUs were shared by treatments. The number of OTUs unique to a particular control or treatment are represented inside the shaded box and the number of OTUs shared between the controls and treatments are represented in the intersections.

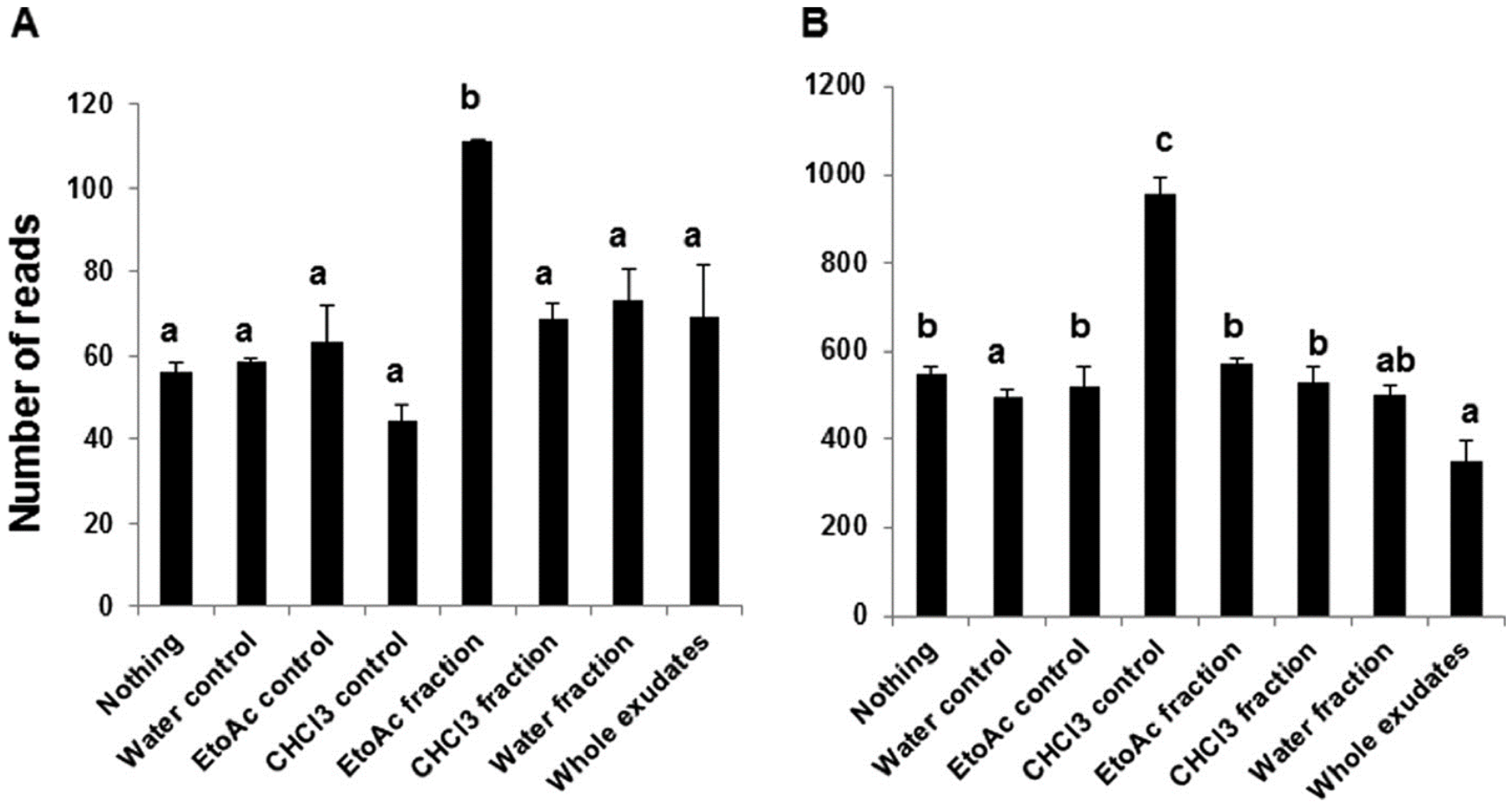

Figure 0-7. Taxonomic to phenotypic mapping based on the biotic habitat. Graph illustrates the number of sequence reads present in the controls and treatments. (A) symbiotic bacteria and (B) free-living bacteria. The bars with different letters are significantly different (p value<0.05) from one another.

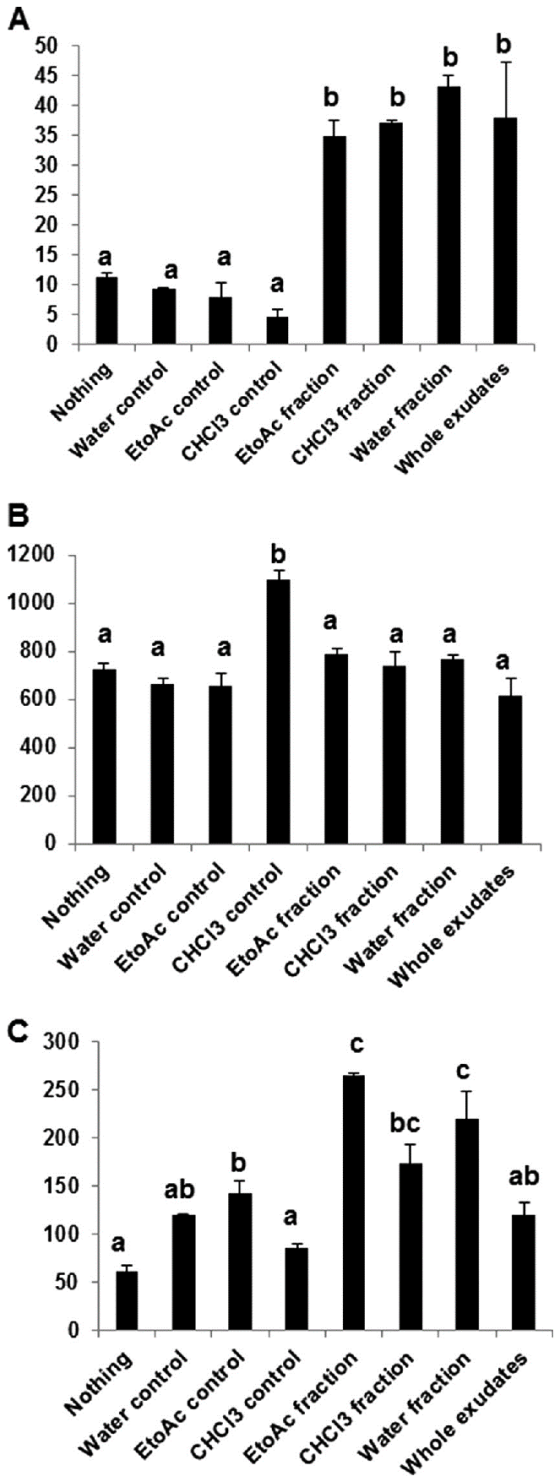

Figure 0-8. Taxonomic to phenotypic mapping based on the metabolism of specific microbial groups. The graph illustrates the number of sequence reads present in the controls and treatments. Carbon fixation (A), nitrite reducing (B), and atrazine degradation (C) are shown. The bars with different letters are significantly different (p value < 0.05) from one another.

REFERENCES

Abdel-Lateif K, Bogusz D, Hocher V (2012). The role of flavonoids in the establishment of plant roots endosymbioses with arbuscular mycorrhiza fungi, rhizobia and Frankia bacteria. *Plant Signal Behav* **7:** 636-641.

Akiyama K, Matsuzaki K, Hayashi H (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature* **435:** 824-827.

Arndt D, Xia J, Liu Y, Zhou Y, Guo AC, Cruz JA *et al* (2012). METAGENassist: a comprehensive web server for comparative metagenomics. *Nucleic Acids Res* **40:** W88-95.

Badri DV, Loyola-Vargas VM, Broeckling CD, De-la-Pena C, Jasinski M, Santelia D *et al* (2008a). Altered profile of secondary metabolites in the root exudates of Arabidopsis ATP-binding cassette transporter mutants. *Plant Physiol* **146:** 762-771.

Badri DV, Loyola-Vargas VM, Du J, Stermitz FR, Broeckling CD, Iglesias-Andreu L *et al* (2008b). Transcriptome analysis of Arabidopsis roots treated with signaling compounds: a focus on signal transduction, metabolic regulation and secretion. *New Phytol* **179:** 209-223.

Badri DV, Quintana N, El Kassis EG, Kim HK, Choi YH, Sugiyama A *et al* (2009a). An ABC transporter mutation alters root exudation of phytochemicals that provoke an overhaul of natural soil microbiota. *Plant Physiol* **151:** 2006-2017.

Badri DV, Vivanco JM (2009). Regulation and function of root exudates. *Plant Cell Environ* **32:** 666-681.

Badri DV, Weir TL, van der Lelie D, Vivanco JM (2009b). Rhizosphere chemical dialogues: plant-microbe interactions. *Curr Opin Biotechnol* **20:** 642-650.

Badri DV, Chaparro JM, Manter DK, Martinoia E, Vivanco JM (2012a). Influence of ATP-Binding Cassette Transporters in Root Exudation of Phytoalexins, Signals, and in Disease Resistance. *Front Plant Sci* **3:** 149.

Badri DV, De-la-Pena C, Lei Z, Manter DK, Chaparro JM, Guimaraes RL *et al* (2012b). Root Secreted Metabolites and Proteins Are Involved in the Early Events of Plant-Plant Recognition Prior to Competition. *PLoS ONE* **7:** e46640.

Bais HP, Park SW, Weir TL, Callaway RM, Vivanco JM (2004). How plants communicate using the underground information superhighway. *Trends Plant Sci* **9:** 26-32.

Bakker MG, Manter DK, Sheflin AM, Weir TL, Vivanco JM (2012). Harnessing the rhizosphere microbiome through plant breeding and agricultural management. *Plant and Soil* **360:** 1-13.

Barberan A, Bates ST, Casamayor EO, Fierer N (2011). Using network analysis to explore co-occurrence patterns in soil microbial communities. *Isme J* **6:** 343-351.

Baudoin E, Benizri E, Guckert A (2003). Impact of artificial root exudates on the bacterial community structure in bulk soil and maize rhizosphere. *Soil Biol and Biochem* **35:** 1183.

Behera B, Wagner GH (1974). Microbial Growth Rate In Glucose-amended Soil. *Soil Sci Soc Am J* **38:** 591-594.

Berendsen RL, Pieterse CM, Bakker PA (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478-486.

Berg G, Smalla K (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* **68:** 1-13.

Broeckling CD, Broz AK, Bergelson J, Manter DK, Vivanco JM (2008). Root exudates regulate soil fungal community composition and diversity. *Appl Environ Microbiol* **74:** 738-744.

Broughton WJ, Zhang F, Perret X, Staehelin C (2003). Signals exchanged between legumes and Rhizobium: agricultural uses and perspectives. *Plant Soil* **252:** 129-137.

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.

Cai T, Cai W, Zhang J, Zheng H, Tsou AM, Xiao L *et al* (2009). Host legume-exuded antimetabolites optimize the symbiotic rhizosphere. *Mol Microbiol* **73:** 507-517.

Castillo UF, Browne L, Strobel G, Hess WM, Ezra S, Pacheco G *et al* (2007). Biologically active endophytic streptomycetes from Nothofagus spp. and other plants in Patagonia. *Microb Ecol* **53:** 12-19.

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

Chen G, Zhu H, Zhang Y (2003). Soil microbial activities and carbon and nitrogen fixation. *Res Microbiol* **154:** 393-398.

Chevrot R, Rosen R, Haudecoeur E, Cirou A, Shelp BJ, Ron E *et al* (2006). GABA controls the level of quorum-sensing signal in Agrobacterium tumefaciens. *Proc Natl Acad Sci U S A* **103:** 7460-7464.

Conn VM, Walker AR, Franco CM (2008). Endophytic actinobacteria induce defense pathways in Arabidopsis thaliana. *Mol Plant Microbe Interact* **21:** 208-218.

de Weert S, Vermeiren H, Mulders IHM, Kuiper I, Hendrickx N, Bloemberg GV *et al* (2002). Flagella-Driven Chemotaxis Towards Exudate Components Is an Important Trait for Tomato Root Colonization by Pseudomonas fluorescens. *Mol Plant Microbe Interact* **15:** 1173-1180.

Dennis PG, Miller AJ, Hirsch PR (2010). Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *FEMS Microbiol Ecol* **72:** 313-327.

Dixon RA (2001). Natural products and plant disease resistance. *Nature* **411:** 843-847.

Doornbos RF, van Loon LC, Bakker P (2012). Impact of root exudates and plant defense signaling on bacterial communities in the rhizosphere. A review. *Agron Sustain Dev* **32:** 227-243.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

Eilers KG, Lauber CL, Knight R, Fierer N (2010). Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. *Soil Biol and Biochem* **42:** 896.

Fiehn O, Wohlgemuth G, Scholz M (2005). Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. In: Ludascher B, Raschid L (eds). *Data Integration In The Life Sciences, Lecture Notes in Computer Science*. Springer-Verlag: Berlin/Heidelberg. pp 224-239.

Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee do Y, Lu Y *et al* (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J* **53:** 691-704.

Gupta Sood S (2003). Chemotactic response of plant-growth-promoting bacteria towards roots of vesicular-arbuscular mycorrhizal tomato plants. *FEMS Microbiol Ecol* **45:** 219-227.

Hartmann A, Schmid M, van Tuinen D, Berg G (2009). Plant-driven selection of microbes. *Plant Soil* **321:** 235-257.

Hosie AHF, Allaway D, Galloway CS, Dunsby HA, Poole PS (2002). Rhizobium leguminosarum has a second general amino acid permease with unusually broad substrate specificity and high similarity to branched-chain amino acid transporters (Bra/LIV) of the ABC family. *J Bacteriol* **184:** 4071-4080.

Janssen PH (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol* **72:** 1719-1728.

Li X, Bergelson J, Chapple C (2010). The ARABIDOPSIS Accession Pna-10 Is a Naturally Occurring sng1 Deletion Mutant. *Mol Plant* **3:** 91-100.

Lin Y-T, Whitman WB, Coleman DC, Chiu C-Y (2012). Comparison of soil bacterial communities between coastal and inland forests in a subtropical area. *App Soil Ecol* **60:** 49.

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.

Mercado-Blanco J, Bakker PA (2007). Interactions between plants and beneficial Pseudomonas spp.: exploiting bacterial traits for crop protection. *Antonie Van Leeuwenhoek* **92:** 367-389.

Micallef SA, Channer S, Shiaris MP, Colon-Carmona A (2009a). Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. *Plant Signal Behav* **4:** 777-780.

Micallef SA, Shiaris MP, Colon-Carmona A (2009b). Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. *J Exp Bot* **60:** 1729-1742.

Neal AL, Ahmad S, Gordon-Weeks R, Ton J (2012). Benzoxazinoids in Root Exudates of Maize Attract Pseudomonas putida to the Rhizosphere. *PLoS One* **7:** e35498.

Nemergut DR, Costello EK, Hamady M, Lozupone C, Jiang L, Schmidt SK *et al* (2011). Global patterns in the biogeography of bacterial taxa. *Environ Microbiol* **13:** 135-144.

Omrane S, Chiurazzi M (2009). A variety of regulatory mechanisms are involved in the nitrogen-dependent modulation of the nodule organogenesis program in legume roots. *Plant Signal Behav* **4:** 1066-1068.

Philippot L, Andersson SG, Battin TJ, Prosser JI, Schimel JP, Whitman WB *et al* (2010). The ecological coherence of high bacterial taxonomic ranks. *Nat Rev Microbiol* **8:** 523-529.

Prithiviraj B, Bais HP, Weir T, Suresh B, Najarro EH, Dayakar BV *et al* (2005). Down regulation of virulence factors of Pseudomonas aeruginosa by salicylic acid attenuates its virulence on Arabidopsis thaliana and Caenorhabditis elegans. *Infect Immun* **73:** 5319-5328.

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J *et al* (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35:** 7188-7196.

Qin S, Li J, Chen HH, Zhao GZ, Zhu WY, Jiang CL *et al* (2009). Isolation, diversity, and antimicrobial activity of rare actinobacteria from medicinal plants of tropical rain forests in Xishuangbanna, China. *Appl Environ Microbiol* **75:** 6176-6186.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12:** 38.

Raaijmakers JM, Paulitz TC, Steinberg C, Alabouvette C, Moenne-Loccoz Y (2009). The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant Soil* **321:** 341-361.

Rudrappa T, Czymmek KJ, Pare PW, Bais HP (2008). Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol* **148:** 1547-1556.

Saleem M, Arshad M, Hussain S, Bhatti AS (2007). Perspective of plant growth promoting rhizobacteria (PGPR) containing ACC deaminase in stress agriculture. *J Ind Microbiol Biotechnol* **34:** 635-648.

Sana TR, Fischer S, Wohlgemuth G, Katrekar A, Jung KH, Ronald PC *et al* (2010). Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen Xanthomonas oryzae pv. oryzae. *Metabolomics* **6:** 451-465.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75:** 7537-7541.

Singh BK, Munro S, Potts JM, Millard P (2007). Influence of grass species and soil type on rhizosphere microbial community structure in grassland soils. *App Soil Ecol* **36:** 147-155.

Somers E, Vanderleyden J, Srinivasan M (2004). Rhizosphere Bacterial Signalling: A Love Parade Beneath Our Feet. *Crit Rev Microbiol* **30:** 205-240.

Speed TP (2003). *Statistical analysis of gene expression microarray data*. CRC: Boca Raton, FL.

Steinkellner S, Lendzemo V, Langer I, Schweiger P, Khaosaad T, Toussaint J-P *et al* (2007). Flavonoids and Strigolactones in Root Exudates as Signals in Symbiotic and Pathogenic Plant-Fungus Interactions. *Molecules* **12:** 1290-1306.

Walker TS, Bais HP, Grotewold E, Vivanco JM (2003). Root exudation and rhizosphere biology. *Plant Physiol* **132:** 44-51.

Wallace RJ (2004). Antimicrobial properties of plant secondary metabolites. *P NUTR SOC* **63:** 621-629.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *App Environ Microb* **73:** 5261-5267.

Yoneyama K, Xie X, Sekimoto H, Takeuchi Y, Ogasawara S, Akiyama K *et al* (2008). Strigolactones, host recognition signals for root parasitic plants and arbuscular mycorrhizal fungi, from Fabaceae plants. *New Phytol* **179:** 484-494.

Zhang J, Subramanian S, Stacey G, Yu O (2009). Flavones and flavonols play distinct critical roles during nodulation of Medicago truncatula by Sinorhizobium meliloti. *Plant J* **57:** 171-183

CHAPTER 5 AN EXPERIMENTAL PIPELINE FOR THE DEVELOPMENT OF SOIL PREBIOTICS OF AGRICULTURAL IMPORTANCE

Synopsis

We report the development of an experimental pipeline to discover natural products that could act as promoters of beneficial microbes for agricultural purposes. We first developed nine chemical libraries by collecting root exudates of *Arabidopsis thaliana* at distinct developmental time points and by additional chemical extraction. These chemical libraries were then supplemented to a natural soil for a period of 2 or 6 weeks, and the 16S ribosomal DNA of bacteria was pyrosequenced. These data sets allowed us to develop 5 distinct models in order to develop highly interactive analyses to correlate compounds with microbial abundances at the Order taxonomic level. To validate and determine which model adequately predicts the dynamics of the soil microbial community twelve compounds from our chemical library were selected and added as distinct artificial blends to the soil for a period of 5 weeks. The 16S ribosomal RNA of bacteria derived from those soils was pyrosequenced in order to identify the active microbial community influenced. Out of 5 statistical models tested, those that implemented machine learning (random forest and boosted decision trees) were able to reliably and precisely model soil microbial community dynamics as measured by the mean square predicted error. Boosted decision trees were then used as a predictive tool to identify compounds that could reliably be used as soil prebiotics.

Introduction

Natural products and synthetic chemicals have traditionally been used as biocides to deter the growth of harmful microorganisms or malignant cells. Chemical biocides are used in distinct

areas ranging from medicine, agriculture, and industry. While the purpose of a microbial biocide is to kill harmful organisms it is usually not target-specific and could be harmful to the beneficial flora of a particular biome (Blaser 2011, Dethlefsen et al 2008, Dethlefsen and Relman 2011, Sjolund et al 2003). Accordingly, biocides have come under scrutiny due to their hazards to humans and the environment (Blaser 2011, Cotter et al 2013). Thus, there is a need to develop pipelines that will allow the screening and identification of chemicals as inducers of beneficial microbes as compared to biocides. For this purpose, the use of a system in which an organism selects beneficial microbes by natural means will be desirable.

Roots of plants secrete compounds into the soil that are essential to attract beneficial microbes involved in defense against pathogens (Rudrappa et al 2008) or in nutrient acquisition (Dimkpa et al 2009). For example, when *Arabidopsis* is under attack by *Pseudomonas syringae* pv tomato they release malic acid to enlist beneficial microbes (Rudrappa et al 2008). Similarly, when nitrogen is limiting the roots of *Medicago truncatula* release flavones and flavonols to directly attract the N-fixer *Sinorhizobium meliloti* (Zhang et al 2009) or they release volatile signals that attract *C. elegans* to the roots which then transfer rhizobium to the plant (Horiuchi et al 2005). Under phosphorus-limiting conditions in the soil certain plants release organic acids to increase phosphorus availability (Dakora and Phillips 2002) and initiate associations with mycorrhizae (Yoneyama et al 2007a, 2007b). Thus, the use of root exudates as a reservoir to identify such chemicals is warranted in order to explore how these chemicals might favor the growth of soil beneficial microbes in the absence of the plant and likewise deter the proliferation of harmful ones.

Arabidopsis thaliana has been found to secrete different blends of compounds at distinct developmental stages (Chaparro et al 2013a) that correlate with specific microbial groups

(Chaparro et al 2014). An initial pilot study tested four chemical libraries that were derived from root exudates taken from one *Arabidopsis thaliana* developmental time point and found that these chemical libraries differentially shaped the soil microbial communities when compared to the controls (Badri et al 2013a). Furthermore, these chemical libraries seemed to culture and attract different types of bacteria as analyzed by 16S ribosomal DNA pyrosequencing. For example, the soils supplemented with two of these chemical libraries were significantly more abundant in bacteria capable of atrazine degradation (Badri et al 2013a). Additionally, one of these libraries significantly enriched symbiotic bacteria in the soil (Badri et al 2013a). In this study, we tested nine distinct chemical libraries that were characterized by GC-MS in order to augment the predictive ability for linking metabolites with microbial OTUs. The libraries were created by collecting root exudates of *Arabidopsis thaliana* at distinct developmental time points (seedling, vegetative, and bolting) and these were further separated by additional chemical extraction. These chemical libraries were then supplemented to a natural soil for a period of 2 or 6 weeks. From this experimental set up, soil microbial DNA was isolated, the 16S rRNA gene was amplified and using next-generation sequencing we identified the soil microbial taxa that were influenced by the addition of these nine distinct natural chemical libraries to the soil. The availability of a large matrix of natural compounds present in the nine chemical libraries and soil microbial OTUs, allowed us to test 5 statistical modeling approaches (partial least squares regression, principle components regression, LASSO, random forest, and boosted decision trees) in order to develop highly interactive analyses to correlate compounds with microbial groups. To validate and determine which of the 5 models more adequately predicted the dynamics of the soil microbial community upon exposure to specific chemical compounds 12 compounds were selected based on results from Badri et al. (2013a). These compounds were then grouped into

chemical classes (amino acids, sugars, or secondary metabolites) and added to the soil along with a mixture of all 12 compounds for a period of 5 weeks. The soil microbial community was then pyrosequenced in order to determine which model most adequately simulates and predicts soil microbial community dynamics.

Materials and methods

Plant growth conditions and collection of root exudates

Arabidopsis thaliana wild type (Col-0) seeds were placed on MS agar media supplemented with 3% sucrose after surface-sterilization for one minute with Clorox® followed by four rinses with sterile distilled water and incubated in a growth chamber (Percival Scientific) at 25°C, with a photoperiod of 16h light/8h dark. Root exudate collection followed the methodology described previously (Badri et al 2008, 2009, 2012, 2013a). Briefly, seven-day-old seedlings were transferred to 6-well culture plates (one plant per well) containing 5 ml of liquid MS (full strength MS salts supplemented with 1% sucrose) media per well, incubated on an orbital shaker at 90 rpm and illuminated under cool white fluorescent light (45 µmol m^{-2} s^{-1}) with a photoperiod of 16h light/8h dark at 25±2°C. Root exudates were collected at three developmental stages: seedling (between 7-10 days), vegetative (between 18-21 days), and bolting (between 25-28 days). For this purpose, we transferred the plants when they were 7, 18, and 25 days old to new 6-well plates containing 5 ml sterile distilled water, after gently washing plants with sterile distilled water to remove the surface adhering exudates, and incubated them for three days on an orbital shaker under the same conditions described above. Sterile distilled water was used to collect the root exudates instead of MS liquid media to prevent the interference of exogenously supplemented salts and sucrose in ensuing GC-MS analyses. No visible toxicity symptoms were observed after the three day period of incubation in sterile

distilled water, consistent with previous results (Badri et al 2013a). The exudates contained in the media were collected from three developmental stages and were filtered using nylon filters of pore size 0.45 µm (Millipore, Billerica, MA, USA) to remove root sheathing and root-border-like cells. For each developmental time point, we collected a total of three liters of root exudates by growing 600 individual *Arabidopsis* plants in 6-well plates and pooling them for further partition analyses. Each three liters of pooled root exudates was collected from 600 individual plants or 600 biological replicates, as discussed in Badri et al. (2013a).

Partition of root exudates

Root exudates collected from 600 individual plants for each developmental time point (1800 plants total) were filtered. Exudates from 300 plants were freeze dried and dissolved in 50 ml of water served as whole exudates. The remaining 300 plants per time point were freeze dried and dissolved in 25 ml of sterile distilled water. The pH was adjusted to two with 1N HCl. The solutions were partitioned with 25 ml of ethyl acetate (EtoAc); the organic phase (EtoAc) was separated and dried under nitrogen gas. The aqueous phase constituted the water phase. We performed the partition two times and then pooled the extractions together. All fractions (EtoAc and water) were collected independently and dried under nitrogen gas. The root exudate partitioning procedure and whole exudates resulted in 9 different chemical libraries (see Table 5-1).

Gas chromatography-mass spectrometry (GC-MS) of the chemical libraries and data analyses

The nine fractions were subjected to GC-MS analyses at the Genome Center Core Services, University of California Davis to identify the compounds present in each library and their relative concentrations. We followed the methods described previously, (Badri et al 2013a, Chaparro et al 2013a, Sana et al 2010). Briefly, samples were derivatized using methoximation

and trimethylsilylation. GC-MS was done on an Agilent 6890 gas chromatograph (Santa Clara, CA) containing a 30 m long, 0.25 mm i.d. rtx5Sil-MS column with an additional 10 m integrated guard column. The instrument was controlled by Leco ChromaTOF software version 2.32 (St. Joseph, MI). Data was processed using the BinBase database (Fiehn et al 2005). The unique ion was used in order to report relative peak height. Retention index, mass spectrum unique ions, apex ions, peak purity, and signal/noise ratio were used to assign the metabolites by the BinBase identifier numbers. All entries in BinBase were matched against the Fiehn mass spectral library (http://fiehnlab.ucdavis.edu/Metabolite-Library). Data normalization was performed by using "total metabolite content" (Fiehn et al 2008); see Table 5-2 for compounds present in each of the 9 chemical libraries and their relative concentrations.

Adding chemical libraries to Arabidopsis co-adapted soil

Arabidopsis co-adapted soil (referred here as natural soil) was collected on July, 2011 from the Michigan Extension Station, Benton Harbor, MI (N42° 05' 34'', W86° 21' 19'' W, elevation 630 feet) where *Arabidopsis* genotype (Pna-10) has grown naturally in a fallow field for more than eight years (Li et al 2010). Soil from this location has been extensively characterized (Badri et al 2009, 2013a, 2013b, Chaparro et al 2013a, 2013b, Zolla et al 2013). Top soil (0-10 cm) was collected from three spots at this location within *Arabidopsis* patches, transported to the laboratory in air tight coolers and stored in a cold room (4°C) until use. All soils collected from the three spots were dried under room temperature, pooled, and thoroughly homogenized by hand before starting the experiment. Cubical pots (2.0L X 1.0W X 2.0D inches) were bottom lined with Whatmann 3MM filter paper to prevent soil loss and filled with soil, incubated in a growth chamber under the photoperiod of 16h light/8h dark at 25±2°C for two weeks and sufficiently watered before supplementing them with the chemical libraries. The

Arabidopsis seedlings that emerged from the existing seed bank present in the natural soil were continuously removed during this two-week period. After completely removing the seedlings, the nine chemical libraries (Table 5-1) were independently supplemented to each of the pots in the absence of any plants. Overall, there were three treatments per developmental time point (when the exudates were collected) along with controls: pots supplemented with whole exudates, EtoAc fraction, water fraction, and three controls (only the solvents: 2% methanol (EtOAc Control), sterile water, and negative control (pots receiving no solvents or water)). For each treatment nine pots (considered as nine biological replicates) were maintained and each pot received 2 ml of the respective chemical library (representing root exudation of two plants) or control. These distinct chemical libraries or controls were added twice a week with an interval of three days over a period of two or six weeks. None of the pots received additional supplementation of water during the experimental period. The soil samples were collected after the second or sixth week of supplementation and stored at -80°C for future extraction of soil DNA and pyrosequencing.

Soil DNA extraction and pyrosequencing

Total DNA was extracted from soil by using a Mo Bio ultraclean soil DNA kit (Mo Bio, Carlsbad, CA) according to the manufacturer's instructions. The DNA was quantified using a Nanodrop (Nanodrop Technologies, Wilmington, DE) spectrophotometer, samples showing an absorbance ratio (A_{260}/A_{280}) between 1.7 and 2.0 were subjected to PCR analyses. The DNA samples that did not show an absorbance ratio between 1.7 and 2.0 were re-extracted. PCR amplification was performed by using the primer pairs 515F (GTGCCAGCMGCCGCGGTAA) and 806R (GGACTACVSGGGTATCTAAT) (Walters et al 2011) with 454 adaptor (454 Life Sciences, Branford, CT, USA) and a ten-base bar-code sequence, not shown here, to amplify the

variable region (V4) of the 16S ribosomal small subunit using the following PCR conditions: the reaction mix (50 µl) contained 0.4 µmoles of each primer, 200 µmoles of dNTPs, 1 X reaction buffer and one unit of Taq DNA polymerase (Takara). PCR included 30 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for one minute in an Applied Biosystems thermal cycler (GeneAmp PCR system 2700). For each chemical library applied to the soil, nine biological replicates were subjected to PCR amplification and three biological replicates were derived from the nine biological replicates by pooling three samples together for pyrosequencing analyses. This was done to improve the coverage of the microbiome and to reduce the variability between the replicates within a given treatment. After pooling, the PCR products were purified using AMPure XP beads (Agencourt) before running the products on an agarose gel electrophoresis. The specific amplicon product (approximately 400 bp) was eluted from the gel by using QIAquick gel extraction kit (Qiagen) according to the manufacturer's instructions. The purified amplicon products were subjected to unidirectional pyrosequencing using a 454 GS FLX Titanium sequencing platform (Roche, Indianapolis, IN). Pyrosequencing was performed under contract from the Roy J. Carver Biotechnology Center, University of Illinois, Urbana-Champaign.

Sequencing analyses

454 pyrosequencing data analysis was performed using Mothur, version 1.32 (Schloss et al 2009). Sequencing reads that had flow lengths between 360 and 720 were retained for analyses and separated by barcode. Additional criteria were used to retain sequences and these were: 1 mismatch to the barcode, up to 2 mismatches for the primer and homopolymer length was capped at 8. Sequences were then de-noised with Mothur's execution of the PyroNoise algorithm (Quince et al 2011). Reads that passed the above quality criteria, had the barcode, and primers

removed, and those reads that again had up to 1 mismatch of barcode or up to 2 mismatches in the primer were removed. Additionally homopolymer length was limited to 8. All retained sequences were aligned to the SILVA bacterial reference database (Pruesse et al 2007). Reads were screened to begin at the same position and those that did not end at which 95% of the reads ended were removed. Chimeras were identified using the UCHIME method within Mothur and excluded from further processing (Edgar et al 2011). Sequences were classified against the ribosomal database using the naïve Baysian classifier (Wang method) within Mothur (Wang et al 2007) with a minimum confidence of 60% for each assignment. Sampling effort was equalized to the depth of the smallest sample (500 reads). Operational taxonomic units (OTUs) were defined at 97% sequence similarity using the furthest neighbor algorithm (Supplementary Material 5-1). Principle coordinate analysis (PCoA) was done using weighted UniFrac distances (Lozupone et al 2007) obtained in Mothur.

Statistical analyses

Relative abundance of each compound within each library was used. Additionally, sequencing data was normalized, in each instance, to the water control using a $\log_2$ ratio.

Statistical modeling

The R project for statistical learning (R Core Team 2014) was used for statistical model implementation. The following statistical packages glmnet (Friedman et al 2010), pls (Mevik and Wehrens 2007), tree (Ripley 2014), randomForest (Liaw and Wiener 2002), and gbm (Ridgeway 2013) were used in order to implement and test the performance of the 5 statistical models, 3 based on linear regression (partial least squares regression, principle components regression, and LASSO) and 2 based on machine learning (random forest and boosted decision trees).

Adding libraries of compounds to Arabidopsis co-adapted soil

In order to validate and determine which statistical model most accurately modeled the behavior of the soil microbial community we selected twelve compounds (4 amino acids, 4 sugars, and 4 secondary metabolites; see Table 5-3) from the original libraries and previously identified in Badri et al. (2013a) to generate four new chemical blends for additional testing. Specifically, we created a mixture of amino acids, sugars, or secondary metabolites (each mixture had 4 compounds) along with a mixture containing all twelve compounds (referred here as artificial blends), and a water control that were added to *Arabidopsis* co-adapted soil in the same manner as described above with a few modifications. Briefly, microcosms containing 5g of *Arabidopsis* co-adapted soil (8 reps per treatment) were supplemented with 2 ml of each mixture twice a week for a period of 5 weeks. The concentration of each compound in the mixture was 0.05 nmoles/ compound/ day. This value was determined based on the results of Badri et al. (2013a) and Larsen et al. (1998). Soil was collected after five weeks of addition (eight reps per treatment) and stored in -80°C for future soil RNA extraction.

Soil RNA extraction

In order to determine what active microbes are directly influenced by the artificial blends, soil RNA extraction was performed as previously published (Chaparro et al 2013a). The Mo Bio RNA PowerSoil® Total RNA isolation kit was used. RNA quality was determined using agarose gel electrophoresis and concentration was determined using a Nanodrop. All samples had an absorbance ratio (A_{260}/A_{280}) between 1.9 and 2.0. One microgram of total RNA underwent a DNAse treatment using DNAse I (Fisher Scientific). Total RNA was subsequently transcribed using Superscript III RT and random primers (Invitrogen, CA) at 45°C for one hour according to the manufacturer's instructions. PCR was performed using the same 16S rRNA universal primer

pair (515F and 806R) (Walters et al 2011) with 454 adaptor (454 Life Sciences, Branford, CT, USA) and a ten-base bar-code sequence as described above except that 2 PCR replicates were performed for each sample and later combined. PCR clean up including AMPure bead purification and gel extraction were carried out as described above. A total of 4 replicates per treatment were obtained for pyrosequencing analysis by combining two of the biological replicates along with their technical replicates together. Pyrosequencing, as described above, was performed under contract from the Roy J. Carver Biotechnology Center, University of Illinois, Urbana-Champaign. Sequences were analyzed as described above using Mothur, version 1.32 (Schloss et al 2009). Operational taxonomic units (OTUs) were defined at 97% sequence similarity using the furthest neighbor algorithm (Supplementary Material 5-2). Principle coordinate analysis (PCoA) was done using weighted UniFrac distances, as described above.

Results

Development of distinct chemical libraries

The 9 chemical libraries obtained through plant development and subsequent fractionation produced distinct changes with respect to the classes of compounds (amino acids, secondary metabolites, sugars, and sugar alcohols) found in each library (Table 5-1). For example, libraries 7 and 9 contained the highest percentage of secondary metabolite compounds (44.33% and 51.80% respectively; Table 5-1). Libraries 4 and 6 were made up mostly of amino acids (39.92% and 38.82%; Table 5-1). Additionally, library 4 contained a high percentage of sugar compounds (31.01%; Table 5-1). Chemical library 3 had the highest abundance of sugar alcohols (6.33%; Table 5-1). Interestingly, library 5 had amino acids, secondary metabolites, and unknown classes of compounds at relatively equal abundance (29.55%, 32.02% and 34.14%; Table 5-1). Libraries 1 and 8 were made up of secondary metabolites and unknown compounds

at the same relative abundance (48.48% and 43.85%; Table 5-1). On the other hand, chemical library 2 was composed mostly of unknown compounds (64.68%; Table 5-1). A closer examination of the compounds in each chemical library revealed that urea accounted for over 7% of library 1 and over 24% of library 7; fructose made up over 26% of chemical library 4, whereas oxoproline was most abundant in libraries 8 and 9 (Table 5-2). Glycerol-3-galactosidase accounted for about 10% of fraction 6 while asparagine was most abundant in library 5 (Table 5-2). On the other hand, libraries 2 and 3 were more abundant in unknowns (19.85% and 8.21% respectively; Table 5-2). These data suggest that each chemical library was unique.

Addition of specific chemical libraries to the soil creates distinct microbial communities

Principle coordinate analysis (PCoA) (Figure 5-1) revealed that the distinct chemical libraries significantly altered soil microbial community composition. The first two axes of PCoA analysis explained a total of 69.9 % of the variability. PCoA 1 explained 43.4% of this variability while PCoA 2 explained 26.5% of the overall variability (Figure 5-1). After the addition of these chemical libraries for 2 weeks, libraries 1, 4, and 6-9 were dissimilar from the controls (nothing added, water, and EtOAc control) (Figure 5-1A). Similarly, the addition of these natural chemical blends for 6 weeks also produced distinct microbial communities (Figure 5-1B). The microbial communities generated by libraries 3, 5, 6, 7, and 9 after 6 weeks of addition were different from the controls (Figure 5-1B). At 6 weeks, libraries 3, 5, and 6 formed distinct microbial communities when compared to each other and the other chemical libraries (Figure 5-1B). It should be noted that at 2 weeks the microbial communities were separated from the controls along PCoA 2; and at 6 weeks the separation between the microbial communities and the controls were along PCoA 1 (Figure 5-1A and B). Overall these data indicate that at both 2 weeks and 6 weeks of chemical library addition distinct changes at various levels of influence

occurred between all the microbial communities. Furthermore, many of the microbial communities established were distinct from that of the solvents alone.

Interestingly, the microbial communities were also influenced by the duration of exposure to a given library (Figure 5-1). Specifically, libraries 3 and 6 produced the most change in microbial communities when comparing 2 (Figure 5-1A) and 6 weeks (Figure 5-1B) of amendment duration. On the other hand, the soil microbial community established after exposure to library 9 did not significantly change irrespective to the duration of exposure. Additionally, for most libraries this prolonged exposure appeared to reduce the variability of the microbial communities established at 6 weeks when compared to 2 weeks for a given treatment.

We analyzed the observed species richness (S_{obs}), Chao estimate of total species richness (S_{chao}), Shannon's diversity (H'), and evenness (E_H) within the soil communities after exposure to a chemical library or control after 2 or 6 weeks. Since there were no significant differences in soil microbial community structure (S_{obs}, S_{chao}, H', and E_H, ANOVA Dunnett post-hoc, $p<0.05$; Table 5-4) between the controls (Nothing, EtoAc, and water control) within time points all treatments were compared to the water control. Within 2 weeks of exposure we observed that S_{obs} and H' for the soil microbial communities exposed to the libraries were not significantly different from the control (water control; Table 5-5). S_{chao} of the soil microbial community exposed to library 8 was significantly different from the control (Table 5-5). Additionally, the E_H of the soil microbial community exposed to library 3 and 6 were significantly lower when compared to the control (ANOVA Dunnett post-hoc $p<0.05$; Table 5-5). Investigating how S_{obs}, S_{chao}, H' and E_H of the soil microbial community were influenced after exposure of the chemical libraries to 6 weeks we observed that E_H was significantly altered when compared to the control (Table 5-5). For example, evenness of soil microbial communities exposed to libraries 2, 3, 5, 6,

8 and 9 for 6 weeks were significantly lower than the control (ANOVA Dunnett post-hoc $p<0.05$; Table 5-5). H' of soil microbial communities exposed to library 1, 3, 6 and 7 were significantly lower than the control at 6 weeks of exposure (Table 5-5). At 6 weeks both S_{chao} and S_{obs} of the microbial community exposed to library 8 for 6 weeks were significantly higher than the control (ANOVA Dunnett post-hoc $p<0.05$; Table 5-5). Interestingly, community evenness significantly decreased with time in microbial communities exposed to libraries 2 and 8 (ANOVA Dunnett post-hoc $p<0.05$; Table 5-5). These observations further indicate that the libraries added to the soil cause changes within soil microbial community structure.

Modeling soil microbial community dynamics

The large matrix of natural compounds present in the nine chemical libraries and soil microbial OTUs, allowed us to test 5 statistical modeling approaches (partial least squares regression, principle components regression, LASSO, random forest, and boosted decision trees) in order to develop highly interactive analyses to link compounds with microbial groups. To control for length of exposure, the water control at each time point was used to normalize the soil microbial communities at each time point since significant differences in microbial community structure were observed in water control communities with time (S_{obs}, S_{chao}, and H', t-test, $p<0.05$; Table 5-4). To evaluate the performance of each model implemented artificial blends derived from the chemical libraries were developed.

Artificial chemical blends added to the soil influence the soil microbiome

To validate and determine which statistical model accurately predicted the behavior of the soil microbial community; soil was exposed to 12 chemical compounds identified previously (Badri et al 2013a) and in our chemical libraries (Table 5-3). These 12 compounds were classified in to 3 categories (amino acids, secondary metabolites, and sugars). Solutions of these

artificial chemical blends were added to the soil along with a mixture of all 12 compounds to determine how the soil microbial community would be influenced by these artificial blends (see Materials and Methods). Analyses of the soil microbial communities exposed to these different classes of compounds revealed that, although subtle, some differences could be detected. PCoA of the weighted UniFrac distance matrix revealed that our control separated from our treatments (mixture, amino acids, secondary metabolites, and sugars) (Figure 5-2). This separation between the treatments and control is due to PCoA 1 which explained 22.28% of the variability observed (Figure 5-2). Furthermore, the solutions containing the three classes of compounds were separated by PCoA 2 which explained 11.88% of the observed variability between the samples (Figure 5-2). Additionally, the solution with all 12 compounds produced a microbial community that appeared to be a transition state between the control and treatments (amino acids, secondary metabolites, and sugar) (Figure 5-2).

Community richness, diversity, and evenness as measured via S_{obs}, S_{chao}, H', and E_H revealed that richness as measured by S_{obs} and S_{chao}, were significantly higher in the amino acid treatment than the control (ANOVA Dunnett post-hoc $p<0.05$; Table 5-5); the evenness (E_H) of the community decreased compared to the control, although this change was not significant. S_{obs} was significantly higher in the sugar treatment compared to the control (ANOVA Dunnett post-hoc $p<0.05$; Table 5-5). While no other significant differences were observed with respect to community structure some trends were noted. For example, our blend of secondary metabolites had the lowest measured richness (S_{obs} and S_{chao}) and diversity (H') when compared to all other soil microbial communities (Table 5-5). The soil microbial community exposed to the sugar blend had the highest H' when compared to all other treatments. Interestingly, the soil microbial community exposed to all 12 compounds had S_{obs}, S_{chao}, H', and E_H measurements that fell

between the other treatments (amino acids, sugars, and secondary metabolites exposed soil microbial communities; Table 5-5). This may indicate that the mixture of compounds nullifies the changes that each chemical class individually produces on the soil microbial community.

Machine learning models outperforms linear models when predicting the dynamics of the soil microbial community

The effectiveness of the three linear based models (partial least squares regression, principle components regression, and LASSO) and two statistical models based on machine leaning (random forest and boosted decision trees) on modeling soil microbial community dynamics was tested. In order to determine each model's goodness of fit the mean predicted squared error (MPSE) was calculated for each bacterial order (Table 5-7). MPSE is a measure of predictive power of a model and how inaccurate the predicted abundance is with respect to the true observed abundance. Our data revealed that the models based on machine learning (random forest and boosted decision trees) outperform the models based on linear models, as determined by MSPE (Table 5-8). Machine learning was able to more accurately predict the dynamics of 75% of the bacterial orders to a better degree (i.e., lower MSPE), as compared to the models based on linear regression (Table 5-7). Eight microbial orders had the lowest MSPE when implementing LASSO, while partial least squares regression and principle components regression had the lowest MSE for only 2 orders (*Bdellovibrionales* and *Methylophilales*) when compared to the other models implemented, yet in each of these cases the machine learning models had comparable results. Overall, boosted decision trees outperformed random forest, as they were able to more accurately predict 16 vs. 14 orders; despite, similar overall average MSPEs (1.96 vs. 1.97; Table 5-8). Thus, boosted decision trees was the model chosen for continued analysis.

Our analysis revealed that the predictive power of forecasting soil microbial community abundance after exposure to the artificial blends (Table 5-3) was dependent on chemical class (Table 5-9 and Table 5-10). For example, the blend composed of secondary metabolites had the lowest MSPE (1.52; Table 5-10). Indicative of the fact that secondary metabolites successfully predict soil microbial community dynamics when compared to the other classes of compounds tested (sugars, amino acids, and a mixture of these; see Table 5-10). Additionally, the model based on secondary metabolites accurately predicted the abundance of 16 microbial orders such as Bacillales, Opitutales, Rhodocyclales, etc when compared to the other artificial blends, as measured by lower MSPEs (Table 5-9). Artificial blends consisting of sugars had the second lowest average MSPE value (1.95; Table 5-10) and was able to accurately predict the changes in 6 microbial orders such as Actinomycetales, Methylophilales, Planctomycetales, Xanthomonadales, etc (Table 5-9). A model derived from the addition of the mixture of all the compounds tested was able to most accurately predict 10 microbial orders such as Acidomicrobiales, Clostridiales, Pseudomonadales, Rhizobiales, etc (Table 5-9). Additionally, the artificial blend consisting of amino acid compounds was able to accurately predict the dynamics of 8 microbial orders (Verrucomicrobiales, Nitrospirales, Burkholderiales, etc; Table 5-9). These data in combination begin to shape our understanding of the soil microbial community and how distinct compounds or mixtures of compounds influence specific microbial orders.

Boosted decision tree modeling identifies potential soil probiotics

Boosted decision tree modeling of the natural compounds present in the nine chemical libraries and soil microbial OTUs provided a measure of importance for each chemical compound (Table 5-11). This measure of importance for each predictor variable, in this case

chemical compound, is based on the how many times a variable is used for separating or splitting the data (Elith et al 2008). Furthermore, a variable with a higher relative importance indicates a strong influence on the response since the relative contribution of each variable is scaled so that the sum of all variables combined is 100 (Elith et al 2008). This enabled the identification of the relative importance of each compound in modeling soil microbial community dynamics (Table 5-11). Seven microbial orders of agricultural importance, ranging from plant growth promotion activity to plant pathogens were selected, and the compounds most responsible for the changes in abundance were identified (Table 5-11). Partial dependence plots were then used to visualize the effect of the important predictor variables (chemical compound) on the response variable (bacterial order) after accounting for the average effect of all other variables in the model (Elith et al 2008). Interestingly, the amino acid methionine was predicted as being very important in determining the abundance of all the bacterial orders selected (Table 5-11). For example, as the relative abundance of methionine increases our model predicts that orders such as Acidomicrobiales, Actinomycetales, Bacillales, Nitrospirales, and Xanthomodales also increase in abundance (Figure 5-3). On the other hand, Pseudomonadales and Rhizobiales appear to behave in a more dynamic matter by which they initially decrease in abundance and then once methionine reaches a specific relative concentration they begin to increase their abundance (Figure 5-3). Increasing the relative abundance of 4-hydroxybutyric acid decreases the abundance of 6 bacterial orders (Figure 5-4). Interestingly, Rhizobiales is predicted to increase with supplementation of 4-hydroxybutyric acid but this only occurs until a specific relative concentration has been reached at which point additional supplementation decreases the abundance of Rhizobiales (Figure 5-4F). Similar phenomena are predicted with Acidomicrobiales. Increasing the abundance of 3-hydroxy-3-methylglutaric acid increases the

abundance of Actinomycetales, Bacillales, and Pseudomonadales, while at low concentrations of 3-hydroxy-3-methylglutaric acid the abundance of Xanthomonadales decreases (Figure 5-5).

Discussion

The importance of the microbial communities associated with all life forms has recently been brought to light (Consortium 2012, Kristin and Miranda 2013, Morrissey et al 2004). Humans benefit from our microbial gut (Bajzer and Seeley 2006, Turnbaugh et al 2007, Zhao 2010) as do plants from their rhizosphere microbiome (Dimkpa et al 2009, Kirankumar et al 2008, Mendes et al 2011, Ramos Solano et al 2008, Rudrappa et al 2008, Selvakumar et al 2012, Yang et al 2009, Zhang et al 2010). Understanding how host microbial communities interact and communicate with their host is of utmost importance in order to employ targeted approaches for selecting and controlling host associated microbial communities. Host associated microbial communities can provide many benefits but they can also be detrimental, and an unbalanced microbial community can lead to disease (Reid et al 2011). Relatively few strategies have been adopted to modulate associated microbial communities in a manner that benefits the human host (Rastall et al 2005). For example, prebiotics such as inulin and fructoologosaccharides have been used to selectively promote the growth of the beneficial bacteria Bifidobacteria in humans (Fukuda et al 2011, Schrezenmeir and de Vrese 2001). This is in contrast to the nonspecific use of antibiotics that result in an overhaul of the gut microbial community which is not always able to fully recover (Modi et al 2014, Reid et al 2011, Willner et al 2011). Similarly the use of pesticides influences both the beneficial and detrimental soil microbial community (Berlec 2012) and subsequent reestablishment of the soil microbial community does not guarantee a community replete with beneficial microbes.

There is mounting evidence identifying the microbial members in the soil that could aid plants in dealing with stress and help in proper development (Chaparro et al 2014, Dimkpa et al 2009, Mendes et al 2011, Palaniyandi et al 2013, Saravanakumar and Samiyappan 2007, Selvakumar et al 2012). Here we created and analyzed 9 chemical libraries derived from root exudates during three important developmental time points (seedling, vegetative, and bolting) that have previously been shown to be distinct with respect to their rhizosphere microbial community (Chaparro et al 2014). We then performed an in depth analysis on how these distinct libraries influenced the soil microbiome after exposing the soil to these libraries. This approach helped in the identification of potential prebiotics that could be used selectively to enhance the growth of beneficial microbes or to deter the growth of pathogenic ones. In order to identify such candidates we tested and validated several statistical modeling approaches in their ability to predict the impact of chemical profiles and thus compounds applied to the soil on the soil microbial community.

Modeling the complex dynamics and multifaceted components that occur in a biological ecosystem, such as soil, is difficult and challenging. Multiple factors and mechanisms need to be taken into account. A successful model requires the ability to incorporate enough complexity to enable robust inferences while simultaneously limiting this complexity for ease of interpretability (Merow et al 2014, Olden et al 2008). Here we selected between 5 distinct modeling approaches to help interpret and tease apart soil microbial community dynamics upon the addition of chemicals. We establish enough complexity within the model selected to accurately predict and forecast soil microbial community abundance. Our criteria, model training, and validation identified machine learning based models as suitable means of ascribing relationships and inference to the data. The linear based models (pcr and pca) implemented had

large error rates (as measured by MSPE; see Table 5-7and Table 5-8) that tend to produce "underfit" models that were inappropriate for our study (Merow et al 2014). The high MSPE found with linear models highlight the complexity of soil microbial community dynamics and indicate that linear relationships are insufficient in explaining soil microbial community abundances at finer taxonomic levels. On the other hand, machine learning models have the ability to model complex, nonlinear relationships, and are capable of handling multifaceted systems containing multiple interacting factors (Elith et al 2008, Olden et al 2008). These abilities are why machine learning models and in this case boosted decision trees were implemented and performed successfully.

Our analysis revealed not only that distinct classes of compounds influenced the soil microbial community but also identified which classes or compounds lend themselves to accurately predict microbial community behavior. We observed that the addition of secondary metabolites to the soil more reliably predicted the behavior of the soil microbes upon exposure. This could be due to the more recalcitrant nature of these compounds which results in the utilization of these compounds in a more predictable manner and thus this predictability is what enables successful modeling. For example, secondary metabolites such as phenolic compounds correlate more readily with soil microbial abundance (Badri et al 2013a). This higher number of significant correlations implicates secondary metabolites with the ability to more readily predict and model these mechanisms. On the other hand, sugars possibly due to the fact that they are more labile and readily used by a multitude of microbes (Chigineva et al 2009) seem to produce fewer predictions with respect to soil microbial community dynamics.

We also observed the effect compounds had on bacteria (*Actinomycetales*, *Bacillales*, *Pseudomonadales* and *Rhizobiales*) shown to be involved in disease suppression (Mendes et al

2011) and plant-growth promotion (Babalola 2010). Additionally, we are also able to identify how compounds or mixes of compounds influence microbial orders replete with plant pathogens, such as Xanthomonadales (Naushad and Gupta 2013, Turner et al 2013). For example, our model predicts that the addition of 3-hydroxy-3-methylglutaric acid increases the abundance of *Actinomycetales*, *Bacillales*, *Pseudomonadales*, and *Rhizobiales* while simultaneous decreasing the abundance of *Xanthomonadales*. These examples highlight the predictive ability of the model and points us in the direction of personalized/boutique agriculture (Stokes and McCourt 2014). Farmers will eventually have the ability to test the soils microbial community, identify areas that need improvement, and add a cocktail of compounds that would target the issue(s) and resolve them.

Here we provide an assessment of how chemical compounds within a matrix and artificial mixtures of these compounds influence the soil microbial community. It is worth noting that we have only tested one soil type and a limited number of natural libraries and artificial mixtures. Depending on a particular starting soil microbial condition the outcome of addition of compounds might be different. Ultimately, the inherent nature of refining and teaching machine learning based models and analyzing how chemical compounds and blends of these interact with distinct soil microbial communities will enable the ability for new, refined, more sophisticated strategies at a higher level of resolution to help manage the soil microbiome for agricultural purposes and mitigate inadequate, unhealthy soil.

Tables

Table 5-1. Summary of the nine chemical libraries obtained. Outlined below is the distribution of the different categories and compounds in each chemical library. Each percentage was calculated by dividing the sum of the normalized peak GC-MS areas for each library with the sum of the peak area for the compounds in each of the given chemical classes. This was done for each chemical library.

Chemical Library	Developmental Time Point	Fraction	% Abundance of the chemical classes in each chemical library				
			Amino Acids	Secondary Metabolites	Sugars	Sugar Alcohols	Unknown
1	Seedling	Whole Exudates	10.16%	35.02%	3.66%	2.66%	48.49%
2	Vegetative	Whole Exudates	9.86%	18.56%	3.10%	3.80%	64.68%
3	Bolting	Whole Exudates	12.68%	21.86%	4.12%	6.33%	55.02%
4	Seedling	Water Fraction	39.29%	13.45%	31.01%	1.40%	14.85%
5	Vegetative	Water Fraction	29.55%	32.02%	2.21%	2.08%	34.14%
6	Bolting	Water Fraction	38.82%	26.24%	3.18%	1.84%	29.92%
7	Seedling	Ethyl Acetate Fraction	16.50%	44.33%	3.23%	2.66%	33.28%
8	Vegetative	Ethyl Acetate Fraction	21.47%	30.24%	1.74%	2.69%	43.85%
9	Bolting	Ethyl Acetate Fraction	27.90%	51.80%	0.80%	0.29%	19.21%

Table 5-2. List of the compounds and their relative concentrations (peak area) in each of the 9 chemical libraries analyzed by GC-MS.

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
xylose	Sugars	6792	48140	188835	2428	120991	144514	1184	9116	586
trehalose	Sugars	16326	35859	1263282	2595	1460	3597	4666	4281	437
maltose	Sugars	15193	687666	1281153	1620	1710	7002	5212	1343	189
levanbiose	Sugars	2037	411284	104501	6077	7388	11496	783	1396	167
glucose	Sugars	1233	18955	89764	54068	66166	57017	2128	1498	3832
glucoheptulose	Sugars	781	184236	402543	97	36110	24745	318	390	48
fucose	Sugars	42291	1528862	751668	1798	244145	124567	5295	6779	626
fructose	Sugars	18498	264206	295293	495584	865611	1512583	48964	14124	8302
cellobiose	Sugars	2342	164849	118408	3688	15590	11327	1400	1155	898
beta-gentiobiose	Sugars	963	173400	12320	262	8563	1717	523	991	87
arabinose	Sugars	2677	94747	144052	6107	165513	269182	2149	4778	371
1-kestose	Sugars	746	5036	8341	199	587	894	1013	816	90
valine	Amino Acids	19578	355380	1587852	35216	1664637	701877	15469	18039	21730
tryptophan	Amino Acids	945	114774	183351	226	642776	7138	2771	2114	931
thymine	Amino Acids	1567	173120	100123	265	7611	8151	2034	1640	469
threonine	Amino Acids	10802	135274	206292	12455	350441	1723855	6036	5960	4358
serine	Amino Acids	6011	103373	42237	28643	1524338	1603194	4869	6973	10482
saccharopine	Amino Acids	892	12730	5478	121	9688	4816	596	777	180
proline	Amino Acids	8354	121116	19946	19684	492722	2330459	2312	4004	3279
phenylalanine	Amino Acids	4638	80116	237975	4473	64862	1093662	3178	3650	6141
oxoproline	Amino Acids	96587	7285219	6410779	382899	3053173	2313524	100422	403819	379438
ornithine	Amino Acids	2425	11250	67495	358	20108	270700	739	1182	358
O-acetylserine	Amino Acids	746	13703	13182	260	2505	10738	421	569	76
N-methylalanine	Amino Acids	10238	13942	55569	5563	8782	31457	3482	3573	1215
methionine	Amino Acids	1098	27628	200943	437	89507	66127	887	1307	381
lysine	Amino Acids	1626	109942	121619	3475	580344	754606	1246	1322	927

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
leucine	Amino Acids	1074	49358	13422	21079	2040	951666	525	604	9140
isoleucine	Amino Acids	30503	71478	77395	28660	322435	3122993	18827	16482	16949
homoserine	Amino Acids	898	35206	2327	190	6487	1138	431	601	58
homocystine	Amino Acids	1749	67404	197487	175	16149	64346	781	1096	97
glycine	Amino Acids	45984	1506591	1337055	50424	2666010	2179122	21747	9158	9668
glutamine	Amino Acids	23241	356283	157325	71697	3140750	2169195	123254	54637	46885
cytidine	Amino Acids	1509	1282	2906	198	3852	372	882	1009	91
cysteine-glycine	Amino Acids	898	145247	166536	182	25349	6460	905	1509	5197
cysteine	Amino Acids	816	3999	12879	118	16549	34652	334	441	1137
cyano-L-alanine	Amino Acids	1591	16548	22022	3369	147523	583001	2340	3171	1253
citrulline	Amino Acids	886	84412	42625	156	360979	578578	1580	2084	365
beta-alanine	Amino Acids	793	2710	3237	2079	233	617004	371	399	788
asparagine	Amino Acids	6240	444986	902583	10011	3251148	1977886	41366	9321	8997
arginine + ornithine	Amino Acids	1691	123243	1214235	3685	880209	1771876	3051	1611	1512
alanine	Amino Acids	21263	26247	947449	41778	1179053	1492453	15534	17464	12563
xylitol	Sugar Alcohols	2348	106736	177301	244	50521	78162	889	947	337
threitol	Sugar Alcohols	4773	732863	519261	429	120618	43665	1981	1998	563
ribitol	Sugar Alcohols	728	9186	56650	1071	41141	21923	435	426	63
pentitol	Sugar Alcohols	2618	423064	548050	387	94751	72317	1087	884	215
inositol-4-monophosphate	Sugar Alcohols	1292	3299	1290	104	912	1694	327	384	39
inositol myo-	Sugar Alcohols	2313	91401	438383	10230	610359	215746	2794	9214	1271

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
galactinol	Sugar Alcohols	30726	986600	4366592	5731	39707	640790	15147	7881	1504
erythritol	Sugar Alcohols	1644	56358	569251	625	136322	68308	1531	697	259
cellobiotol	Sugar Alcohols	2254	78479	70273	6652	4615	15236	33315	36053	284
arabitol	Sugar Alcohols	804	10614	938	50	44268	12002	272	393	46
2-deoxyerythritol	Sugar Alcohols	30333	1935197	412015	428	305238	82308	2863	13141	1104
xylonolactone NIST	Secondary Metabolites	2436	244377	116267	460	27201	58271	2248	1742	415
xylonic acid	Secondary Metabolites	2002	9548	2564	124	19619	750	288	926	194
urea	Secondary Metabolites	231903	702075	841624	29449	1313938	71993	568705	263104	50780
threonic acid	Secondary Metabolites	63713	469810	648552	7540	280262	79132	5652	2924	1382
succinic acid	Secondary Metabolites	26916	84103	58917	5111	426486	165588	14735	16901	47489
salicylic acid	Secondary Metabolites	2383	179386	61002	356	26469	3975	1861	2269	1722
saccharic acid	Secondary Metabolites	816	13523	8284	74	8678	22969	329	432	95
pyrazine 2,5-dihydroxy NIST	Secondary Metabolites	916	28491	9665	250	9515	4494	797	1495	154
phthalic acid	Secondary Metabolites	25496	6256646	103574	937	835392	196171	10557	19858	1826
phosphoric acid	Secondary Metabolites	5912	3551870	3324	5478	413050	262752	17544	33687	2948

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
parabanic acid NIST	Secondary Metabolites	3417	289620	39980	1913	49905	17516	4726	6151	15548
oxalic acid	Secondary Metabolites	185226	814722	35580	16513	65210	5197	36995	32082	5584
N-acetyl-glutamic acid	Secondary Metabolites	2154	24195	38175	322	3032668	1987539	1064	1206	880
myristic acid	Secondary Metabolites	5154	17871	16569	760	7749	5533	4383	4123	576
methylmaleic acid	Secondary Metabolites	710	2040	722	69	1691	155	516	491	390
methylhexadecanoic acid	Secondary Metabolites	6634	40755	28584	2115	5602	17535	10734	10486	1179
methionine sulfoxide	Secondary Metabolites	2571	54243	85736	520	190198	513840	1246	1530	1412
mannonic acid NIST	Secondary Metabolites	1268	90293	38985	522	15487	47457	419	631	265
malic acid	Secondary Metabolites	10514	3999	10954	26252	1079479	744338	8507	12287	156836
maleimide	Secondary Metabolites	3411	201565	21066	3030	40765	18459	8705	15335	1481
lauric acid	Secondary Metabolites	21615	235342	551042	2844	45239	152387	21793	13725	3168
lactic acid	Secondary Metabolites	35293	34973	20387	10071	140470	13116	31572	49412	6346
kynurenine	Secondary Metabolites	746	17148	16268	155	7266	20033	428	753	126
itaconic acid	Secondary Metabolites	18175	117508	237156	3752	57150	135926	11294	25318	4970
isonicotinic acid	Secondary Metabolites	30703	669411	84797	813	63273	91472	2727	1911	4293

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
isocitric acid	Secondary Metabolites	2818	3002	4404	168	46196	93056	442	488	3109
idonic acid NIST	Secondary Metabolites	904	239	17054	74	8678	22969	240	598	95
glycolic acid	Secondary Metabolites	27544	33434	57411	2529	75588	17401	8081	19310	1523
glycerol-alpha-phosphate	Secondary Metabolites	2841	2909	975	667	7205	29057	2602	1819	237
glycerol-3-galactoside	Secondary Metabolites	13027	207644	7701842	1053	1902447	6517643	66513	2912	707
glycerol	Secondary Metabolites	45491	313098	241128	31384	971358	172945	25379	49653	5222
glyceric acid	Secondary Metabolites	31472	60339	119375	6380	150848	939963	4254	7970	14287
glutamic acid	Secondary Metabolites	3105	1034886	4993096	14552	1789456	1013786	7651	13156	15345
glucuronic acid	Secondary Metabolites	2642	36745	8373	741	10483	40747	295	453	1952
galactonic acid	Secondary Metabolites	4133	309339	11384	527	51365	48626	963	715	321
galactinol	Secondary Metabolites	30726	986600	4366592	5731	39707	640790	15147	7881	1504
GABA	Secondary Metabolites	8688	13972	105843	21244	2966578	840200	11411	5797	20488
fumaric acid	Secondary Metabolites	8765	739391	132399	4159	1533389	560655	17093	17911	362398
ferulic acid	Secondary Metabolites	622	6499	4337	104	784	708	306	640	176
erythronic acid lactone	Secondary Metabolites	29587	185774	417206	6889	253878	521297	7572	18006	2972

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
citric acid	Secondary Metabolites	2712	36494	19252	1158	1071913	659494	2404	5216	189846
citramalic acid	Secondary Metabolites	2436	17259	57977	216	11329	8789	707	2307	1032
capric acid	Secondary Metabolites	13027	36348	24912	1396	8100	24440	6897	7955	1340
beta-sitosterol	Secondary Metabolites	2143	17848	14039	206	812	652	1149	1319	900
azelaic acid	Secondary Metabolites	2560	94554	102646	325	6371	6030	2128	4388	5123
aspartic acid	Secondary Metabolites	6910	1531788	1290685	22849	2407414	711763	11690	18030	15573
aminomalonic acid	Secondary Metabolites	611	9419	7900	178	483	1947	444	491	62
alpha ketoglutaric acid	Secondary Metabolites	1186	43727	30936	354	14855	5239	820	887	7664
adipic acid	Secondary Metabolites	17218	40551	42817	1055	27006	3636	6688	8991	2320
aconitic acid	Secondary Metabolites	1245	4488	5694	145	39592	45395	560	581	2511
4-hydroxybutyric acid	Secondary Metabolites	2055	36593	326324	103	85749	154	2674	1099	997
4-hydroxybenzoate	Secondary Metabolites	3223	30572	138983	679	27248	5695	2865	23805	10166
3-hydroxypropionic acid	Secondary Metabolites	12974	24848	169172	931	134563	49388	6331	8357	3351

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
3-hydroxy-3-methylglutaric acid	Secondary Metabolites	2448	24306	117935	502	108040	13420	1393	2867	12229
3-aminoisobutyric acid	Secondary Metabolites	611	15085	33320	85	4407	14532	290	470	45
2-hydroxyvaleric acid	Secondary Metabolites	50345	359687	343372	1646	56880	11156	18799	49739	2940
2-hydroxyglutaric acid	Secondary Metabolites	24298	1197141	736585	1519	264706	230184	8484	5671	14585
2-5-diketopiperazine NIST	Secondary Metabolites	710	4226	1506	80	1338	825	357	491	33
2,3-dihydroxybutanoic acid NIST	Secondary Metabolites	1039	25285	15525	74	2678	5512	544	521	131
214152	Unknowns	1110	363231	81833	170	256962	80332	1004	664	199
268610	Unknowns	16273	830250	4297797	4134	23975	613982	2395	7160	697
357788	Unknowns	2219	1071362	2052	75	57542	582	332	357	112
288866	Unknowns	939	22283	24603	150	522593	783222	387	619	277
214165	Unknowns	1004	362806	129245	340	458864	144075	877	1244	634
226848	Unknowns	2730	984829	28608	77	52784	3767	389	485	108
213160	Unknowns	6675	19544	4066	130	113760	87848	808	393	99
213148	Unknowns	1162	28001	5124	106	86441	1442	716	613	75
216832	Unknowns	3599	1703854	1481954	480	63640	116754	5246	2257	306
373725	Unknowns	998	19357	34488	2211	228926	16680	935	1066	471
352812	Unknowns	851	1426859	316104	140	73311	10207	477	712	118
201005	Unknowns	1139	2244	63770	1574	1363076	574715	997	1349	3870
385048	Unknowns	704	3206	2558	124	224240	147739	827	706	427

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
223830	Unknowns	652	4616	5892	339	107644	166570	532	530	167
371568	Unknowns	945	29593	9219	115	30035	3828	401	414	76
362008	Unknowns	9358	430670	792245	305	56041	59141	1034	1426	169
213179	Unknowns	3311	869979	297224	209	183377	135099	769	2534	346
367939	Unknowns	9880	5631942	2434517	2790	689647	345509	5108	5186	4991
213188	Unknowns	810	8545	6784	90	283898	230370	415	679	167
217809	Unknowns	2430	781090	524452	443	148157	224702	2529	1036	511
267904	Unknowns	986	123500	8610	174	96630	12144	562	932	133
199203	Unknowns	16461	23143232	6602637	16263	2564884	1079905	48969	95546	5393
228018	Unknowns	611	4681	5284	347	99909	192965	398	613	333
238134	Unknowns	1362	39921	16524	1510	129315	23269	866	917	680
327468	Unknowns	839	2780	22196	342	2201936	1060429	562	726	1976
385107	Unknowns	687	2617	2881	227	184085	122831	557	878	328
357841	Unknowns	1286	8766	13516	175	84199	16901	755	792	233
362010	Unknowns	1104	12252	8547	125	110459	31373	659	581	249
269256	Unknowns	928	7921	5472	84	41975	4835	408	411	100
362028	Unknowns	933	24184	139355	148	108149	55628	979	753	325
321749	Unknowns	2184	170550	225879	163	37902	32516	364	551	76
367944	Unknowns	3763	732863	519261	298	120618	147556	924	1998	563
267692	Unknowns	1209	76351	4066	150	70354	73008	580	566	116
362023	Unknowns	922	19427	16795	311	79431	36374	537	744	196
307912	Unknowns	611	2576	5130	81	57242	115480	368	917	179
219507	Unknowns	1297	1858537	9294272	4745	2691884	2293558	53519	14249	9257
207188	Unknowns	1174	48979	10112	118	46902	83403	444	697	245
200490	Unknowns	675	46857	103902	72	19234	10207	336	456	82
328420	Unknowns	740	5362	3842	264	114166	83531	491	491	247
226841	Unknowns	1021	35730	7431	83	33075	4054	362	584	85
200906	Unknowns	669	4681	4586	456	99909	192965	316	613	333
218492	Unknowns	1497	32309	4135	128	70354	18368	580	515	155

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
224843	Unknowns	4332	365936	144321	380	34085	20182	815	756	86
330992	Unknowns	4538	4302097	434901	1825	191009	46238	1041	2224	603
299185	Unknowns	1104	56696	431	226	74	777	2057	1375	96
362130	Unknowns	1104	56696	431	226	1098796	777	2057	1375	96
268345	Unknowns	869	110344	93159	147	22528	27427	567	1453	103
352980	Unknowns	746	414350	471084	88	56028	17334	859	575	247
269249	Unknowns	757	4715	2202	99	32488	11669	412	470	74
213141	Unknowns	986	13610	7427	133	17464	2505	544	744	99
224632	Unknowns	710	6528	5128	72	21072	100784	260	366	51
385023	Unknowns	681	105693	28414	88	11008	1354	304	360	26
359697	Unknowns	1450	27372	8582	135	48877	36910	790	610	151
269294	Unknowns	886	61896	12571	147	43784	43976	534	819	134
385021	Unknowns	605	5036	11748	83	5907	3036	327	393	32
200610	Unknowns	8354	372441	309700	795	36159	25286	1290	1012	215
310063	Unknowns	8354	372441	309700	795	36159	22925	1290	1012	215
310897	Unknowns	1338	11448	12177	393	35566	5671	691	720	514
199177	Unknowns	2560	165420	144149	216	23945	129784	912	777	100
268353	Unknowns	646	2646	475	65	10821	393	304	369	49
228911	Unknowns	922	65714	99702	111	12580	15878	327	488	60
213154	Unknowns	1473	50623	20744	561	131721	106806	871	1736	735
362086	Unknowns	5354	378713	267708	288	58769	30382	1467	1524	210
381469	Unknowns	2524	295274	537711	7227	893133	1619609	2432	1718	3881
268565	Unknowns	945	16915	86100	140	42043	241192	451	664	176
225327	Unknowns	881	92217	52028	111	32000	2619	924	515	147
385058	Unknowns	652	17673	5344	90	11626	894	514	357	57
385112	Unknowns	5148	4867	4839	232	733	883	274	393	43
303838	Unknowns	1620	1045489	2160935	2125	227608	682584	1193	2468	719
385028	Unknowns	928	38715	95696	75	21587	1624	419	866	104
280930	Unknowns	1720	100284	27878	134	12191	11829	666	506	159
285340	Unknowns	2824	163672	51731	192	3365	3362	1467	595	68

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
224589	Unknowns	1051	16495	2287	131	29875	1197	647	1054	270
280945	Unknowns	787	45313	18354	79	7194	6180	389	399	130
352849	Unknowns	1168	3684	1533	96	1229	4176	504	536	55
385024	Unknowns	663	4238	1246	91	756	118	332	375	26
307669	Unknowns	916	41483	270818	118	14465	87322	368	423	60
218748	Unknowns	746	359705	64747	106	21356	310	507	3570	824
268583	Unknowns	3105	25029	42767	148	22105	6180	318	551	54
303956	Unknowns	869	431760	501979	323	126013	230806	1720	1551	900
208897	Unknowns	969	23921	65824	98	13629	19452	465	366	100
250380	Unknowns	1708	313733	414251	2308	63889	22056	882	709	107
385034	Unknowns	2765	22371	208682	1001	45008	291273	1522	1176	249
223191	Unknowns	1532	14234	35499	355	18632	10199	666	902	109
310871	Unknowns	687	8289	6818	73	7174	6011	281	378	82
199205	Unknowns	2783	437997	867318	1349	53220	222053	4115	1060	398
202899	Unknowns	2002	10445	255820	336	40509	366962	1025	1227	156
208686	Unknowns	3217	196529	359263	301	61628	110440	5387	2007	1765
352777	Unknowns	1145	158700	364393	180	24570	39282	1094	869	101
362109	Unknowns	1039	7041	129404	190	73917	45375	601	846	322
367914	Unknowns	1802	94852	78665	6471	1158474	235953	9841	8815	15582
211945	Unknowns	2507	38586	181001	289	8054	52933	1039	1995	1020
245705	Unknowns	1338	265879	936463	9003	183842	74046	2918	3221	619
385065	Unknowns	1045	46793	12779	196	3254	829	781	738	111
200509	Unknowns	14976	339421	134595	2597	6253	11969	1442	932	134
213227	Unknowns	986	23041	5684	234	20256	9666	698	1030	220
310888	Unknowns	822	7117	14862	6323	35454	4893	926	1066	532
231350	Unknowns	622	114005	353274	107	18450	29994	541	506	74
237415	Unknowns	916	28491	9665	250	9515	4324	797	1495	154
359567	Unknowns	951	28380	5577	170	22271	13859	474	840	222
224818	Unknowns	669	28852	32877	118	3557	4543	309	333	47
303992	Unknowns	851	44328	20354	79	3920	101	297	393	56

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
281187	Unknowns	1920	147124	256178	462	26764	134337	1338	1480	381
213253	Unknowns	4291	77756	34377	1938	213188	476941	4820	3239	2041
208538	Unknowns	12164	50162	731473	1334	79228	195704	5251	5311	732
267751	Unknowns	1286	57938	321107	4259	15656	56002	624	1102	340
310757	Unknowns	3869	313075	483903	302	40578	73844	1925	1012	377
240551	Unknowns	928	51013	91897	174	21218	38809	458	741	308
267691	Unknowns	933	26766	17157	186	17178	30603	474	584	63
267649	Unknowns	1039	65375	103811	280	6924	6064	550	4588	280
310162	Unknowns	2225	264393	274492	347	46704	52310	1511	1902	410
385117	Unknowns	6052	8044	5338	365	22580	972	1564	2569	585
236810	Unknowns	1802	10445	92989	244	33614	171720	951	1227	211
202599	Unknowns	1773	30706	17821	1458	18012	18146	8178	1212	489
228619	Unknowns	716	70464	31564	140	11273	44340	396	414	207
202573	Unknowns	1579	35486	37875	676	589604	151049	610	604	1843
227816	Unknowns	675	8428	3472	84	3690	2059	355	456	52
225540	Unknowns	704	1259	683	75	6866	336	244	372	28
310875	Unknowns	875	4657	4418	292	14910	9779	431	866	141
324275	Unknowns	1139	3515	22081	621	71438	23424	848	735	249
362073	Unknowns	875	4657	4418	292	14910	16096	431	866	141
199242	Unknowns	646	20815	47686	119	2419	9943	638	464	64
236605	Unknowns	1368	297267	58339	287	118305	40050	838	1024	206
218734	Unknowns	2554	594913	1542321	276	91221	309273	4189	4641	1322
231254	Unknowns	1362	43727	30936	354	14855	3940	820	887	7625
281132	Unknowns	804	7076	50269	73	4735	11866	359	456	152
284607	Unknowns	1110	3526	1262	70	4165	871	322	560	111
375029	Unknowns	1714	85572	165729	171	13272	10620	576	4912	1744
212022	Unknowns	3393	1675951	45116	3725	80717	6034	6838	6910	566
227598	Unknowns	787	76986	6715	81	6078	1089	590	598	52
307924	Unknowns	775	3521	6226	69	4050	8697	302	330	34
215445	Unknowns	1732	1114693	977096	4697	186572	143690	3332	2831	619

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
267696	Unknowns	746	29056	34339	76	4358	1656	352	473	77
202572	Unknowns	1080	414192	243437	253	29318	51213	986	658	2493
208397	Unknowns	10567	99695	539963	785	1842	75540	3666	3787	429
213143	Unknowns	881	11652	5452	91	21244	1100	246	339	45
374786	Unknowns	2143	320955	518485	989	57761	56126	3740	3635	2140
323686	Unknowns	2700	348595	449013	1406	148335	152959	3639	5145	6031
200710	Unknowns	1755	34017	25846	358	5734	6831	633	765	100
212761	Unknowns	1961	91844	69082	188	10069	523	723	1923	147
212016	Unknowns	2061	225212	219255	245	46078	92297	1110	2933	2075
200557	Unknowns	910	6015	9550	121	628	1115	355	450	63
233408	Unknowns	1427	22178	63279	83	8375	3315	431	464	223
353047	Unknowns	869	45062	23498	135	3179	3621	829	9107	784
280546	Unknowns	939	28258	24101	125	4156	1779	412	384	95
228680	Unknowns	1162	5555	11601	102	917	179	221	393	29
268365	Unknowns	1427	7414	9415	4019	63685	51876	4086	3546	808
238267 trisaccharide	Unknowns	1380	12870	12731	138	851	220	329	426	52
269250	Unknowns	1303	23583	34458	124	20549	9830	500	2778	1065
235449	Unknowns	1368	6954	9390	206	15469	5776	417	515	99
215397	Unknowns	141321	1225901	745379	12970	181597	38534	90335	46819	7547
296119	Unknowns	605	17049	13150	199	3052	2135	329	405	53
237606	Unknowns	834	16455	15350	94	8822	48363	670	700	75
224811	Unknowns	728	28287	22020	104	2113	3675	281	447	177
227652	Unknowns	2630	111201	573811	205	25977	5905	2957	1060	784
214416	Unknowns	881	24819	8424	85	2262	898	811	494	102
237520	Unknowns	740	15411	15685	70	2821	4281	375	470	46
224635	Unknowns	1051	9594	26198	240	10554	6718	288	429	138
385085	Unknowns	2154	586613	100661	255	17822	1793	1108	7205	5607
310053	Unknowns	798	10649	14094	398	3429	2979	442	688	435
200624	Unknowns	793	21776	275	67	2210	4472	401	447	31

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
212177	Unknowns	1620	102633	150027	466	10263	17582	1057	1280	674
205672	Unknowns	12445	105571	95201	1347	21700	44686	6006	6550	1276
205674	Unknowns	24093	306640	23387	2186	38566	83273	5859	77668	3782
237605	Unknowns	652	15866	10804	63	4829	20193	357	613	44
232755	Unknowns	4280	89419	36480	207	26462	3895	1092	813	786
240436	Unknowns	31801	669411	84797	813	63273	88630	2727	1911	4347
310380	Unknowns	1139	113160	290256	408	13370	23745	599	777	761
227728	Unknowns	1409	14584	757350	2423	8060	12902	1755	1754	1045
213243	Unknowns	928	473669	649897	419	68324	57102	3040	6457	4098
296071	Unknowns	892	6009	4362	129	892	8043	403	548	444
385030	Unknowns	875	142997	19994	112	11579	5145	822	887	896
374402	Unknowns	693	63924	380826	54	6865	2290	778	497	586
211896	Unknowns	3757	968368	1307316	18985	126738	115996	2651	3055	624
200615	Unknowns	863	10468	12391	75	1637	3263	362	503	37
231248	Unknowns	822	250923	93555	412	4225	7518	1034	709	292
238938	Unknowns	675	8061	6246	69	895	3046	408	518	45
233790	Unknowns	963	24976	35449	133	3060	3337	583	9256	2329
212261	Unknowns	2342	87805	934091	685	60626	104091	4813	12248	5235
310581	Unknowns	751	3118	13050	152	6908	19716	424	500	100
331031	Unknowns	1315	82722	83524	231	11123	5076	804	2135	909
267654	Unknowns	1333	4098	253050	157	14928	6446	944	1691	1200
227822	Unknowns	969	10101	3088	110	5144	3061	371	581	93
268671	Unknowns	599	34092	5452	71	3533	289	260	310	869
231098	Unknowns	980	13255	14824	89	3139	1718	357	848	99
322204	Unknowns	2507	98215	80450	4687	43361	45277	1541	1885	315
200702	Unknowns	1732	273993	977096	889	171226	143690	5437	5975	868
224322	Unknowns	7080	91605	8705	942	18638	19730	4770	4719	264
208714	Unknowns	2360	6592	7718	161	1201	1889	408	616	43
232946	Unknowns	687	24761	12191	70	562	1985	345	509	165
228164	Unknowns	3264	137961	135038	443	23095	18673	1674	4942	1523

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
304391	Unknowns	886	7869	12919	88	277	289	410	607	49
384918	Unknowns	46712	846850	153968	996	136306	227258	6009	18786	1977
268506	Unknowns	1004	15598	10945	355	4357	11542	562	1027	303
318770	Unknowns	1585	29190	54034	4069	63439	61813	1955	1021	748
211921	Unknowns	945	11698	115422	88	1821	218	322	1081	45
202083	Unknowns	1092	299412	326833	395	6353	54317	2757	1268	417
241312	Unknowns	828	7059	2956	185	3182	6996	615	965	127
385104	Unknowns	1039	9722	14814	87	15900	19885	435	464	108
310006	Unknowns	3916	179351	108897	190	10444	25992	1465	893	391
310367	Unknowns	951	51287	40882	166	7331	3602	465	1316	234
237333	Unknowns	622	33014	27088	85	1477	662	311	518	96
310448	Unknowns	2912	114570	192095	194	11123	9506	889	2135	582
353091	Unknowns	710	30409	25862	2247	38370	29503	868	616	774
231576	Unknowns	1016	17690	7589	91	14211	2875	304	595	64
241168	Unknowns	3194	5578	2861	213	1227	277	541	1125	403
220122	Unknowns	1773	93208	133693	180	11672	35479	820	1096	117
231260	Unknowns	3194	5578	2861	249	1379	277	541	1125	403
237392	Unknowns	810	26585	15900	96	2182	1010	265	598	63
211916	Unknowns	3833	22056	6093	379	7775	105188	1126	1337	164
237652	Unknowns	1068	21263	28234	88	2035	2011	670	458	186
385042	Unknowns	687	13768	1480	58	1415	87	269	429	623
268712	Unknowns	1239	82430	18091	75	10812	751	903	735	6316
214201	Unknowns	1785	8259	4924	158	2124	6550	873	610	45
310193	Unknowns	1039	20389	15170	105	1740	1556	431	664	57
269776	Unknowns	58388	456416	394795	5887	27091	6714	24030	17717	2878
199562	Unknowns	2736	85176	25614	447	4233	227717	732	723	243
206309	Unknowns	3610	87752	170881	289	11327	7731	10288	1834	219
229277	Unknowns	3957	6761	7201	438	667	879	1239	1914	196
385075	Unknowns	693	91791	183009	62	6695	119	728	515	170
206965	Unknowns	6123	21496	48305	2423	14908	37766	1013	884	139

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
213960	Unknowns	746	1405	412	97	1354	334	686	768	35
231210	Unknowns	769	20511	30303	107	2374	5326	523	435	49
241387	Unknowns	3282	96600	196903	289	14052	7731	7950	1834	219
303966	Unknowns	2906	63773	48194	331	3445	23615	2151	3826	130
233471	Unknowns	616	4057	8972	81	381	138	313	432	39
267880	Unknowns	710	3637	5530	74	194	206	343	396	35
233289	Unknowns	957	17906	19299	229	1176	1382	352	607	49
211910	Unknowns	24556	235342	551042	3255	39621	146319	21793	16237	3494
319168	Unknowns	36749	322587	52926	8620	80546	55771	28032	36012	3886
311041	Unknowns	793	18716	11129	236	1593	248	332	557	385
305637	Unknowns	1462	15662	3367	110	732	178	412	646	69
225867	Unknowns	264	27658	35989	62	2066	5911	514	316	329
238549	Unknowns	2225	23403	34834	292	6430	3818	415	429	1075
239332	Unknowns	6745	335550	374774	456	68151	127866	1283	3406	4456
374356	Unknowns	722	23845	15742	99	7967	30835	470	3924	567
211636	Unknowns	115062	508222	191446	1532	31368	240048	8457	45634	5612
211590	Unknowns	175211	707053	286663	1118	37077	196358	12650	76626	9497
322260	Unknowns	1104	10772	1496	78	4814	149	339	846	90
199942	Unknowns	15921	20191	24184	1381	11570	11685	7609	7895	660
385006	Unknowns	1245	869081	38431	109	2245	684	2496	6335	25525
227658	Unknowns	816	1609	813	84	274	247	304	1364	45
227923	Unknowns	722	3153	4542	245	1288	536	647	676	37
238550	Unknowns	1391	81725	39802	136	7228	4678	2538	2138	1005
226256	Unknowns	581	3713	7296	63	211	1168	309	375	31
271049	Unknowns	13549	131695	44411	1974	43636	39832	16439	12578	708
200486	Unknowns	2548	81725	39802	464	7228	4678	2137	2138	938
236965	Unknowns	10332	25349	7629	129	2954	6232	2913	3004	422
299487	Unknowns	710	962	621	68	122	96	348	524	33
385055	Unknowns	763	143242	143094	202	15857	632	739	1831	6267
218694	Unknowns	10221	263315	116853	617	14160	32250	2508	3284	4130

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
203592	Unknowns	11764	40178	28078	1097	5577	8807	8141	8839	960
199246	Unknowns	3135	13534	23598	318	1857	909	1363	2129	229
200905	Unknowns	6657	331657	23699	4535	34843	45122	3478	17631	2640
373752	Unknowns	2154	101327	228308	341	8182	5395	2151	1587	1668
309617	Unknowns	17159	33894	15459	377	3474	1741	2312	8747	708
200567	Unknowns	1350	33387	98624	6077	4605	16108	898	1253	316
308219	Unknowns	1920	12205	6505	316	2737	897	1870	1655	125
268438	Unknowns	1256	11162	1527	126	20	164	555	801	75
219512	Unknowns	18598	733096	933990	87	7647	2571	8364	157989	59793
211919	Unknowns	1086	27477	13920	140	1983	7013	702	1051	673
299441	Unknowns	957	9513	40644	236	850	607	668	158	415
385120	Unknowns	992	66337	98885	74	10941	7748	725	655	2804
224627	Unknowns	2301	7875	40733	790	4390	321	1603	5273	638
206022	Unknowns	16279	26387	31680	1501	7993	7908	8583	9631	1113
224574	Unknowns	781	10399	1163	264	1013	1068	1476	1474	101
303839	Unknowns	3217	105040	300670	19210	37668	155118	3579	4171	3111
238506	Unknowns	793	123698	403914	79	13625	5895	1345	655	9146
234717	Unknowns	17958	41757	115499	2568	16911	27314	13489	15844	2050
201042	Unknowns	12727	72183	119258	1161	8656	5267	5813	11578	10919
294511	Unknowns	975	3975	6618	150	418	1442	751	569	88
267987	Unknowns	5307	95545	46142	343	11744	5819	2204	2596	4451
268437	Unknowns	1749	2926	18075	236	1160	3197	859	837	247
235327	Unknowns	5777	39035	38541	1505	5405	2638	12240	9604	500
384992	Unknowns	40712	85059	37954	4036	13957	13988	25964	28450	2786
213191	Unknowns	975	7834	8258	121	742	1897	3597	1346	63
357685	Unknowns	14982	37549	13991	427	4413	69379	4086	6737	1077
241881	Unknowns	2014	857	1587	205	106	348	801	280	37
288810	Unknowns	4796	19241	5817	815	1998	3281	3151	3668	315
368156	Unknowns	4602	7939	2964	471	1149	315	2227	2727	272
357502	Unknowns	693	52779	6022	152	6014	9861	702	545	1964

Name	Group	Library 1	Library 2	Library 3	Library 4	Library 5	Library 6	Library 7	Library 8	Library 9
200540	Unknowns	1427	300764	884934	3049	32747	23526	3275	1902	20459
224849	Unknowns	2372	6604	6568	357	1186	1710	2043	1685	259
216428	Unknowns	63889	99730	25535	5783	12030	3571	12795	41129	1430
327143	Unknowns	1309	2448	1428	32	291	207	624	950	20

Table 5-3. The 12 selected compounds composing the artificial chemical blends added to the soil in order to validate the 5 statistical models.

Compound	Chemical Class
Glycine	Amino Acid
Oxoproline	Amino Acid
Phenylalanine	Amino Acid
Tryptophan	Amino Acid
citramalic acid	Secondary Metabolite
fumaric acid	Secondary Metabolite
Ethyl-4-hydroxybenzoate	Secondary Metabolite
Maleimide	Secondary Metabolite
Fructose	Sugar
Fucose	Sugar
Maltose	Sugar
Xylose	Sugar

Table 5-4. Observed species richness (Sobs), Chao1 estimate of total species richness Schao, Shannon diversity (H'), and Shannon Evenness (E_H) of soil microbial community controls (Water, Nothing, and EtOAc) after a period of 2 and 6 weeks. * indicates statistically significant differences between water control soil microbial communities after 2 and 6 weeks.

	Control	**2 weeks**	**6 weeks**
S_{obs}	Water*	503.67 ± 56.22	1054.67 ± 128.53
	EtOAc	319.33 ± 31.42	1190.67 ± 157.35
	Nothing	1012.00 ± 210.97	1681.33 ± 261.02
S_{chao}	Water*	826.80 ± 99.42	1858.75 ± 197.43
	EtOAc	618.21 ± 84.47	2091.50 ± 256.55
	Nothing	1702.66 ± 357.37	2889.74 ± 459.53
H'	Water*	5.76 ± 0.09	6.34 ± 0.10
	EtOAc	5.35 ± 0.10	6.36 ± 0.14
	Nothing	6.25 ± 0.23	6.73 ± 0.10
E_H	Water	0.93 ± 0.005	0.91 ± 0.003
	EtOAc	0.92 ± 0.008	0.90 ± 0.005
	Nothing	0.91 ± 0.001	0.91 ± 0.006

Table 5-5. The observed species richness (Sobs), Chao estimate of total species richness Schao, Shannon diversity (H'), and Shannon Evenness (E_H) for soil microbial communities, which have been exposed to a given chemical library for 2 or 6 weeks. * indicates statistically significant differences between water control soil microbial communities within a given week (ANOVA Dunnett post-hoc, $p<0.05$). # indicates statistically significant differences between the soil microbial communities after exposure to a given chemical library at 2 and 6 weeks (t-test Bonferroni correction, $p<0.001$).

	Library	2 weeks	6 weeks
S_{obs}	1	368.50 ± 213.50	318.00 ± 9.00
	2	737.33 ± 293.42	1220.33 ± 290.89
	3	677.33 ± 195.65	1276.33 ± 196.71
	4	639.00 ± 147.60	560.00 ± 103.59
	5	508.67 ± 65.51	1192.33 ± 126.24
	6	974.00 ± 429.49	625.33 ± 81.24
	7	507.67 ± 65.70	525.33 ± 97.05
	8	1055.33 ± 166.25	2078.33 ± 379.31*
	9	698.00 ± 133.51	1335.33 ± 114.62
S_{chao}	1	470.94 ± 273.16	477.36 ± 0.68
	2	1889.38 ± 692.44	3252.66 ± 574.36
	3	1307.82 ± 429.15	2725.19 ± 463.12
	4	867.62 ± 197.99	881.18 ± 150.25
	5	1386.96 ± 201.88	3043.46 ± 302.16
	6	1659.79 ± 678.55	1452.58 ± 267.41
	7	689.60 ± 106.67	801.22 ± 152.46
	8	2506.66 ± 343.32*	4766.85 ± 1025.15*
	9	1373.12 ± 193.87	3086.33 ± 299.31
H'	1	5.30 ± 0.60	5.43 ± 0.07*
	2	6.03 ± 0.32	6.17 ± 0.21
	3	5.31 ± 0.13	5.64 ± 0.18*
	4	5.91 ± 0.24	5.79 ± 0.18
	5	5.74 ± 0.06	6.16 ± 0.12
	6	5.47 ± 0.27	4.87 ± 0.19*
	7	5.67 ± 0.18	5.47 ± 0.20*
	8	6.31 ± 0.19	6.41 ± 0.15
	9	5.79 ± 0.12	6.30 ± 0.10
E_H	1	0.93 ± 0.003	0.94 ± 0.007
	2#	0.94 ± 0.007	0.88 ± 0.003*
	3	0.83 ± 0.047*	0.79 ± 0.013*
	4	0.93 ± 0.004	0.92 ± 0.003
	5	0.92 ± 0.017	0.87 ± 0.004*
	6	0.83 ± 0.046*	0.76 ± 0.016*
	7	0.91 ± 0.011	0.88 ± 0.013
	8#	0.91 ± 0.006	0.84 ± 0.002*
	9	0.89 ± 0.011	0.88 ± 0.005*

Table 5-6. The observed species richness (Sobs), Chao estimate of total species richness Schao, Shannon diversity (H'), and Shannon Evenness (E_H) for soil microbial communities, which have been exposed to artificial chemical blends. * indicates statistically significant differences between water control soil microbial communities (ANOVA Dunnett post-hoc, $p<0.05$).

Treatments	Richness		Shannon	
	S_{obs}	**Chao (S_{Chao})**	**Diversity (H')**	**Evenness (E_H)**
Water Control	437.00 ± 61.99	1201.04 ± 144.42	5.44 ± 0.18	0.90 ± 0.02
Amino Acids	707.75 ± 68.92*	1894.28 ± 62.45*	5.58 ± 0.29	0.85 ± 0.04
Mixture	481.75 ± 18.86	1437.93 ± 52.32	5.45 ± 0.27	0.88 ± 0.04
Secondary Metabolites	320.00 ± 18.25	981.77 ± 29.48	5.00 ± 0.29	0.87 ± 0.05
Sugars	711.75 ± 109.66*	1744.43 ± 303.39	5.87 ± 0.10	0.90 ± 0.02

Table 5-7. Mean square predicted error (MSPE) of each of the 5 statistical models implemented for each microbial order (pcr- partial components regression, pls - partial least squares regression, LASSO- LASSO, random forest-random forest, and boosted - boosted decision trees).

Microbial Orders	MSPE pcr	MSPE pls	MSPE LASSO	MSPE random forest	MSPE boosted
Acidimicrobiales	5.68	5.68	0.72	1.17	1.07
Acidobacteria Gp1 order incertae sedis	244.74	244.74	2.40	1.08	0.90
Acidobacteria Gp16 order incertae sedis	26.27	26.27	1.34	2.93	2.29
Acidobacteria Gp17 order incertae sedis	49.35	49.35	1.28	2.39	2.43
Acidobacteria Gp3 order incertae sedis	8.79	8.79	1.92	1.40	1.22
Acidobacteria Gp4 order incertae sedis	116.37	116.37	3.10	1.45	1.59
Actinomycetales	67.49	67.49	1.08	1.28	1.19
Armatimonadetes gp5 order incetae sedis	7.82	7.82	2.52	1.05	1.29
Bacillales	244.42	244.42	1.31	0.82	0.70
Bacteroidetes incertae sedis order incertae sedis	3.20	3.20	3.74	1.02	1.02
Bdellovibrionales	1.95	1.95	2.00	2.08	2.01
Burkholderiales	184.60	184.60	5.98	2.28	2.21
Caulobacterales	26.70	26.70	3.94	3.81	3.14
Clostridiales	0.72	0.72	3.92	0.65	0.66
Desulfuromonadales	1.17	1.17	2.18	0.48	0.48
Flavobacteriales	50.84	50.84	1.96	1.09	1.15
Legionellales	1.92	1.92	4.05	0.31	0.39
Methylophilales	6.37	6.37	15.72	8.12	8.29
Myxococcales	239.79	239.79	1.62	1.53	1.51
Nitrospirales	516.68	516.68	5.36	3.21	3.92
Opitutales	5.90	5.90	0.96	0.94	1.19
Planctomycetales	303.24	303.24	2.92	1.83	1.65
Pseudomonadales	640.18	640.18	0.13	3.94	4.65
Rhizobiales	2.04	2.04	4.80	0.75	0.67
Rhodocyclales	288.93	288.93	1.61	1.08	1.24
Rhodospirillales	6.14	6.14	3.03	1.30	1.29
Solirubrobacterales	21.72	21.72	5.34	2.33	2.07
Spartobacteria order incertae sedis	217.95	217.95	1.67	1.32	1.43
Sphingobacteriales	124.17	124.17	7.87	6.07	5.73

Microbial Orders	MSPE pcr	MSPE pls	MSPE LASSO	MSPE random forest	MSPE boosted
Sphingomonadales	18.15	18.15	5.06	3.12	2.86
Subdivision3 order incertae sedis	50.45	50.45	5.46	2.99	3.14
unclassified Actinobacteria	22.43	22.43	1.91	2.59	2.58
unclassified Bacteria	35.50	35.50	3.58	2.98	2.80
unclassified Bacteroidetes	118.42	118.42	1.17	1.79	1.81
unclassified Chloroflexi	35.29	35.29	4.70	1.67	1.71
unclassified Firmicutes	115.62	115.62	1.02	1.20	1.38
unclassified Planctomycetes	11.35	11.35	6.62	1.90	2.00
unclassified Proteobacteria	104.35	104.35	3.96	1.66	1.49
Verrucomicrobiales	5.76	5.76	2.26	0.42	0.41
Xanthomonadales	210.79	210.79	1.85	0.63	0.79

Table 5-8. Average and standard error obtained for overall MSPE of each statistical model employed.

	pls	**pcr**	**LASSO**	**random forest**	**boosted**
Average	103.58	103.58	3.30	1.97	1.96
Standard Error	22.80	22.80	0.43	0.24	0.24

Table 5-9. Mean square predicted error (MSPE) after boosted decision tree modeling for each of the artificial blends added to the soil for each microbial order.

Microbial Orders	Mixture	Amino Acids	Secondary Metabolites	Sugars
Acidimicrobiales	0.22	0.95	0.42	2.52
Acidobacteria Gp1 order incertae sedis	0.70	1.53	0.35	0.89
Acidobacteria Gp16 order incertae sedis	3.12	2.69	1.50	1.64
Acidobacteria Gp17 order incertae sedis	1.76	1.39	3.78	3.13
Acidobacteria Gp3 order incertae sedis	1.42	1.21	0.76	1.37
Acidobacteria Gp4 order incertae sedis	0.79	1.36	3.43	1.24
Actinomycetales	0.67	1.42	2.51	0.49
Armatimonadetes gp5 order incetae sedis	3.16	1.03	0.43	0.31
Bacillales	0.31	1.56	0.19	0.61
Bacteroidetes incertae sedis order incertae sedis	0.21	2.74	0.69	0.34
Bdellovibrionales	1.53	4.17	0.91	1.16
Burkholderiales	2.34	1.60	3.39	1.83
Caulobacterales	4.09	3.45	0.77	3.67
Clostridiales	0.39	0.76	0.89	0.66
Desulfuromonadales	0.46	0.44	0.34	0.65
Flavobacteriales	2.08	0.79	0.57	1.01
Legionellales	0.91	0.04	0.07	0.47
Methylophilales	11.52	6.96	10.00	5.12
Myxococcales	1.05	2.64	1.13	1.11
Nitrospirales	6.31	0.57	6.14	3.22
Opitutales	2.37	0.93	0.10	1.09
Planctomycetales	1.76	2.13	2.08	0.73
Pseudomonadales	0.77	11.68	1.82	3.64
Rhizobiales	0.11	0.21	1.24	1.27
Rhodocyclales	2.79	1.20	0.04	0.64
Rhodospirillales	1.27	0.84	1.38	1.69
Solirubrobacterales	1.02	1.51	0.27	5.03
Spartobacteria order incertae sedis	1.45	1.38	1.98	1.04
Sphingobacteriales	5.53	7.20	0.38	8.48
Sphingomonadales	1.19	4.92	0.48	4.26
Subdivision3 order incertae sedis	6.08	1.52	3.09	1.85
unclassified Actinobacteria	3.54	1.43	3.27	2.24
unclassified Bacteria	3.32	4.59	0.23	2.40
unclassified Bacteroidetes	0.32	1.63	1.73	3.54
unclassified Chloroflexi	0.57	1.58	1.11	3.43
unclassified Firmicutes	0.74	2.23	1.05	1.43
unclassified Planctomycetes	2.18	2.56	1.45	1.68
unclassified Proteobacteria	1.26	2.70	0.41	1.32
Verrucomicrobiales	0.56	0.06	0.22	0.77
Xanthomonadales	2.37	0.27	0.25	0.15

Table 5-10. Average and standard error obtained for overall MSPE of each artificial blend employed after boosted decision tree modeling.

	Mixture	**Amino Acids**	**Secondary Metabolites**	**Sugars**
Average	2.06	2.20	1.52	1.95
Standard Error	0.35	0.36	0.30	0.27

Table 5-11. Relative importance of each compound as determined by boosted decision tree modeling for predicting the abundance of microbial Orders found in the soil. The relative contribution of each variable is scaled so that the sum of all variables combined within a bacterial order is 100, higher numbers indicate stronger influences on respective bacterial order relative abundance.

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
methionine	1.68	2.53	2.29	1.77	1.34	1.61	1.66
erythronic acid lactone	2.03	1.57	1.70	1.65	1.31	1.88	1.69
3-hydroxy-3-methylglutaric acid	1.89	2.43	1.80	1.82	2.01	2.67	1.79
4-hydroxybutyric acid	1.60	1.68	1.79	2.35	1.39	1.62	2.14
threonic acid	1.82	2.07	1.61	1.76	1.25	1.60	1.96
Unk206965	1.90	1.47	1.52	1.68	1.16	1.84	1.60
leucine	1.27	1.29	1.33	1.62	1.70	1.29	1.37
Unk218748	1.49	1.33	1.17	0.51	0.64	1.61	0.58
Unk224635	1.17	1.10	1.28	1.32	1.49	1.29	1.58
Unk232755	1.27	1.06	1.70	1.60	1.00	1.06	1.42
2-hydroxyglutaric acid	1.72	1.13	1.29	1.26	1.21	1.37	1.55
Unk238134	1.48	2.65	1.35	1.33	0.78	1.20	1.08
citrulline	1.41	1.43	1.29	1.17	1.10	1.30	1.20
fucose	1.03	0.83	0.92	0.82	0.91	0.99	0.99
isonicotinic acid	1.12	1.19	1.19	1.25	1.47	1.36	1.23
ornithine	0.97	0.93	1.45	1.15	1.11	0.97	1.11
Unk199562	1.02	1.08	1.45	0.67	1.10	1.76	1.08
Unk212261	1.09	2.81	0.92	0.98	0.81	1.35	0.71
citramalic acid	0.76	0.77	0.70	0.86	1.08	0.90	0.78
cytidine	0.44	0.37	0.55	0.54	1.44	0.50	0.46
fructose	1.55	1.09	1.14	0.90	1.28	1.33	0.93

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
galactonic acid	0.92	1.15	1.01	1.07	1.07	1.76	1.04
glucoheptulose	0.89	0.82	0.76	0.85	1.14	0.90	0.91
glycerol-3-galactoside	0.70	0.35	1.49	0.35	0.75	0.68	0.44
homocystine	0.57	0.49	0.85	0.56	1.44	0.75	0.46
levanbiose	1.07	0.97	0.93	0.82	1.11	0.93	1.02
oxoproline	1.07	0.83	0.99	0.86	0.60	1.17	0.98
tryptophan	1.04	0.96	0.99	0.76	0.78	0.94	0.89
Unk211945	1.01	0.99	0.78	1.15	0.67	0.85	0.99
Unk212022	0.84	0.71	1.14	1.00	1.18	0.82	0.88
Unk214165	0.67	0.88	0.43	0.53	0.20	0.48	0.51
Unk227728	1.58	0.99	1.07	1.08	0.97	0.99	1.04
Unk245705	0.68	0.53	0.62	0.64	0.52	0.54	0.64
Unk384918	0.87	0.84	1.02	1.08	0.88	1.07	0.89
Unk385034	1.13	0.80	1.90	0.53	0.90	1.19	0.55
xylonic acid	0.70	0.66	0.74	0.97	1.29	0.78	0.60
aconitic acid	0.01	0.01	0.01	0.02	0.01	0.01	0.01
adipic acid	0.02	0.02	0.03	0.03	0.03	0.02	0.02
alanine	0.60	0.48	0.46	0.66	0.48	0.55	0.75
alpha ketoglutaric acid	0.13	0.15	0.13	0.18	0.05	0.10	0.29
aminomalonic acid	0.04	0.03	0.03	0.04	0.04	0.03	0.02
arabinose	0.88	0.79	0.77	0.92	0.55	0.67	1.15
arabitol	0.49	0.41	0.45	0.41	0.69	0.45	0.42
arginine + ornithine	0.25	0.22	0.25	0.35	0.34	0.20	0.28
asparagine	1.01	0.97	1.37	0.88	0.94	1.07	0.89
aspartic acid	0.54	0.47	0.36	0.51	0.52	0.42	0.39
azelaic acid	0.35	0.38	0.29	0.36	0.29	0.48	0.36
beta-alanine	0.16	0.22	0.21	0.24	0.24	0.21	0.17
beta-gentiobiose	0.70	0.62	0.88	0.83	1.18	0.75	0.81

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
beta-sitosterol	0.04	0.05	0.05	0.06	0.06	0.05	0.07
capric acid	0.03	0.02	0.03	0.04	0.07	0.03	0.02
cellobiose	0.45	0.48	0.48	0.50	0.45	0.44	0.50
cellobiotol	0.26	0.18	0.14	0.18	0.14	0.15	0.17
citric acid	0.27	0.26	0.28	0.34	0.25	0.28	0.37
cyano-L-alanine	0.31	0.38	0.34	0.42	0.59	0.24	0.43
cysteine	1.45	1.28	1.18	1.17	1.24	1.32	1.16
cysteine-glycine	0.28	0.39	0.32	0.32	0.44	0.35	0.30
erythritol	0.37	0.25	0.53	0.41	0.40	0.35	0.39
ferulic acid	0.02	0.02	0.02	0.03	0.02	0.01	0.02
fumaric acid	0.43	0.38	0.42	0.40	0.56	0.44	0.51
GABA	0.09	0.07	0.10	0.12	0.04	0.07	0.11
galactinol1	0.51	0.50	0.62	0.68	0.64	0.64	0.66
galactinol2	0.00	0.00	0.00	0.00	0.00	0.00	0.00
glucose	0.66	0.59	0.51	0.73	0.54	0.48	0.66
glucuronic acid	0.30	0.32	0.38	0.30	0.36	0.36	0.32
glutamic acid	0.13	0.14	0.14	0.15	0.22	0.10	0.12
glutamine	0.42	0.28	0.28	0.27	0.18	0.28	0.20
glyceric acid	0.25	0.30	0.27	0.28	0.25	0.27	0.32
glycerol	0.65	0.68	0.61	0.64	0.77	0.52	0.56
glycerol-alpha-phosphate	0.14	0.16	0.19	0.23	0.28	0.20	0.24
glycine	0.74	0.95	0.74	0.68	0.51	1.03	0.74
glycolic acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00
homoserine	0.16	0.16	0.24	0.24	0.33	0.21	0.26
idonic acid NIST	0.31	0.27	0.37	0.26	0.30	0.26	0.35
inositol myo-	0.19	0.20	0.17	0.25	0.26	0.17	0.22
inositol-4-monophosphate	0.60	0.64	0.62	0.76	0.51	0.57	0.64
isocitric acid	0.30	0.32	0.31	0.40	0.34	0.32	0.31

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
isoleucine	0.76	0.72	0.64	0.62	0.67	0.72	0.59
itaconic acid	0.48	0.44	0.40	0.57	0.63	0.53	0.46
kynurenine	0.26	0.21	0.25	0.30	0.23	0.25	0.33
lactic acid	0.03	0.03	0.05	0.03	0.14	0.05	0.03
lauric acid	0.28	0.27	0.29	0.35	0.20	0.24	0.36
lysine	0.52	0.41	0.41	0.67	0.59	0.42	0.67
maleimide	0.23	0.22	0.25	0.28	0.26	0.26	0.26
malic acid	0.06	0.07	0.07	0.09	0.07	0.06	0.09
maltose	0.89	0.93	0.91	0.81	0.70	0.85	0.85
mannonic acid NIST	0.94	1.09	0.92	1.05	1.13	1.30	1.16
methionine sulfoxide	1.21	1.16	1.12	1.05	1.14	1.12	1.02
methylhexadecanoic acid	0.06	0.03	0.04	0.07	0.08	0.05	0.03
methylmaleic acid	0.09	0.10	0.16	0.10	0.29	0.10	0.14
myristic acid	0.08	0.09	0.12	0.12	0.18	0.09	0.06
N-acetyl-glutamic acid	0.84	0.86	0.83	0.97	0.79	0.92	0.92
N-methylalanine	0.47	0.52	0.35	0.35	0.48	0.37	0.29
O-acetylserine	0.89	0.74	1.01	0.97	0.92	0.75	0.93
oxalic acid	0.06	0.06	0.06	0.07	0.07	0.04	0.06
parabanic acid NIST	0.15	0.13	0.19	0.19	0.25	0.14	0.20
pentitol	0.09	0.06	0.07	0.10	0.06	0.09	0.12
phenylalanine	0.82	0.86	0.74	1.03	0.86	0.85	0.98
phosphoric acid	0.46	0.47	0.46	0.47	0.40	0.46	0.50
phthalic acid	0.25	0.29	0.23	0.21	0.23	0.24	0.19
proline	0.47	0.49	0.35	0.40	0.38	0.40	0.30
pyrazine 2,5-dihydroxy NIST	0.15	0.13	0.12	0.19	0.17	0.12	0.20
ribitol	0.15	0.10	0.12	0.15	0.14	0.14	0.17
saccharic acid	0.29	0.21	0.27	0.31	0.25	0.27	0.34

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
saccharopine	0.88	0.61	0.91	0.90	1.12	0.81	0.84
salicylic acid	0.35	0.25	0.27	0.40	0.27	0.23	0.47
serine	0.31	0.32	0.41	0.37	0.31	0.33	0.38
succinic acid	0.23	0.21	0.26	0.40	0.59	0.25	0.30
threitol	0.54	0.47	0.45	0.66	0.48	0.70	0.71
threonine	0.69	0.60	0.49	0.45	0.54	0.72	0.40
thymine	0.64	0.71	0.62	0.68	0.63	0.57	0.66
trehalose	0.83	0.89	0.89	1.10	1.00	0.78	0.89
Unk199177	0.03	0.02	0.04	0.03	0.06	0.05	0.03
Unk199203	0.27	0.20	0.21	0.31	0.21	0.23	0.24
Unk199205	0.14	0.07	0.12	0.12	0.17	0.13	0.13
Unk199242	0.02	0.03	0.03	0.01	0.01	0.01	0.02
Unk199246	0.05	0.03	0.05	0.05	0.05	0.03	0.07
Unk199942	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk200486	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Unk200490	0.17	0.12	0.08	0.08	0.13	0.17	0.15
Unk200509	0.16	0.46	0.07	0.10	0.04	0.14	0.07
Unk200540	0.16	0.19	0.16	0.17	0.25	0.15	0.18
Unk200557	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk200567	0.73	0.65	0.62	0.59	0.54	0.55	0.51
Unk200610	0.17	0.21	0.20	0.27	0.20	0.25	0.29
Unk200615	0.00	0.01	0.02	0.01	0.01	0.00	0.01
Unk200624	0.10	0.09	0.09	0.10	0.09	0.05	0.15
Unk200702	0.08	0.08	0.07	0.06	0.07	0.07	0.05
Unk200710	0.00	0.00	0.00	0.01	0.01	0.01	0.01
Unk200905	0.88	0.86	0.53	0.94	0.66	0.75	0.89
Unk200906	0.15	0.15	0.14	0.17	0.10	0.15	0.23
Unk201005	0.08	0.13	0.10	0.14	0.08	0.09	0.12
Unk201042	0.02	0.02	0.02	0.04	0.04	0.03	0.04
Unk202083	0.21	0.35	0.19	0.21	0.15	0.22	0.22

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk202572	0.33	0.31	0.24	0.38	0.62	0.26	0.29
Unk202573	0.21	0.20	0.22	0.24	0.20	0.18	0.27
Unk202599	0.09	0.14	0.07	0.08	0.06	0.05	0.10
Unk202899	0.07	0.08	0.19	0.07	0.13	0.11	0.09
Unk203592	0.01	0.00	0.02	0.00	0.02	0.01	0.01
Unk205672	0.00	0.01	0.01	0.00	0.01	0.01	0.01
Unk205674	0.06	0.06	0.06	0.07	0.07	0.03	0.08
Unk206022	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk206309	0.01	0.01	0.03	0.03	0.01	0.00	0.01
Unk207188	0.80	0.66	0.62	0.58	0.52	0.68	0.67
Unk208397	0.07	0.07	0.08	0.06	0.05	0.03	0.06
Unk208538	0.31	0.26	0.53	0.21	0.27	0.29	0.20
Unk208686	0.33	0.28	0.26	0.39	0.25	0.33	0.30
Unk208714	0.12	0.06	0.08	0.08	0.07	0.07	0.07
Unk208897	0.00	0.00	0.01	0.00	0.01	0.01	0.00
Unk211590	0.01	0.01	0.02	0.01	0.03	0.01	0.02
Unk211636	0.07	0.05	0.05	0.06	0.04	0.03	0.04
Unk211896	0.03	0.03	0.03	0.03	0.05	0.05	0.02
Unk211910	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk211916	0.02	0.01	0.02	0.01	0.01	0.03	0.03
Unk211919	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Unk211921	0.01	0.01	0.02	0.02	0.01	0.01	0.01
Unk212016	0.47	0.90	0.42	0.47	0.30	0.46	0.49
Unk212177	0.02	0.02	0.02	0.03	0.03	0.02	0.02
Unk212761	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk213141	0.02	0.02	0.03	0.03	0.01	0.00	0.04
Unk213143	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk213148	0.11	0.31	0.16	0.20	0.05	0.17	0.20
Unk213154	0.10	0.05	0.09	0.07	0.05	0.09	0.05
Unk213160	0.23	0.24	0.28	0.24	0.14	0.37	0.27

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk213179	0.03	0.04	0.03	0.02	0.02	0.05	0.03
Unk213188	0.08	0.07	0.07	0.09	0.09	0.10	0.06
Unk213191	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk213227	0.01	0.01	0.01	0.00	0.02	0.02	0.03
Unk213243	0.02	0.05	0.03	0.04	0.04	0.02	0.03
Unk213253	0.01	0.00	0.01	0.01	0.00	0.01	0.00
Unk213960	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Unk214152	0.34	0.37	0.21	0.41	0.31	0.31	0.38
Unk214201	0.02	0.02	0.03	0.06	0.03	0.02	0.05
Unk214416	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Unk215397	0.02	0.01	0.02	0.01	0.03	0.02	0.03
Unk215445	0.04	0.05	0.04	0.06	0.05	0.05	0.05
Unk216428	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk216832	0.16	0.12	0.20	0.15	0.17	0.20	0.20
Unk217809	0.02	0.01	0.02	0.01	0.07	0.02	0.00
Unk218492	0.09	0.13	0.09	0.11	0.08	0.18	0.12
Unk218694	0.06	0.04	0.06	0.05	0.04	0.06	0.05
Unk218734	0.83	1.00	0.62	0.55	0.32	0.53	0.65
Unk219507	0.19	0.19	0.16	0.15	0.09	0.18	0.18
Unk219512	0.01	0.01	0.00	0.01	0.01	0.00	0.01
Unk220122	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk223191	0.01	0.02	0.03	0.04	0.04	0.01	0.01
Unk223830	0.13	0.10	0.16	0.13	0.15	0.13	0.13
Unk224322	0.01	0.02	0.01	0.03	0.03	0.03	0.03
Unk224574	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk224589	0.00	0.01	0.00	0.02	0.01	0.00	0.01
Unk224627	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk224632	0.02	0.02	0.02	0.04	0.02	0.02	0.02
Unk224811	0.02	0.02	0.03	0.03	0.04	0.02	0.03
Unk224818	0.03	0.04	0.04	0.02	0.04	0.03	0.05

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk224843	0.20	0.13	0.16	0.18	0.15	0.18	0.22
Unk224849	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk225327	0.10	0.06	0.09	0.11	0.09	0.10	0.13
Unk225540	0.02	0.02	0.02	0.02	0.05	0.02	0.02
Unk225867	0.13	0.14	0.12	0.11	0.09	0.16	0.10
Unk226256	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Unk226841	0.06	0.06	0.08	0.07	0.07	0.05	0.05
Unk226848	0.19	0.14	0.17	0.18	0.20	0.21	0.15
Unk227598	0.02	0.03	0.02	0.01	0.01	0.02	0.02
Unk227652	0.19	0.17	0.17	0.21	0.12	0.14	0.19
Unk227658	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk227816	0.03	0.03	0.09	0.05	0.09	0.03	0.05
Unk227822	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Unk227923	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Unk228018	0.35	0.35	0.31	0.36	0.33	0.23	0.45
Unk228164	0.01	0.03	0.02	0.02	0.06	0.02	0.03
Unk228619	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Unk228680	0.14	0.10	0.12	0.11	0.11	0.07	0.10
Unk228911	0.06	0.05	0.03	0.04	0.04	0.03	0.04
Unk229277	0.01	0.01	0.01	0.01	0.03	0.01	0.01
Unk231098	0.01	0.02	0.01	0.02	0.01	0.01	0.01
Unk231210	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk231248	0.01	0.02	0.01	0.01	0.03	0.01	0.02
Unk231254	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk231260	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk231350	0.05	0.05	0.02	0.04	0.05	0.05	0.05
Unk231576	0.02	0.02	0.02	0.04	0.03	0.03	0.02
Unk232946	0.00	0.00	0.01	0.01	0.00	0.00	0.01
Unk233289	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk233408	0.02	0.03	0.05	0.02	0.04	0.04	0.05

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk233471	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk233790	0.10	0.11	0.11	0.12	0.13	0.11	0.13
Unk234717	0.00	0.01	0.00	0.01	0.01	0.01	0.01
Unk235327	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk235449	0.01	0.03	0.02	0.03	0.13	0.02	0.05
Unk236605	0.00	0.00	0.01	0.01	0.01	0.01	0.01
Unk236810	0.02	0.02	0.03	0.02	0.04	0.03	0.05
Unk236965	0.02	0.01	0.01	0.02	0.02	0.00	0.01
Unk237333	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Unk237392	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk237415	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk237520	0.03	0.03	0.05	0.04	0.02	0.03	0.05
Unk237605	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk237606	0.04	0.03	0.07	0.02	0.01	0.08	0.02
Unk237652	0.01	0.02	0.02	0.01	0.00	0.00	0.01
Unk238267 trisaccharide	0.01	0.03	0.03	0.02	0.03	0.01	0.03
Unk238506	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk238549	0.06	0.09	0.08	0.11	0.13	0.06	0.11
Unk238550	0.00	0.00	0.00	0.00	0.00	0.01	0.01
Unk238938	0.00	0.01	0.00	0.01	0.00	0.00	0.01
Unk239332	0.06	0.09	0.04	0.07	0.07	0.04	0.04
Unk240436	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk240551	0.01	0.01	0.01	0.01	0.02	0.02	0.01
Unk241168	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk241312	0.02	0.01	0.02	0.02	0.03	0.02	0.02
Unk241387	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk241881	0.02	0.03	0.01	0.00	0.00	0.00	0.00
Unk250380	0.12	0.13	0.09	0.17	0.25	0.09	0.12
Unk267649	0.03	0.04	0.04	0.04	0.04	0.01	0.02

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk267654	0.04	0.07	0.11	0.10	0.12	0.13	0.13
Unk267691	0.05	0.04	0.05	0.03	0.07	0.14	0.02
Unk267692	0.07	0.11	0.10	0.10	0.10	0.21	0.07
Unk267696	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk267751	0.48	0.36	0.44	0.48	0.95	0.52	0.41
Unk267880	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk267904	0.13	0.10	0.09	0.18	0.14	0.11	0.15
Unk267987	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk268345	0.40	0.46	0.51	0.53	0.44	0.29	0.38
Unk268353	0.05	0.06	0.07	0.05	0.12	0.04	0.04
Unk268365	0.14	0.14	0.08	0.13	0.13	0.11	0.13
Unk268437	0.04	0.05	0.08	0.07	0.08	0.05	0.06
Unk268438	0.02	0.01	0.01	0.01	0.01	0.01	0.02
Unk268506	0.05	0.04	0.04	0.07	0.05	0.04	0.05
Unk268565	0.17	0.19	0.25	0.14	0.23	0.17	0.14
Unk268583	0.04	0.09	0.07	0.03	0.04	0.05	0.02
Unk268610	0.15	0.13	0.17	0.17	0.22	0.18	0.18
Unk268671	0.20	0.18	0.14	0.17	0.22	0.18	0.19
Unk268712	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk269249	0.02	0.04	0.02	0.04	0.06	0.02	0.02
Unk269250	0.28	0.31	0.29	0.27	0.39	0.48	0.31
Unk269256	0.05	0.05	0.02	0.05	0.03	0.03	0.04
Unk269294	0.29	0.23	0.25	0.35	0.29	0.26	0.28
Unk269776	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk271049	0.04	0.02	0.07	0.05	0.10	0.06	0.06
Unk280546	0.02	0.02	0.03	0.02	0.03	0.02	0.03
Unk280930	0.10	0.07	0.08	0.08	0.06	0.08	0.10
Unk280945	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk281132	0.12	0.10	0.14	0.15	0.14	0.10	0.15
Unk281187	0.00	0.01	0.02	0.02	0.03	0.02	0.01

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk284607	0.01	0.02	0.04	0.03	0.05	0.02	0.02
Unk285340	0.05	0.04	0.05	0.05	0.05	0.05	0.05
Unk288810	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk288866	0.18	0.18	0.15	0.14	0.09	0.14	0.18
Unk294511	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk296071	0.13	0.17	0.12	0.15	0.19	0.10	0.16
Unk296119	0.00	0.02	0.02	0.02	0.03	0.01	0.01
Unk299185	0.05	0.07	0.05	0.07	0.07	0.05	0.07
Unk299441	0.82	0.84	0.91	1.08	0.51	0.59	1.04
Unk299487	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk303838	0.04	0.08	0.05	0.08	0.12	0.07	0.09
Unk303839	0.20	0.26	0.27	0.25	0.34	0.24	0.18
Unk303956	0.04	0.04	0.05	0.04	0.07	0.06	0.05
Unk303966	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk303992	0.06	0.09	0.08	0.11	0.14	0.04	0.12
Unk304391	0.00	0.01	0.00	0.01	0.01	0.00	0.00
Unk305637	0.00	0.01	0.00	0.00	0.00	0.00	0.01
Unk307669	0.00	0.01	0.01	0.01	0.01	0.02	0.01
Unk307912	0.04	0.01	0.04	0.04	0.04	0.03	0.03
Unk307924	0.24	0.22	0.32	0.09	0.11	0.29	0.17
Unk308219	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk309617	0.00	0.01	0.01	0.02	0.01	0.00	0.01
Unk310006	0.04	0.05	0.05	0.04	0.02	0.04	0.04
Unk310053	0.12	0.12	0.11	0.14	0.17	0.11	0.15
Unk310063	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk310162	0.04	0.02	0.04	0.03	0.04	0.03	0.03
Unk310193	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk310367	0.01	0.02	0.00	0.01	0.02	0.01	0.01
Unk310380	0.20	0.18	0.15	0.20	0.37	0.16	0.20
Unk310448	0.00	0.00	0.01	0.01	0.01	0.01	0.00

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk310581	0.02	0.04	0.04	0.05	0.03	0.02	0.04
Unk310757	0.04	0.03	0.04	0.03	0.04	0.05	0.02
Unk310871	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Unk310875	0.04	0.05	0.04	0.06	0.13	0.03	0.06
Unk310888	0.15	0.10	0.10	0.13	0.21	0.11	0.11
Unk310897	0.02	0.02	0.01	0.02	0.02	0.01	0.03
Unk311041	0.06	0.03	0.03	0.05	0.04	0.02	0.04
Unk318770	0.11	0.11	0.10	0.10	0.12	0.09	0.12
Unk319168	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Unk321749	0.04	0.05	0.04	0.05	0.08	0.05	0.04
Unk322204	0.01	0.01	0.00	0.01	0.00	0.00	0.00
Unk322260	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Unk323686	0.14	0.20	0.12	0.12	0.24	0.12	0.12
Unk324275	0.07	0.08	0.12	0.07	0.10	0.07	0.12
Unk327143	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk327468	0.01	0.00	0.01	0.01	0.00	0.01	0.01
Unk328420	0.02	0.02	0.05	0.03	0.06	0.03	0.03
Unk330992	0.56	0.50	0.66	0.50	0.44	0.62	0.96
Unk331031	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Unk352777	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk352812	0.01	0.00	0.00	0.00	0.02	0.00	0.00
Unk352849	0.09	0.06	0.08	0.06	0.09	0.08	0.07
Unk352980	0.07	0.08	0.08	0.08	0.08	0.04	0.07
Unk353047	0.06	0.04	0.07	0.08	0.08	0.07	0.12
Unk353091	0.04	0.04	0.04	0.07	0.06	0.04	0.05
Unk357502	0.05	0.10	0.06	0.10	0.05	0.04	0.16
Unk357685	0.06	0.06	0.06	0.05	0.05	0.07	0.05
Unk357788	0.41	0.98	0.60	0.63	0.30	0.68	0.77
Unk357841	0.09	0.11	0.06	0.10	0.09	0.09	0.10
Unk359567	0.02	0.02	0.03	0.04	0.04	0.02	0.02

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk359697	0.06	0.07	0.09	0.07	0.08	0.10	0.05
Unk362008	0.02	0.04	0.04	0.02	0.08	0.06	0.04
Unk362010	0.07	0.05	0.06	0.02	0.02	0.04	0.01
Unk362023	0.04	0.02	0.05	0.06	0.18	0.04	0.04
Unk362028	0.15	0.08	0.14	0.13	0.16	0.15	0.17
Unk362073	0.00	0.01	0.00	0.02	0.01	0.02	0.01
Unk362086	0.02	0.03	0.02	0.02	0.01	0.03	0.03
Unk362109	0.04	0.03	0.05	0.03	0.08	0.06	0.03
Unk362130	0.19	0.19	0.20	0.28	0.17	0.17	0.22
Unk367914	0.10	0.11	0.10	0.14	0.08	0.07	0.15
Unk367939	0.14	0.13	0.12	0.12	0.07	0.12	0.16
Unk367944	0.01	0.01	0.02	0.01	0.03	0.01	0.00
Unk368156	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk371568	0.14	0.17	0.19	0.16	0.14	0.17	0.23
Unk373725	0.60	0.60	0.64	0.66	1.02	0.65	0.78
Unk373752	0.02	0.03	0.03	0.05	0.04	0.02	0.04
Unk374356	0.33	0.36	0.35	0.39	0.30	0.39	0.43
Unk374402	0.16	0.12	0.15	0.14	0.20	0.13	0.20
Unk374786	0.17	0.20	0.17	0.09	0.15	0.23	0.12
Unk375029	0.36	0.31	0.23	0.29	0.19	0.28	0.30
Unk381469	0.06	0.06	0.05	0.09	0.05	0.05	0.08
Unk384992	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk385006	0.01	0.01	0.01	0.01	0.01	0.01	0.02
Unk385021	0.33	0.28	0.31	0.40	0.65	0.36	0.52
Unk385023	0.06	0.05	0.08	0.07	0.06	0.10	0.08
Unk385024	0.02	0.02	0.01	0.02	0.03	0.02	0.01
Unk385028	0.14	0.09	0.13	0.16	0.12	0.13	0.15
Unk385030	0.08	0.07	0.05	0.07	0.06	0.05	0.09
Unk385042	0.06	0.04	0.03	0.06	0.03	0.05	0.08
Unk385048	0.87	0.74	0.85	0.84	0.73	0.85	0.75

Compound Name	Acidimicrobiales	Actinomycetales	Bacillales	Nitrospirales	Pseudomonadales	Rhizobiales	Xanthomonadales
Unk385055	0.03	0.05	0.03	0.03	0.06	0.03	0.02
Unk385058	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unk385065	0.02	0.05	0.03	0.03	0.02	0.03	0.06
Unk385075	0.00	0.00	0.01	0.01	0.00	0.01	0.01
Unk385085	0.12	0.06	0.06	0.05	0.05	0.06	0.09
Unk385104	0.02	0.02	0.05	0.04	0.02	0.06	0.04
Unk385107	0.10	0.10	0.11	0.12	0.14	0.09	0.10
Unk385112	0.02	0.02	0.01	0.01	0.01	0.02	0.02
Unk385117	0.02	0.03	0.03	0.04	0.05	0.02	0.03
Unk385120	0.01	0.00	0.01	0.02	0.05	0.01	0.02
urea	0.23	0.25	0.21	0.33	0.40	0.33	0.34
valine	0.58	0.49	0.51	0.75	0.50	0.44	0.70
1-kestose	0.41	0.44	0.39	0.52	0.47	0.45	0.46
2,3-dihydroxybutanoic acid NIST	0.11	0.17	0.06	0.13	0.06	0.09	0.21
2-5-diketopiperazine NIST	0.01	0.01	0.01	0.00	0.00	0.00	0.00
2-deoxyerythritol	0.42	0.45	0.35	0.50	0.31	0.39	0.36
2-hydroxyvaleric acid	0.07	0.05	0.05	0.08	0.06	0.06	0.09
3-aminoisobutyric acid	0.10	0.09	0.21	0.11	0.14	0.10	0.06
3-hydroxypropionic acid	0.19	0.16	0.17	0.25	0.37	0.27	0.14
4-hydroxybenzoate	0.20	0.23	0.13	0.18	0.29	0.29	0.22
xylitol	0.37	0.34	0.44	0.33	0.98	0.49	0.39
xylonolactone NIST	0.65	0.64	1.16	0.61	1.08	0.82	0.67
xylose	0.87	0.84	0.81	0.91	0.60	0.83	1.06

Figures

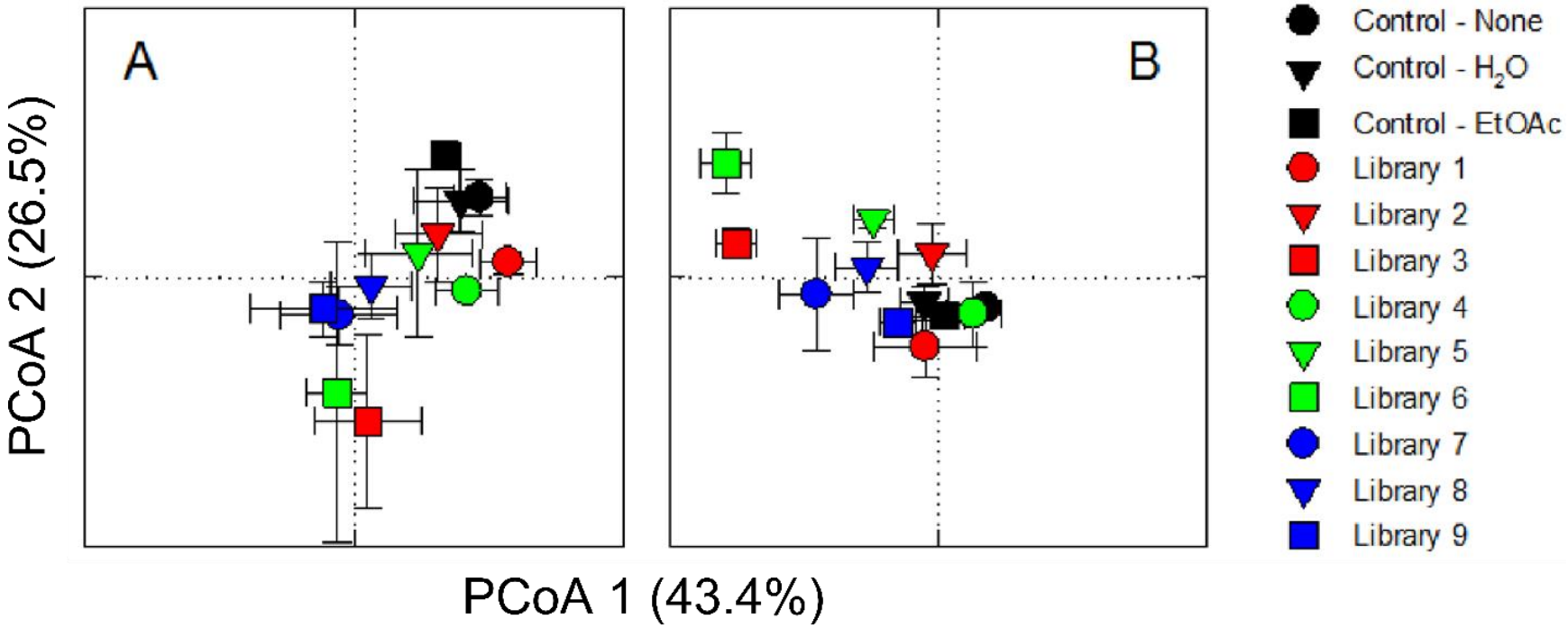

Figure 5-1. Principle coordinates analysis (PCoA) of the weighted UniFrac values for the visualization of the soil microbial community after exposure to the chemical libraries analyzed by 454 pyrosequencing. A-After 2 weeks of exposure. B- After 6 weeks of exposure.

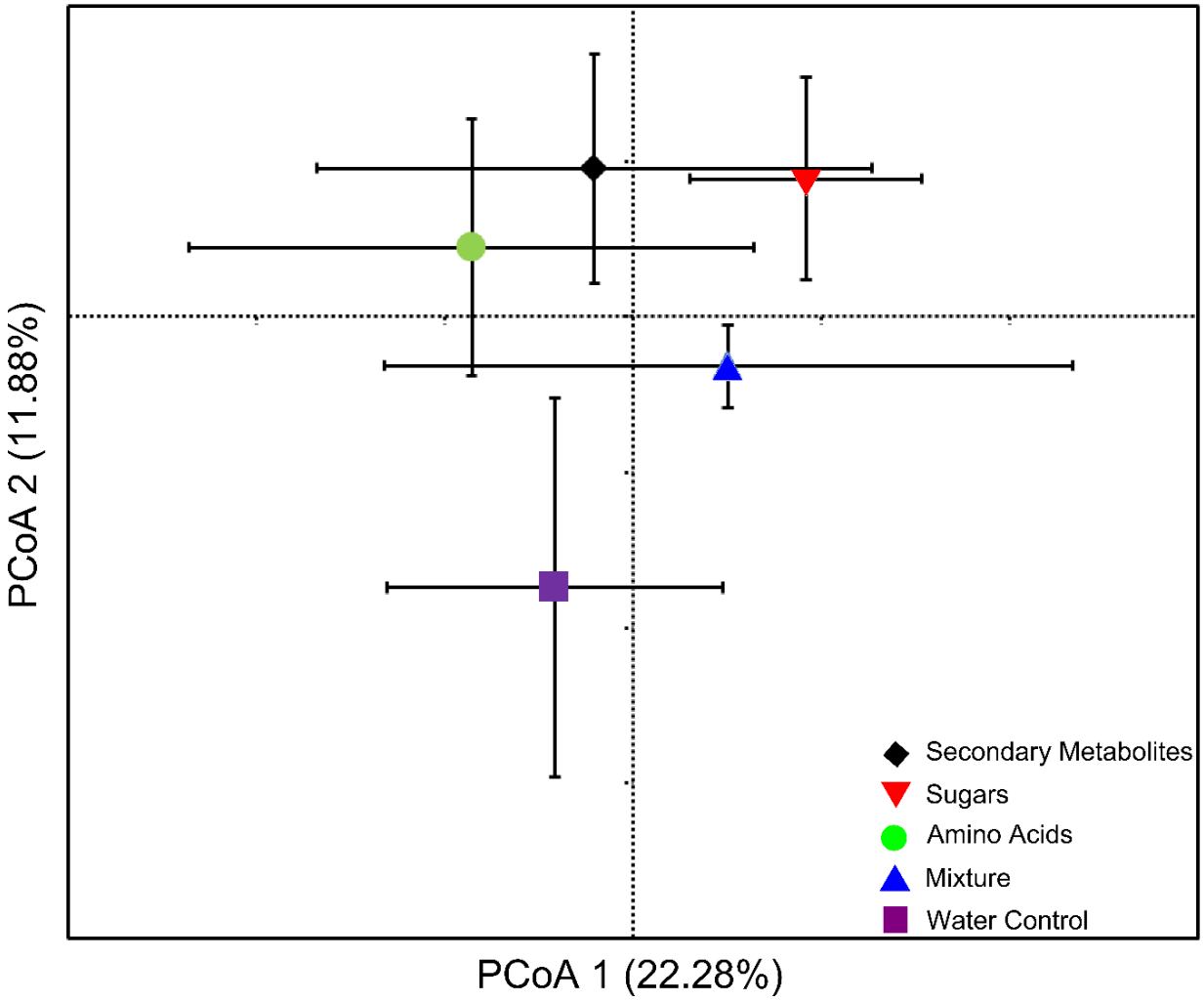

Figure 5-2. Principle coordinates analysis (PCoA) of the weighted UniFrac values for the visualization of the soil microbial community after exposure to artificial blends analyzed by 454 pyrosequencing.

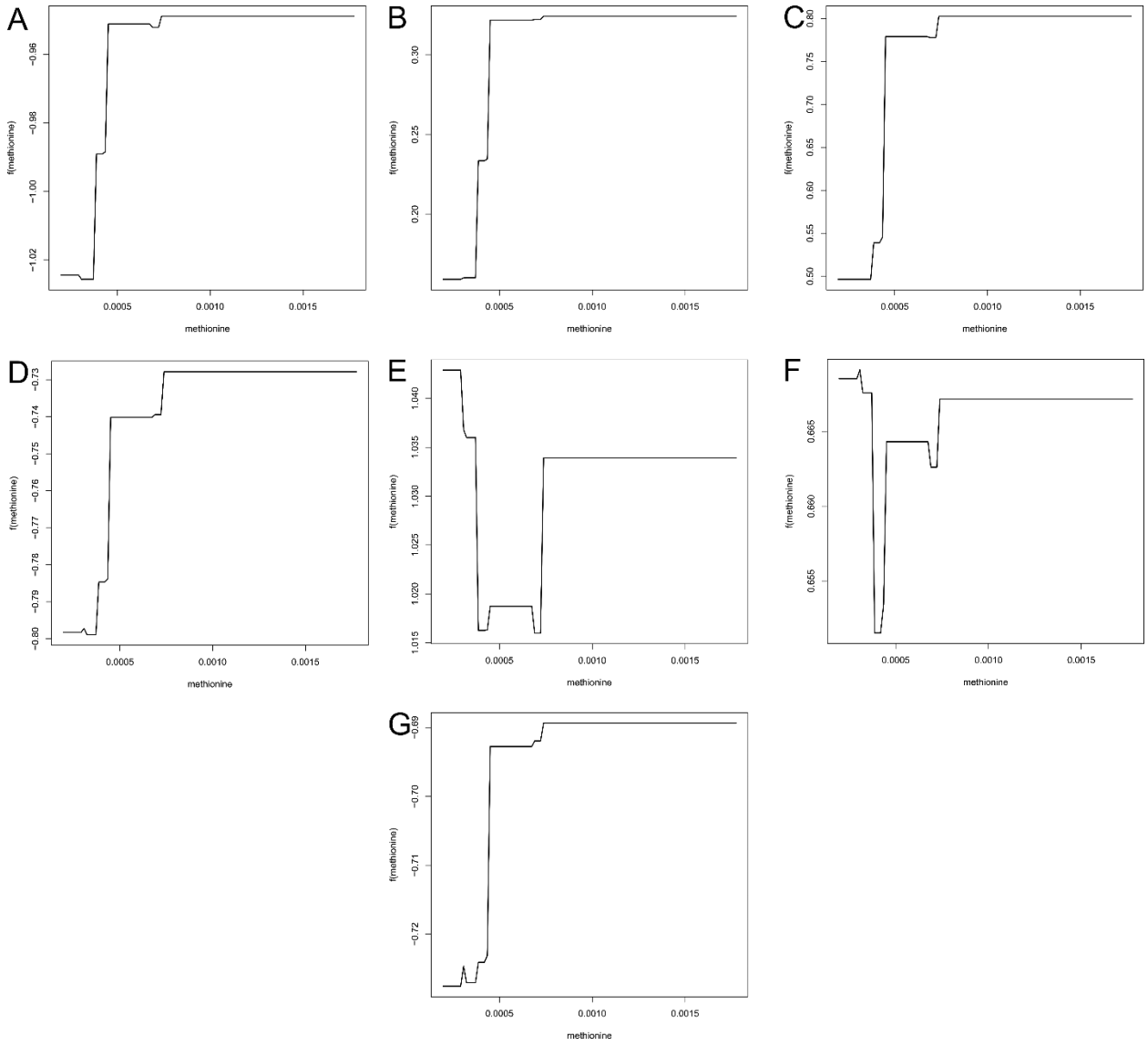

Figure 5-3. Partial dependency plots for boosted decision tree analyses identifies how methionine, after accounting for the average effect of all other variables in the model, influences bacterial relative abundance of A- Acidobacteria, B- Actinobacteria, C- Bacillales, D- Nitrospirales, E- Pseudomonadales, F- Rhizobiales, and G- Xanthomonadales.

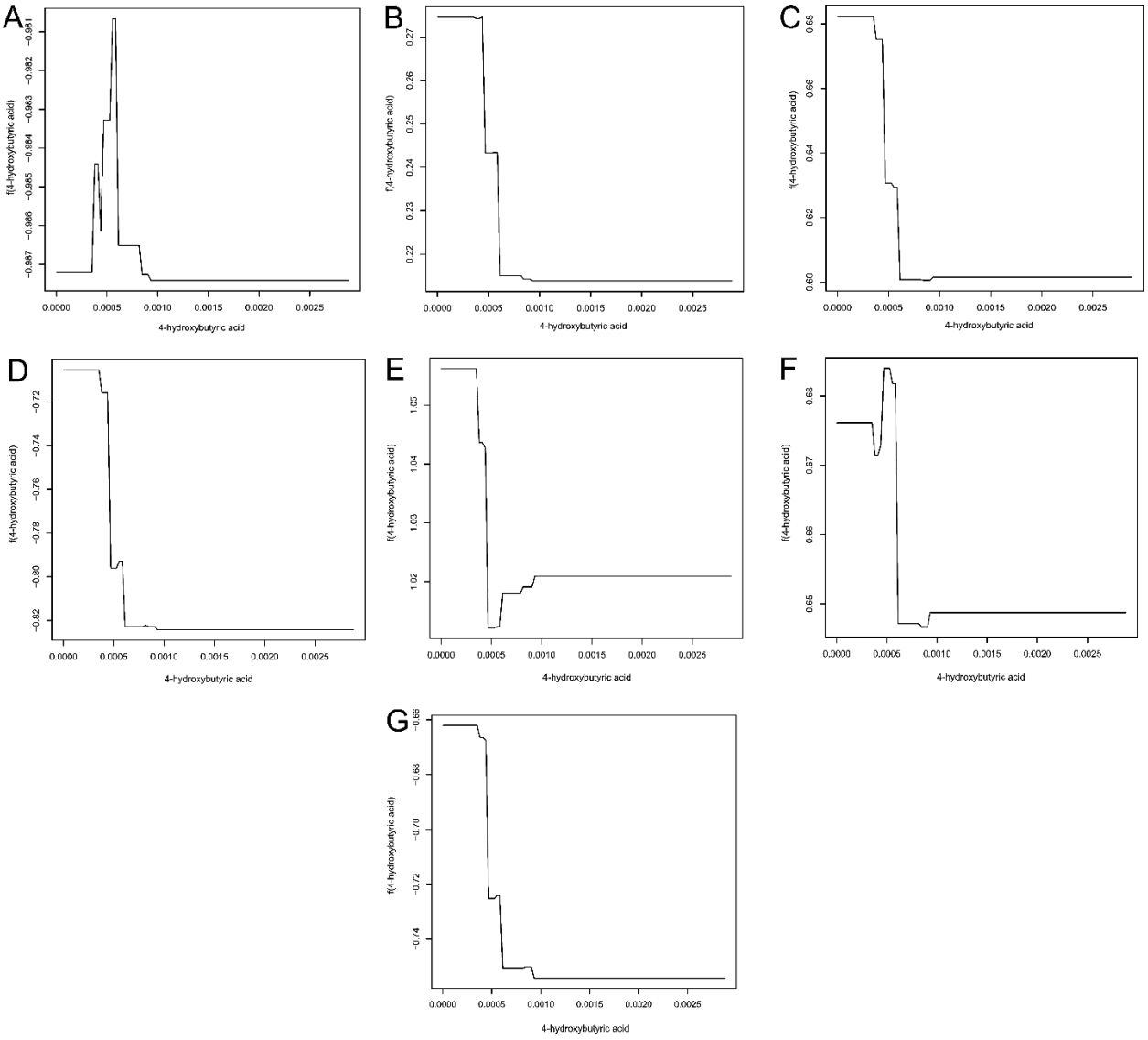

Figure 5-4. Partial dependency plots for boosted decision tree analyses identifies how 4-hydroxybutyric acid, after accounting for the average effect of all other variables in the model, influences bacterial relative abundance of A- Acidobacteria, B- Actinobacteria, C- Bacillales, D- Nitrospirales, E- Pseudomonadales, F- Rhizobiales, and G- Xanthomonadales.

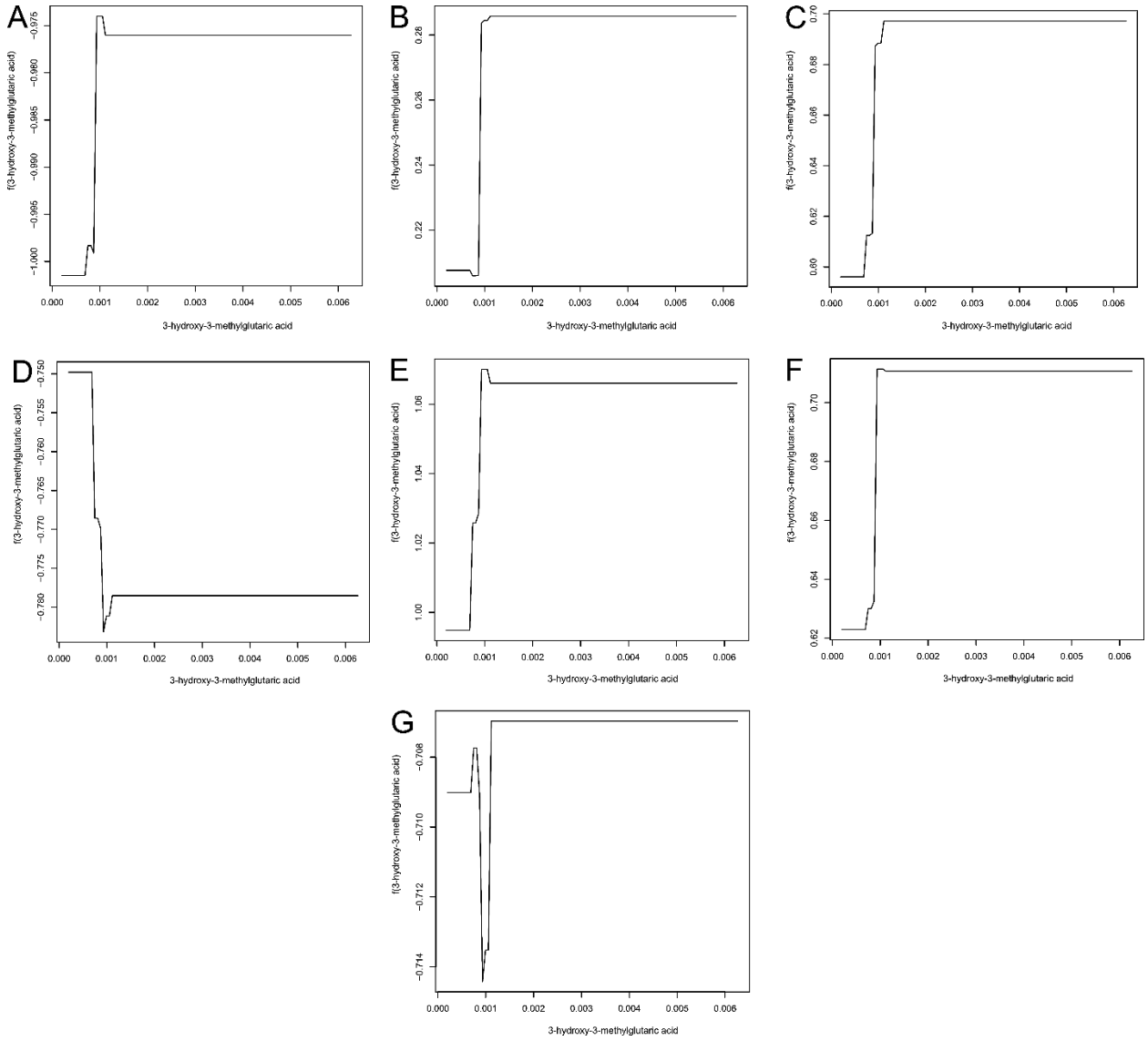

Figure 5-5. Partial dependency plots for boosted decision tree analyses of the identifies how 3-hydroxy-3-methylglutaric acid, after accounting for the average effect of all other variables in the model, influences bacterial relative abundance of A- Acidobacteria, B- Actinobacteria, C- Bacillales, D- Nitrospirales, E- Pseudomonadales, F- Rhizobiales, and G- Xanthomonadales.

REFERENCES

Babalola OO (2010). Beneficial bacteria of agricultural importance. *Biotechnol Lett* **32:** 1559-1570.

Badri DV, Loyola-Vargas VM, Broeckling CD, De-la-Pena C, Jasinski M, Santelia D *et al* (2008). Altered profile of secondary metabolites in the root exudates of Arabidopsis ATP-binding cassette transporter mutants. *Plant Physiol* **146:** 762-771.

Badri DV, Quintana N, El Kassis EG, Kim HK, Choi YH, Sugiyama A *et al* (2009). An ABC transporter mutation alters root exudation of phytochemicals that provoke an overhaul of natural soil microbiota. *Plant Physiol* **151:** 2006-2017.

Badri DV, De-la-Pena C, Lei Z, Manter DK, Chaparro JM, Guimaraes RL *et al* (2012). Root Secreted Metabolites and Proteins Are Involved in the Early Events of Plant-Plant Recognition Prior to Competition. *PLoS ONE* **7:** e46640.

Badri DV, Chaparro JM, Zhang R, Shen Q, Vivanco JM (2013a). Application of natural blends of phytochemicals derived from the root exudates of Arabidopsis to the soil reveal that phenolic related compounds predominantly modulate the soil microbiome. *J Biol Chem* **288:** 4502-4512.

Badri DV, Zolla G, Bakker MG, Manter DK, Vivanco JM (2013b). Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. *New Phytol*.

Bajzer M, Seeley RJ (2006). Physiology: obesity and gut flora. *Nature* **444:** 1009-1010.

Berlec A (2012). Novel techniques and findings in the study of plant microbiota: Search for plant probiotics. *Plant Sci* **193:** 96-102.

Blaser M (2011). Antibiotic overuse: Stop the killing of beneficial bacteria. *Nature* **476:** 393-394.

Chaparro JM, Badri DV, Bakker MG, Sugiyama A, Manter DK, Vivanco JM (2013a). Root exudation of phytochemicals in Arabidopsis follows specific patterns that are developmentally programmed and correlate with soil microbial functions. *PLOS ONE* **8:** e55731.

Chaparro JM, Dayakar BV, Vivanco JM (2013b). Rhizosphere microbiome assemblage is affected by plant development. *Isme J*.

Chaparro JM, Badri DV, Vivanco JM (2014). Rhizosphere microbiome assemblage is affected by plant development. *ISME J* **8:** 790-803.

Chigineva NI, Aleksandrova AV, Tiunov AV (2009). The addition of labile carbon alters litter fungal communities and decreases litter decomposition rates. *Applied Soil Ecology* **42:** 264-270.

Consortium THMP (2012). Structure, function and diversity of the healthy human microbiome. *Nature* **486:** 207-214.

Cotter PD, Ross RP, Hill C (2013). Bacteriocins - a viable alternative to antibiotics? *Nat Rev Microbiol* **11:** 95-105.

Dakora F, Phillips D (2002). Root exudates as mediators of mineral acquisition in low-nutrient environments. *Plant Soil* **245:** 35-47.

Dethlefsen L, Huse S, Sogin ML, Relman DA (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol* **6:** e280.

Dethlefsen L, Relman DA (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci U S A* **108 Suppl 1:** 4554-4561.

Dimkpa C, Weinand T, Asch F (2009). Plant-rhizobacteria interactions alleviate abiotic stress conditions. *Plant Cell Environ* **32:** 1682-1694.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

Elith J, Leathwick JR, Hastie T (2008). A working guide to boosted regression trees. *J Anim Ecol* **77:** 802-813.

Fiehn O, Wohlgemuth G, Scholz M (2005). Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. In: Ludascher B, Raschid L (eds). *Data Integration In The Life Sciences, Lecture Notes in Computer Science*. Springer-Verlag: Berlin/Heidelberg. pp 224-239.

Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee do Y, Lu Y *et al* (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J* **53:** 691-704.

Friedman J, Hastie T, Tibshirani R (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* **33:** 1-22.

Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K *et al* (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469:** 543-547.

Horiuchi J, Prithiviraj B, Bais HP, Kimball BA, Vivanco JM (2005). Soil nematodes mediate positive interactions between legume plants and rhizobium bacteria. *Planta* **222:** 848-857.

Kirankumar R, Jagadeesh KS, Krishnaraj PU, Patil MS (2008). Enhanced growth promotion of tomato and nutrient uptake by plant growth promoting rhizobacterial isolates in presence of tobacco mosaic virus pathogen. *Karnataka J Agric Sci* **21:** 309-311.

Kristin A, Miranda H (2013). The root microbiota—a fingerprint in the soil? *Plant and Soil* **370:** 671-686.

Larsen PB, Degenhardt J, Tai CY, Stenzler LM, Howell SH, Kochian LV (1998). Aluminum-resistant Arabidopsis mutants that exhibit altered patterns of aluminum accumulation and organic acid release from roots. *Plant Physiol* **117:** 9-18.

Li X, Bergelson J, Chapple C (2010). The ARABIDOPSIS Accession Pna-10 Is a Naturally Occurring sng1 Deletion Mutant. *Mol Plant* **3:** 91-100.

Liaw A, Wiener M (2002). Classification and Regression by randomForest. *R News* **2:** 18-22.

Lozupone CA, Hamady M, Kelley ST, Knight R (2007). Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Appl Environ Microbiol* **73:** 1576-1585.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332:** 1097-1100.

Merow C, Smith MJ, Edwards TC, Guisan A, McMahon SM, Normand S *et al* (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography* **37:** 1267-1281.

Mevik B-H, Wehrens R (2007). The pls Package: Principal Component and Partial Least Squares Regression in R. *J Stat Softw* **18:** 1-24.

Modi SR, Collins JJ, Relman DA (2014). Antibiotics and the gut microbiota. *The Journal of clinical investigation* **124:** 4212-4218.

Morrissey JP, Dow JM, Mark GL, O'Gara F (2004). Are microbes at the root of a solution to world food production? Rational exploitation of interactions between microbes and plants can help to transform agriculture. *EMBO Rep* **5:** 922-926.

Naushad HS, Gupta RS (2013). Phylogenomics and molecular signatures for species from the plant pathogen-containing order xanthomonadales. *PLoS One* **8:** e55216.

Olden JD, Lawler JJ, Poff NL (2008). Machine learning methods without tears: A primer for ecologists. *Q Rev Biol* **83:** 171-193.

Palaniyandi SA, Yang SH, Zhang L, Suh JW (2013). Effects of actinobacteria on plant disease suppression and growth promotion. *Appl Microbiol Biotechnol* **97:** 9621-9636.
Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J *et al* (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35:** 7188-7196.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12:** 38.

R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria.

Ramos Solano B, Barriuso Maicas J, Pereyra de la Iglesia MT, Domenech J, Gutierrez Manero FJ (2008). Systemic Disease Protection Elicited by Plant Growth Promoting Rhizobacteria Strains: Relationship Between Metabolic Responses, Systemic Disease Protection, and Biotic Elicitors. *Phytopathology* **98:** 451-457.

Rastall RA, Gibson GR, Gill HS, Guarner F, Klaenhammer TR, Pot B *et al* (2005). Modulation of the microbial ecology of the human colon by probiotics, prebiotics and synbiotics to enhance human health: an overview of enabling science and potential applications. *FEMS Microbiol Ecol* **52:** 145-152.

Reid G, Younes JA, Van der Mei HC, Gloor GB, Knight R, Busscher HJ (2011). Microbiota restoration: natural and supplemented recovery of human microbial communities. *Nat Rev Microbiol* **9:** 27-38.

Ridgeway G (2013). gbm: Generalized Boosted Regression Models, R package version 2.1 edn.

Ripley B (2014). tree: Classification and regression trees, R package version 1.0-35 edn.

Rudrappa T, Czymmek KJ, Pare PW, Bais HP (2008). Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol* **148:** 1547-1556.

Sana TR, Fischer S, Wohlgemuth G, Katrekar A, Jung KH, Ronald PC *et al* (2010). Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen Xanthomonas oryzae pv. oryzae. *Metabolomics* **6:** 451-465.

Saravanakumar D, Samiyappan R (2007). ACC deaminase from Pseudomonas fluorescens mediated saline resistance in groundnut (Arachis hypogea) plants. *J Appl Microbiol* **102:** 1283-1292.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75:** 7537-7541.

Schrezenmeir J, de Vrese M (2001). Probiotics, prebiotics, and synbiotics--approaching a definition. *The American journal of clinical nutrition* **73:** 361S-364S.

Selvakumar G, Panneerselvam P, Ganeshamurthy AN, Maheshwari DK (2012). Bacterial Mediated Alleviation of Abiotic Stress in Crops. In: Maheshwari DK (ed). *Bacteria in Agrobiology: Stress Management*. Springer: New York, NY. pp 205-224.

Sjolund M, Wreiber K, Andersson DI, Blaser MJ, Engstrand L (2003). Long-term persistence of resistant Enterococcus species after antibiotics to eradicate Helicobacter pylori. *Ann Intern Med* **139:** 483-487.

Stokes ME, McCourt P (2014). Towards personalized agriculture: what chemical genomics can bring to plant biotechnology. *Front Plant Sci* **5:** 344.

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI (2007). The human microbiome project. *Nature* **449:** 804-810.

Turner TR, Ramakrishnan K, Walshaw J, Heavens D, Alston M, Swarbreck D *et al* (2013). Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *Isme J*.

Walters WA, Caporaso JG, Lauber CL, Berg-Lyons D, Fierer N, Knight R (2011). PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics* **27:** 1159-1161.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *App Environ Microb* **73:** 5261-5267.

Willner D, Furlan M, Schmieder R, Grasis JA, Pride DT, Relman DA *et al* (2011). Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity. *Proc Natl Acad Sci U S A* **108 Suppl 1:** 4547-4553.

Yang J, Kloepper JW, Ryu CM (2009). Rhizosphere bacteria help plants tolerate abiotic stress. *Trends Plant Sci* **14:** 1-4.

Yoneyama K, Xie X, Kusumoto D, Sekimoto H, Sugimoto Y, Takeuchi Y *et al* (2007a). Nitrogen deficiency as well as phosphorus deficiency in sorghum promotes the production and exudation of 5-deoxystrigol, the host recognition signal for arbuscular mycorrhizal fungi and root parasites. *Planta* **227:** 125-132.

Yoneyama K, Yoneyama K, Takeuchi Y, Sekimoto H (2007b). Phosphorus deficiency in red clover promotes exudation of orobanchol, the signal for mycorrhizal symbionts and germination stimulant for root parasites. *Planta* **225:** 1031-1038.

Zhang F, Shen J, Zhang J, Zuo Y, Li L, Chen X (2010). Chapter One - Rhizosphere Processes and Management for Improving Nutrient Use Efficiency and Crop Productivity: Implications for China. In: Sparks DL (ed). *Advances in Agronomy*. Academic Press. p 1.

Zhang J, Subramanian S, Stacey G, Yu O (2009). Flavones and flavonols play distinct critical roles during nodulation of Medicago truncatula by Sinorhizobium meliloti. *Plant J* **57:** 171-183.

Zhao L (2010). Genomics: The tale of our other genome. *Nature* **465:** 879-880.

Zolla G, Badri DV, Bakker MG, Manter DK, Vivanco JM (2013). Soil microbiomes vary in their ability to confer drought tolerance to Arabidopsis. *App Soil Ecol* **68:** 1.

CHAPTER 6 CONCLUSIONS AND FUTURE WORK

Conclusions

Host associated microbial communities are fundamental for host health (Caporaso et al 2011, Chaparro et al 2012, Consortium 2012, Dethlefsen and Relman 2011, Dimkpa et al 2009, Marcobal et al 2013, Selvakumar et al 2012, Yang et al 2009). Furthermore, a compromised associated microbial community may result in disease (Bakker et al 2012, Berendsen et al 2012, Consortium 2012, Mercado-Blanco and Bakker 2007, Sugiyama et al 2010, Turnbaugh et al 2007, van Loon et al 1998). Plant associated microbial communities, specifically those in the rhizosphere, are increasingly important in helping overcome biotic (Berendsen et al 2012, Mendes et al 2011) and abiotic stresses and increasing productivity. In this study, analysis of the fundamental mechanisms of plant-microbiome interactions was undertaken. This research was accomplished to understand how to exploit soil microbial communities and provide plants with healthy soils to increase plant health and productivity. Furthermore, results from these studies may be extended to crops of agricultural importance.

Understanding how plants and the soil microbiome interact naturally without the addition of biotic or abiotic stresses is crucial to our understanding of the essential mechanisms at play in plant-microbiome interactions. Our initial foray into the interactions between plants and their rhizosphere microbiome identified root exudation as significant in initiating, establishing, and enhancing these exchanges. Furthermore, results indicate that there are patterns in root exudation. Qualitative and quantitative changes were found in root exudation that are associated with plant development. Significant correlations were also observed between root exudation

patterns and the functional capacity of the microbiome. Overall these results suggest that root exudation profiles follow a developmental pattern that is genetically programmed.

Analysis of the members making up the *Arabidopsis* rhizosphere microbiome reveal that *Arabidopsis* develops a core rhizosphere microbiome. Certain members that make up this core microbiome change with plant development, which suggests that the plant can select a subset of microbes through development for specific functions. For example, we observed that Actinobacteria were significantly more abundant at early developmental stages while bacteria classified as Cyanobacteria were more abundant at later developmental stages. This is noteworthy as plants are more susceptible to disease early in development and Actinobacteria are associated with disease suppression. On the other hand, Cyanobacteria are known plant growth promoters and are an important source of inorganic nitrogen at a point in development when the plant needs the most nitrogen. Overall, these results suggest that the plant maintains control over the rhizosphere microbiome enabling it to enhance and repress specific interactions as necessary. Furthermore, the plant secretes blends of compounds and specific phytochemicals that are differentially produced at distinct stages of plant growth to help orchestrate these activities. In other words, plants and the rhizomicrobiome are in constant communication through the exchange of signals.

Chapters 2 and 3 explores how through the use of root exudation, plants shape and modulate both the functionality and identity of their rhizosphere microbiome at their behest. In Chapter 4, root exudates are added to the soil in the absence of the plant to determine how these compounds on their own influence the soil microbial community. Root exudates were fractionated to obtain distinct blends of root exudate compounds. The addition of these blends to the soil caused distinct changes in the soil microbial community. Furthermore, one

phytochemical can culture different microbial species, mixtures of these phytochemicals can synergistically culture groups of microbial species and the same phytochemical can enhance or deter the abundance of different microbial species. Additionally, phenolic compounds more readily modulate the soil microbial community. This suggests that secondary metabolites could be used as candidate compounds for controlling and changing soil microbial communities.

The knowledge gained from studying the interaction of plants and their rhizosphere microbial communities was used to model these interactions and complexities. Nine chemical libraries derived from root exudates were supplemented to the soil. The resulting soil microbial communities, after exposure to these chemical libraries, were used in statistical modeling. Five statistical models were evaluated to identify a modeling approach that could be used to accurately predict soil microbial community dynamics upon exposure to natural chemical compounds. After model training and validation, models based on machine learning reliably predicted soil microbial community dynamics. This modeling approach was able to identify specific compounds and classes of compounds (secondary metabolites) that could be used as a means to correctly forecast soil microbial community behavior and thus identify compounds that could be used to enhance and create healthy soils for increased crop production.

Future directions

The ability to manipulate and enhance soil microbial communities can only be achieved once there is a comprehensive understanding of the dynamics that occur between plants and their rhizosphere microbiome. In this study, the model organism *Arabidopsis thaliana* and its associated rhizosphere microbial community was used to provide a framework to understand the mechanisms and complexities that occur in this niche. The knowledge gained was then used to produce a model that would help identify compounds that could be used to modulate the soil

microbial community in a predictive manner. Similar studies need to be conducted on different plant species, soil types, and soil microbial communities to determine whether these results and predictions could be extended to other soil types and plant species. Recent developments in the identification of the rhizosphere microbial communities of important agricultural crops such as rice (Edwards et al 2015), maize (Peiffer et al 2013), and barley (Bulgarelli et al 2015), provide a means of evaluating the microbial species that are important for increased crop health and productivity. Utilizing this information and applying machine learning will produce new models tailored to each plant species and soil type. This is essential to understand which microbial species need to be altered and/or manipulated with respect to each crop studied.

Additional studies should also focus on the functional capacity of the soil microbiome since the functional redundancy of the microbiome may provide a more reliable way to predict plant health and productivity (Chaparro et al 2012, Nannipieri et al 2003, 2008). Such redundancy would allow potential models and discoveries to be implemented and extended to a wide array of plant species.

One of the benefits of using statistical modeling based on machine learning is the ability to teach and refine your model as new data is acquired. Analyzing how new chemical compounds and blends of these interact with distinct soil microbial communities will enable the implementation of new, refined, and more sophisticated strategies at a higher level of resolution to help manage the soil microbiome. Additionally, combining all the information gathered across plant species, soil type, and soil microbial communities will result in a more robust model that can be utilized across the world to help remediate poor soils enabling the growth of crops in marginal lands.

REFERENCES

Bakker MG, Manter DK, Sheflin AM, Weir TL, Vivanco JM (2012). Harnessing the rhizosphere microbiome through plant breeding and agricultural management. *Plant and Soil* **360:** 1-13.

Berendsen RL, Pieterse CM, Bakker PA (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478-486.

Bulgarelli D, Garrido-Oter R, Munch PC, Weiman A, Droge J, Pan Y *et al* (2015). Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell host & microbe*.

Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J *et al* (2011). Moving pictures of the human microbiome. *Genome Biol* **12:** R50.

Chaparro JM, Sheflin AM, Manter DK, Vivanco JM (2012). Manipulating the soil microbiome to increase soil health and plant fertility. *Biol Fertil Soils* **48:** 489-499.

Consortium THMP (2012). Structure, function and diversity of the healthy human microbiome. *Nature* **486:** 207-214.

Dethlefsen L, Relman DA (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci U S A* **108 Suppl 1:** 4554-4561.

Dimkpa C, Weinand T, Asch F (2009). Plant-rhizobacteria interactions alleviate abiotic stress conditions. *Plant Cell Environ* **32:** 1682-1694.

Edwards J, Johnson C, Santos-Medellin C, Lurie E, Podishetty NK, Bhatnagar S *et al* (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A* **112:** E911-920.

Marcobal A, Kashyap PC, Nelson TA, Aronov PA, Donia MS, Spormann A *et al* (2013). A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *Isme J*.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332:** 1097-1100.

Mercado-Blanco J, Bakker PA (2007). Interactions between plants and beneficial Pseudomonas spp.: exploiting bacterial traits for crop protection. *Antonie Van Leeuwenhoek* **92:** 367-389.

Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G (2003). Microbial diversity and soil functions. *Eur J Soil Sci* **54:** 655-670.

Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G *et al* (2008). Effects of root exudates in microbial diversity and activity in rhizosphere soils. In: Nautiyal CS, Dion P (eds). *Molecular Mechanisms of Plant and Microbe Coexistence*. Springer-Verlag: Berlin Heidelberg. pp 339-365.

Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL *et al* (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci U S A* **110:** 6548-6553.

Selvakumar G, Panneerselvam P, Ganeshamurthy AN, Maheshwari DK (2012). Bacterial Mediated Alleviation of Abiotic Stress in Crops. In: Maheshwari DK (ed). *Bacteria in Agrobiology: Stress Management*. Springer: New York, NY. pp 205-224.

Sugiyama A, Vivanco JM, Jayanty SS, Manter DK (2010). Pyrosequencing Assessment of Soil Microbial Communities in Organic and Conventional Potato Farms. *Plant Dis* **94:** 1329-1335.

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI (2007). The human microbiome project. *Nature* **449:** 804-810.

van Loon LC, Bakker PA, Pieterse CM (1998). Systemic resistance induced by rhizosphere bacteria. *Annu Rev Phytopathol* **36:** 453-483.

Yang J, Kloepper JW, Ryu CM (2009). Rhizosphere bacteria help plants tolerate abiotic stress. *Trends Plant Sci* **14:** 1-4.