UNDERSTANDING THE INFLUENCE OF RHIZOSPHERE MICROBIOMES ON HORTICULTURAL CROP TRAITS AND PRODUCTION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Joshua Andrew Garcia

August 2022

UNDERSTANDING THE INFLUENCE OF RHIZOSPHERE MICROBIOMES ON HORTICULTURAL CROP TRAITS AND PRODUCTION

Joshua Andrew Garcia, Ph. D.

Cornell University 2022

Microbial consortia in the rhizosphere have a tremendous impact on crop traits and production but are often overlooked in agricultural research. Here, our goal was to utilize a variety of research frameworks and methodologies to develop a better understanding of the influence rhizosphere microbiomes can have on horticultural crop traits and production. Specifically, we sought to 1) search the literature for concepts and methodologies from other disciplines that may help understand rhizosphere interactions; 2) investigate group selection for enhanced plant productivity in the rhizosphere of *Brassica rapa*; 3) utilize plant transcriptomics to study the effects of beneficial microbial consortia derived from vermicompost on tomato foliar and fruit traits; and 4) determine the effects of intercropping on soil properties and production outcomes in horticultural cropping systems.

In our group selection experiment, we showed seed yield and nitrogen agronomic use efficiency can be altered in *B. rapa* by repeatedly selecting for rhizosphere microbiomes associated with increased aboveground biomass production, which may be mediated by group-level microbial behaviors such as nitrogen cycling. Our greenhouse study of vermicompost microbiomes revealed microbial consortia derived from vermicompost beneficially influence tomato growth and development by altering host gene expression. Our intercropping study revealed soil chemical and biological properties can be altered in response to crop diversification, which may lead to shifts in crop production outcomes and ecosystem-level properties.

Broadly, our work shows crop traits can be altered in response to microbiome manipulation, beneficial microbiomes can influence host traits via changes in plant gene expression, and aboveground management practices in horticultural systems can alter soil and plant properties relevant to production. Overall, this research demonstrates the usefulness of applying a variety of research frameworks and methodologies in developing a better understanding of the influence of rhizosphere microbiomes on crop traits.

BIOGRAPHICAL SKETCH

Joshua Garcia is originally from Phoenix, Arizona and attended the University of California, Davis for his undergraduate degree. There, he majored in Biological Science and minored in Chicanx Studies. At UC Davis, Joshua was selected to participate in the McNair Scholars Program, a program that prepares students from marginalized backgrounds for graduate studies. Joshua completed his McNair Scholars thesis project with Dr. Amelie Gaudin's lab in agroecology, where he developed a strong interest in plant-soil-microbe interactions in agriculture. In 2017, Joshua graduated from UC Davis and received the National Science Foundation Graduate Research Fellowship and the McNair SUNY Diversity Fellowship to begin his PhD work in horticultural biology at Cornell University. Joshua joined Dr. Jenny Kao-Kniffin's lab in rhizosphere ecology, where his dissertation research focused on the influence microbiomes in the rhizosphere have on horticultural crop traits and production. Joshua's work was featured in several high-profile academic journals, including *Frontiers in Microbiology*, *Functional Ecology*, and *Communications Biology*. In addition to his research, Joshua was involved in leadership, service, and diversity, equity, and inclusion (DEI) efforts through a variety of organizations both on- and off-campus, including the Latinx Graduate Student Coalition, the Cornell Prison Education Program, the Cornell Center for Teaching Innovation, and Black Hands Universal. Joshua was later honored as an Edward A. Bouchet Scholar in recognition of his scholarship, leadership, and service to the community. Joshua will continue his research in plant-soil-microbe interactions as a postdoctoral researcher at UC Davis with Dr. Cristina Lazcano and Dr. Mallika Nocco.

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Jenny Kao-Kniffin, for believing in me as a graduate student and for all her help and mentorship throughout the PhD process. I would also like to thank my dissertation committee members, Dr. Jim Giovannoni and Dr. Jed Sparks, for their assistance with various aspects of my work and for supporting my growth as a scientist. I also appreciate the support and help from all of my lab mates in the Kao-Kniffin lab and my undergraduate mentee, Micaela Moravek. Finally, I thank my various friends and mentors from throughout the years that have helped me in multiple ways, including Dr. Amélie Gaudin, Dr. Sarah Perrault, everyone at the UC Davis McNair Scholars Program, Dr. Sara Hernandez, Dr. Derina Samuel, and Dr. Lynn Marie Johnson.

I would like to thank the various agencies that supported my work financially, including the United States Department of Agriculture National Institute of Food and Agriculture Grant [2016-67013-24414], the National Science Foundation Graduate Research Fellowship Program [DGE-1650441], the McNair SUNY Diversity Fellowship, and the Cornell Atkinson Sustainable Biodiversity Fund.

Finally, thank you to my family and loved ones, who have always been there for me and who I can always count on.

TABLE OF CONTENTS

Chapter 1: Microbial Group Dynamics in Plant Rhizospheres and their Implications on Nutrient Cycling

Abstract

Plant rhizospheres encompass a dynamic zone of interactions between microorganisms and their respective plant hosts. For decades, researchers have worked to understand how these complex interactions influence different aspects of plant growth, development, and evolution. Studies of plant-microbial interactions in the root zone have typically focused on the effect of single microbial species or strains on a plant host. These studies, however, provide only a snapshot of the complex interactions that occur in the rhizosphere, leaving researchers with a limited understanding of how the complex microbiome influences the biology of the plant host. To better understand how rhizosphere interactions influence plant growth and development, novel frameworks and research methodologies could be implemented. In this chapter, we propose applying concepts in evolutionary biology to microbiome experiments for improved understanding of group-to-group and community-level microbial interactions influencing soil nutrient cycling. We also put forth simple experimental designs utilizing -omics techniques that can reveal important changes in the rhizosphere impacting the plant host. A greater focus on the components of complexity of the microbiome and how these impact plant host biology could yield more insight into previously unexplored aspects of host-microbe biology relevant to crop production and protection.

1.1 Introduction

With growing interest in host-microbe biology, a mixture of new and re-established terms have been developed in recent years to describe complex associations of organisms that are heritable. The concept of the "holobiont" and its "hologenome" (i.e., a host and all of its microbial symbionts, and consequently, the collective genomes of the adaptive unit)[1] signify the potential for natural selection to act, not only on the individual, but also on its suite of associations with its microbial members. While much of the discussion on the holobiont– hologenome theory has focused on animals and their microbiomes, the concept appropriately extends to plants and their associated microbiomes.

There is an extensive history of research detailing the co-evolutionary forces dually acting on plants and their microbial symbionts. For example, research on vesicular-arbuscular mycorrhiza (VAM) has suggested that the highly beneficial symbiosis between the plant root and fungal symbiont has driven the diversification of plant root morphology as well as VAM structure and function[2]. In addition, decades of legume research suggest that interactions, such as plant sanctions against nitrogen-fixing rhizobia, are largely what help stabilize the easily compromised legume-rhizobium mutualism[3]. In more recent years, the advent of next generation sequencing and other -omics tools has generated interest in better understanding how more complex associations between plants and their microbiome, termed the "phytobiome"[4], play a role in plant fitness, as well as plant health, growth, and development.

Similar to animal microbiomes, different compartments of a plant can harbor

functionally and taxonomically distinct sets of microbiomes—these include the phyllosphere microbiome encompassing the aboveground parts of plants, root and shoot endophytes (microbiomes colonizing root and shoot tissues), and the rhizosphere microbiome that inhabits soil surrounding plant roots and adhering to root surfaces[5]. Although the rhizosphere is not a plant organ or a physically intact compartment of the plant, this narrow band of soil surrounding roots harbors a tremendous diversity of microorganisms that are free-living or intricately linked to their plant hosts[6]. The root zone is an environment that is heavily enriched in compounds that are secreted by both plants and microorganisms, and play a key role in maintaining plant-microbe interactions[7]. The exudates include sugars, complex polysaccharides, amino acids, proteins, and a multitude of secondary metabolites[8]. An important component of these compounds includes signaling of host-microbe and microbe-microbe interactions[9]. With hundreds of different microbial taxa inhabiting a plant's rhizosphere, the possibilities of interactions shaping nutrient dynamics impacting plant growth are expansive. In this chapter, we discuss the background research on community evolution and genetics that are applicable to rhizosphere controls on nutrient cycling. We then propose conceptual ideas for experiments that explicitly test group-level dynamics in rhizospheres that can be targeted to alter nutrient capture and utilization processes impacting plant performance.

1.2 A novel framework for understanding rhizosphere microbiomes

While it appears overwhelming to study complex interactions between a microbiome and its host plant, there is growing interest in broadening the concept of

heritability to be inclusive of a host and its microbiome[10]. It is conceivable that a plant's rhizosphere microbiome is an ecological unit with heritable traits, although it is explicitly a physical extension of the host because it involves the soil habitat surrounding roots and not an intact unit of a plant. This concept of an "extended phenotype" is not new, but is derived from the concept proposed by Dawkins[11] that states an organism's phenotype should extend from its cellular components to its environment. Examples of extended phenotypes include manipulation of an organism's physical environment and alterations to behaviors, both of which can start at the gene-level. Other concepts in evolutionary biology propose a broader view that heritability is shaped at multiple levels beyond the individual, such that natural selection acts on ecological units beyond the individual. Perhaps one of the most salient among them is the concept of multilevel selection, which is often referred to as group selection.

Multilevel selection can be defined as natural selection acting on whole groups of organisms in addition to individuals, as initially proposed by Darwin[12]. During the latter half of the 20th century, many evolutionary biologists applied this concept to study the evolution of organisms, such as insects, small mammals, and humans[12]. While the framework typically has been used to study animals, very little work has been done to examine how the concept may extend to plants and their associated microbiomes. Applying the concept to the plant rhizosphere could help to elucidate the complex interactions occurring in microbe-microbe and microbe-plant networks that can be acted upon by natural and artificial selection at multiple levels. The application of this concept to microbiome science may be especially worthwhile given the types of

selective pressures that dictate microbial density-dependent rhizosphere processes, such as nutrient cycling.

A key tenet of multilevel selection theory is that selection not only acts on individuals, but can act on multiple levels of organization across biological units to influence the observed phenotype (Figure 1.1)[12]. A classic example includes cooperative sharing of a common resource such as food within a group. Evolutionary biologists have long questioned how cooperative resource sharing could arise within a group, given that cooperation can lower one's fitness according to the classic selfish-gene view[13]. In the case of resource sharing among members of a group, those who utilize the resource in an uncooperative manner, sometimes referred to as "cheaters," can circumvent paying the price of cooperation while reaping the benefits of utilizing the resource, thereby increasing their own fitness[14]. Such a scenario could ultimately lead to a "tragedy of the commons"[15]. Studies of microbial populations, however, have shown that groups of cooperative individuals could arise and outcompete cheaters if certain group-level selective pressures determine fitness outcomes at the population or community level[16,17,18].

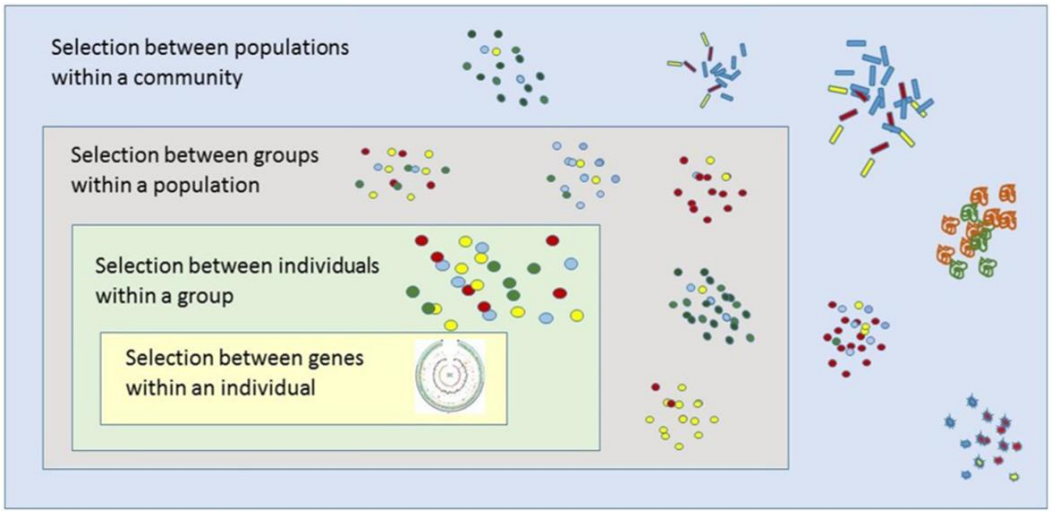


Figure 1.1 Applying group-level concepts in evolutionary biology to the study of microbiomes. A key tenet of multilevel selection theory is that selection not only acts on individuals, but can act on multiple levels of organization across biological units to influence the observed phenotype. For microbiota inhabiting rhizospheres, selection can act on different levels of organization (including simultaneously) consisting of genes, individuals, cells, groups, and the holobiont (i.e., plant host with its extended microbiome). The figure is adapted from Wilson and Wilson[13].

In the case of the plant rhizosphere, individual and group-level selective pressures constantly interact to shape the phenotype of the rhizosphere, although this is often overlooked when focusing on classical symbiosis studies. If perceptions of symbiosis expand to encompass multiple partnerships, including dozens of microbial taxa, it is then suitable to test the effect of group dynamics on nutrient cycling in the rhizosphere influencing plant growth. A plant growing in nitrogen-limited soils could gain a fitness advantage over competitors by enriching its rhizosphere for microbial communities that enhance nutrient capture and utilization capabilities[19,20,21].

For example, actively growing roots could signal for microorganisms that are capable of producing extracellular enzymes that release nitrogen bound in soil organic matter[22,23]. For prokaryotes, these mineralization processes are density-dependent and need a quorum of producers to sufficiently access key nutrients in soil[24]. The groups of

producers are not kin-based and can be formed from taxonomically diverse microbiota if they have the ability to biosynthetically produce the specific enzymes that are secreted into the soil. In this scenario, selection could favor microhabitats in rhizospheres that promote coordinated group behaviors that enhance plant access to nitrogen or phosphorus upon cell turnover, while the microorganisms benefit from having an abundant supply of carbon and other nutrients from plant roots (Figure 1.2). While the example illustrates how group dynamics at multiple levels of organization impact the fitness of the plant host, the microbial community, and specific individuals or related cells, the interaction has posed a significant challenge to test experimentally until recent years.

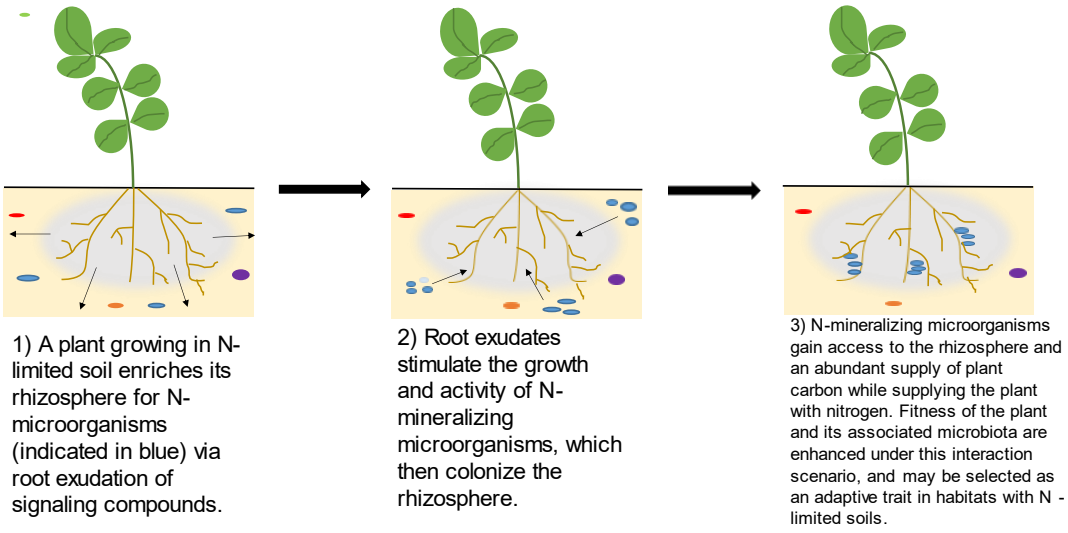


Figure 1.2 Enrichment of microbial groups in plant rhizospheres grown under nitrogen-limited soils. The illustrations depict a scenario in which bacteria form groups within the rhizosphere, with members that produce a specific extracellular enzyme (e.g., N-Acetyl glucoseaminidase, NAGase) targeting a particular substrate (e.g., chitin). Nitrogen that was bound to soil organic matter comprised of chitin is mineralized and captured by nearby microorganisms. Upon microbial cell turnover, these hotspots of inorganic nitrogen can be intercepted by actively growing plant roots. It is conceivable that plants and groups of microorganisms have formed complex associations that can be adaptive under resource limitation. Applying multilevel selection theory to microbiome studies could reveal novel cross-kingdom signaling mechanisms that challenge traditionally defined units of heritability.

There are many more components to nutrient cycling that are microbially mediated, with potential group dynamics influencing plant growth. For example, microorganisms such as rhizobia and certain free-living diazotrophs can form groups that fix atmospheric nitrogen through metabolic processes and make the nutrient readily available for plant uptake[25,26,27]. In addition, certain microbial taxa are known to mobilize phosphorus in the soil via the production of extracellular compounds and enzymes, which increases phosphorus availability for plant use[28,29,30]. Other important plant nutrients, such as iron, can be made into plant available forms by groups of unrelated microbial taxa, through the production of compounds such as

siderophores[30,31]. Coordinated groups of microorganisms may also enhance plant nutrient uptake through more indirect mechanisms, such as stimulating root growth via phytohormone production, which can increase plant access to essential nutrients[32].

1.3 Observing multiple levels of selection in the plant rhizosphere

The application of multilevel selection theory to rhizosphere processes, such as nutrient cycling, offers researchers an opportunity to understand how multiple levels of selection shape a plant's microbiome. Experimental tests of group dynamics in the rhizosphere impacting a plant phenotype are needed to help us understand the role of complex associations of microorganisms in determining plant growth, development, and fitness. In this section, we propose three concepts for experimental designs that could aid in elucidating the role of the rhizosphere microbiome in impacting various plant phenotypes via alterations in nutrient cycling processes.

1.3.1 Microbial Experimental Systems and Network Analysis

Microbial experimental systems have allowed researchers to document shifts in population structure or adaptive changes in organisms over time[33]. The experiments were typically conducted under highly controlled conditions involving paired comparisons of cultivated bacterial strains or lower eukaryotes, to reveal outcomes of interactions. While the greatest utility of microbial experimental systems is the ability to control for environmental and biological variables that mask potential microbe-microbe interactions, hierarchies of complexity can be introduced into the system to reveal more dynamic interactions. In recent years, a small number of these

experimental systems were developed to evaluate how changes in complex soil microbiomes can influence the biology of a plant following generations of selection for microbiota associated with the plant trait. For example, work by Swenson et al.[34] demonstrated that the aboveground biomass of *Arabidopsis thaliana* could be enhanced by artificially selecting for rhizosphere microbiomes that are known to increase plant biomass production. In a similar study using *A. thaliana*, flowering time was modified by soil microbiota through repeated selections of rhizosphere microbiomes, which was predicated based on the idea that selection can be performed on a plant's phenotype while the agents of selection are comprised of the rhizosphere microbiome[21]. The experiment resulted in taxonomic and functional shifts in the microbiome across early versus delayed flowering plants, but at the time of the study, group-level behaviors were challenging to measure when studying soil microbiomes.

Since then, advances in sequencing technology and bioinformatic analysis have provided an opportunity to examine networks of associations across multiple microbial partners that may indicate coordination across taxa[35]. An experiment can be designed to test for associations among microbial taxa relevant to group-level processes, such as mineralization dynamics in the soil. Given that certain plant community traits could be predictive of soil microbiome composition[36], plant-soil systems that show greater nutrient capture or use efficiency over time could indicate more associations among microbial taxa, including positive, negative, or neutral, that may aid in modeling complex interactions of microbial partners in a soil system. Network analysis involving soil microbiomes could indicate specific taxa that co-occur consistently in an experimental system, which could signal targets for

assembling synthetic microbial communities in rhizosphere studies[37]. Top-down approaches involving the use of microbial experimental systems for generating plant trait associated microbiomes, followed by bottom-up approaches that use the resulting sequencing and bioinformatic data to design synthetic microbial communities, could help identify consortia that influence plant growth via alterations to nutrient cycling. These consortia of isolates could then be tested in various controlled environment and field settings, which would yield insight into interactions that alter plasticity of plant host traits.

1.3.2 Observing Microbiome Controls Over Observed Phenotypes of the Plant Using -Omics Techniques

Advances in sequencing technologies provide opportunities to examine how microbiomes or individual strains of microorganisms impact plant traits through changes in plant gene expression[38]. While it is widely acknowledged that soil microorganisms alter plant developmental processes[39], sequencing techniques such as transcriptomics, offer researchers an opportunity to observe these alterations at previously unexplored scales[40]. Microbial influences on its plant host can be observed at the gene level using plant transcriptional analysis methods, and then combined with network analysis to indicate individual genes or entire metabolic pathways that are impacted when a specific microbial partner or consortia of partners interacts with its host. Experiments can be designed to observe the effect of specific microorganisms on the development of plant traits, while observing changes in a host's transcriptome over time. From these experiments, the effect of specific microbial associations in the

rhizosphere, including effects relevant to plant nutrition requirements, could be determined.

In addition to plant transcriptomics, other -omics techniques could be applied to plant-soil systems to better understand how rhizosphere microbiomes influence host biology under different conditions. For example, exometabolomics techniques could be used to characterize microbial metabolite utilization in the rhizosphere environment[41]. This data could then be used to determine microbial exometabolic niches in the rhizosphere, which may help link metabolite composition to microbial community structure[42]. Other techniques, such as proteomics, could be applied to the microbiome studies to determine how different functions of the microbiome influence the plant host[43]. By integrating these multiple -omics techniques, researchers can develop a more robust understanding of how rhizosphere microorganisms influence soil processes and plant host biology[43].

1.3.3 Genome-Editing Techniques to Uncover Plant Host Controls Over Microbiome Composition and Function

The development of genome-editing tools such as clustered regulatory interspaced short palindromic repeats (CRISPR)- Cas can be useful in modifying rhizosphere processes through changes in plant structure and physiology[44]. The CRISPR-Cas method works by using short RNA sequences to guide a Cas enzyme to a target site for alteration[45]. Systematic alterations in plant architecture or physiology could aid in mechanistic studies of how plants control microbiome assembly and function in the rhizosphere. For example, disruption of the mutualistic interactions

between microorganisms that aid in nutrient capture and plants that provide carbon sources for microbial growth could be designed through modification of root exudation processes via genetic engineering and gene-editing[46]. Plants that no longer provide readily available carbon sources to soil microorganisms involved in mineralization dynamics may show a distinct microbial community that specializes in more complex polysaccharides found in soil organic matter. In addition, root exudation processes could be modified to produce specific families of compounds in the rhizosphere. Taxonomic shifts in the rhizosphere community could then be documented with these shifts in root exudate composition, revealing plant selection for specific microbial taxa. Additionally, researchers could study plant selection for functionally distinct microbial taxa by altering host physiology via transcription factors, similar to the methods used by Su et al.[47]. These alterations could cause functional and taxonomic shifts in the rhizosphere, which could help further our understanding of selection for specific groups of microorganisms.

1.4 Concluding Remarks and Future Directions

Currently, there is a growing interest in developing a broader understanding of host-microbe biology. Similar to human and other animals, different plant compartments harbor distinct microbiomes, which could evolve and adapt with their host to influence the observed phenotypes. In essence, fitness is both influenced by and shared among multiple levels—the individual plant host and groups comprising its microbiome. Thus, the application of evolutionary frameworks, such as multilevel selection, to plant microbiomes could be useful in developing a more robust

understanding of host-microbe interactions. Utilizing –omics techniques is key to uncovering potential mechanisms underlying group-level interactions in the rhizosphere. It is conceivable that cross kingdom signaling dominates rhizosphere processes, which suggests that the definition of heritability should be inclusive of an individual and its associated microbiome. The plant host may influence the composition and function of the resulting microbiome, but microorganisms have the ability to modify plant traits.

An applied outcome of studying group-level dynamics in the rhizosphere is the ability to incorporate concepts of the holobiont into plant breeding. Selection efforts that consider rhizosphere microbiomes as an extended phenotype of a plant could help identify potential mechanisms that enrich for subpopulations of the microbiome. These "plant cultivated" members of the rhizosphere could play essential roles in supporting the development of specific phenotypes of the plant that improve plant growth under biotic or abiotic stress.

However, plant genetics may not play a significant role in influencing the microbiome, but instead the plant may be highly susceptible to microbiome effects on plant traits. A greater understanding of how microbiomes influence the observed phenotype of a plant can help to tease apart the effect of environmental variables from biotic factors. For example, in studies involving genotype by environment interactions (G × E), phenotypic variation is assumed to be a result of plant genetics influenced by varying environmental conditions across field sites. In the near future, we expect that the low cost of microbiome sequencing methods will result in the adoption of rapid microbiome diagnostics revealing the role of microbiome variability across field sites

in influencing plasticity of the plant phenotypes. In essence, there will likely be a shift toward analysis of genotype by environment by microbiome interactions (G × E × M) in the coming years.

Finally, we believe that the current industry focus on examining single microbial isolate effects on plant traits will be replaced with more emphasis on complex interactions involving multiple players. The recent popularity of examining synthetic communities comprised of multiple microbial strains helps to advance microbiome science forward, but it would be beneficial to move beyond cultivation-dependent methods. Applying selective filters to reduce the diversity of complex microbiomes associated with a plant trait could enable more top-down and bottom-up approaches comprised of cultivation dependent and independent multi-player interaction studies. While teasing apart the complexity of the rhizosphere will be incredibly challenging, such research could ultimately help develop a better understanding of how rhizosphere microbiomes influence plant growth, development, and fitness.

1.5 References

1. Rosenberg, E., & Zilber-Rosenberg, I. Microbes Drive Evolution of Animals and Plants: the Hologenome Concept. *MBio* 7:2, e01395 (2016).

2. Brundrett, M. C. Coevolution of roots and mycorrhizas of land plants. *New Phytol*. 154(2), 275-304 (2002).

3. Masson-Boivin, C., & Sachs, J. L. Symbiotic nitrogen fixation by rhizobia-the roots of a success story. *Curr. Opin. Plant Biol.* 44, 7-15 (2017).

4. Leach, J. E., Triplett, L. R., Argueso, C. T., & Trivedi, P. Communication in the Phytobiome. *Cell* 169(4), 587-596 (2017).

5. Turner, T. R., James, E. K., & Poole, P. S. The plant microbiome. *Genome Biol.* 14(6), 209 (2013).

6. Mendes, R., Garbeva, P., & Raaijmakers, J. M. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol. Rev.* 37, 634–663 (2013).

7. Badri, D. V., Weir, T. L., van der Lelie, D., & Vivanco, J. M. (2009). Rhizosphere chemical dialogues: plant-microbe interactions. *Curr. Opin. Biotech.* 20(6), 642-650 (2009).

8. Badri, D. V., & Vivanco, J. M. Regulation and function of root exudates. *Plant, Cell and Environ. 32*(6), 666-681 (2009).

9. Venturi, V., & Keel, C. Signaling in the Rhizosphere. *Trends in Plant Sci.* 21(3), 187-198 (2016).

10. van Opstal, E. J., & Bordenstein, S. R. Rethinking heritability of the microbiome. *Science* 349(6253), 1172-1173 (2015).

11. Dawkins, Richard (1982). The Extended Phenotype. Oxford: Oxford University Press.

12. Wilson, D. S., & Wilson, E. O. Evolution "for the good of the group". *Am. Sci.* 96(5), 380-389 (2008).

13. Sachs, J. L., Mueller, U. G., Wilcox, T. P., & Bull, J. J. The evolution of cooperation. *Q. Rev. Biol.* 79(2), 135-160 (2004).

14. Riehl, C., & Frederickson, M. E. Cheating and punishment in cooperative animal societies. *Phil. Trans. R. Soc. Lon. B.* 371(1687), 20150090 (2016).

15. Waite, A. J., & Shou, W. Adaptation to a new environment allows cooperators to purge cheaters stochastically. *Proc. Natl. Acad. Sci. U S A* 109(47), 19079- 19086 (2012).

16. Allen, B., & Nowak, M. A. Cooperation and the Fate of Microbial Societies. *PLOS Biol.* 11(4), e1001549 (2013).

17. Boyle, R. A., Williams, H. T., & Lenton, T. M. Natural selection for costly nutrient recycling in simulated microbial metacommunities. *J. Theor. Biol.* 312, 1-12 (2012).

18. MacLean, R. C., & Gudelj, I. Resource competition and social conflict in experimental populations of yeast. *Nature* 441(7092), 498-501 (2016).

19. DeAngelis, K. M. et al. Selective progressive response of soil microbial community to wild oat roots. *ISME J.* 3(2), 168-178 (2009).

20. Kao-Kniffin, J., & Balser, T.C. Soil microbial composition and nitrogen cycling in a disturbed wet prairie restoration (Wisconsin). *Ecol. Restor.* 28:20-22 (2010).

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

22. Jacoby, R., Peukert, M., Succurro, A., Koprivova, A., & Kopriva, S. The Role of Soil Microorganisms in Plant Mineral Nutrition—Current Knowledge and Future Directions. *Front. Plant Sci.* 8, 1617 (2017).

23. Lemanceau, P. et al. "Plant communication with associated microbiota in the spermosphere, rhizosphere and phyllosphere." in How Plants Communicate with Their Biotic Environment, ed. G. Becard (London: Academic Press Ltd- Elsevier Science Ltd.) 101-133 (2017).

23. DeAngelis, K. M., Lindow, S. E., & Firestone, M. K. Bacterial quorum sensing and nitrogen cycling in rhizosphere soil. *FEMS Microbiol. Ecol.* 66(2), 197-20t (2008).

24. Santi, C., Bogusz, D., & Franche, C. Biological nitrogen fixation in non-legume plants. *Ann. Bot.* 111, 743–767 (2013).

25. Oldroyd, G. E. D., & Dixon, R. Biotechnological solutions to the nitrogen problem. *Curr. Opin. Biotechnol*. 26, 19–24 (2014).

26. Mus, F. et al. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Appl. Environ. Microbiol.* 82, 3698–3710 (2016).

27. Adnan, M. et al. Phosphate-solubilizing bacteria nullify the antagonistic effect of soil calcification on bioavailability of phosphorus in alkaline soils. *Sci. Rep.* 7:16131 (2017).

28. Alori, E. T., Glick, B. R., & Babalola, O. O. Microbial phosphorus solubilization and its potential for use in sustainable agriculture. *Front. Microbiol.* 8:971 (2017).

29. Vessey, J. K. Plant growth promoting rhizobacteria as biofertilizers. *Plant Soil* 255, 571–586 (2003).

30. Radzki, W. et al. Bacterial siderophores efficiently provide iron to iron-starved tomato plants in hydroponics culture. *Antonie Van Leeuwenhoek* 104, 321–330 (2013).

31. Adesemoye, A. O., & Kloepper, J. W. Plant-microbes interactions in enhanced fertilizer-use efficiency. *Appl. Microbiol. Biotechnol.* 85, 1–12 (2009).

32. Jessup, C.M., Forde, S.E., & B. Bohannan. Microbial Experimental Systems in Ecology. *Adv. Ecol. Res.* 37, 273-307 (2005).

33. Swenson, W., Wilson, D. S., & Elias, R. Artificial ecosystem selection. *Proc. Natl. Acad. Sci. U.S.A.* 97(16), 9110-9114 (2000).

34. Faust, K., & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10, 538–550 (2002).

35. Leff, J. W. et al. Predicting the structure of soil communities from plant community taxonomy, phylogeny, and traits. *ISME J.* (2018).

36. Großkopf, T., & Soyer, O.S. Synthetic microbial communities. *Curr. Opin. Microbiol.* 18, 72-77 (2014).

37. Schenk, P. M., Carvalhais, L. C., & Kazan, K. Unraveling plant-microbe interactions: can multi-species transcriptomics help? *Trends Biotechnol.* 30, 177–184 (2012).

38. van Loon, L. C. Plant responses to plant growth-promoting rhizobacteria. *Eur. J. Plant Pathol.* 119, 243–254 (2007).

39. Wolf, J. B. W. Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Mole. Ecol. Resour.* 13(4), 559-572 (2013).

40. Swenson, T. L. et al. Linking soil biology and chemistry in biological soil crust using isolate exometabolomics. *Nat. Commun.* 9(1), 19 (2018).

41. Baran, R. et al. Exometabolite niche partitioning among sympatric soil bacteria. *Nat. Commun.* 6, 8289 (2015).

42. White, R. A. et al. The state of rhizospheric science in the era of multi-omics: A practical guide to omics technologies. *Rhizosphere*, 3, 212-221 (2017).

43. Bortesia, L. & Fischerab, R. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechol. Adv.* 33, 41-52 (2015).

44. Nemudryi, A. A., Valetdinova, K. R., Medvedev, S. P., & Zakian, S. M. (2014). TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery. *Acta Naturae* 6(3), 19-40 (2014).

45. Ahkami, A. H., White, R.A., Handakumbura, P. P., & Jansson, C. Rhizosphere engineering: Enhancing sustainable plant ecosystem productivity. *Rhizosphere* 3, 233-243 (2017).

46. Su, J. et al. Expression of barley SUSIBA2 transcription factor yields high-starch low-methane rice. *Nature* 523, 602 (2015).

Chapter 2: Selection pressure on the rhizosphere microbiome can alter nitrogen use efficiency and seed yield in *Brassica rapa*

Abstract

Microbial experimental systems provide a platform to observe how networks of groups emerge to impact plant development. We applied selection pressure on the rhizosphere microbiome for increased biomass of a non-mycorrhizal plant host, *Brassica rapa*, across iterative cycles to examine adaptive bacterial group dynamics under soil nitrogen limitation. In the final and 9th generation of the experiment, selection pressure enhanced *B. rapa* seed yield and nitrogen use efficiency compared to our no selection (control) treatment, with changes in bacterial community profiles and interaction networks compared to the control and random selection treatment. Soil bacterial diversity declined in response to selection for high *B. rapa* biomass production suggesting a possible ecological filtering mechanism to remove bacterial taxa. Distinct sub-groups of interactions emerged among bacterial phyla such as *Proteobacteria* and *Bacteroidetes* in response to selection. Extended Local Similarity Analysis (eLSA) and quantification of bacterial community rewiring and change showed greater connectivity of the bacterial community, with more edges, shorter path lengths, and altered modularity through the course of selection for enhanced plant biomass. In contrast, bacterial communities under random selection and no selection showed less complex interaction profiles of bacterial taxa. These results suggest that group-level bacterial interactions could be modified to collectively shift microbiome

functions impacting the growth of the host plant under soil nitrogen limitation.

2.1 Introduction

In recent years, numerous plant microbiome studies have highlighted the intricate links between a plant host and its associated microbiomes[1,2,3]. In the rhizosphere, microbiomes play a key role in processes that affect plant fitness, such as nutrient cycling[4,5] and disease suppression[6,7]. Given their influence on plant growth and health, microbial consortia in the rhizosphere have become targets for altering plant traits that may help improve agricultural productivity[8,9]. Several studies have demonstrated the ability to alter host traits via selection (i.e. directed microbiome manipulation) for microbial consortia in the root zone. For example, experimental evolution work by Swenson et al.[10] demonstrated that plant biomass could be enhanced in *Arabidopsis thaliana* following inoculation with rhizosphere microbiomes selected based on their association with increased host biomass production over multiple generations. A similar study demonstrated microbiome modification of flowering time in *A. thaliana*[11]. Considering these findings, the creation of robust plant growth promoting rhizosphere microbiomes could be possible through careful conveyance of group-level selective processes (i.e. selective processes that affect whole groups of organisms rather than individuals alone) for rhizosphere microbial consortia associated with enhanced productivity and fitness.

Microbiome sequencing data can provide a window into the changing communities and dynamics of group-level processes such as nutrient cycling in microbiome experiments over time. Microbial experimental systems that entail

simulations of both microbe-microbe and plant-microbe interactions can be designed as empirical studies to track the outcome of selective pressure on a group-level trait[12]. Resulting sequence data of the microbiome can be used to observe the emergence of interaction networks that could signify group-level processes developing through time. The significance of this network profile is the potential record of group selection processes unfolding within a system[13, 14]. A key tenet of group selection is that selective pressures at the group- and individual-levels constantly interact with one another to produce the observed phenotype of a population[15]. In the context of host-microbe interactions, the microbiome can be shaped heavily by plants when the microbial environment and resource pool are largely defined by the host[16]. The resulting microbiome, in turn, alters different aspects of host growth and development[4]. A better understanding of the plant-microbiome feedback processes in the rhizosphere could reveal how group-level interactions collectively shape host phenotypes.

For microbial experimental systems involving time series sequence data, it is important to consider what types of microbiome interactions can be interpreted in a network profile. Microbial taxa that are positively associated could indicate cooperative or antagonistic behaviors. For example, resource sharing could be extremely costly for individuals given that members in the community may utilize the resource in an uncooperative manner[17], thereby increasing their own fitness while avoiding the cost of cooperation. Network models may not distinguish between these forms of interactions unless the experimental design explicitly controls for contrasting scenarios that determine if more cooperative communities can outcompete others

when certain group-level selective pressures are acting on the system[18,19,20]. In the case of the rhizosphere and its root-associated microbiome, group-level selection could be ubiquitous[4,21,22] and may help explain the high rates of carbon losses from the plant via root exudation[23] as an adaptive strategy to harbor beneficial microbiota in the root zone. Given that these forms of selective pressures are known to influence the microbiome, directed evolution in the rhizosphere for microbial communities that confer an ecosystem-level trait, such as enhanced aboveground biomass production, could be a possibility.

Here, we examined if plant productivity (i.e. aboveground biomass and seed yield) could be enhanced in the non-mycorrhizal host *Brassica rapa* through repeated selection for rhizosphere microbiota associated with increased aboveground biomass production over nine generations of plantings. The objective of this experiment is to examine changes in rhizosphere bacterial group dynamics across generations of selection for high biomass growth of a non-mycorrhizal plant under soil nitrogen limiting conditions. Arbuscular mycorrhizal fungi form associations with most terrestrial plant species that support nitrogen and phosphorus uptake of the plant host[24,25], but the Brassica lineage of plants is dominated by species that do not serve as mycorrhizal hosts[26]. The loss of key symbiosis genes in non-host Brassicas occurred early in the lineage of the Brassicales order[27], which suggests that these plants have evolved different strategies to acquire nitrogen and phosphorus from soil. We hypothesized that selection pressure on rhizosphere microbiomes for increased biomass of a non-mycorrhizal plant host growing under nitrogen-limiting conditions will lead to the assembly of distinct, highly-connected and interactive rhizosphere

bacterial groups associated with plant host nitrogen uptake strategies. Additionally, we hypothesize that enhanced soil N cycling would be a key mechanism by which the rhizosphere microbiome of a non-mycorrhizal plant host would enhance biomass growth and seed yield.

The selection treatment for high biomass plants was compared with two types of control treatments: plants that were chosen through random selection (random treatment) and plants that were grown with cryopreserved soil microbial inocula that served as controls against microbial adaptation (control treatment). All three selection treatments in the experiment received the same bacterial inoculant in the first generation, derived from organic farm soils. Thus, the initial bacterial community was the same for all three selection treatments but diverged over time across generations of the selection experiment. The microbial inocula consisted of rhizosphere soil dispersed with sterile water into autoclaved soils. Each selection treatment consisted of 15 replicated pots with identical autoclaved soil and *B. rapa* seeds across treatments and generations of plantings, which ensured that the microbial inocula were the dominant variable. The experimental design optimized soil bacterial passage across selection generations through use of the liquid dispersal medium and rapid generation times of the host plant (10-day growth of *B. rapa* Wisconsin Fast Plants). For a detailed illustration of the experimental design, see Figure 1. We used amplicon sequencing of 16S rRNA genes to observe shifts in bacterial community composition and interaction profiles across taxa. For bacterial community function, we assessed changes in plant nutrient use efficiency by measuring isotopic nitrogen ($\delta^{15}N$) of plant tissue to indicate

potential changes in soil N processing and nutrient metabolism within plants, which is a major determinant of plant productivity and yield.

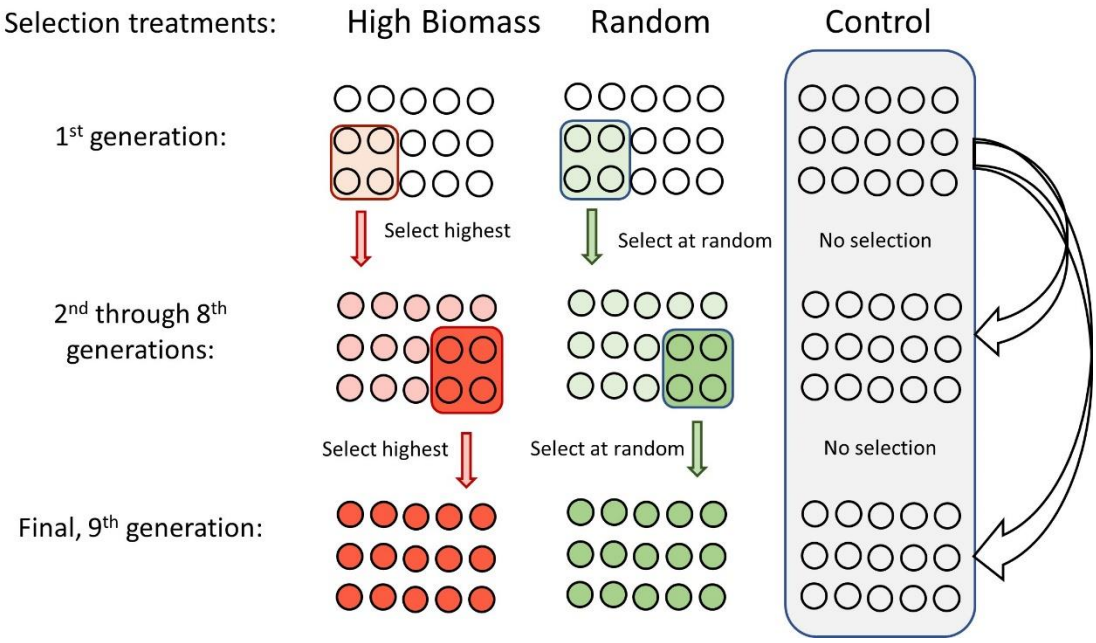


Figure 2.1 Experimental design of selection for rhizosphere microbiomes that enhance plant host yield. Selection pressure is placed on the rhizosphere microbiome for phenotypic changes (enhanced yield) of the host plant. For the 1st through 8th generations of the high biomass selection treatment, the rhizosphere soil from the four highest yielding units (from n=15 replicates) is composited into microbial inoculants for the subsequent generation of plantings. For the random selection treatment, four units (from n=15 replicates) are selected at random with rhizosphere harvested and used as microbial inoculants for the next planting. The control treatment consists of cryopreserved microbial inoculants that are identical to the start of the 1st generation of plantings and represent non-adaptive microbiomes (n=15 replicates). For all three selection treatments, seeds were collected to determine seed yield and nitrogen use efficiency (NUE) in the final (9th) generation. Units in all treatments and generations of plantings were comprised of the same *B. rapa* seed pool to minimize genetic diversity of the plant host. All units of the three selection treatments in the 1st generation received identical microbial inoculants derived from a mixture of organic farm soils.

2.2 Results

2.2.1 Plant, soil, and microbiome samples

In the 9th and final generation of the experiment, the plants were allowed to set seed to determine selection pressure outcomes on *B. rapa* seed yield and nitrogen use efficiency. Only the harvested aboveground biomass is reported for plant data from the previous 1st through 8th generations, followed by seed yield in the 9th generation. Likewise, rhizosphere bacteria data are shown for the 1st through 8th generations, and the 9th generation sampling was avoided due to the prolonged growth of seed-bearing plants and low moisture soil conditions (and eventual desiccation) required for seed collection.

2.2.2 Aboveground dry biomass production and seed yield

To assess if our selection process enhanced plant productivity in the high biomass selection treatment, we compared the total dry aboveground biomass production and seed yield of the three selection treatments in the 9th generation. Analysis of aboveground dry biomass production (g) revealed significant effects of selection treatment ($p=3.28e-05$). Pot units in the high biomass and random selection treatments had 36% and 38% greater *B. rapa* aboveground dry biomass production compared to pots in the control treatment at the $p=0.002$ and $p=0.0001$ significance levels, respectively (Figure 2.2A). Aboveground dry biomass production between the high biomass and random selection treatments did not differ ($p=0.98$). Analysis of seed yield revealed significant effects of selection treatment ($p=0.032$). The seed yield, measured as the total g of seeds produced per pot unit, in the high biomass selection

treatment was 62% higher compared to the control treatment at the p=0.025 significance level (Figure 2.2B). In contrast, the seed yield of the random selection treatment did not differ from the high biomass selection treatment (p=0.52) or the control (p=0.25).

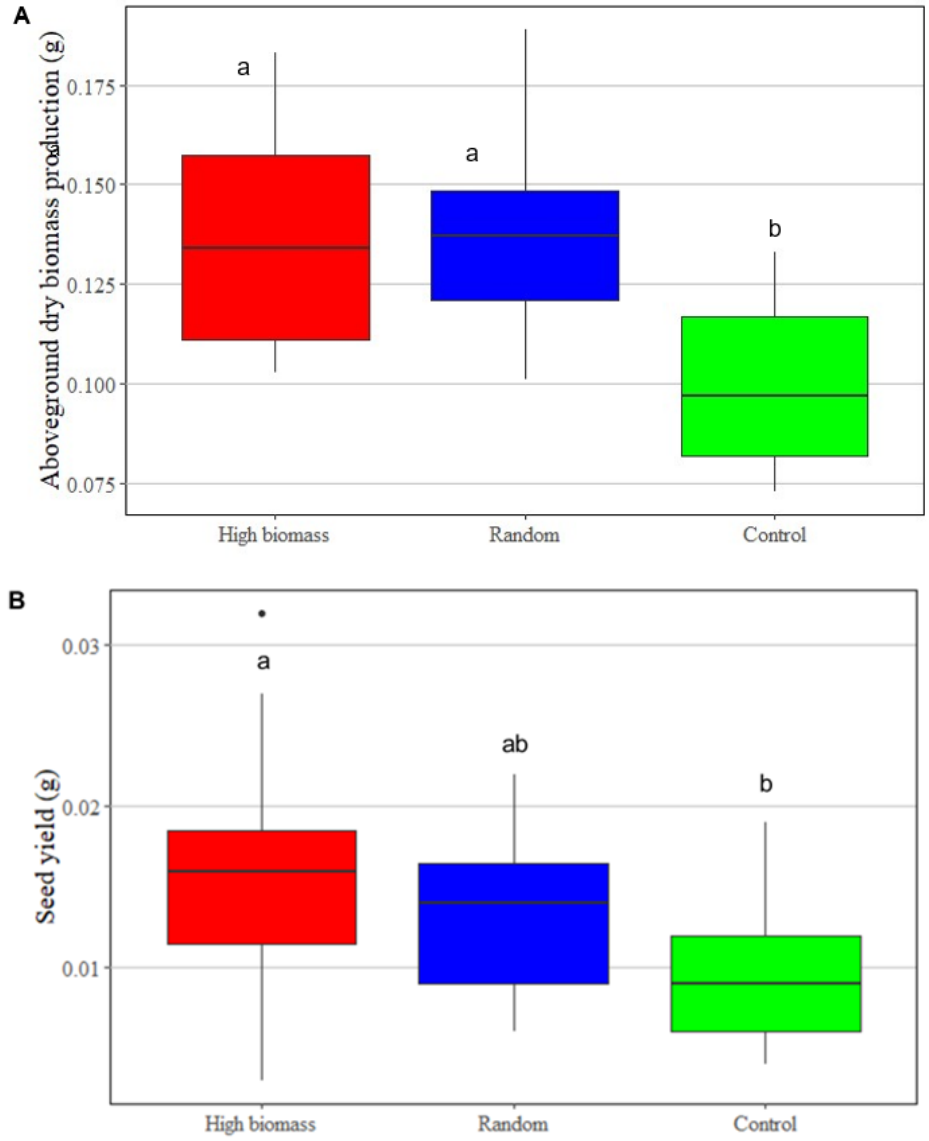


Figure 2.2 A.) Box plot displaying the average aboveground biomass production (g) of the pots in each selection treatment after the 9th generation of selection. B.) Box plot displaying the average seed yield (g) of the pots for each selection treatment after the 9th generation of selection.

2.2.3 Analysis of plant tissue N content

Plant N uptake is one of the largest determinants of productivity[28]. We expected that increased soil N cycling would be a key mechanism by which the rhizosphere microbiome of a non-mycorrhizal plant host would enhance biomass growth and seed yield. Thus, we analyzed foliar N data from the 9th generation of the experiment to determine how selection may have altered microbial nutrient cycling activity. Analysis of total foliar N revealed effects of selection treatment in the 9th generation of the experiment at the p=0.031 significance level (Table 2.1). Plants in the high biomass selection treatment had 57% more total foliar N compared to plants in the control treatment at the p=0.028 significance level, while the total foliar N content of plants in the random selection treatment did not differ from the high biomass (p=0.6) or control (p=0.12) treatments. Analysis of foliar δ^{15}N values in the 9th generation revealed further effects of selection treatment (p=0.0013). Plants in the high biomass selection treatment had 43% lower foliar δ^{15}N values compared to plants in the random selection treatment, and 50% lower foliar δ^{15}N values compared to plants in the control treatment at the p=0.0077 and p=0.0014 significance levels, respectively (Table 2.1). The foliar δ^{15}N values of the random selection and control treatments did not differ from one another (p=0.46). Foliar % N did not differ between any selection treatments (p=0.22). Selection treatment also had a weak effect on N agronomic use efficiency (NUE), defined as seed yield (g) g^{-1} nitrogen added[28], at a p=0.08 significance level. Plants in the high biomass selection treatment had a similarly weak relationship exhibiting a 113% greater NUE compared to the control treatment in the 9th generation (p=0.071) (Table 2.1). The NUE of the random

selection treatment did not differ significantly from the high biomass (p=0.6) and control (p=0.12) selection treatments. To ensure differences observed in plant phenotype between selection treatments in the 9th generation were a result of our selection process, we analyzed aboveground dry biomass production, % foliar N, foliar $\delta^{15}N$, and total foliar N from the 1st generation (Table 2.1), which showed no effect of selection treatment.

Table 2.1 Group means and ANOVAs for total foliar N, foliar $\delta^{15}N$ vs At. Air, % foliar N, and N agronomic efficiency for the three selection treatments at the 9th generation and group means and ANOVAs for total foliar N, foliar $\delta^{15}N$ vs At. Air, % foliar N, and aboveground dry biomass production, for the 1st generation. N data are from analysis of four replicates chosen at random from each selection treatment in the 9th and 1st generation. Average aboveground dry biomass in the 1st generation was calculated using 15 replicates from each selection treatment.

9th Generation	Total foliar N (g)	Foliar $\delta^{15}N$ vs. At. Air	% foliar N	N agronomic efficiency (g seed/g N added)
High biomass	0.0013 ± 0.00014^{a}	1.47 ± 0.12^{a}	0.89 ± 0.035^{a}	1.17 ± 0.17^{a}
Random	0.0011 ± 0.00013^{ab}	2.58 ± 0.21^{b}	0.80 ± 0.052^{a}	0.95 ± 0.21^{ab}
Control	0.00083 ± 0.000057^{b}	2.92 ± 0.24^{b}	0.97 ± 0.082^{a}	0.55 ± 0.13^{b}
ANOVA Selection line	0.031**	0.0013**	0.22	0.08*

1st Generation	Total foliar N (g)	Foliar $\delta^{15}N$ vs. At. Air	% foliar N	Aboveground dry biomass production (g)
High biomass	0.0028 ± 0.65^{a}	5.35 ± 0.65^{a}	2.48 ± 0.50^{a}	0.11 ± 0.019^{a}
Random	0.0028 ± 0.31^{a}	5.39 ± 0.31^{a}	2.04 ± 0.38^{a}	0.13 ± 0.021^{a}
Control	0.0039 ± 0.56^{a}	5.91 ± 0.56^{a}	2.50 ± 0.35^{a}	0.14 ± 0.023^{a}
ANOVA Selection line	0.81	0.71	0.68	0.55

*Significance at the p=0.1 probability level.
**Significance at the p=0.05 probability level.
***Significance at the p<0.01 probability level.

To link our foliar N measurements to plant productivity, we performed Pearson correlations using data from the 9th generation of the experiment. Total foliar N was found to have a strong positive correlation with aboveground dry biomass production (p=1e-04) (Figure 2.3A) and seed yield (p=0.0051) (Figure 2.3B).

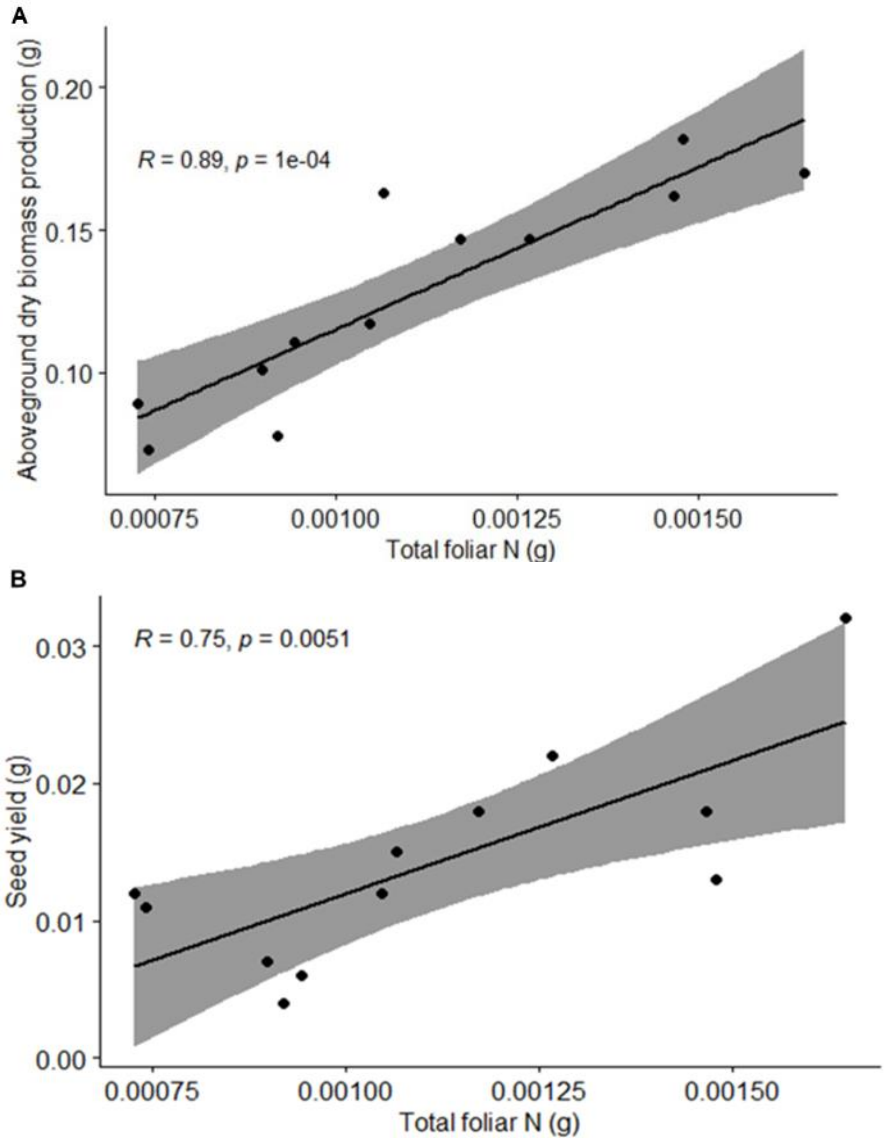


Figure 2.3 Pearson correlations between plant productivity and total foliar nitrogen (N) measurements in the 9th generation of the experiment. A.) Scatterplot and Pearson correlation of aboveground dry biomass production (g) vs. total foliar N (g) using plant phenotype data from the 9th generation. B.) Scatterplot and Pearson correlation of seed yield (g) vs. total foliar N (g) using plant phenotype data from the 9th generation.

2.2.4 Analysis of the rhizosphere bacterial community

We analyzed bacterial community composition through interpretations of 16S rRNA gene sequences to assess changes in the composition and interaction profiles of microbial taxonomic groups and specific operational taxonomic units (OTU). Plotting of the Bray-Curtis distances in an ordination followed by PERMANOVA revealed significant effects of Generation ($p<0.001$), Selection Treatment ($p<0.001$), and the interaction between the two factors ($p<0.001$) on bacterial community composition. Clustering patterns in our PCoA show the control samples from the 1st through the 8th generation cluster into one large group, while the high biomass and random selection samples separate from this group together as the generations of plantings progress (Figure 2.4). Differences in bacterial community composition between the selection treatments as the generations advanced are also displayed in our heatmap constructed using Euclidean distances and average linkage clustering (Supplemental Figure 2.1).

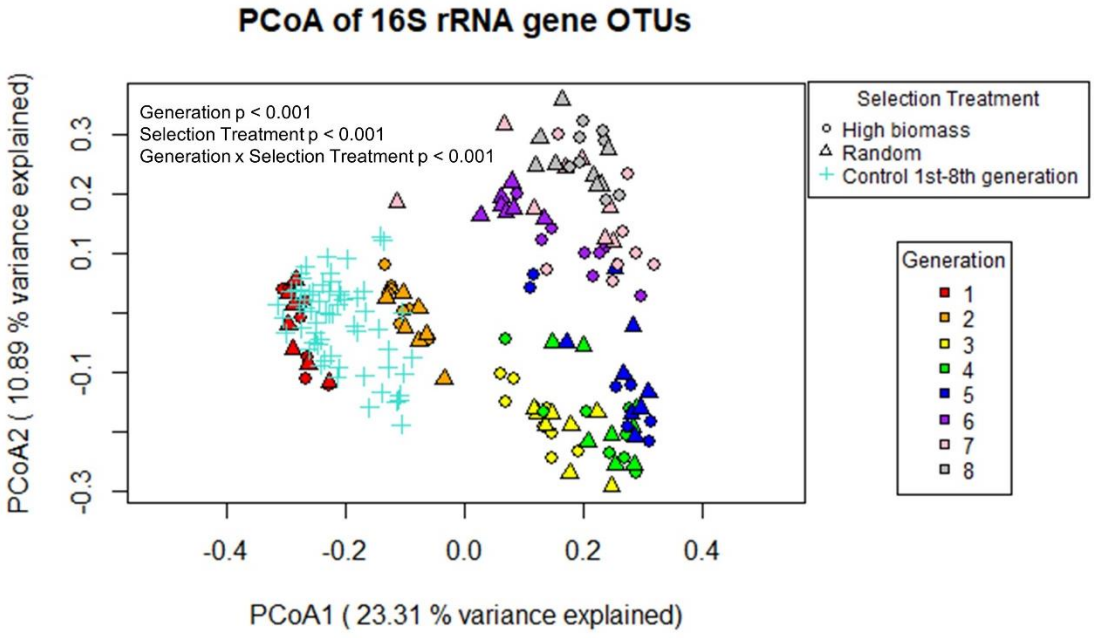


Figure 2.4 Principal coordinates analysis (PCoA) of the rhizosphere microbial community based on 16S rRNA gene sequencing on MiSeq Illumina. Circles indicate microbiomes from the high biomass selection treatment and triangles indicate those derived from random selection. The crosses represent the control selection treatment, which are used to measure against genetic drift and growth chamber adaptation. For the high biomass and random selection treatments, the different generations are represented by color: 1st generation is red, 2nd generation is orange, 3rd generation is yellow, 4th generation is green, 5th generation is blue, 6th generation is purple, 7th generation is pink, and 8th generation is grey. For the control selection treatment, all generations (1st through 8th) are represented in teal to highlight a distinct cluster of control samples. PERMANOVA revealed significant effects of Generation ($p<0.001$), Selection Treatment ($p<0.001$), and the interaction between the two factors ($p<0.001$) on bacterial community composition.

In order to assess the effect of Generation and Selection Treatment on abundances of specific bacterial taxa, which may help explain functional shifts in the rhizosphere bacterial community, we ran ANOVAs on the relative abundances of the top seven most abundant bacterial families. For our ANOVA models, we included Generation, Selection Treatment, and the interaction between the two as model factors, then report here pairwise comparisons of the abundances of different bacterial families in the 8th generation between the selection treatment. In other words, the ANOVAs and pairwise comparisons reported here are from analysis of the Generation x Selection Treatment interaction term. Our analysis shows shifts in the quantities of bacterial families such as *Chitinophagaceae* ($p=0.022$) (Supplementary Figure 2.2). Members of this family, which contains plant growth promoting species such as *Arachidicoccus rhizosphaerae*[29], increased in relative abundance in the high biomass and random selection treatments by 378.1% ($p=0.0007$) and 329.5% ($p=0.026$) compared to the control treatment in the 8th generation (Supplementary Figure 2.2). Our analysis of *Flavobacteriaceae* ($p =2.98e-7$), which is a member of the widespread

Bacteroidetes phylum, and *Oxalobacteraceae* (p=1.6e-8), belonging to the functionally diverse *Proteobacteria* phylum, revealed further shifts in abundances between selection treatments in the 8th generation. The abundance of *Flavobacteriaceae* was 83.1% lower in the high biomass selection treatment ($p<0.001$) and 84.2% lower in the random selection treatment ($p<0.001$) compared to the control treatment. The abundance of *Oxalobacteraceae* was 51.1% lower in the high biomass selection treatment (p=0.0082) and 60.6% lower in the random selection treatment (p=0.0006), compared to the control treatment. Analysis of *Sphingobacteriaceae*, also a member of the *Bacteroidetes* phylum, revealed further shifts between selection treatments in the 8th generation (p=8.69e-9). When analyzing *Sphinobacteriacaea* abundance between selection treatments in the 8th generation, we found the high biomass selection treatment had 52.3% lower abundance compared to the random selection treatment ($p<0.001$) and 37.6% lower abundance compared to the control treatment (p=0.0136).

Analysis of Shannon diversity indexes revealed differences in bacterial diversity between all three selection treatments (p=3.14e-08) (Figure 2.5). The control treatment had the greatest bacterial diversity, with a diversity index 3.7% greater than the random selection treatment (p=0.018) and 7.8% greater than the high biomass selection treatment ($p<0.001$). In addition, the random selection treatment had a diversity index 3.9% greater than the high biomass selection treatment (p=0.016). This decrease in diversity could be indicative of ecological filtering[30] or a similar process in the high biomass selection treatment.

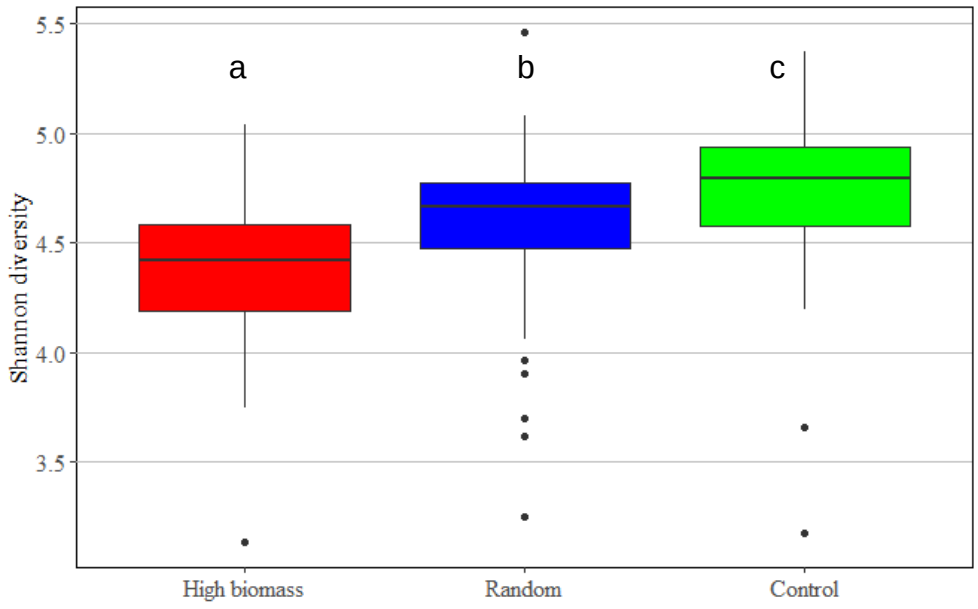


Figure 2.5 Analysis of mean Shannon diversity index for the three selection treatments using microbial data from all eight generations of soil samples sequenced.

2.2.5 Microbiome network analysis

To compare how bacterial group dynamics shifted between the different selection treatments in our experiment, we performed an Extended Local Similarity Analysis (eLSA) using bacterial sequence data from the 1st through 8th generation. Briefly, an eLSA is a modified version of Local Similarity Analysis (LSA), which measures local and potentially time-delayed co-occurrence patterns in time series data[31]. Unlike LSA, eLSA can reveal statistically significant local and potentially time-delayed association patterns in replicated time series data, such as the bacterial sequence data from this experiment. Results from the analysis can then be used to construct an association network consisting of nodes (representing bacterial Operational Taxonomic Units, or OTUs), edges (representing potential relationships between nodes), and modules (clusters of nodes), which together can be used to

visualize changes in microbial interaction networks (i.e. microbial ecological relationships).

The eLSA revealed that the three selection treatments differ in edge connections with the same node set, which included OTUs that had a relative abundance above 1% in at least 1 sample. The high biomass selection treatment showed the greatest number of edge connections, followed by the random selection treatment, then the control treatment (Figure 2.6A). Notably, the high biomass selection treatment visually showed a densely connected cluster of OTUs not observed in the random or control selection treatments when the three networks are stacked (Figure 2.6B). When plotting the degree distribution (number of edges connected to a specific node) of all three selection treatments, the high biomass selection treatment had substantially more nodes that are more connected when the degree is larger than 12 (Supplementary Figure 2.3A). Changing the alpha threshold for the eLSA analysis, which determines if a correlation between two OTUs is significant, yields consistent results where the network of the high biomass selection treatment is always the most dense and indicates robustness of eLSA on this data set (Supplementary Figure 2.3B).

In addition to differences in edge connections between the three selection treatments, we found decreased modularity of the bacterial community under high biomass selection, which is indicative of a more highly organized network topology displaying tightly connected bacterial OTUs. The high biomass bacterial communities showed the lowest modularity of 0.546, followed by 0.562 in the random selection group and 0.741 in the control group. Module size, taxonomic composition, and the interaction strength (connectivity, as indicated by the node size) differed across the

selection treatments and control. Figure 2.6C shows the two largest group formations within the network and their taxonomic profile. The greater density of the two major groups in the high biomass selection treatment compared to those generated from random selection shows stronger interconnection within the groups.

To further quantify these visual differences, we used our sample edge lists to run a publicly available network model tool, known as NetShift[32], which quantifies bacterial community rewiring and changes between two data sets. With the analysis, we sought to perform pairwise comparisons of global graph properties for the three selection treatments. Our analysis revealed the high biomass selection treatment overall had higher network density (portion of potential connections in a network that are actual connections), lower average pathlength (fewer edges between nodes), more total edges (more connections), and more exclusive edges than the random selection and control treatments (Figure 2.6D). The results indicate a more connected bacterial community in the high biomass selection treatment compared to the random and control treatments, which was also visually observed in our microbial network constructed from the eLSA.

Figure 2.6 Extended Local Similarity Analysis (eLSA) networks representing microbial taxa (OTU) interactions. A.) Networks of the three different selection treatments comprised of sequencing data from all eight generations. Size of the node is proportional to the degree of the node (the number of connections a node has). B.) Network of the combined three selection treatments. C.) Two largest modules (clusters) from each selection treatment and control using our network data. The networks represent OTU interactions for different selection treatments using sequencing data from all eight generations. Size of the node is proportional to the level of connectivity and colors represent taxonomic groups of bacteria. D.) Comparisons of network density, average path length, total edges, and exclusive edges between selection treatments using NetShift. We used the same edge lists generated for our LSA to run our NetShift analysis.

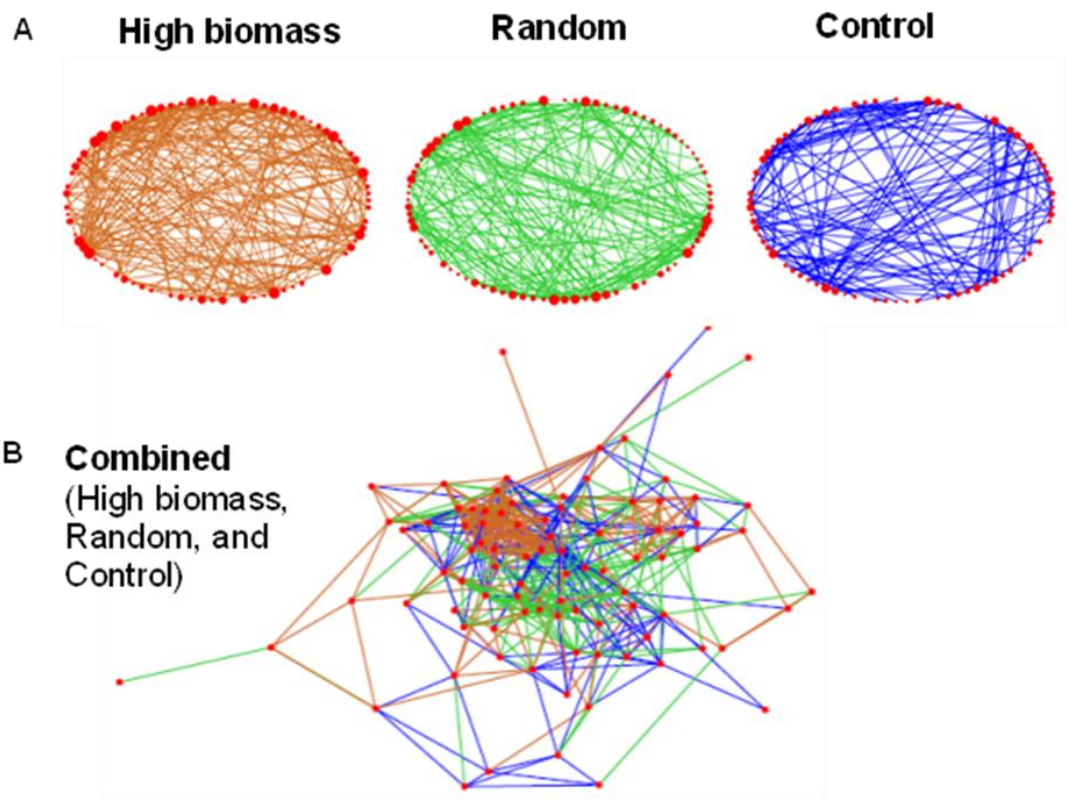
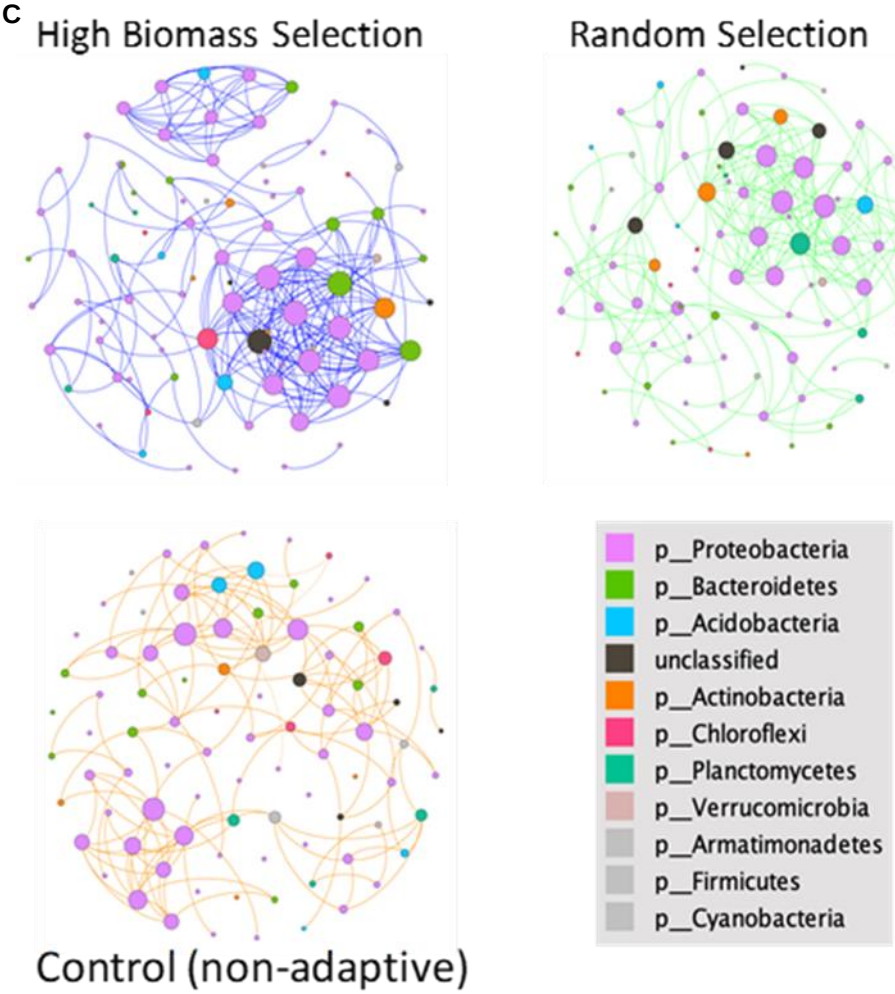

A
High biomass
Random
Control
B
Combined
(High biomass,
Random, and
Control)

Figure 2.6 (Continued)

C



D

	Density	Average path length	Total edges	Exclusive edges
Sel. Trt. 1 vs. Sel. Trt. 2	Sel. Trt. 1 vs. Sel. Trt. 2	Sel. Trt. 1 vs. Sel. Trt. 2	Sel. Trt. 1 vs. Sel. Trt.2	Sel. Trt. 1 vs. Sel. Trt. 2
High biomass vs. Random	0.090 vs. 0.066	3.9 vs. 5.5	175 vs. 128	148 vs. 101
High biomass vs. Control	0.11 vs. 0.060	3.2 vs. 3.8	145 vs. 92	134 vs. 81
Random vs. Control	0.077 vs. 0.064	5.0 vs. 3.7	118 vs. 118	106 vs. 106

2.3 Discussion

The objective of the experiment was to examine changes in rhizosphere bacterial group dynamics across generations of selection for high biomass growth of a non-mycorrhizal plant under soil nitrogen limiting conditions. In essence, we used a specific phenotype of the host plant (high aboveground biomass) to report the collective actions of a bacterial community that are modifiable over the course of selection[4]. We used an ecological filtering approach[30] in this experiment to remove rhizosphere microbiota that were not associated with the four highest performing units for each generation of plantings in the high biomass selection treatment. The plants used in this experiment did not undergo selection pressure. Instead, seeds of the same parental origin were used across generations of plantings, which means the genetic profile of plants from the 1st through 9th generations were the same. The PCoA indicated divergence in the bacterial communities between the high biomass and random treatments starting in the 6th generation. The bacterial network analysis using eLSA and NetShift indicate that bacteria associated with increased aboveground biomass may be coordinating group-level behaviors that impact fitness traits, such as the higher seed yield in the high biomass selection treatment compared to the control treatment. Altogether, the stronger connectivity of specific taxonomic groups forming within the rhizosphere of the high biomass plants over time, along with the observed changes in bacterial community composition, may have contributed to shifts in plant seed yield and plant N dynamics.

The eLSA network model revealed that the selection process had a significant effect on bacterial group dynamics in each treatment over the course of the

experiment. A notable outcome of the experiment was the formation of bacterial groups (modules) that differed in composition and interaction intensity across the selection treatments. Modules or clusters are biologically relevant units whose interaction patterns provide information in understanding the function of the network[33]. For example, past research has suggested highly modular microbiota could play a key role in plant immune responses to pathogens[34]. The significance of these groups is how they may reveal emergent properties of microbiomes associated with alterations in the host's phenotype, such as seed yield. In this study, the selective pressure on plant biomass resulted in lower diversity and higher connectivity in the high biomass selection treatment, suggesting a highly organized bacterial community formed in the rhizosphere.

Analysis using NetShift provided quantitative data confirming changes in bacterial community dynamics across treatments. NetShift indicated a more interconnected bacterial community: lower average pathlength, which describes the average number of steps connecting one node to the next, and total edges, which represents connectedness among community members[32]. Previous work has highlighted the influence of external pressures on bacterial community properties, such as density. Faust et al. showed that microbial networks in bulk soils are generally less dense than host-associated networks[35]. Considering empirical results such as this, it is conceivable the selection pressure for increased plant biomass production resulted in highly connected assemblies of bacterial groups that coordinate complex functions that are beneficial to plant growth and N use efficiency.

It is well known that the rhizosphere bacterial community regulates several

plant activities, such as nutrient uptake, through a variety of mechanisms. For example, rhizobia form nodules in legume roots and fix atmospheric nitrogen for their host[36], while free-living microbes, such as phosphorus solubilizing bacteria, regulate plant P availability through collective production of extracellular enzymes[37]. It is possible that the changes observed in plant tissue N content in this experiment were a result of changes in the nutrient cycling activities of the rhizosphere bacterial community between the selection and control treatments. Our sequencing data indicate that the bacterial communities of the high biomass selection treatment showed distinct characteristics from the bacterial communities of the random selection and control treatments. Among these distinct characteristics were a decrease in bacterial diversity, which could indicate ecological filtering[30] for more efficient N cycling microorganisms. An example is the shift in the relative abundance of *Chitinophagaceae* bacteria which have previously been shown to be associated with increased N cycling activity[38].

By repeatedly selecting for rhizosphere bacterial communities associated with greater aboveground biomass production, it is conceivable the rhizosphere bacterial communities in the high biomass selection treatment became enriched with taxa that aid in the mineralization of N, which is an essential nutrient for seed production and development[39, 40,41]. It is important to note that the plants in this study were grown under nitrogen limitation, specifically 20 ml of 100 ppm N fertilizer as one dose in each generation. In nitrogen-limiting conditions such as this, plants may enrich the rhizosphere for N capturing microorganisms via root exudation[42], which could impact soil N availability or N use efficiency in the host.

Stable isotopes provide an opportunity to assess potential contributions of the rhizosphere microbiome to plant N utilization[43,44]. Changes in foliar $\delta^{15}N$ result from corresponding alterations in the $\delta^{15}N$ of the soil solution or some fundamental change in within-plant N metabolism, which is often interpreted as modifications in soil processes, especially nitrification and denitrification[45]. In this study, it is challenging to determine the soil or metabolic processes that were impacted between the 1st and 9th generations. However, the observed changes in foliar $\delta^{15}N$ (Table 2.1) suggest some alteration in soil N cycling by the rhizosphere microbial community or metabolism in the high biomass selection treatment. There are several potential ways in which the rhizosphere could change soil N cycling. For example, differences in foliar $\delta^{15}N$ values among the treatments could be a result of differential rates of nitrification and denitrification between groups[46]. Similarly, differences in $\delta^{15}N$ values could indicate contrasts in fertilizer- versus soil-N use by plants across treatments[47], which could be mediated by the microbiome. Plant fractionation of nitrogen isotopes may have differed across selection treatments[48]. The increases in total foliar N and N agronomic use efficiency, and the $\delta^{15}N$ data altogether suggest alterations in rhizosphere N cycling and plant N utilization in the high biomass selection treatment. As shown in our regression analysis, this alteration in available N could be a major mechanism by which bacterial communities in the high biomass selection treatment altered the yield and NUE phenotypes of *B. rapa*.

In addition to potentially promoting N cycling, the rhizosphere bacterial community in the high biomass selection treatment could have also promoted the uptake of other important plant nutrients, such as phosphorus[49,50], but plant tissue P

and soil extractable orthophosphate were not analyzed in this study. It is also possible the shifts in the high biomass bacterial communities over the course of selection represent the elimination of organisms that inhibit plant productivity[51], as we observed a decline in diversity in the high biomass selection treatment compared to the other two treatments. Our analysis revealed lower abundances of taxa belonging to the family *Sphingobacteriaceae* in the high biomass treatment. This lower abundance may have benefited plants in the high biomass treatment as members of this family have been associated with plant and animal diseases[52,53]. We expected ecological filtering of the bacterial community, and thus a decline in diversity as observed in our experiment, since our selection for high biomass production likely altered bacterial community assembly processes[54,55].

We have shown in this experiment that it is possible to develop rhizosphere bacterial communities associated with desirable plant traits in agriculturally relevant crops, such as *B. rapa,* through artificial group selection over multiple generations. Our results show seed yield and plant tissue N may have been altered by the microbiomes assembled under high biomass selection. Random selection appears to also have effects on plant and soil responses. However, the more interconnected microbial network along with enrichment for N cycling bacterial taxa and changes to N cycling dynamics observed in the high biomass selection treatment suggest significant effects of high biomass selection that may ultimately drive enhanced seed yield and NUE compared to the no selection control, which was not observed in the random selection treatment. Our work demonstrates that group-level behaviors within the rhizosphere can be altered under selection and may serve a key role in modifying

plant host phenotypes related to productivity. This project contributes to the growing area of research showing that rhizosphere microbiomes influence crop traits and have the potential to promote plant productivity and yield[56,57] without the use of genetic engineering or gene-editing techniques. As this experiment focused on bacterial communities, future studies could be designed to analyze how microbiome composition or behaviors can be modified to alter plant fitness and other phenotypes. These types of microbial experimental systems could ultimately help develop a better understanding of plant-microbe interactions that can be utilized for regenerative agriculture.

2.4 Materials & methods

2.4.1 Growth chamber conditions

All plants in this experiment were grown in a growth chamber set at 350 µmol of light, 10% humidity, alternating between 28ºC for 16 hours and 23ºC for 8 hours for every generation (Percival-Cornell University Weill Hall Life Sciences Growth Chamber Facility, Ithaca, NY, USA).

2.4.2 Selection on rhizosphere microbiomes

Bacterial inoculants for the high biomass and random selection treatment were created using an iterative selection process over nine generations adapted from Swenson et al.[10] and Panke-Buisse et al.[11]. In the 1st generation, 0.5 L pots were filled with LM-111 container soil (Lambert Peat Moss, Inc., Riviere-Ouelle, Quebec, Canada) that was autoclaved twice with a 24-hour resting period in between cycles.

After sterilization, the pots were inoculated with 20 mL of a bacterial inoculant created from field soils gathered from diverse organic farms in Ithaca, New York (42.4440° N, 76.5019° W). This initial bacterial inoculant was created by compiling and homogenizing the field soils, then adding 55 mL of the mix to 500 mL of sterilized deionized water. The mix was then shaken with 30 mL of sterilized glass beads at 180 opm for one hour. The resulting slurry was then filtered through four layers of sterilized food preparation cloth, which retains a large portion of the soil bacterial community in the filtered slurry while eliminating large soil particles. Aliquots of this mix were frozen at -20ºC to serve as bacterial inoculants for the control selection treatment in the following generations. In every generation, all three selection treatments (high biomass, random, control) had 15 replicate pots total. The high biomass and random selection inoculants were kept frozen at -20ºC for two days and thawed, along with the control inoculants, prior to the start of each generation. This additional process of freezing and thawing at each generation was set in place to minimize potential confounding differences between the control samples and the selection treatments.

A single *B. rapa* seed pool consisting of a single genotype was gathered for use throughout the entire experiment. This prevented any genetic changes at the plant host level, helping to ensure any changes in plant phenotype observed in our experiment were a result of changes in the bacterial community. Before planting, seeds were surface sterilized with 2.5% bleach for 15 minutes and then rinsed with sterile water at least five times before seeding. Eight seeds were planted into each pot to ensure adequate germination. Pots were then thinned to five plants per pot upon emergence.

Pots were placed randomly into trays covered with sterile clear plastic domes to help prevent outside contamination. The trays were arranged randomly across two growth chamber benches. Plants were watered daily with filter sterilized water and received one dose of 20 mL filter sterilized Jacks 21-5-20 + Epsom salt fertilizer, at a concentration of 100 ppm N. Plants were grown for 10 days, then the stem and leaf aboveground biomass of all plants in each pot were harvested by cutting at the base of each plant's cotyledon leaves. The composited plant samples from each pot were dried at 65ºC for 24 hours to record total aboveground dry biomass production (i.e, the biomass reported for each pot is the biomass production of all plants, not an average).

Rhizosphere soils were collected and stored at 4ºC upon harvest. Soil was collected by gently breaking up the root ball, holding the roots, and gently shaking away unattached soil. For the high biomass selection treatment, we kept soils from the eight pots with the highest plant biomass production. Rhizosphere soil from the four plants with the greatest total aboveground biomass production in this selection treatment was used to create the high biomass inoculant for the following generation. After creating a homogenous mix of the four soils, 100 mL of the soil mix was dissolved into 500 mL of sterile DI water. The slurry was then shaken with 30 mL of sterile glass beads for one hour at 180 opm and filtered through four layers of sterile cheesecloth. The resulting mix was then used as the bacterial inoculant for plants in the successive generation of the high biomass selection treatment with each pot receiving 20 mL of the inoculant. For the random selection treatment, the rhizosphere soil from eight random plants was kept and soils from four random plants within this subgroup were used to create bacterial inoculants for the next generation of plants in

this selection treatment. The four randomly chosen soils were homogenized and 100 mL was then mixed with 500 mL of sterile DI water, shaken at 180 opm with 30 mL of sterile glass beads for one hour, then filtered through four layers of sterile cheese cloth. This filtrate was then used to inoculate plants in the random selection treatment in the following generation with each plant receiving 20 mL of the mix. Soils from eight random plants were kept from the control selection treatment, which were only used for analysis of the rhizosphere bacterial community (i.e. they were not used to make bacterial inoculants for the following generations). We repeated this exact selection process for a total of 9 generations (Figure 2.1).

At the final generation (9th generation), we grew the plants to seed in order to assess the effect of our selection process on *B. rapa* seed yield, which required plants to be grown for 31 days. Upon flowering (~18 days), plants were hand pollinated daily by transferring pollen to neighbor *B. rapa* plants. In order to initiate seed ripening, we ceased watering the plants two days before harvesting. After harvesting, stem and leaf tissue was collected and dried at 65ºC for 24 hours then weighed to record total dry aboveground biomass production. Seeds were collected from all plants every pot and weighed to assess seed yield. Because the plants required dried soil to set seed in this generation, we were not able to collect soil samples for sequencing the rhizosphere bacterial community for this final time point.

2.4.3 Bacterial 16S rRNA gene sequencing

Bacterial DNA was extracted from frozen rhizosphere soil samples collected from the 1st through 8th generation using a PowerSoil DNA Isolation Kit (Qiagen,

Hilden, Germany) following the manufacturers protocol for more absorbent soils. Approximately 150 to 200 mg of soil from each sample was used for extraction. 16S rRNA gene sequences were amplified using PCR primers described in Herlemann et al.[58], which target the bacterial/archeal 16S rRNA gene variable region for pair-end Illumina (Illumina, Inc., San Diego, CA, USA) barcoded sequencing. After an initial clean-up of the samples using the HighPrep PCR Clean-up System (MAGBIO Genomics, Gaithersburg, MD, USA), unique Index Primers were attached to amplicons in each sample with a second PCR cycle. The indexed samples were cleaned and normalized using the SequalPrep Normalization Plate Kit (Thermo Fisher Scientific, Waltham, MA, USA). Sample normalization was followed by pooling 5 ul of each sample into one composite sample. The pooled sample was then ran on a 1.2% agarose gel with SyberSafe added and the target band was excised. The DNA was then extracted from the gel using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). The final pooled sample was then sent to the Cornell Genomics Facility (Cornell University, Ithaca, NY, USA) for sequencing on the Illumina MiSeq platform.

2.4.4 Plant tissue N analysis

N content and isotope ratio of plant tissue in this experiment was analyzed at the Cornell Stable Isotope Laboratory (COIL) using a Thermo Delta V isotope ratio mass spectrometer (IRMS) interfaced to an NC2500 elemental analyzer. Tissue from four random samples from each Generation x Selection Treatment group were submitted to COIL for analysis. Measurements of %N were used along with

aboveground dry biomass production to calculate total plant N.

2.4.5 Statistics

All statistical analyses were performed using R statistical software (Rproject.org). To analyze the effect of our selection process, we analyzed plant biomass production, seed yield, and N data from the 9th generation using linear modeling. We focused on plant data from this generation and analyzed it separate from other generations given it was the final generation of the experiment and was grown for a longer period than the generations before. One-way ANOVAs were performed for each model to determine the effect of selection treatment on each outcome in the 9th generation. Least squared means was then used to perform pairwise comparisons of group means between selection treatments. Normality of residuals and homoscedasticity were verified for each model using plotting methods. To show our selection process was responsible for any differences observed in our ANOVAs at the 9th generation, we performed these same procedures on data from the 1st generation, which showed no differences in treatment means for any response variable (Table 2.1). Data transformations were performed as necessary for all analyses to meet ANOVA assumptions.

Pearson correlations were performed using foliar N data, aboveground dry biomass production, and seed yield data from the 9th generation to assess potential linkages between plant nitrogen uptake and productivity. Data were first visualized in scatterplots using the ggscatter function to verify linear relationships between variables. Normality for each variable was then verified using the Shapiro-Wilk

normality test. We then ran our correlation analysis using the cor.test function in R, which provided Pearson correlation coefficients along with p-values to indicate if a relationship was significant.

2.4.6 Analysis of microbiome sequences

The pipeline from the Brazilian Microbiome Project (http://www.brmicrobiome.org/) was modified to process our sequences. Mothur v. 1.3613 was used to merge paired-end sequence (make.contigs), trim off primers (trim.seqs, pdiffs = 2, maxambig = 0), remove singletons (unique.seqs → split.abund, cutoff = 1) and classify sequences (97% similarity). OTUs that were suspected to not be of bacterial origin were removed (remove.lineage). Qiime v. 1.9.114 was used to cluster OTUs and create an OTU table.

Following sequence processing, the dataset was rarefied to 2988 sequences per sample (i.e. the minimum number of reads per sample in this dataset). We then converted OTU counts in each sample to percentage, then calculated Bray-Curtis distances using vegdist. We then performed a PCoA analysis on the Bray-Curtis distances calculated using cmdscale. We then calculated % variance explained by the first two axes, then plotted the calculated Bray-Curtis distances in an ordination to visualize the effect of Generation and Selection Treatment on community composition. The adonis package was used to perform a PERMANOVA on our sequence data and assess the effects of Generation, Selection Treatment, and the interaction between the two factors on community composition. Due to a lack of adequate post hoc tests in adonis, we were not able to do pairwise comparisons.

PERMANOVA requires homogeneity of variance, which we analyzed using the betadisper function. Shannon diversity indices we calculated for each sample using the diversity function. We then created a linear model assessing the effect of Selection Treatment alone on the Shannon diversity indices of our samples (i.e. Generation and an interaction term were not included in the model), then ran an ANOVA on the model. We then used least squares means to do pairwise comparisons of diversity indices.

Changes in the relative abundances (%) of different bacterial taxa across generations and selection treatment were visualized by creating stacked barplots. We first aggregated our OTU data, taxonomic data, and metadata using the phyloseq function, then identified the seven most abundant bacterial families using the sort function. We then calculated % abundance for each bacterial family in the different Generation x Selection Treatment groups then plotted the abundances in a stacked barplot. ANOVAs were then performed on the relative abundances of each taxonomic group using linear modeling and group means were compared using least squared means to quantitatively assess differences in taxonomic abundances across generations and selection treatments. Normality of residuals and homoscedasticity were verified for each model using plotting methods, and transformations were performed as necessary to meet ANOVA assumptions. A heatmap was also created to further visualize changes in the relative abundances of bacterial taxa. We first identified the maximum abundance of each individual OTU in every sample, then filtered out OTUs with maximum abundances below 7.5%, which we found was an appropriate cutoff to visually see shifts in OTU abundances across generations and selection treatments in

our heatmaps. After filtering, we used the pheatmap function in R to visualize shifts in OTU relative abundances across generations and selection treatment using euclidean distances of OTU relative abundances and average linkage clustering.

2.4.7 Extended local similarity analysis (eLSA) and network construction

In order to study the effect of our selection process on bacterial community co-occurrence patterns, which can represent important ecological interactions and processes[59], we constructed dynamic networks for each selection treatment using bacterial sequence data from the 1st through 8th generation. To construct our networks, we employed Extended Local Similarity Analysis (eLSA), a similarity-based method that uses dynamic programming to build association networks from time series data[31,60]. The algorithm used is summarized in Supplementary Table 2.1. In a standard LSA method, we are provided sequence data in which only one sample is available for each time step for each sequence we have. We first used pairs of OTU time series data of the same length as inputs. Positive and negative correlation scores between each pair were then calculated, which we used to determine local similarity scores. Larger local similarity scores indicated stronger potential relationships between the two paired OTUs. However, this is too stringent in our case where multiple samples are available within one sequence for each time step. Hence, instead of using LSA, we applied the eLSA method in this work where a sufficient statistic is computed for each time step given all the samples observed at that time step and then applied the standard LSA method to compare the local similarity score. Here, we used the simple average method to summarize the replicated time series data[31]. For example, if we have 10

samples at each time step, we can then compute the average value of the 10 samples at the same time step and apply LSA on this new sequence data where the data at each time step is the average value for all samples. In this paper, we take the mean of all the samples observed at every time step.

Next, we used our LSA scores between OTUs to construct a network representing the strong relationships within bacterial communities. We defined a graph G (V, Ɛ) where V = {v1, v2, ..., vm} represents different OTUs and (vi, vj) ∈ Ɛ if there is a strong correlation between OTUs represented by vi, vj. Strong correlations here are defined using the significance level α.

To compute p-values for two OTU time series data O1 and O2, we first randomly permuted O1 and O2 and computed LS (O1, O2). P-value here was defined as the probability that this local similarity value is at least as large as the local similarity value with the non-permuted data. We computed an upper bound for the p-value[61]. After establishing the upper bound for the p-value between O1 and O2, we defined a global significance level α and filtered the upper bounds with α. Specifically, if the p-value upper bound between OTU time series represented by vi and vj was larger than α, we drew an edge between vi and vj. For our analysis, only OTUs with a relative abundance of at least 1% in any of the samples were used, which gave us a network with 90 nodes representing different OTUs. We used a time delay of 1 and the significance level α to be 1.

2.5 References

1. Hawkes, C. V., Bull, J.J., & Lau, J.A. Symbiosis and stress: how plant microbiomes affect host evolution. *Philos. T. R. Soc. B.* 375(1808): 20190590 (2020).

2. Leopold, D. R. & Busby, P.E. Host Genotype and Colonist Arrival Order Jointly Govern Plant Microbiome Composition and Function. *Curr. Biol.* 30(16): 3260-3266.e3265 (2020).

3. Morella, N. M. et al. Successive passaging of a plant-associated microbiome reveals robust habitat and host genotype-dependent selection. *Proc. Natl. Acad. Sci. U.S.A.* (2): 1148-1159 (2020).

4. Garcia, J. & Kao-Kniffin, J. Microbial Group Dynamics in Plant Rhizospheres and Their Implications on Nutrient Cycling. *Front. Plant Sci.* 9(1516) (2018).

5. Marschner P. Plant-Microbe Interactions in the Rhizosphere and Nutrient Cycling. In: Marschner P., Rengel Z. (eds) *Nutrient Cycling in Terrestrial Ecosystems.* Soil Biology, vol 10. Springer, Berlin, Heidelberg (2007).

6. Vannier, N., Agler, M., & Hacquard, S. Microbiota-mediated disease resistance in plants. *PLoS Pathog.* 15(6): e1007740 (2019).

7. Wei, Z. et al. Initial soil microbiome composition and functioning predetermine future plant health. *Sci. Adv.* 5(9): eaaw0759 (2019).

8. Bender, S. F., Wagg, C., & van der Heijden, M.G.A. An Underground Revolution: Biodiversity and Soil Ecological Engineering for Agricultural Sustainability. *Trends Ecol. Evol.* 31(6): 440-452 (2016).

9. Dessaux, Y., Grandclemént, C., & Faure, D. Engineering the Rhizosphere. *Trends Plant Sci*. 21(3): 266-278 (2016).

10. Swenson, W., Wilson, D. S., & Elias, R. Artificial ecosystem selection. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9110–9114 (2000).

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

12. van den Bergh, B., et al. Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution. *Microbiol. Mol. Biol. Rev.* 82(3) (2018).

13. Garcia, J, & Kao-Kniffin, J. Can dynamic network modelling be used to identify adaptive microbiomes? *Funct Ecol.* (2020) 34: 2065– 2074.

14. Faust, K., & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10(8): 538-550 (2012).

15. Wilson, D. & Wilson, E. Evolution "for the good of the group". *Am. Sci.* 96, 380–389 (2008).

16. de la Fuente Cantó, C. et al. An extended root phenotype: the rhizosphere, its formation and impacts on plant fitness. *Plant J.* 103: 951-964 (2020).

17. Sachs, J., Mueller, U., Wilcox, T., & Bull, J. The evolution of cooperation. *Q. Rev. Biol.* 79, 135–160 (2004).

18. Harrington, K. & Sanchez, A. Eco-evolutionary dynamics of complex social strategies in microbial communities. *Commun. Integr. Biol.* 7(1) (2014).

19. Rauch, J., Kondev, J., & Sanchez, A. Cooperators trade off ecological resilience and evolutionary stability in public goods games. *J. R. Soc. Interface* 14(127) (2017).

20. Sexton, D. & Schuster, M. Nutrient limitation determines the fitness of cheaters in bacterial siderophore cooperation. *Nat. Commun.* 8, 230 (2017).

21. Schmidt, J.E., Kent, A.D., Brisson, V.L., & Gaudin, A.C.M. Agricultural management and plant selection interactively affect rhizosphere microbial community structure and nitrogen cycling. *Microbiome* 7, 146 (2019).

22. Turner, T. et al. Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *ISME J.* 7, 2248–2258 (2013).

23. Jones, D.L., Nguyen, C. & Finlay, R.D. Carbon flow in the rhizosphere: carbon trading at the soil–root interface. *Plant Soil* 321, 5–33 (2009).

24. George, E., Marschner, H., & Jakobsen, I. Role of Arbuscular Mycorrhizal Fungi in Uptake of Phosphorus and Nitrogen From Soil. *Crit. Rev. Biotechnol.* 15(3-4): 257-270 (1995).

25. Hodge, A., & Fitter Alastair, H. Substantial nitrogen acquisition by arbuscular mycorrhizal fungi from organic material has implications for N cycling. *Proc. Natl. Acad. Sci. U.S.A.* 107(31): 13754-13759 (2010).

26. Lambers, H., & Teste, F. P. Interactions between arbuscular mycorrhizal and non-mycorrhizal plants: do non-mycorrhizal species at both extremes of nutrient availability play the same game. *Plant Cell Environ* 36(11), 1911-1915 (2013).

27. Delaux, P.M. et al. Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. *PLoS Genet.* 10: e1004487 (2014).

28. Anas, M. et al. Fate of nitrogen in agriculture and environment: agronomic, eco-physiological and molecular approaches to improve nitrogen use efficiency. *Biol. Res.* 53, 47 (2020).

29. Madhaiyan, M. et al. *Arachidicoccus rhizosphaerae* gen. nov., sp. nov., a plant-growth-promoting bacterium in the family *Chitinophagaceae* isolated from rhizosphere soil. *Int. J. Syst. Evol. Microbiol.* 65(Pt 2): 578-586 (2015).

30. Song, H. et al. Environmental filtering of bacterial functional diversity along an aridity gradient. *Sci. Rep.* 9, 866 (2019).

31. Xia, L.C. et al. Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. *BMC Syst. Biol.* 5, S15 (2011).

32. Kuntal, B.K., Chandrakar, P., Sadhu, S., & Mandhi, S.S. 'NetShift': a methodology for understanding 'driver microbes' from healthy and disease microbiome datasets. *ISME J*. 13, 442–454 (2019).

33. Layeghifard, M., Hwang, D.M., & Guttman, D.S. Disentangling Interactions in the Microbiome: A Network Perspective. *Trends Microbiol*. 25(3): 217-228 (2017).

34. Hu, Q. et al. Network analysis infers the wilt pathogen invasion associated with non-detrimental bacteria. *NPJ Biofilms Microbiomes* 6, 8 (2020).

35. Faust, K. et al. Cross-biome comparison of microbial association networks. *Front. Microbiol.* 6(1200) (2015).

36. Rengel, Z. & Marschner, P. Nutrient availability and management in the rhizosphere: exploiting genotypic differences. *New Phytol.* 168: 305-312 (2005).

37. Marschner, P. The Role of Rhizosphere Microorganisms in Relation to P Uptake by Plants. In P. White & J. Hammond, *The Ecophysiology of Plant-Phosphorus Interactions* (pp. 165-167). Springer (2008).

38. Repert, D., Underwood, J., Smith, R., & Song, B. Nitrogen cycling processes and microbial community composition in bed sediments in the Yukon River at Pilot Station, *J. Geophys. Res. Biogeosci.* 119, 2328– 2344 (2014).

39. Rolletschek, H., et al. Ectopic expression of an amino acid transporter (VfAAP1) in seeds of Vicia narbonensis and pea increases storage proteins. *Plant Physiol.* 137(4): 1236-1249 (2005).

40. Sanders, A. et al. AAP1 regulates import of amino acids into developing Arabidopsis embryos. *Plant J.* 59: 540–552 (2009).

41. Carter, A.M. & Tegeder, M. Increasing nitrogen fixation and seed development in soybean requires complex adjustments of nodule nitrogen metabolism and partitioning processes. *Curr. Biol.* 26: 2044–2051 (2016).

42. Meier, I.C. et al. Root exudation of mature beech forests across a nutrient availability gradient: the role of root morphology and fungal activity. *New Phytol.* 226: 583-594 (2020).

43. Xu, Y., He, J., Cheng, W., Xing, X., & Li, L. Natural ^{15}N abundance in soils and plants in relation to N cycling in a rangeland in Inner Mongolia. *J. Plant Ecol*. 3(3): 201-207 (2010).

44. Henneron, L. et al. Rhizosphere control of soil nitrogen cycling: a key component of plant economic strategies. *New Phytol.* 228: 1269-1282 (2020).

45. Hobbie, E. A., & Högberg, P. Nitrogen isotopes link mycorrhizal fungi and plants to nitrogen dynamics. *New Phytol*. 196(2): 367-382 (2012).

46. Zhou, S. et al. Assessing nitrification and denitrification in a paddy soil with different water dynamics and applied liquid cattle waste using the ^{15}N isotopic technique. *Sci. Total Environ.* 2012 Jul 15;430:93-100 (2012).

47. Fuertes-Mendizábal, T. et al. ^{15}N Natural Abundance Evidences a Better Use of N Sources by Late Nitrogen Application in Bread Wheat. *Front. Plant Sci.* 9(853) (2018).

48. Yoneyama, T., Omata, T., Nakata, S., & Yazaki, J. Fractionation of Nitrogen Isotopes during the Uptake and Assimilation of Ammonia by Plants. *Plant Cell Physiol.* 32(8): 1211-1217 (1991).

49. Vacheron, J. et al. Plant growth-promoting rhizobacteria and root system functioning. *Front. Plant Sci*. 4(356) (2013).

50. Granada, C., Passaglia, L., de Souza, E., & Sperotto, R. Is Phosphate Solubilization the Forgotten Child of Plant Growth-Promoting Rhizobacteria? *Front. Microbiol.* 9, 2054 (2018).

51. Compant, S., Duffy, B., Nowak, J., Clément, C., & Barka, E. Use of Plant Growth-Promoting Bacteria for Biocontrol of Plant Diseases: Principles, Mechanisms of Action, and Future Prospects. *Appl. Environ. Microbiol.* 71(9): 4951 (2005).

52. Berges, J.A., & Mulholland, M.R. Enzymes and nitrogen cycling. In *Nitrogen in the marine environment* (pp. 1385-1444). Amsterdam, The Netherlands: Elsevier (2008).

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

54. Evans, S., Martiny, J. & Allison, S. Effects of dispersal and selection on stochastic assembly in microbial communities. *ISME J.* 11, 176–185 (2017).

55. Ron, R., Fragman-Sapir, O., & Kadmon, R. (2018). Dispersal increases ecological selection by increasing effective community size. *Proc. Natl. Acad. Sci. U.S.A.* 115(44): 11280-11285.

56. Busby, P. et al. Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol*. 15(3) (2017).

57. Sergaki, C., Lagunas, B., Lidbury, I., Gifford, M., & Schäfer, P. Challenges and Approaches in Microbiome Research: From Fundamental to Applied. *Front. Plant Sci.* 9(1205) (2018).

58. Herlemann et al.. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* Oct;5(10):1571-9 (2011).

59. Ruan, Q. et al. Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* 22(20): 2532-2538 (2006).

60. Pollet, T. et al. Prokaryotic community successions and interactions in marine biofilms: the key role of Flavobacteria. *FEMS Microbiol. Ecol.* 94(6) (2018).

61. Durno, W.E. et al. Expanding the boundaries of local similarity analysis. *BMC Genomics* 14, S3 (2013).

Chapter 3: Transcriptome of Tomato (*Solanum lycopersicum*) Reveals Impacts of Root Microbiomes on Host Interactions

Abstract

We are interested in developing a better understanding of interactions between a host plant and its associated rhizosphere microbiome. The effects of the rhizosphere microbiome on the expression of specific genes and regulatory pathways of a host remain largely unknown. We inoculated tomato (*Solanum lycopersicum*) with microbiomes from vermicompost to study plant transcriptome responses. We hypothesized that tomatoes grown with living vermicompost microbiomes will display transcriptome changes in key nutrient pathways compared to heat-treated vermicompost and control (i.e. no compost) treatments. Illumina DNA sequencing indicated the rhizosphere bacterial community profile was distinct between the three treatments at harvest. Plant phenotypes remained largely unchanged between treatments, except for increases in aboveground biomass production and foliar $\delta^{15}N$ values in the vermicompost and heat-treated vermicompost treatments at harvest and decreases in red ripe fruit phytoene content in the vermicompost and heat-treated vermicompost treatments. Our analysis of foliar and fruit tissue RNA sequences revealed distinct changes in gene expression in the vermicompost treatment, including upregulation of nutrient transporter genes compared to both the heat-treated vermicompost and control treatments. Overall, plant transcriptome analyses indicate bacterial communities in vermicompost influenced the expression of important tomato genes and regulatory pathways while phenotypic outcomes were harder to detect,

likely due to high levels of soil nutrients in this experiment. Nevertheless, these results highlight discrete molecular effects that rhizosphere microbiomes can have on tomato gene expression and regulatory pathways and the usefulness of molecular tools such as RNA sequencing in identifying molecular responses in phytobiome studies.

3.1 Introduction

Microbiomes are defined as the collection of microbial cells in a given environment[1]. They can be found in a variety of settings, including in and around multicellular eukaryotic organisms such as plants. Each compartment of a plant from the roots to the leaves harbors taxonomically and functionally distinct microbiomes[2,3]. Similar to animals, plant microbiomes determine a variety of host traits and can either be detrimental or beneficial to the host[4,5,6]. As a result, there is growing appreciation for the need to develop a molecular understanding of the responses and effects rhizosphere microbiomes have on plant hosts and associated traits. Several trials involving the experimental manipulation of rhizosphere microbial communities have demonstrated how microbial consortia can alter plant traits such as aboveground biomass production and flowering time[7,8], which may be a result of microbial nutrient cycling, phytohormone production, and other processes that influence plant productivity[9]. While much progress has been made in understanding the effect of the microbiome on host phenotypic traits, the effects microbial consortia in the rhizosphere can have on the expression of genes and regulatory pathways of a plant host remain largely unknown and likely represent some of the initial responses of plant tissues to microbiomes and changes in microbiome composition.

Prior studies have highlighted the influences of environmental factors on plant growth and development, including gene expression and regulatory pathways. For example, Neugart et al.[10] showed the expression of genes associated with the production of secondary metabolites in kale plants is dependent on temperature and light intensity, while transcriptomic sequencing of tobacco plants by Yan et al.[11] showed biochar application after topping resulted in increased amino acid and lipid synthesis. In addition to such abiotic factors, biotic components of the environment such as pathogens have also been shown to alter plant transcriptomes. For example, research by Meng et al.[12] demonstrated tobacco infection with *Phytophthora nicotianae* elicits the expression of genes associated with disease-resistance proteins along with several other proteins. Other studies have also highlighted similar effects of plant pathogens on host transcriptomes[13,14], yet the effect of whole plant-associated microbial consortia has received less attention. This is likely due at least in part to the complexity of microbiome communities[15]. Nevertheless, given the effect of the environment on gene expression and rhizosphere microbiome research showing effects of experimentally manipulated communities on host plant phenotypes, it is reasonable to hypothesize that plant transcriptomes will shift in response to changes in the rhizosphere microbiome. As they are likely to reflect early steps in the chain of events connecting microbiome-plant interactions to microbiome-influenced plant traits, understanding how the plant transcriptome changes in response to whole microbial consortia is necessary in developing a better understanding of plant-microbiome interactions.

Here we analyze the effects of rhizosphere microbiomes on tomato (*Solanum lycopersicum* L.) foliar and fruit responses with an emphasis on gene expression and regulatory pathway changes. We performed a greenhouse trial in which tomato plants were grown to maturity in sterilized potting soil treated with one of three microbial inoculants: a microbial inoculant created from vermicompost (V), a microbial inoculant created from heat-treated vermicompost (HT), which served as a control for increased nutrient levels in the vermicompost inoculant, or a no-compost control inoculant made only from sterilized microbial growth broth (C). Each treatment had five replicate tomato plants. Given the beneficial effects vermicompost has been shown to have on plant performance[16], we hypothesized that tomatoes grown with living vermicompost microbiomes will display transcriptome responses, especially for key nutrient-associated pathways and possibly with measurable phenotypic outcomes as well. Our objectives for this study were thus to: (i) assess the influence of rhizosphere microbiomes derived from vermicompost on the host transcriptome and (ii) record any effects of the rhizosphere microbiome derived from vermicompost on tomato foliar and fruit phenotypes.

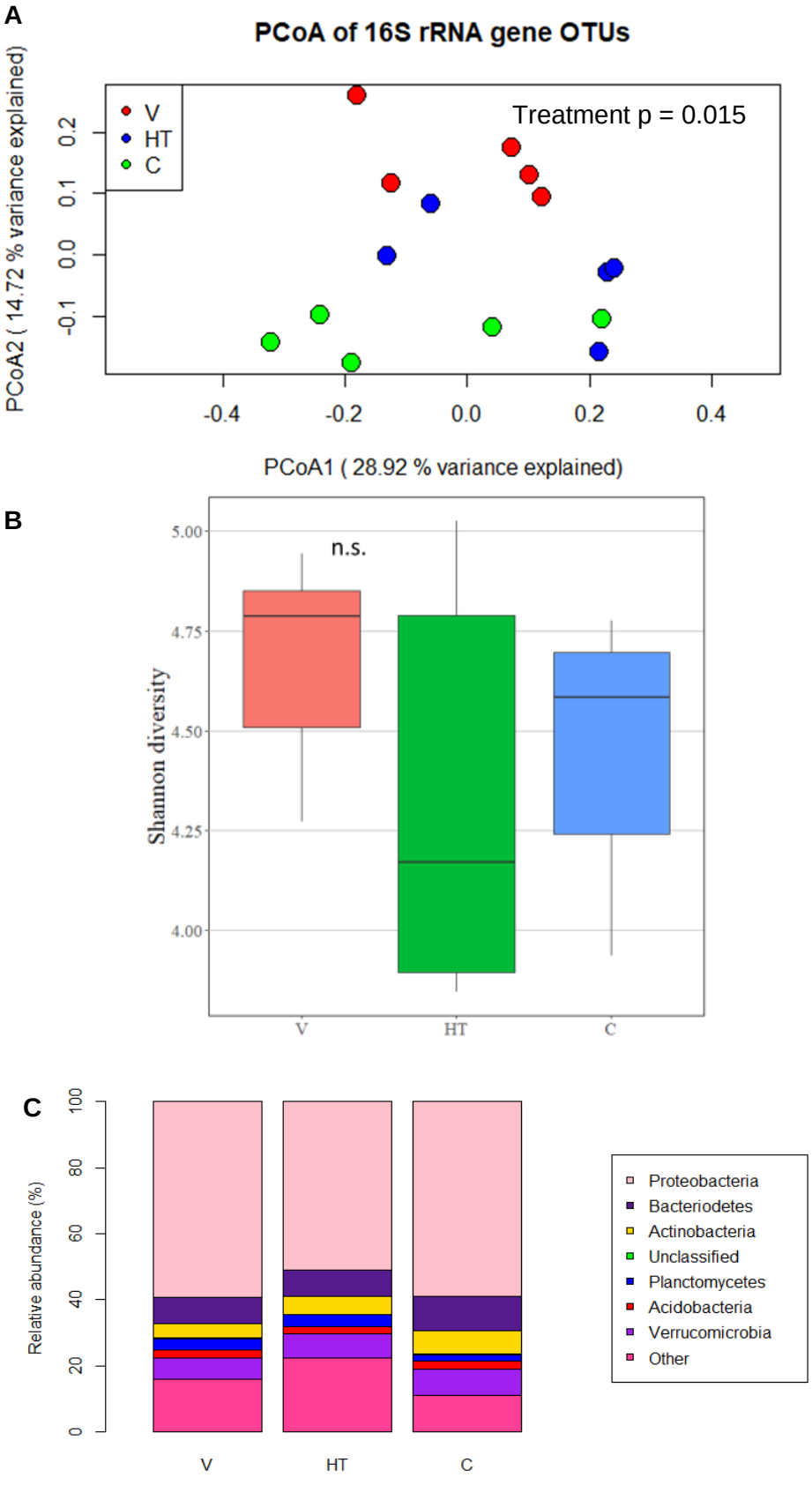
3.2 Results

3.2.1 Soil microbiome composition

In order to assess the effect of treatment on rhizosphere microbial community composition, we sequenced bacterial and fungal DNA from rhizosphere soils collected at harvest. Principle coordinate analysis (PCoA) of bacterial sequence data showed samples clustered based on treatment (Figure 3.1A). This suggests a significant effect

of treatment on bacterial community composition, which was verified using PERMANOVA ($p=0.015$). Treatment did not have a significant effect on bacterial Shannon diversity indexes ($p=0.44$) or relative abundances of the most abundant bacterial phyla (Figure 3.1BC). Analysis of rhizosphere fungal communities revealed no significant effect of treatment on fungal community composition ($p=0.39$), fungal Shannon diversity ($p=0.88$), or relative abundances of the most abundant fungal phyla (Supplementary Figure 3.1ABC).

Figure 3.1 Comparisons of bacterial community sequencing data between treatments A.) Principle coordinate analysis (PCoA) comparing bacterial community composition between the three treatments. PERMANOVA shows significant effect of treatment on community composition ($p=0.015$). B.) Boxplot comparing Shannon diversity indexes between treatments. C.) Stack barplot displaying relative abundances of the most abundant bacterial phyla across the treatments.


A
PCoA of 16S rRNA gene OTUs
PCoA2 (14.72 % variance explained)
V
HT
C
Treatment p = 0.015
-0.1
0.0
0.1
0.2
-0.4
-0.2
0.0
0.2
0.4
PCoA1 (28.92 % variance explained)
B
n.s.
Shannon diversity
5.00
4.75
4.50
4.25
4.00
V
HT
C
C
Relative abundance (%)
0
20
40
60
80
100
V
HT
C
Proteobacteria
Bacteriodetes
Actinobacteria
Unclassified
Planctomycetes
Acidobacteria
Verrucomicrobia
Other

3.2.2 Tomato foliar and fruit phenotypes

We compared a suite of foliar and fruit phenotypes between treatments in order to link rhizosphere microbial community composition to host plant traits. Analysis of foliar phenotypes revealed plants in the vermicompost and heat-treated vermicompost treatment had significantly greater aboveground dry biomass production compared to plants in the control treatment at the $p=0.0099$ and $p=0.0088$ significance levels, respectively (Table 3.1). Aboveground dry biomass production did not differ between plants treated in the vermicompost and heat-treated vermicompost treatments ($p=0.99$). There were no significant differences in percent foliar nitrogen or total foliar nitrogen between the three treatments (Table 3.1). Foliar $\delta^{15}N$ vs. atm. air values, which is a proxy for soil nitrogen cycling activity, was significantly higher in the vermicompost ($p=0.0003$) and heat-treated vermicompost ($p<0.0001$) treatments compared to the control treatment (Table 3.1). Foliar $\delta^{15}N$ differed between the vermicompost and heat-treated vermicompost treatments at the $p=0.0785$ significance level, with plants in the heat-treated vermicompost treatment showing an increase in foliar $\delta^{15}N$ values.

Table 3.1 Group means and ANOVAs for total aboveground dry biomass (g), % foliar N, total foliar N (g), foliar $\delta^{15}N$ vs. at. air values, and average fruit size (g) measured at harvest.

	Total aboveground dry biomass (g)	% foliar N	Total foliar N (g)	Foliar $\delta^{15}N$ vs. At. air	Average fruit size (g)
V	$59.4^{+/-}2.3^{a}$	$1.92^{+/-}0.14^{a}$	$1.15^{+/-}0.09^{a}$	$5.39^{+/-}0.19^{b}$	$19.0^{+/-}0.74^{a}$
HT	$59.7^{+/-}4.1^{a}$	$1.88^{+/-}0.11^{a}$	$1.12^{+/-}0.11^{a}$	$5.96^{+/-}0.14^{a}$	$19.6^{+/-}0.87^{a}$
C	$43.9^{+/-}2.3^{b}$	$1.84^{+/-}0.17^{a}$	$0.82^{+/-}0.11^{a}$	$4.10^{+/-}0.17^{c}$	$21.2^{+/-}0.81^{a}$
ANOVA					
Treatment	0.004**	0.89	0.08	0.001***	0.33

*Significance at the p = 0.05 level.
**Significance at the p = 0.01 level.
***Significance at the p < 0.001 level.

Average tomato fruit size did not differ between treatments (Table 3.1). Analysis of fruit ethylene production revealed no significant effects of treatment (Supplementary Figure 3.2). Analysis of fruit carotenoid production revealed significant effects of treatment on the production of phytoene at the red ripe developmental stage (p=0.028), with plants in the control treatment having significantly greater phytoene compared to the vermicompost and heat-treated vermicompost treatments at the p=0.054 and p=0.024 significance levels, respectively (Figure 3.2). The production of phytofluene, lycopene, beta-carotene, and lutein did not differ between treatments (Supplementary Figure 3.3ABCD). Additionally, fruit chlorophyll a and b content did not differ between treatments at any of the developmental stages (Supplementary Figure 3.4AB).

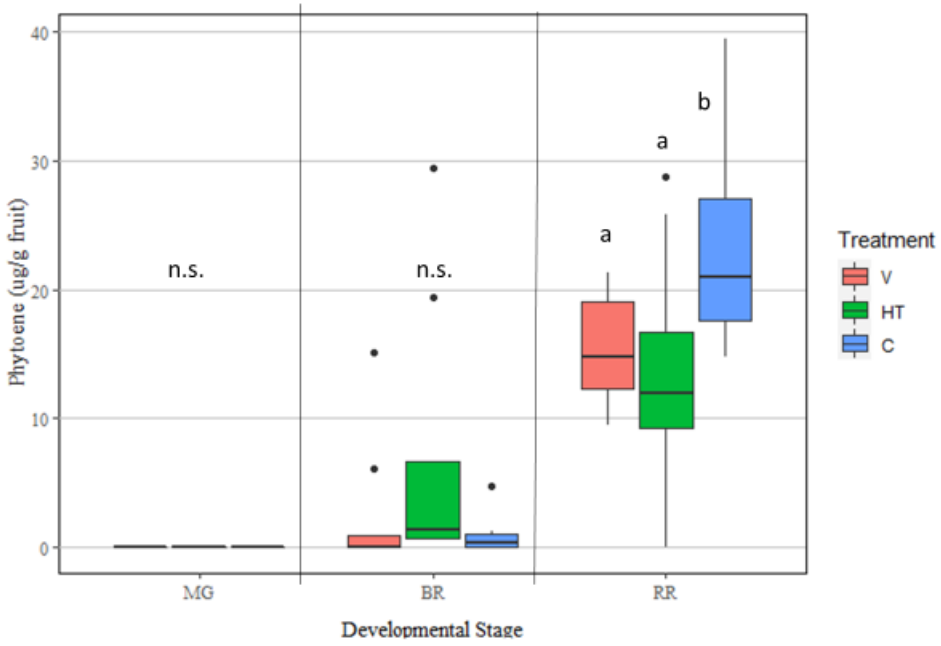


Figure 3.2 Boxplot comparing tomato fruit phytoene content (µg/g fruit) for each treatment at the mature green, breaker, and red ripe stages.

3.2.3 Analysis of Differentially Expressed Genes (DEGs)

To gain insight into microbial influences on host plant physiological processes, we performed RNA sequencing on tomato foliar and fruit tissue. Foliar tissue was collected from each plant at four weeks old and at maturity, while mature green, breaker, and red ripe fruits were harvested from each plant to analyze effects of treatment on fruit ripening processes. Analysis of RNA sequences revealed numerous differentially expressed genes (DEGs) between treatment groups (Supplementary Table 3.1). Observation of the 1000 genes showing the greatest change in expression at the p=0.05 significance level showed that in mature foliar tissue at harvest, copper, zinc, vacuolar iron ion, and phosphate transporter genes were upregulated in the vermicompost treatment compared to both the heat-treated vermicompost and control treatment. Several WRKY transcription factor genes were also upregulated in the

vermicompost treatment compared to the heat-treated vermicompost and control treatments. In foliar tissue harvested at four weeks, genes associated with phosphate transporters, chitinases, and various kinases were upregulated in the vermicompost and heat-treated vermicompost treatments compared to the control treatment. In red ripe fruits, genes associated with chlorophyll binding proteins, rubisco related genes, and cell wall protein X77373 were upregulated in the vermicompost treatment compared to the heat-treated vermicompost and control treatments.

Repeating the analysis to examine the 1000 genes showing the greatest change in expression with a significance level below p=0.01 revealed similar results (Supplementary Table 3.2). In mature foliar tissue at harvest, genes associated with phosphate, nitrate, copper, zinc, sugar, peptide, heavy metal, and vacuolar iron ion transporters were upregulated in the vermicompost treatment compared to the heat-treated vermicompost and control treatments. Several phosphatase and kinase genes were also upregulated in the vermicompost treatment compared to the other two treatments. Meanwhile, defensin, sesquiterpene synthase, and terpene synthase genes were downregulated in the vermicompost treatment compared to the sterile vermicompost and control treatments. In foliar tissue harvested at four weeks, heavy metal transporter and WRKY transcription factor genes were upregulated in the vermicompost and heat-treated vermicompost treatments, while genes associated with the RADIALIS transcription factor and fatty acid hydrolases were downregulated in the vermicompost treatment. In red ripe fruits, several chlorophyll binding proteins, cell wall protein X77373, and RuBP carboxylase small subunits genes were upregulated in the vermicompost treatment compared to the heat-treated

vermicompost and control treatments.

3.3 Discussion

Here we report the effects of rhizosphere microbiomes from vermicompost on tomato foliar and fruit phenotypes and transcriptomes. Our analysis of soil microbiomes in this experiment showed bacterial community composition was significantly different between our three microbial treatments, which was verified using PERMANOVA. Given results from past experiments with vermicompost[16], we hypothesized microbial communities derived from vermicompost would significantly alter tomato gene expression especially as related to nutrient metabolism, transport and associated regulatory pathways in addition to possibly altering plant phenotypes as compared to control treatments. Although microbiomes from vermicompost did not markedly alter plant phenotypes compared to the heat-treated vermicompost and control treatments, analysis of foliar and fruit transcriptomes revealed transcriptome alterations among the vermicompost inoculant and controls. These changes included the upregulation of key nutrient transporters in foliar tissue and the upregulation of chlorophyll binding protein genes in red ripe fruits. Together, our results suggest vermicompost microbiomes may alter different aspects of tomato growth and development and demonstrate the usefulness of whole genome transcriptome profiling in studying plant-microbe interactions.

Our sequence data of rhizosphere bacterial communities showed community composition was significantly different between the three treatments, indicating robustness of our inoculation method. Rhizosphere bacterial communities in this

experiment showed overall shifts in community profile, but not differences in diversity. Previous rhizosphere microbiome research has shown ecological filtering (i.e. decreases in diversity) in the microbiome may lead to changes in microbiome function[17], which could influence plant phenotypes. Such a pattern could be indicative of host plants selecting for microbial consortia that confer a fitness advantage[15]. Given the fitness benefits microorganisms in vermicompost likely have on host plant traits, we hypothesized a similar pattern may emerge in this experiment in the vermicompost treatment, but did not observe shifts in bacterial or fungal diversity in any of the treatments. This may be due in large part to our experimental set up. Tomato plants in this experiment were fertilized three times a week with 150 ppm N Peter's excel fertilizer to assure plant survival. Under more nutrient-limited conditions, plant selection for microbial consortia conferring a fitness advantage is expected to become stronger[18]. However, with our higher N fertilizer regime and with additional nutrients from the microbial inoculants, ecological filtering in the rhizosphere by tomato plants in this experiment may not have been as strong, resulting in shifts in community composition but weaker ecological filtering in the rhizosphere.

Fungi play a central role in a variety of rhizosphere processes, including nutrient solubilization, plant nutrient capture, and disease suppression[19,20,21]. Previous research in tomato has highlighted the beneficial effects of genera such as arbuscular mycorrhizal fungi (AMF) on plant growth[22]. Given the growth promoting effects vermicompost has been shown to have, we expected the abundance of growth-promoting fungi such as AMF to increase in the vermicompost treatment. However, no observed aspects of fungal community composition differed between any

treatments in this experiment. Previous experiments have shown bacterial and fungal communities in the rhizosphere can adapt and shift composition independent of one another[23]. Further, these types of community shifts could still influence plant performance and soil processes. Our data suggest this type of pattern may have occurred in this experiment, with bacterial communities in vermicompost being more significant in determining host plant traits such as gene expression.

Plant phenotypes demonstrating statistical change in this experiment were aboveground biomass production and foliar $\delta^{15}N$, a proxy for soil nitrogen cycling with increased values generally representing enhanced nitrogen cycling activity[24], in the vermicompost and heat-treated vermicompost treatments as well as a decrease in fruit phytoene at the red ripe stage in the vermicompost and heat-treated vermicompost treatments. We hypothesized microorganisms within vermicompost promote soil nutrient cycling and plant nutrient uptake would indeed enhance aboveground biomass production and nitrogen capture. As fruit set and yield can be inversely correlated with vegetative growth in high nutrient environments[25,26], this may be the reason fruit yield was not increased. The increase in aboveground biomass production and in foliar $\delta^{15}N$ in the vermicompost and heat-treated vermicompost treatment may have been due to growth promotion via two different mechanisms. In the vermicompost treatment, plants could have captured more nutrients via microbially-mediated upregulation of nutrient transporter genes, as shown in our transcriptome data with the upregulation of genes such as nitrate transporter 1:2 (Solyc05g006990) and high-affinity nitrate transporter-like protein (Solyc03g112100). On the other hand, plants in the heat-treated vermicompost treatment could have more

readily accessed nutrients released during microbial cell lysis as a result of the heating process, which is supported by the observed increase in foliar $\delta^{15}N$.

Fruit maturation and quality traits such as ethylene production and carotenoid content remained largely unchanged between the treatments. However, we did observe a statistically significant change in phytoene content at the red ripe developmental stage, with fruits from the vermicompost and heat treated vermicompost treatment showing a decrease in phytoene compared to the control treatment. The carotenoid pathway is a complex multistep process that is regulated by numerous genetic, hormonal, and external factors[27]. The accumulation of the predominant downstream carotenoids that accumulate in tomato fruit (lycopene and beta-carotene) remained largely unchanged. The observed difference in phytoene accumulation, the first committed product in carotonogenesis, which in many plant tissues including ripening tomato fruit accumulates only transiently, suggests some effect of treatment on carotenoid pathway flux. While phytoene is the first product in the carotenoid pathway[27], effects of vermicompost microbiomes on carotenoid production could be through reduction of early pathway activities and/or acceleration of later activities prior to lycopene and beta-carotene synthesis or in downstream carotenoid catabolism.

Analysis of foliar and fruit transcriptomes revealed several effects of the vermicompost microbiome that were not directly captured through plant trait analyses but possibly related to increased vegetative growth. Indeed, at both stringency levels, we observed the upregulation of several important nutrient transporter genes in the foliar tissue of plants in the vermicompost treatment compared to the heat-treated vermicompost and control treatments. This suggest microbiomes from vermicompost

altered plant nutrient uptake processes, which could have been achieved through microbial cycling of soil nutrients or changes in plant biology via phytohormone production, which is supported through our transcriptome data showing the upregulation of growth hormone genes such as gibberellin 20-oxidase-3 (Solyc11g072310). It is again worth considering the growth conditions in this experiment. As previously discussed, plants were fertilized three times a week with 150 ppm N Peter's excel fertilizer, a more liberal fertilizer regime, in order to ensure plants grew to maturity, produced fruit, and survived until harvest. Additionally, no other stressors were applied during the course of the experiment. Considering this, it is conceivable that the upregulation of nutrient transporter genes by the microbiomes derived from vermicompost may confer a fitness advantage if stress is imposed on the plants. Previous research has shown these type of interactions occurring in microbiome studies[28,29,30], with growth promotion activities by microbiomes becoming more significant under stressed conditions.

In fruit tissue, chlorophyll binding proteins and cell wall protein X77373 genes were upregulated in the red ripe fruits of vermicompost at both the p=0.05 and p=0.01 stringency levels, suggesting effects of vermicompost microbiomes later in growth. Previous research has found cell wall protein X77373 may be involved with disease resistance in fruit[31], suggesting vermicompost microbiomes may activate disease resistance genes along with plant nutrient transporter genes. The upregulation of chlorophyll binding protein genes in the vermicompost treatment suggest an effect on fruit developmental processes and specifically a delay in the chloroplast to chromoplast transition. This result along with our analysis of fruit phytoene content at

the red ripe stage further suggests vermicompost microbiomes alter earlier steps of carotenoid production in fruits.

3.4 Conclusion

Here, we report the effects of vermicompost rhizosphere microbiomes on tomato foliar and fruit traits compared to a heat-treated and sterile control. Rhizosphere microbiome sequencing showed differences between all three microbial treatments, indicating robustness of our inoculation method and potential shifts in not only microbiome composition but also corresponding influence on the plant host. We hypothesized plants grown with vermicompost microbiomes would show increased productivity compared to plants grown with the two control inoculants, and indeed observed changes in N uptake and vegetative growth. Analysis of foliar and fruit transcriptome data further revealed vermicompost microbiomes upregulated genes associated with nutrient transport and disease resistance, in addition to genes associated with altering the transition in fruit maturation to ripening, together highlighting the growth promoting capabilities of vermicompost microbiomes. We speculate under more stressed environmental conditions the growth promoting effects of vermicompost microbiomes may be more significant. Our results highlight the effects of rhizosphere microbiomes on host transcriptomes and regulatory pathways and demonstrate the usefulness of molecular techniques such as plant transcriptomics in capturing these effects at a finer level to both explain observed phenotypic differences and informing development of hypotheses to be tested in future studies.

3.5 Materials & methods

3.5.1 Creation of microbial inoculants

The rhizosphere microbiomes in this experiment were derived from microbial inoculants created in lab. For the vermicompost inoculant, 150 g of Worm Power (Avon, NY, USA) vermicompost was dissolved in 300 ml of sterile 10% Luria Broth, which was used in place of sterile water to help stimulate microbial growth. Sterilized glass beads were then added and the mixture was homogenized using a shaker set at 180 opm for 1 hour. After, the mix was then filtered through 4 layers of sterilized cheese cloth and the resulting slurry was used as our vermicompost inoculant. For the sterile vermicompost inoculant, we followed the same procedure using Worm Power vermicompost that was autoclaved twice with a 48-hour resting period between cycles, which significantly reduces the microbial community. This treatment was included to control for the effects of the plant nutrients in Worm Power likely remaining in the vermicompost inoculant. For the control inoculant, sterile 10% Luria Broth was used with nothing added.

3.5.2 Greenhouse trial

Tomato plants (Ailsa Craig) were grown in a plant growth facility at Boyce Thompson Institute, Ithaca, NY, USA. To start, we filled 15 magenta boxes with sterilized potting soil (50% metromix 360, 50% cornell mix + osmocote) to grow our seedlings. Each magenta box was randomly assigned to one of the 3 bacterial treatments, giving us 5 plants per treatment. Before planting, each box received 20 ml of a solution made of 16 ml of one of the microbial inoculants and 4 ml of 500 ppm

Peter's excel fertilizer (15-5-15). The solution was then mixed into the soil to keep the medium moist and evenly distribute the fertilizer and inoculant solution. We then planted one pre-germinated Ailsa Craig seed into each magenta box and watered each seedling with 10 ml of sterile water.

After planting, the magenta boxes were placed in a growth chamber set at 250 µmol of light, 10% humidity, alternating between 28ºC for 16 hours and 23ºC for 8 hours for three weeks. During this phase, plants were hand watered with sterile water as needed. Seedlings were then hardened off in the growth chamber, transplanted into 6-inch pots filled with sterilized potting soil, then placed in a seedling greenhouse to be hardened off before the rest of the experiment. The seedling greenhouse had day temperatures between 25ºC and 27ºC, night temperatures between 18ºC and 20ºC, and natural light and humidity. The seedlings were hand watered every day with tap water during this phase. After 1 week, tomatoes were transplanted into 1-gallon pots filled with sterilized potting soil and placed in a new greenhouse with day temperatures between 24ºC and 26ºC, night temperatures between 21ºC and 23ºC, and natural light and humidity. Plants were watered daily with tap water and fertilized with 150 ppm N Peter's excel fertilizer three times a week. The plants were kept in this greenhouse for the remainder of the experiment. Aboveground plant matter and rhizosphere soil samples were harvested three months after transplanting.

3.5.3 Analysis of tomato fruit phenotypes

To analyze the effect of microbiome treatment on fruit ethylene production, we harvested 1-2 fruits at the breaker stage from each plant during the trial and analyzed

ethylene production based on modified procedures outlined in Nguyen et al.[32]. To allow wound ethylene to subside, fruits were left on the lab bench overnight before analysis. The following day, fruits were placed into sealed 250 ml air-tight mason jars and incubated for 3 hours. We then drew 1 ml air samples from each mason jar and analyzed the sample using an Agilent 6850 GC System equipped with a flame ionization detector. Ethylene concentrations were then calculated by comparison with a standard of known concentration and normalizing of fruit mass. We recorded one measurement per fruit everyday for five days.

To analyze how fruit carotenoid production may change in response to the rhizosphere microbiome, we harvested 2 mature green, 2 breaker, and 2 red ripe fruits from each plant and performed carotenoid extractions. After being harvested and remaining on the lab bench overnight, fruits were cut, frozen in liquid nitrogen, and ground to a fine powder. We then aliquoted ~100 mg of frozen ground tissue into 2 ml tubes with glass beads and lyophilized the samples. Each tube then received 50 µl of magnesium carbonate ($MgCO_3$) at a 0.003M concentration followed by 500 µl of 100% ethyl acetate. The tubes were then shaken for 30 seconds, incubated at 4ºC for 15 minutes, then shaken again. Samples were then centrifuged at max speed for 5 minutes, then the supernatant was transferred to a clean 2 ml tube. The samples were then dried down completely in a rotovap and then analyzed as reported previously[33.] Fruit size was recorded for all harvested fruits and used to compare average fruit size between treatments across the fruit developmental stages.

3.5.4 Foliar and fruit tissue RNA extractions

In order to assess the effect of the rhizosphere microbiome on growth and developmental processes, we performed RNA extractions on tomato foliar and fruit tissue. We gathered foliar tissue from plants at four weeks old and at harvest. Foliar tissue was sampled from random leaves on each plant and the composite sample was then immediately frozen, ground in liquid nitrogen, and stored at -80ºC until extraction. Fruit tissue was gathered from 2 mature green, 2 breaker, and 2 red ripe fruits from each plant. Upon harvesting, fruits were left on the lab bench overnight to allow wound ethylene to subside. Fruits were then cut, frozen, and ground in liquid nitrogen. Powdered fruit samples were then stored at -80ºC until extraction.

RNA was extracted from foliar and fruit tissue by adding 800 µl of a modified RLT buffer and 800 µl of chloroform to our samples. The samples were then vortexed, centrifuged at 12000 rpm for 10 minutes, then the supernatant was transferred to a new tube. The supernatant was then mixed with half the volume of 100% ethanol, then transferred to spin columns. The samples were then centrifuged for 1 minute at 8000 rpm, then washed twice with 80% ethanol. Spin columns were then spun at 12000rpm for one minute, then transferred to new 1.5 ml tubes. We then added 55 µl of DEPC water to each column, incubated the columns for 2 minutes, then spun them down for 1 minute at 12000rpm. Samples were then quantified and ran on a 1% agarose gel to check the RNA quantity and quality of each sample. Total RNA from foliar and fruit tissue was then used to construct strand-specific RNA libraries as described in Zhong et al.[34].

3.5.5 Analysis of tomato foliar phenotypes

After harvest, all plants were placed in a drying oven set at 60ºC for one week and total aboveground dry biomass was then recorded. Dried foliar tissue from each plant was used for analysis of foliar tissue nitrogen content at the Cornell Stable Isotope Laboratory (Cornell University, Ithaca, NY, USA) using a Thermo Delta V isotope ratio mass spectrometer (IRMS) interfaced to an NC2500 elemental analyzer. The analysis provided information on $\delta^{15}N$ as well as %N, which was used along with total aboveground dry biomass to calculate total foliar N.

3.5.6 Soil microbiome DNA extractions

Bacterial and fungal DNA was extracted from rhizosphere soil samples collected at harvest using a PowerSoil DNA Isolation Kit (Qiagen, Hilden, Germany) following the manufacturers protocol. Approximately 200-250 mg of frozen rhizosphere soil from each pot was used for each extraction. We then amplified the 16S rRNA and ITS gene sequences to capture bacterial and fungal taxonomic data, respectively[35]. Following amplification, we performed an initial clean-up of the samples using a HighPrep PCR Clean-up System (MAGBIO Genomics, Gaithersburg, MD, USA) and then attached unique index primers to our amplicons with a second PCR cycle. Indexed samples were then cleaned and normalized with a SequalPrep Normalization Plant Kit (Thermo Fisher Scientific, Waltham, MA, USA) then pooled into one composite sample. The pooled sample was then ran on a 1.2% agarose gel with SyberSafe added and then the target band was excised. The pooled DNA was then extracted from the excised gel with a Wizard SV Gel and PCR Clean-Up System

(Promega, Madison, WI, USA) and then submitted to the Cornell Genomic Facility (Cornell University, Ithaca, NY, USA) for Illumina (Illumina, Inc., San Diego, CA, USA) barcoded sequencing.

3.5.7 Statistics

Plant phenotype data were analyzed using R statistical software (Rproject.org). To determine the effect of our microbiome treatment on total aboveground dry biomass, foliar nitrogen content, and fruit size, we utilized linear models with one-way ANOVAs. Normality of residuals and homoscedasticity were verified for each model using plotting methods. Data transformations were performed as necessary in order to meet ANOVA assumptions. Following significant ANOVA outcomes, we used least squared means to perform pairwise comparisons of group means between microbiome treatments.

Mixed effect modeling with the lmer function in R was used to determine the effect of treatment and time on fruit ethylene production. Treatment and day were treated as mixed effects, while plant and fruit replicates were treated as random effects. ANOVA assumptions were verified using plotting methods. The effect of treatment and fruit ripening stage on carotenoid and chlorophyll content was determined using linear mixed effect modeling with the lme function. Lme allowed us to separate variance components for each developmental stage, which we found produced the lowest Akaike Information Criterion (AIC) value for our model. Following significant outcomes, post-hoc analyses using the emmeans function were performed to do pairwise comparisons of group means.

3.5.8 Foliar and fruit tissue RNA sequence processing and analysis

Raw RNA-Seq reads were processed to remove adaptors and low-quality sequences using Trimmomatic (version 0.36) with default parameters[36]. Cleaned reads shorter than 40 bp were discarded. The remaining cleaned reads were aligned to the ribosomal RNA database[37] using bowtie (version 1.1.2)[38] allowing up to three mismatches, and those aligned were discarded. The remaining high-quality cleaned reads were aligned to the tomato reference genome (SL3.0 and ITAG3.2) using HISAT2 (version 2.1.0)[39]. Based on the alignments, raw read counts for each gene were calculated and then normalized to reads per kilobase of exon model per million mapped reads (RPKM). Raw read counts were then fed to DESeq2[40] to identify differentially expressed genes (DEGs), with a cutoff of adjusted P value < 0.05 and fold change > 2.

Next, we performed gene ontology (GO) enrichment analyses to identify sets of genes that were over and under expressed when comparing tissue from different treatments. For each pairwise GO enrichment analysis, we first filtered out genes with expression p-values above the 0.05 and 0.01 threshold levels, respectively. We then compiled a list of the 1000 genes that showed the greatest change in gene expression. We then used agriGO v2[41] to identify enriched GO terms. Based on the enriched GO terms, we then identified specific genes that were over and under expressed between treatments.

3.5.9 Soil microbiome DNA sequence processing and analysis

In order to process our soil microbiome sequences, we used a modified version of the pipeline from the Brazilian Microbiome Project (http://www.brmicrobiome.org/). Mothur v. 1.3613 was used to merge paired-end sequence (make.contigs), trim off primers (trim.seqs, pdiffs = 2, maxambig = 0), remove singletons (unique.seqs → split.abund, cutoff = 1) and classify sequences (97% similarity). OTUs that were suspected to not be of fungal or bacterial origin were removed (remove.lineage). Qiime v. 1.9.114 was then used to cluster OTUs and create our bacterial and fungal OTU tables for analysis.

After obtaining bacterial and fungal OTU tables, the bacterial dataset was rarefied to 10705 sequences per sample and the fungal dataset was rarefied to 37 sequences per sample. We then calculated differences in community composition between samples using Bray-Curtis distances. These values were then utilized to create an ordination to visualize differences in microbial community fingerprints between microbiome treatments. We then utilized the adonis package in R to perform a PERMANOVA, which allowed us to assess the effect of treatment on bacterial and fungal community composition. Shannon diversity indexes were calculated for each treatment and compared using ANOVA and stacked relative abundance plots were created to further visualize shifts in microbial community composition. ANOVAs were then utilized to determine the effect of treatment on relative abundances of the most abundant bacterial and fungal phyla.

3.6 References

1. Berg, G. et al. Microbiome definition re-visited: old concepts and new challenges. *Microbiome* 8(1), 103 (2020).

2. Brown, S., Grillo, M., Podowski, J., & Heath, K. Soil origin and plant genotype structure distinct microbiome compartments in the model legume Medicago truncatula. *Microbiome* 8(1), 139 (2020).

3. Gopal, M., & Gupta, A. Microbiome Selection Could Spur Next-Generation Plant Breeding Strategies. *Front. Microbiol.* 7(1971) (2016).

4. Babalola, O. et al.The Nexus Between Plant and Plant Microbiome: Revelation of the Networking Strategies. *Front. Microbiol.* 11(2128) (2020).

5. Compant, S., Samad, A., Faist, H., & Sessitsch, A. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *J. Adv. Res.* 19, 29-37 (2019).

6. Ulrich, D. et al. Plant-microbe interactions before drought influence plant physiological responses to subsequent severe drought. *Sci. Rep.* 9, 249 (2019).

7. Swenson, W., Wilson, D., & Elias, R. Artificial ecosystem selection. *Proc. Natl. Acad. Sci. U.S.A.* 97(16): 9110-9114 (2000).

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

9. Hakim, S., et al. Rhizosphere Engineering With Plant Growth-Promoting Microorganisms for Agriculture and Ecological Sustainability. *Front. Sustain. Food Syst.* 5(16) (2021).

10. Neugart, S., Krumbein, A., & Zrenner, R. Influence of Light and Temperature on Gene Expression Leading to Accumulation of Specific Flavonol Glycosides and Hydroxycinnamic Acid Derivatives in Kale (Brassica oleracea var. sabellica). *Front. Plant Sci.* 7(326) (2016).

11. Yan, S., Niu, Z., Yan, H., Zhang, A., & Liu, G. Transcriptome sequencing reveals the effect of biochar improvement on the development of tobacco plants before and after topping. *PLoS ONE* 14(10): e0224556 (2019).

12. Meng, H. et al. Comparative transcriptome analysis reveals resistant and susceptible genes in tobacco cultivars in response to infection by Phytophthora nicotianae. *Sci. Rep.* 11(1): 809 (2021).

13. Nobori, T. et al. Transcriptome landscape of a bacterial pathogen under plant immunity. *Proc. Natl. Acad. Sci. U.S.A.* 97115(13), E3055–E3064 (2018).

14. Willig, C., Duan, K., & Zhang, Z. Transcriptome Profiling of Plant Genes in Response to Agrobacterium tumefaciens-Mediated Transformation. *Curr Top Microbiol Immunol*. 418:319-348 (2018).

15. Garcia, J., & Kao-Kniffin J. Microbial Group Dynamics in Plant Rhizospheres and Their Implications on Nutrient Cycling. *Front. Microbiol.* Jul 11;9:1516 (2018).

16. Azarmi R, Ziveh PS, Satari MR. Effect of vermicompost on growth, yield and nutrition status of tomato (Lycopersicum esculentum). *Pak. J. Biol. Sci.* Jul 15;11(14):1797-802 (2008).

17. Song, H. et al. Environmental filtering of bacterial functional diversity along an aridity gradient. *Sci. Rep.* 9, 866 (2019).

18. Liu, H.,Brettell, L.E., Qiu, Z., & Singh, B.K. Microbiome-Mediated Stress Resistance in Plants. *Trends Plant Sci.* 25(8): 733-743 (2020).

19. Singh, H., & Reddy, M. Effect of inoculation with phosphate solubilizing fungus on growth and nutrient uptake of wheat and maize plants fertilized with rock phosphate in alkaline soils. *Eur. J. Soil Biol.* 47(1): 30-34 (2011).

20. Moreau, D., Bardgett, R.D., Finlay, R.D., Jones, D.L., & Philippot, L. A plant perspective on nitrogen cycling in the rhizosphere. Funct Ecol. 2019; 33: 540– 552 (2019).

21. Qiao, Q. et al. Characterization and variation of the rhizosphere fungal community structure of cultivated tetraploid cotton. *PLoS ONE.* 14(10):e0207903 (2019).

22. Begum, N. et al. Role of Arbuscular Mycorrhizal Fungi in Plant Growth Regulation: Implications in Abiotic Stress Tolerance. *Front Plant Sci.* 10(1068) (2019).

23. Wang, Z., Zhang, J., Wu, F., Zhou, X. Changes in rhizosphere microbial communities in potted cucumber seedlings treated with syringic acid. *PLoS ONE* 13(6): e0200007 (2018).

24. Xu, Y., He, J., Cheng, W., Xing, X., & Li, L. Natural ^{15}N abundance in soils and plants in relation to N cycling in a rangeland in Inner Mongolia. *J. Plant Ecol.* 3(3): 201-207 (2010).

25. Trinklein, D. Understanding Tomato Fruit Set. In: Integrated Pest Management. University of Missouri. 2013. https://ipm.missouri.edu/MPG/2013/4/Understanding-Tomato-Fruit-Set/

26. Bonomelli, C. et al.Ammonium Excess Leads to Ca Restrictions, Morphological Changes, and Nutritional Imbalances in Tomato Plants, Which Can Be Monitored by the N/Ca Ratio. *Agronomy* 11(7): 1437 (2021).

27. Liu, L., Shao, Z., Zhang, M., & Wang, Q. Regulation of Carotenoid Metabolism in Tomato. *Mol. Plant* 8(1): 28-39 (2015).

28. Backer, R. et al. Plant Growth-Promoting Rhizobacteria: Context, Mechanisms of Action, and Roadmap to Commercialization of Biostimulants for Sustainable Agriculture. *Front Plant Sci.* 9(1473) (2018).

29. Khan, N. et al. Insights into the Interactions among Roots, Rhizosphere, and Rhizobacteria for Improving Plant Growth and Tolerance to Abiotic Stresses: A Review. *Cells* 10(6), 1551 (2021).

30. Lazcano, C. et al. The rhizosphere microbiome plays a role in the resistance to soil-borne pathogens and nutrient uptake of strawberry cultivars under field conditions. *Sci. Rep.* 11, 3188 (2021).

31. Liu, X. et al. Tomato SlPUB24 enhances resistance to Xanthomonas euvesicatoria pv. perforans race T3. *Hortic. Res.* 8, 30 (2021).

32. Nguyen, C. et al. Tomato GOLDEN2-LIKE Transcription Factors Reveal Molecular Gradients That Function during Fruit Development and Ripening. *Plant Cell*. 26(2), 585-601 (2014).

33. Shi, Y., Vrebalov, J., Zheng, H., & Giovannoni J. A tomato LATERAL ORGAN BOUNDARIES transcription factor, SlLOB1, predominantly regulates cell wall and softening components of ripening. *Proc. Natl. Acad. Sci. U.S.A.* 118(33): e2102486118 (2021).

34. Zhong, S. et al. High-throughput illumina strand-specific RNA sequencing library preparation. *Cold Spring Harb. Protoc.* 2011: 940–949 (2011).

35. Caporaso, J. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6: 1621–1624 (2012).

36. Bolger, A.M., Lohse, M., & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30:2114–2120 (2014).

37. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41:D590–D596 (2013).

38. Langmead B. Aligning short sequencing reads with Bowtie. *Curr Protoc. Bioinformatics.* 32:11.17.11–11.17.14 (2010).

39. Kim, D., Paggi, J.M., Park, C., Bennett, C., & Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37: 907–915 (2019).

40. Love, M.I., Huber, W., & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550 (2014).

41. Tian, T. et al. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* (2017).

Chapter 4: Spatial Crop Diversity Alters Soil Properties and Influences Fruit Yield of Tomatoes

Abstract

In recent years, interest has grown in developing a better understanding of how biodiversity in agriculture may promote multiple ecosystem functions and sustainability. Management practices such as intercropping, which involves the cultivation of multiple crop species in a single field at a time, have been shown to have beneficial effects on aboveground agroecosystem functions such as pest regulation. However, few studies have investigated the effects of intercropping on belowground ecosystem functions in agriculture. Here, we performed a field trial in which we sought to assess how intercropping may alter soil properties and influence horticultural crop productivity. The experiment consisted of growing tomato plants in either monocrops or with tobacco and kale in intercropped plots. We found that plots where tomato was intercropped with tobacco had significantly higher soil nitrate compared to monocrop and kale intercrop plots, while plots where tomato was intercropped with kale had higher soil phosphorus compared to tobacco plots and lower soil $\delta^{13}C$ values compared to both tobacco intercrop and monocrop plots. Further, analysis of soil microbial communities showed significant effects of intercropping on community composition, including the abundances of key nutrient cycling microbial taxa such as acidobacteria. Analysis of tomato foliar transcriptomes showed altered transcriptional profiles later in the season, while analysis of fruit production showed tomatoes

intercropped with tobacco had higher average fruit yield compared to tomato monocrops. Ultimately, these results show intercropping can influence ecosystem functions in agriculture and alter several soil properties, but the effects are dependent on the crops that are cultivated.

4.1 Introduction

In conventional agroecosystems, management practices often prioritize achieving high crop yields at the cost of other important ecosystem functions. One example is the usage of synthetic nitrogen fertilizers, which promote crop productivity but can reduce soil biodiversity, long-term soil fertility, and lead to leaching and increased nitrous oxide emissions[1,2,3]. With threats to agricultural productivity such as climate change and growing inaccessibility to agronomic resources such as phosphate[4], interest has grown in developing more sustainable cultivation methods that can promote multiple ecosystem functions while also maintaining or enhancing crop yields.

Increasing biodiversity in cropping systems may be a useful strategy to promote multiple ecosystem functions in agriculture. Diversified cropping systems, those that involve the cultivation of different species/varieties of crops at different spatial/temporal scales[5], draw on the biodiversity-ecosystem functioning relationship in which increased biodiversity is associated with enhanced ecosystem functioning[6]. While most studies of the relationship between biodiversity and enhanced ecosystem functioning have focused on natural ecosystems, there is strong evidence that the relationship exists in agricultural settings as well and could be leveraged in developing

more sustainable management practices. For example, diversified crop rotations (i.e. temporal crop diversity) have been shown to decrease pest and pathogen pressure[7], enhance drought stress tolerance in crop plants[8], and under certain conditions increase crop yield[9].

Intercropping, which involves imposing spatial crop diversity by cultivating two or more commodity crops in a single field at the same time, may be another useful strategy for increasing ecosystem functions in agriculture. In recent years, several studies have shown intercropping can enhance multiple ecosystem functions and influence crop production outcomes. For example, a meta-analysis by Zhang et al.[10] showed intercropping cereal with faba beans reduced plant disease incidence, which was likely driven by greater host functional divergence. In addition, work by Dong et al.[11] showed that intercropping can promote yield in four different intercropping systems, which was likely driven by niche differentiation between crop species. Although intercropping has been shown to promote these aboveground ecosystem functions, the effects on belowground ecosystem functions and how these effects may translate into changes in crop productivity remain to be explored.

Previous research in both managed and unmanaged terrestrial ecosystems suggest spatial plant biodiversity can drive a variety of soil ecosystem functions. For example, work by Lange et al.[12] showed enhanced plant species richness can promote carbon sequestration in grasslands by diversifying carbon species available for soil microbial metabolism. Further, meta-analyses by Shackelford et al.[13] showed that cover cropping can promote several soil ecosystem functions, including soil organic matter formation and weed suppression. Considering findings such as these, imposing

spatial crop diversity via intercropping may alter several soil ecosystem functions and produce measurable differences in soil properties that can influence crop productivity, yet few studies have explored these potential linkages.

In the present study, we sought to investigate how intercropping may promote multiple ecosystem functions by characterizing the effects of intercropping on soil and crop properties. The study consisted of growing tomato plants in replicated microplots subjected to either a monocrop treatment or one of two intercrop treatments. The first intercrop treatment consisted of tomato intercropped with tobacco (a close genetic relative), while the second treatment consisted of tomato intercropped with kale (a more distant genetic relative). We hypothesized intercropping would alter several soil ecosystem functions and produce measurable differences in soil chemical and biological properties, which would improve crop performance. A suite of soil properties, including soil nutrition, stable isotope ratios, and bacterial and fungal community composition, were recorded to assess how intercropping might alter belowground ecosystem functions. To identify potential linkages between belowground changes and horticultural crop production outcomes, we recorded average aboveground biomass production and fruit yield of tomato plants and performed RNA sequencing on tomato foliar tissue to understand how plants respond to diversification at the physiological level.

4.2 Results

4.2.1 Soil nutrition and stable isotope analysis

In order to assess the effect of intercrop treatment on soil nutrition, we buried plant root simulator (PRS) probes manufactured by Western Ag Innovations (Saskatoon, Saskatchewan, CA) in each experimental microplot for four weeks during the middle of our growing season. Analysis of PRS probes after burial showed several effects of intercrop treatment on soil nutrition. We observed a significant effect of treatment on soil nitrate content ($p=0.004$), with tobacco intercrop plots having significantly greater soil nitrate compared to monocrop and kale intercrop plots at the $p=0.06$ and $p=0.003$ significance levels, respectively (Figure 4.1A). Despite changes in soil nitrate, we found soil ammonium content remained unchanged between the three treatments ($p=0.97$) (Figure 4.1B). Further analysis revealed intercrop treatment also had a significant effect on soil phosphorus ($p=0.047$), with kale intercrop plots having significantly higher soil phosphorus compared to tobacco intercrop plots ($p=0.04$) (Figure 4.1C). Meanwhile, monocrop plots showed no differences in soil phosphorus between the tobacco ($p=0.71$) or kale intercrop plots ($p=0.15$). Treatment was also found to have a significant effect on soil iron, copper, lead, and boron (Supplementary Table 4.1).

Analysis of soil carbon and nitrogen stable isotope ratios from aggregated soil cores collected at harvest revealed a significant effect of treatment on soil $\delta^{13}C$ ($p=0.033$). We found kale intercrop plots had significantly lower $\delta^{13}C$ compared to tobacco intercrop ($p=0.034$) and monocrop plots ($p=0.098$) (Figure 4.1D). Meanwhile, treatment had no effect on soil %N, %C, or $\delta^{15}N$ (Supplementary Table 4.1).

Figure 4.1 A.) Comparison of soil nitrate content (mg / mg^2 ion exchange membrane SA) between treatments. B.) Comparison of soil ammonium content (mg / mg^2 ion exchange membrane SA) between treatments. C.) Comparsion of soil phosphorus content (mg / mg^2 ion exchange membrane SA) between treatments. D.) Comparison of soil $\delta^{13}C$ vs. VPDB values between treatments.

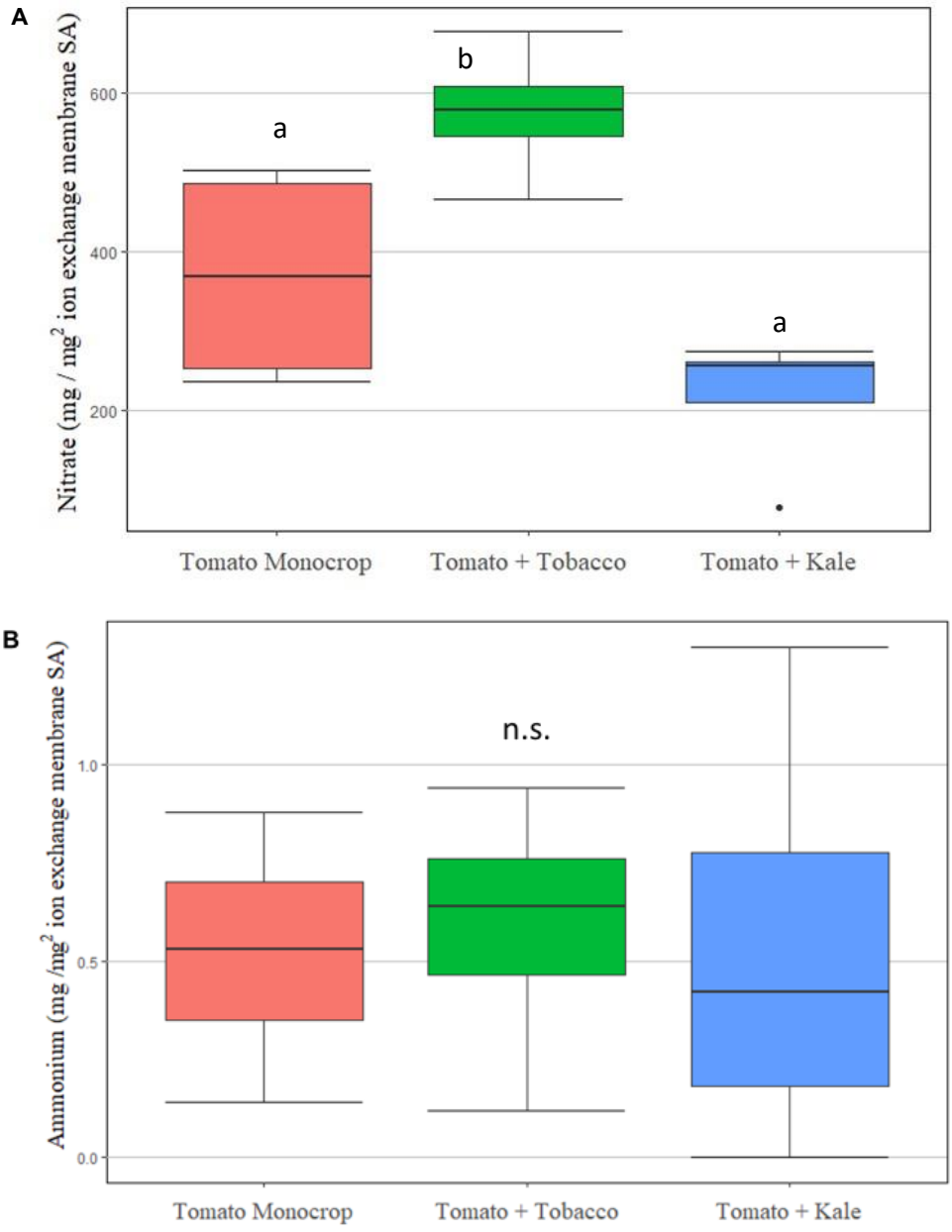
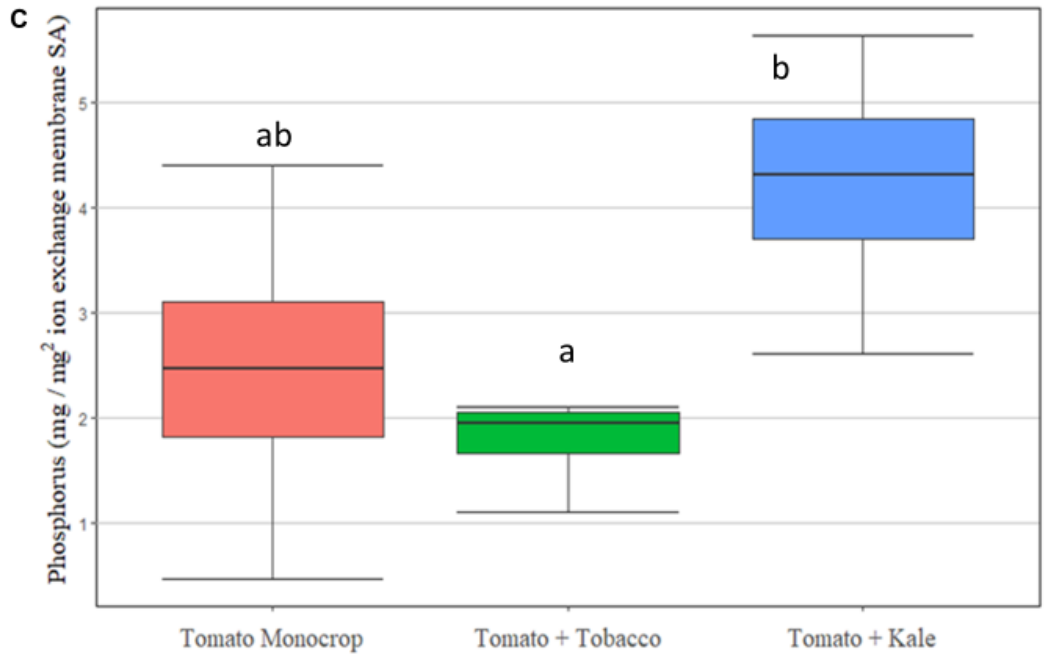
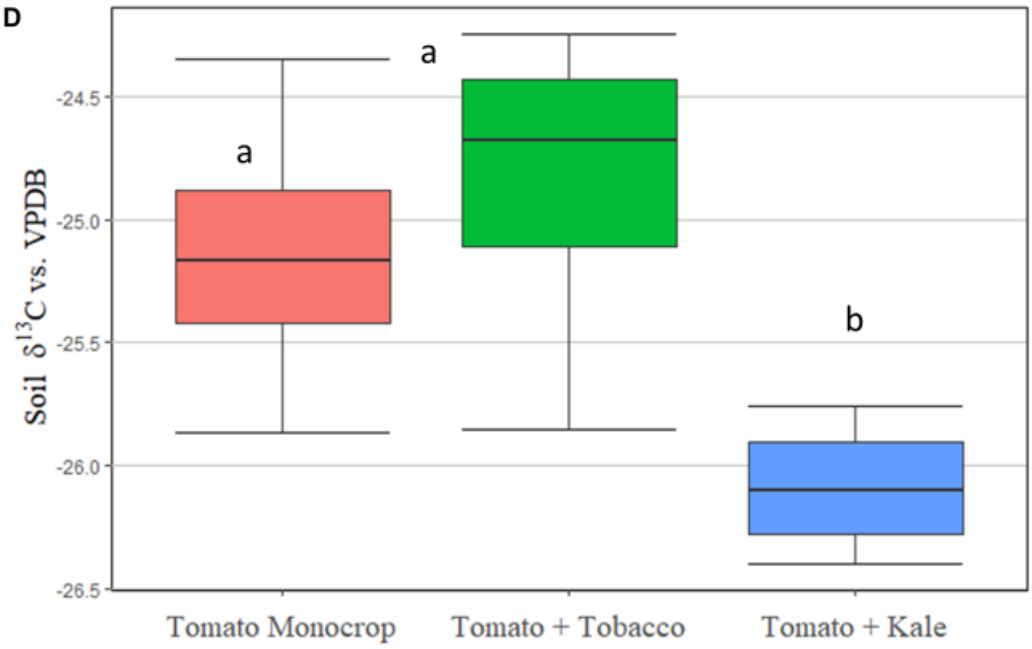

A
Nitrate (mg / mg^2 ion exchange membrane SA)
600
400
200
a
b
a
Tomato Monocrop
Tomato + Tobacco
Tomato + Kale
B
Ammonium (mg /mg^2 ion exchange membrane SA)
1.0
0.5
0.0
n.s.
Tomato Monocrop
Tomato + Tobacco
Tomato + Kale

Figure 4.1 (Continued)





4.2.2 Soil microbial community analysis

To assess the effect of intercrop treatment on soil microbiome composition, we analyzed bacterial and fungal sequencing data collected three weeks after planting and at harvest using aggregated soil cores from each microplot. Analysis of 16S sequences using ordination followed by PERMANOVA revealed a significant effect of treatment ($p=0.042$) and sampling date ($p=0.05$) but not the interaction between the two factors ($p=0.91$) on bacterial community composition (Figure 4.2A). The Shannon diversity index did not differ between sampling dates ($p=0.54$), treatment ($p=0.75$), or sampling date x treatment groups ($p=0.42$) (Supplementary Figure 4.1). Meanwhile, analysis of the relative abundances of the top eight most abundant bacterial phyla revealed significant effects of sampling date ($p=0.063$) and treatment ($p=0.03$) on the abundance of acidobacteria (Figure 4.2B), with members of the phyla becoming enriched in the kale intercrop treatment compared to the tobacco intercrop treatment ($p=0.023$). No effects of treatment were detected on the abundances of other bacterial taxa.

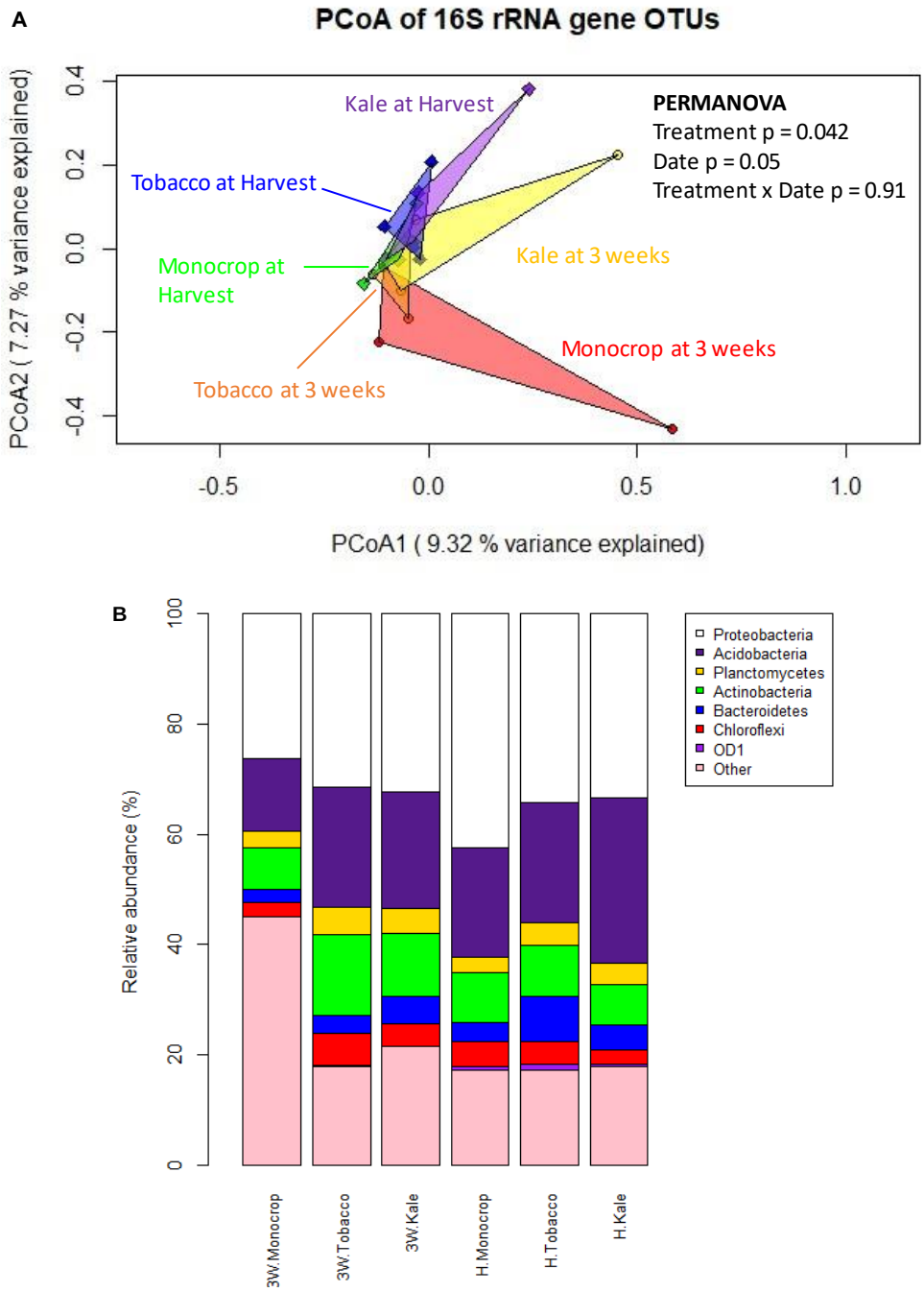
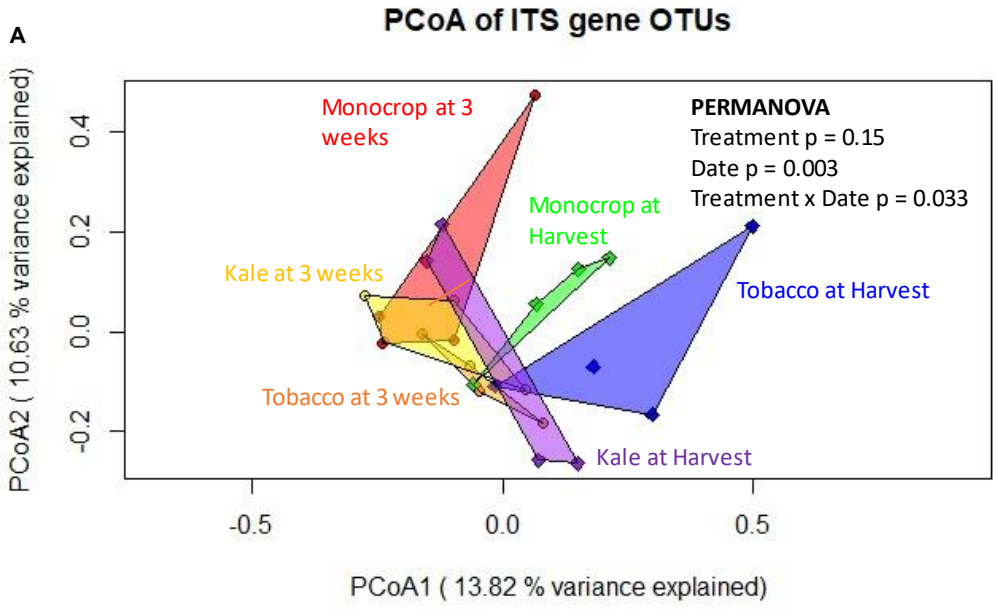


Figure 4.2 A.) Principle coordinates analysis (PCoA) of soil bacterial communities based on 16S rRNA gene sequencing on MiSeq Illumina. Sampling Date x Treatment groups are represented by colored ellipses. Monocrop three weeks after planting is represented in red, Tobacco three weeks after planting is represented in orange, Kale at three weeks after planting is represented in yellow, Monocrop at harvest is represented in green, Tobacco at harvest is represented in blue, and Kale at harvest is represented in purple. B.) Stacked barplot displaying relative abundances of the top 8 most abundant bacterial phyla across the different Sampling Date x Treatment groups.

Analysis of ITS sequences using ordination followed by PERMANOVA revealed a significant effect of sampling date (p=0.003) but no significant effects of treatment (p=0.15) or the interaction between sampling date and treatment (p=0.33) on fungal community composition (Figure 4.3A). Additionally, we found sampling date (p=0.89), treatment (p=0.9), and the interaction between the two factors (p=0.89) did not significantly affect the Shannon diversity indexes of soil fungal communities (Supplementary Figure 4.2). Analysis of the top eight most abundant fungal phyla revealed date (p=0.0014) and treatment (p=0.056) had significant effects on the relative abundance of mortierellomycota, with the abundance of mortierellomycota increasing in the kale intercrop plots compared to the monocrop plots (p=0.054) (Figure 4.3B).



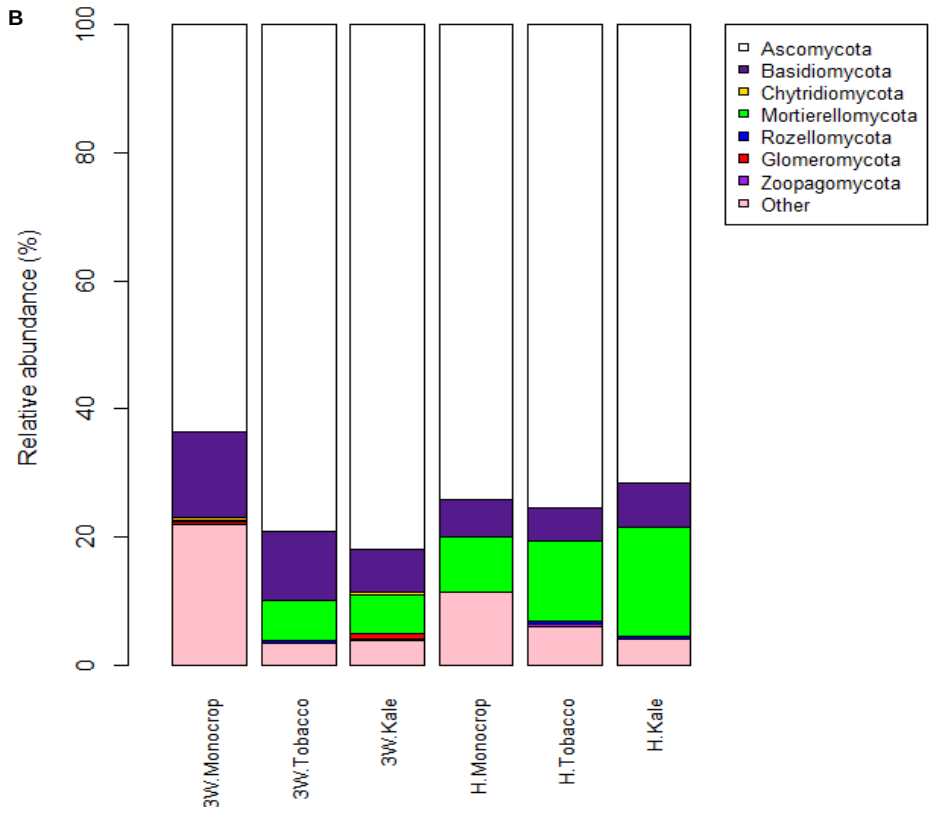


Figure 4.3 A.) Principle coordinates analysis (PCoA) of soil fungal communities based on ITS gene sequencing on MiSeq Illumina. Sampling Date x Treatment groups are represented by colored ellipses. Monocrop three weeks after planting is represented in red, Tobacco three weeks after planting is represented in orange, Kale at three weeks after planting is represented in yellow, Monocrop at harvest is represented in green, Tobacco at harvest is represented in blue, and Kale at harvest is represented in purple. B.) Stacked barplot displaying relative abundances of the top 8 most abundant fungal phyla across the different Sampling Date x Treatment groups.

Several of the most abundant fungal phyla in the dataset were only present in very low relative abundances and completely absent in many of the treatment x sampling date groups. Given this data structure, we report here our observations without statistical modeling. We found chytridiomycota decreased in abundance in the kale intercrop and monocrop plots at harvest, while the abundance of glomeromycota increased in abundance in the kale plots at harvest. Rozellomycota decreased in

abundance in the kale and monocrop plots, while the abundance of zoopagomycota decreased in kale plots at harvest and increased in tobacco plots at harvest.

4.2.3 Tomato RNA sequencing

To assess the effect of intercropping treatment on tomato plant physiology, we performed RNA sequencing on tomato foliar tissue collected from a single representative plan from each treatment across four sampling dates during the growing season. Analysis of tomato foliar RNA sequences from the three intercrop treatments showed gene expression remained mostly unchanged between treatments during the first two sampling dates (Supplementary Table 4.2). In the later sampling dates, tomato foliar transcriptomes showed more Differentially Expressed Genes (DEGs) between treatments. At harvest, we observed the ammonium transporter AF118858 (Solyc03g045070.1.1) gene was upregulated in tomato plants from the kale and tobacco intercrop treatments, while a different ammonium transporter gene (Solyc10g076480.2.1) was upregulated in tomatoes from the monocrop treatment at the p=0.05 significance level (Supplementary Table 4.3). In addition, we observed an upregulation of the bidirectional sugar transporter SWEET gene (Solyc05g024260.3.1) in plants from the tobacco and kale intercrop treatments.

4.2.4 Plant phenotypic analysis

Foliar carbon and nitrogen stable isotope ratios were measured using dried tissue samples collected from four representative plants from each plot four times throughout the growing season to find potential linkages between soil properties and

plant nutrition. Analysis of foliar isotope data revealed no effect of treatment on %N, $\delta^{15}N$, %C, $\delta^{13}C$ (Supplementary Figure 4.4). Additionally, we found treatment had no effect on average aboveground biomass production per plant at harvest (p=0.33) (Supplementary Figure 4.4). Meanwhile, analysis of the effects of treatment and fruit ripening stage (non-ripe vs. ripe fruit) on average fruit yield per plant at harvest revealed significant effects of treatment (p=0.047) and the interaction between treatment and ripening stage (p=0.0033). We found average fruit yield of tomatoes grown with tobacco was significantly higher than tomatoes that were grown in monocrop (p=0.043), while average fruit yield of tomato plants intercropped with kale did not differ from the tobacco intercrop (p=0.24) or monocrop treatments (p=0.75) (Figure 4.3). Analysis of the interaction between treatment and ripening stage revealed average non-ripe fruit yield being significantly greater in tobacco intercrop plots compared to kale intercrop (p=0.0065) and monocrop plots (p=0.0007).

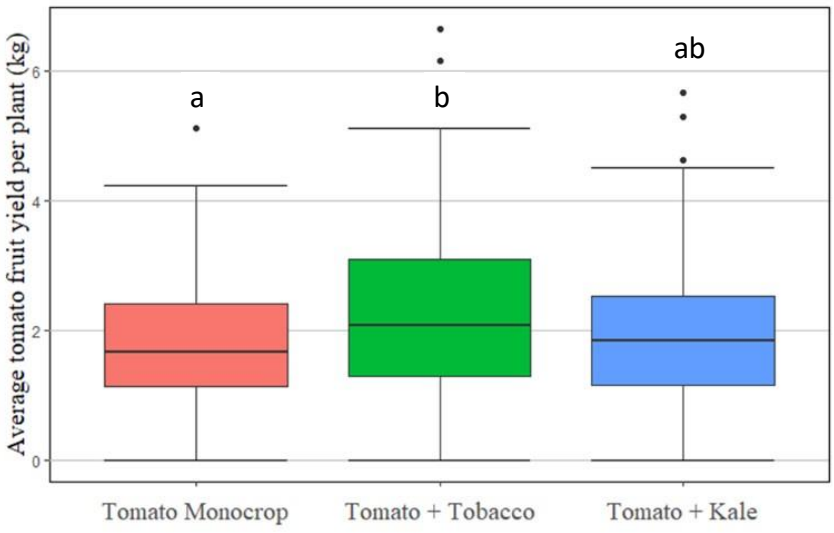


Figure 4.3 Average total tomato fruit yield per plant between the three treatments at harvest. Letters represent pairwise comparisons of average yield between treatments at each developmental stage (i.e. non-ripe and ripe). Treatments are color coded, with monocrop in red, tobacco in green, and kale in blue.

4.3 Discussion

Diversified cropping systems show great promise for promoting sustainability in agriculture, but much remains unknown about how practices such as intercropping may impact belowground ecosystem functions in agricultural production. The goal of this experiment was to investigate how intercropping may alter soil chemical and biological properties in horticultural systems, which we hypothesized would influence tomato plant traits such as yield. Results from our experiment showed intercropping had several effects on soil chemistry and biology as well as tomato fruit yield and other physiological traits. However, these effects were dependent on what plant tomato was intercropped with. In plots where tomato was intercropped with tobacco, also a solanaceous crop, we observed an increase in soil nitrate, abundances of microbial phyla such as glomeromycota, and average tomato fruit yield compared to tomato monocrops. In plots where tomato was intercropped with kale, a more distantly related brassica crop, we observed an increase in soil phosphorus compared to tobacco intercrop plots, a decrease in soil $\delta^{13}C$ values compared to both monocrop and tobacco intercrop plots, and an increase in the abundance of microbial phyla such as acidobacteria and glomeromycota compared to monocrop and tobacco plots. Ultimately, these results show intercropping could alter soil ecosystem functions in agriculture and produce measurable changes in soil properties relevant to crop production, but the effects are dependent on what commodities are cultivated together.

The observed differences in soil nutrition in the intercropped treatments may have been driven by differences in plant physiology. Several brassica plant species such as kale do not form mycorrhizal associations and must mine for soil phosphorus via other mechanisms. One of the primary mechanisms through which brassica plants mine for phosphorus is through the production of root exudates, which includes the release of organic acids and nutrient cycling enzymes that contribute to soil phosphorous cycling and eventual plant uptake[14]. The release of these exudates may have not only enhanced soil phosphorus, but also stimulated soil organic matter mineralization in kale plots[15], which is supported by the observed decrease in $\delta^{13}C$ in these plots[16].

The increased soil nitrate in tobacco plots could be explained by reduced nitrate uptake by tobacco plants. Tobacco plants have shallow root systems that mainly occupy higher soil layers and are generally inefficient at absorbing nutrients[17]. Further, due to its inefficient nutrient absorption, previous research has shown nitrogen application rates must be carefully planned in tobacco in order to prevent leaching and volatilization[17]. In the context of intercropping, slower nutrient uptake by tobacco plants could lead to higher soil nitrate levels and absorption of nitrogen by the tobacco's companion crop. In this experiment, tobacco plants senesced weeks before the tomato plants were harvested, which may have additionally reduced tomato competition for soil nitrate. Considering this, tobacco intercropping might be useful in developing novel soil nitrogen management techniques.

The observed shifts in soil microbial community composition were also likely driven by differences in the plant species cultivated in each plot. In plots where tomato

was intercropped with kale, we observed an increase in the abundance of microbial phyla such as acidobacteria. Members of this bacterial phyla are abundant in soil environments and play a key role in ecological processes such as decomposition and mineralization[18]. Previous research has shown acidobacteria are sensitive to pH and are particularly abundant in more acidic environments[18]. As kale plants could have acidified the soil environment for phosphorus acquisition, they may have also contributed to enhanced abundances of acidobacteria. Although it is unknown what specific species of acidobacteria were enriched in this study and the functional significance of acidobacteria remains to be explored[19], this result nonetheless represents a potential functional shift in the soil microbiomes of kale plots.

In addition to an increase in the abundance of acidobacteria in kale plots, we also observed an increase in the abundance of glomeromycota in both the kale and tobacco plots. Members of this phyla form arbuscular mycorrhizas with terrestrial plants and aid in the acquisition of important soil resources such as water and phosphorus[20]. Previous agricultural research has shown management practices can affect abundances of glomeromycota in the soil and influence root colonization rates[21]. Results from our study indicate intercropping can influence glomeromycota abundances, likely via shifts in root exudation and soil chemistry as observed in this study[22,23].

RNA sequencing of field grown plant tissue could be a useful method through which to assess how plants respond to changes in management practices at a physiological level[24]. In this experiment, we performed RNA sequencing of tomato foliar tissue to assess how hypothesized changes in soil properties due to intercropping

may affect plant physiological processes such as nutrient transport. We found tomato foliar gene expression remained unchanged earlier in the growing season, but later in the growing season we observed differential expression of several genes, including the expression of ammonium transporters. The ammonium transporter AF118858 (Solyc03g045070.1.1) gene was upregulated in tomato plants intercropped with kale and tobacco, while in tomatoes grown in monocrops we saw the ammonium transporter (Solyc10g076480.2.1) gene was upregulated. As ammonium transporters are involved in a variety of physiological processes[25], we hypothesize these differences in ammonium transporter expression represent differing plant responses to the soil environment in our treatments. In addition to changes in ammonium transporter gene expression, we also saw an upregulation of the bidirectional sugar transporter SWEET (Solyc05g024260.3.1) gene in tomato plants grown in tobacco and kale plots. Bidirectional sugar transporter SWEET genes are involved in a number of physiological processes[26] and may represent other physiological changes in tomato plants intercropped with tobacco and kale.

We hypothesized in this experiment that intercropping would enhance ecosystem functioning and lead to several changes in soil biological and chemical properties. Further, these changes in soil properties would alter tomato plant traits, including fruit yield. While average aboveground biomass remained unchanged between treatments, analysis of average ripe and non-ripe fruit yields showed significant effects of intercrop treatment on average yield. Specifically, tomato plants intercropped with tobacco had higher non-ripe and total fruit yield compared to the other two treatments. The enhanced soil nitrate in tomato plots could explain this

difference in plant phenotypes. Fruit production represents a significant nitrogen demand for tomato plants[27]. As nitrate is more readily utilized by tomatoes, it is possible that greater soil nitrate led to enhanced average fruit production. The observed differences in non-ripe fruit yield also indicate changes to soil properties could alter tomato fruit developmental processes. Fruit production and development are complex processes that can be influenced by several environmental factors[28]. While the effects of soil chemistry and biology on fruit traits and ripening are not well understood, our data suggest they can be important drivers of fruit production and future research should investigate linkages between soil properties and fruit traits.

Ultimately, our experiment shows that imposing spatial crop diversity through intercropping could promote ecosystem functions in agriculture and alter soil properties relevant to production. However, the effects of intercropping on soil and crop properties are dependent on the crops cultivated. Data from this experiment shows intercropping tomatoes with tobacco, a close genetic relative, can enhance soil nitrate levels while intercropping tomatoes with kale, a more distant genetic relative, can enhance soil phosphorus. These changes in soil properties could lead to changes in plant physiology and phenotypic outcomes such as fruit yield, as observed in the tobacco plots in this experiment. Future work should focus on teasing apart the specific mechanisms through which intercropped plants alter ecosystem functions, what the effect of intercropping with other crop species may be, and how intercropping may be optimized or combined with other practices such as crop rotations to enhance multiple ecosystem services in agricultural settings.

4.4 Materials and Methods

4.4.1 Cultivation of crop seedlings

All crops used in this experiment were first cultivated in a greenhouse until planting. Tomato (M82), tobacco (*Nicotiana quadrivalvis*, The Plant Good Seed Company, Ojai, CA, USA), and kale (Darkibor, Johnny's Selected Seeds, Winslow, ME, USA) seeds were planted into seedling trays filled with 50% metromix 360, 50% cornell mix + osmocote container soil. Trays were then kept in a seedling greenhouse with day temperatures between 25ºC and 27ºC, night temperatures between 18ºC and 20ºC, and natural light and humidity. All plants were hand watered with tap water and fertilized with Peter's excel fertilizer (N-P-K) during this phase. After 5 weeks in the greenhouse, seedlings were then hardened off and transplanted into the field.

4.4.2 Field trial

The field trial was performed at Caldwell Field at Cornell University. The field plot consisted of 12 microplots, each randomly assigned to 1 of the 3 treatments. The first treatment consisted of tomato plants grown in monocrops, which served as our control treatment. The second treatment consisted of tomato intercropped with tobacco and the final treatment consisted of tomato intercropped with kale. We chose tobacco and kale as companion crops to assess if growing tomato with a close and distant genetic relative has any significant effect on soil ecosystem functions and properties. Each microplot consisted of 5 rows of plants. The outer two rows of each microplot were buffer rows consisting of tomato plants only, while the inner two rows were

planted according to the microplot's assigned treatment. In monocrop plots only tomatoes were planted, while in intercrop plots tomato plants were planted with either tobacco or kale in a pattern within a single row.

Plastic mulch was placed over each row for weed control. A drip irrigation line was used throughout the growing season and programmed to water plants everyday. The program was adjusted during periods of high soil moisture. We used Neptune's Harvest (Gloucester, MA, USA) organic fertilizer (2-4-1 N-P-K) to fertilize our plots once every two weeks following the manufacturer's recommendation for outdoor plants. In line with organic management practices, we refrained from any pesticide usage throughout the experiment.

4.4.3 Soil and plant sampling

5 soil cores were gathered from each microplot using an ethanol sterilized soil corer every three weeks until harvest starting 3 weeks after transplanting in order to assess the effect of date and intercrop treatment on soil microbiome composition. Cores were collected from a diagonal transect across the middle bed of each plot to capture bulk and rhizosphere soil. The 5 cores from each plot were aggregated into a plastic bag, mixed, sieved, then kept at -20ºC for downstream analysis.

Plant root simulator (PRS) probes manufactured by Western Ag Innovations (Saskatoon, Saskatchewan, CA) were buried 5 weeks after transplanting following the manufacturer's recommendations in order to assess the effect of intercropping on soil nutrient dynamics. Four pairs of anion and cation PRS probes we buried in the center of each microplot. Probes were left buried for 4 weeks, then harvested and washed

with deionized water. The probes were then sent to Western Ag Innovation for analysis of ammonium and nitrate using colorimetrical methods as well as soil P, K, S, Ca, Mg, Fe, Mn, Cu, Zn, B, Al, Pb, and Cd using inductively coupled plasma optical emission spectrometry (ICP-OES).

Tomato foliar tissue was harvested from a single representative plant from each plot four times throughout the growing season in order to assess the effect of date and intercrop treatment on foliar carbon and nitrogen stable isotope ratios and the foliar transcriptome. For transcriptomic analysis, ~0.5 g of tissue was gathered from each representative plant, then immediately placed on dry ice and submerged in liquid nitrogen in the lab. Frozen foliar samples were then kept at -80ºC for downstream RNA extractions. For foliar isotopic measurements, tissue was dried at 60ºC for at least 3 days, then ground using a mortar and pestle for analysis. Foliar samples gathered throughout the growing season and soil samples gathered at harvest were sent to the Cornell Stable Isotope Laboratory (COIL) for analysis on a Thermo Delta V isotope ratio mass spectrometer (IRMS) interfaced to an NC2500 elemental analyzer.

4.4.4 Soil microbiome sequencing

Soil microbial DNA was extracted from aggregated soil samples using a PowerSoil DNA Isolation Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. Approximately 200 mg of soil from each sample was used for extraction. 16S rRNA and ITS gene sequences were amplified using PCR primers described in Caporaso et al.[29] for barcoded sequencing. After an initial clean-up of the samples using the HighPrep PCR Clean-up System (MAGBIO Genomics,

Gaithersburg, MD, USA), unique Index Primers were attached to amplicons in each sample with a second PCR cycle. The indexed samples were cleaned and normalized using the SequalPrep Normalization Plate Kit (Thermo Fisher Scientific, Waltham, MA, USA). Sample normalization was followed by pooling 5 ul of each sample into one composite sample. The pooled sample was then ran on a 1.2% agarose gel with SyberSafe added and the target band was excised. The DNA was then extracted from the gel using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). The final pooled sample was then sent to the Cornell Genomics Facility (Cornell University, Ithaca, NY, USA) for sequencing on the Illumina MiSeq platform.

4.4.5 Tomato foliar RNA sequencing

RNA was extracted from frozen tissue samples by first adding 800 µl of a modified RLT buffer and 800 µl of chloroform to our samples. The samples were then vortexed, centrifuged at 12000 rpm for 10 minutes, then the supernatant was transferred to a new tube. The supernatant was then mixed with half the volume of 100% ethanol, then transferred to spin columns. The samples were then centrifuged for 1 minute at 8000 rpm, then washed twice with 80% ethanol. Spin columns were then spun at 12000rpm for one minute, then transferred to new 1.5 ml tubes. We then added 55 µl of DEPC water to each column, incubated the columns for 2 minutes, then spun them down for 1 minute at 12000rpm. Samples were then quantified and ran on a 1% agarose gel to check the RNA quantity and quality of each sample. Total RNA from foliar and fruit tissue was then used to construct strand-specific RNA libraries as

described in Zhong et al.[30].

4.4.6 Statistical analyses of soil and plant measurements

The effects of intercropping on soil properties and plant phenotypes were analyzed using R statistical software. To assess the effect of intercropping on soil chemistry and CN stable isotope ratios, one-way ANOVAs on linear models with treatment as the sole factor were used. Two-way ANOVAs with sampling date and treatment as factors were used for analysis of foliar CN stable isotope ratios. Following significant outcomes, pairwise comparisons of group means was performed using the emmeans package in R. Mixed-effect modeling was used to analyze potential differences in average tomato aboveground biomass production in each plot and average fruit yield of ripe and non-ripe tomato fruits after harvest. All ANOVA assumptions were verified using plotting methods. Data transformations were performed as necessary in order to meet assumptions.

4.4.7 Processing and analysis of soil microbiome sequences

A modified version of the pipeline from the Brazilian Microbiome Project (http://www.brmicrobiome.org/) was utilized to process soil microbiome sequences. Mothur v. 1.3613 was used to merge paired-end sequence (make.contigs), trim off primers (trim.seqs, pdiffs = 2, maxambig = 0), remove singletons (unique.seqs → split.abund, cutoff = 1) and classify sequences (97% similarity). OTUs that were suspected to not be of fungal or bacterial origin were removed (remove.lineage). Qiime v. 1.9.114 was then used to cluster OTUs and create our bacterial and fungal

OTU tables for analysis.

After processing, we found the reads for bacterial and fungal OTUs from our second sampling date were low quality, thus decided to only examine microbial data from the first and final sampling dates. The bacterial OTU table was rarefied to 163 reads per sample and the fungal OTU table was rarefied to 66 reads per sample, which was the minimum number of reads for both datasets. We then calculated differences in community composition between samples using Bray-Curtis distances. These values were then utilized to create an ordination to visualize differences in microbial community fingerprints between microbiome treatments. We then utilized the adonis package in R to perform a PERMANOVA, which allowed us to assess the effect of sampling date, treatment, and the interaction between the two on bacterial and fungal community composition. Shannon diversity indexes were calculated for each treatment and compared using a two-way ANOVA with sampling date, treatment, and the interaction between the two as factors. Relative abundances of bacterial and fungal phyla were compared using two-way ANOVAs with date, treatment, and the interaction between the two as factors. Some fungal phyla had low relative abundances with many factor groups having average abundances of zero. Given the data structure for these phyla, we report absolute values without statistics for these phyla.

4.4.8 Processing and analysis of tomato foliar tissue RNA sequences

Raw RNA-Seq reads were processed to remove adaptors and low-quality sequences using Trimmomatic (version 0.36) with default parameters[31]. Cleaned reads shorter than 40 bp were discarded. The remaining cleaned reads were aligned to the

ribosomal RNA database[32] using bowtie (version 1.1.2)[33] allowing up to three mismatches, and those aligned were discarded. The remaining high-quality cleaned reads were aligned to the tomato reference genome (SL3.0 and ITAG3.2) using HISAT2 (version 2.1.0). Based on the alignments, raw read counts for each gene were calculated and then normalized to reads per kilobase of exon model per million mapped reads (RPKM). Raw read counts were then fed to DESeq2 to identify differentially expressed genes (DEGs), with a cutoff of adjusted P value < 0.05 and fold change > 2.

4.5 References

1. Tilman, D., Cassman, K., Matson, P., Naylor, R., & Polasky, S. Agricultural sustainability and intensive production practices. *Nature* 418: 671 (2002).

2. Su, Y., Wang, F., Suo, D., Zhang, Z., & Du, M. Long-term effect of fertilizer and manure application on soil-carbon sequestration and soil fertility under the wheat–wheat–maize cropping system in northwest China. *Nut. Cycl. Agroecosystems* 75(1): 285-295 (2006).

3. Zeng, J. et al. Nitrogen fertilization directly affects soil bacterial diversity and indirectly affects bacterial community composition. *Soil Biol. Biochem.* 92: 41-49 (2016).

4. Alewell, C. et al. Global phosphorus shortage will be aggravated by soil erosion. *Nat. Commun.* 11, 4546 (2020).

5. Kremen, C., Iles, A., & Bacon, C. Diversified farming systems: an agroecological, systems-based alternative to modern industrial agriculture. *Ecol. Soc.* 17(4): 44 (2012).

6. Curtright, A., & Tiemann, L. Intercropping increases soil extracellular enzyme activity: A meta-analysis. *Agriculture, Ecosystems & Environment v.* 319: pp. 107489--102021 v.107319 (2021).

7. Jalli, M., et al. Effects of Crop Rotation on Spring Wheat Yield and Pest Occurrence in Different Tillage Systems: A Multi-Year Experiment in Finnish Growing Conditions. *Front. Sustain. Food Syst.* 5 (2021).

8. Renwick, L., et al. Long-term crop rotation diversification enhances maize drought resistance through soil organic matter. *Environ. Res. Lett*. 16 084067 (2021).

9. Gaudin, A. C. M., et al. Wheat improves nitrogen use efficiency of maize and soybean-based cropping systems. *Agric. Ecosyst. Environ. AGR* 210: 1-10 (2015).

10. Zhang et al. Intercropping cereals with faba bean reduces plant disease incidence regardless of fertilizer input; a meta-analysis. *Eur. J. Plant Pathol*. (2019).

11. Dong, N. et al. Temporal Differentiation of Crop Growth as One of the Drivers of Intercropping Yield Advantage. *Sci Rep* 8, 3110 (2018).

12. Lange, M., et al. Plant diversity increases soil microbial activity and soil carbon storage. *Nat. Commun.* 6, 6707 (2015).

13. Shackelford, G., et al. Effects of cover crops on multiple ecosystem services: Ten meta-analyses of data from arable farmland in California and the Mediterranean. *Land Use Policy* 88: 104204 (2019).

14. Hunter, P., Teakle, G., & Bending, G. Root traits and microbial community interactions in relation to phosphorus availability and acquisition, with particular reference to Brassica." *Front. Plant Sci.* 5 (2014).

15. Dijkstra, F.A., Zhu, B. & Cheng, W. Root effects on soil organic carbon: a double-edged sword. *New Phytol.*, 230: 60-65 (2021).

16. Volk, M., Bassin, S., Lehmann, M. F., Johnson, M. G., & Andersen, C. P. 13C isotopic signature and C concentration of soil density fractions illustrate reduced C allocation to subalpine grassland soil under high atmospheric N deposition. *Soil Biol. Biochem.* 178–184 (2018).

17. Soares, T., Coelho, F., de Oliveira, V., Pontes, O., & Pavinato, P. Soil nitrogen dynamics under tobacco with different fertilizer management in southern Brazil. *Geoderma Reg.* 21: e00282 (2020).

18. Kalam, S., et al. Recent Understanding of Soil Acidobacteria and Their Ecological Significance: A Critical Review. *Front. Microbiol.* 11 (2020).

19. Kielak, A. M., et al. The Ecology of Acidobacteria: Moving beyond Genes and Genomes. *Front. Microbiol.* 7 (2016).

20. Willis, A., Rodrigues, B., & Harris, P. The Ecology of Arbuscular Mycorrhizal Fungi. *CRC Crit. Rev. Plant Sci.* 32:1, 1-20 (2013).

21. Vasilikiotis, C., et al. Orchard management practices affect arbuscular mycorrhizal fungal root colonisation of almond. *Biol. Agric. Hortic.* 36(4): 230-248 (2020).

22. Wang, J. et al. Effect of root exudates on beneficial microorganisms—evidence from a continuous soybean monoculture. *Plant Ecol.* 213, 1883–1892 (2012).

23. Schappe, T. et al. (2017), The role of soil chemistry and plant neighbourhoods in structuring fungal communities in three Panamanian rainforests. *J Ecol*, 105: 569-579 (2017).

24. Bowles, T., Hollander, A., Steenwerth, K., & Jackson, L. Tightly-Coupled Plant-Soil Nitrogen Cycling: Comparison of Organic Farms across an Agricultural Landscape. *PLOS ONE* 10(6): e0131888 (2015).

25. Hao, D., et al. Function and Regulation of Ammonium Transporters in Plants. *Int. J. Mol. Sci.* 21(10) (2020).

26. Chen, L.-Q. SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol.* 201:1150-1155 (2014).

27. Trandel, M., Vigardt, A., Walters, S., Lefticariu, M., & Kinsel, M. Nitrogen Isotope Composition, Nitrogen Amount, and Fruit Yield of Tomato Plants Affected by the Soil-Fertilizer Types. *ACS omega*, 3(6), 6419–6426 (2018).

28. Quinet, M., et al. Tomato Fruit Development and Metabolism. *Front. Plant Sci.* 10 (2019).

29. Caporaso, J. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6: 1621–1624 (2012).

30. Zhong, S. et al. High-throughput illumina strand-specific RNA sequencing library preparation. *Cold Spring Harb. Protoc.* 2011: 940–949 (2011).

31. Bolger, A.M., Lohse, M., & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30:2114–2120 (2014).

32. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41:D590–D596 (2013).

33. Langmead, B. Aligning short sequencing reads with Bowtie. *Curr Protoc. Bioinformatics.* 32:11.17.11–11.17.14 (2010).