Investigating Structure and Function of Rhizosphere Associated Microbial Communities in Natural and Managed Plant Systems

Richard Rosario Rodrigues

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Genetics, Bioinformatics, and Computational Biology

Mark A. Williams
Bingyu Zhao
Roderick V. Jensen
David R. Bevan

March 16, 2016
Blacksburg, Virginia

Keywords: Metagenomics, Biological nitrogen fixation, Associative nitrogen fixers, Rhizosphere, Switchgrass, Invasive plants

Investigating Structure and Function of Rhizosphere Associated Microbial Communities in Natural and Managed Plant Systems

Richard Rosario Rodrigues

ABSTRACT

Many plants, especially grasses, have Nitrogen (N) as their growth-limiting nutrient. Large amounts of N fertilizer (>100 kg N ha^{-1}) are used in managed systems to maximize crop productivity. However, the plant captures less than 50% of the (~12 million tons per year, U.S.) applied N-fertilizer. The remaining mobile N lost through leaching and denitrification accumulates in waterways and the atmosphere, respectively. Losses of fertilizers create environmental and economic concerns globally and create conditions that support the invasion of exotic plants in the natural landscapes. There is thus a need to come up with biological solutions to better manage nitrogen for plant growth and ecosystem sustainability. Microbial communities in the rhizosphere are known to potentially have beneficial effects on plant growth. Diazotrophs, for example, are bacteria that can convert the atmospheric nitrogen to ammonia, a process called "nitrogen fixation." Utilizing the natural process of associative nitrogen fixation to support most of the plant's N needs would substantially reduce fertilizer use and thus reduce production and environmental costs. *The goal of this dissertation was to determine the structure and function of root-zone microbial communities for increasing productivity of native plants.* Towards this end, we study the root-zone bacterial and fungal communities of native and exotic invasive plants. This study identifies that shifts in rhizosphere microbial communities are associated with invasion and highlights the importance of rhizosphere associated structure and function of microbes. A study of root-zone associated microbes in switchgrass (*Panicum virgatum* L.) - a U.S. native, warm-season, perennial, bioenergy crop indicates that high biomass yield

and taller growth are associated with increased plant N-demand and supportive of bacteria with greater rates of N_2-fixation in the rhizosphere. Another crucial outcome of the thesis is a better description of the core and cultivar-specific taxa that comprise the switchgrass root-zone associated microbiome. The work in this dissertation has brought us closer to designing N supply strategies by utilizing the natural microbial communities to balance the N-cycle in agroecosystems and support a sustainable environment.

Investigating Structure and Function of Rhizosphere Associated Microbial Communities in Natural and Managed Plant Systems

Richard Rosario Rodrigues

General Audience Abstract

The capability to grow crops while also taking care of the environment is a major scientific challenge for global sustainability. Plants, especially grasses, cannot directly utilize the nitrogen gas, which is abundant in our atmosphere. The resulting nitrogen limitation serves as a bottleneck in plant production. Extreme amounts of chemical nitrogen-based fertilizers added to increase plant yield pollute land, air, and water. These in turn, affect human and marine health and come with massive economic costs of more than 500 billion U.S. dollars, annually. One major advance to achieve sustainable plant production would be to harness and further develop plant-microbial technologies.

This dissertation was aimed towards identifying bacteria that grow in the plant's root-zone and promote plant growth. Towards this end, we have characterized the microbes in the root-zones of natural and managed plant systems that may help to meet the plant's changing nitrogen requirements. Indeed, such "diazotrophic" bacteria, for example, can serve as alternatives to chemical fertilizers, help to produce high-quality, nutritious food and are ultimately responsible for the large majority of protein found on planet Earth. The research is thus an important step forward in the production of healthy plants, soils, and clean water for a sustainable global environment.

By better describing the complex, highly diverse, plant associated bacterial communities the scientific community is closer to managing the activities of these crop-associated bacteria. Using such plant-microbial associations will allow for growth on low-fertility land, restore soil health, and provide other ecosystem benefits, such as the production

of high-quality forage for livestock and feedstock for making biofuels. Yield of other grasses (e.g. rice, maize, wheat), like switchgrass, are limited by nitrogen; hence it is possible to transfer our research insights to food crops as well. Increasing crop production using sustainable methods can help to satisfy the world food and fuel demands, while simultaneously enhancing the quality of life and health of the environment.

Dedicated to my loving family!

Nathubai (Aajimai), Bavatis (Shepai), Philomena (Nana) & Francis (Pappa)

Maria (Mummy) & Rosario (Daddy)

Charles (Brother) & Neelam (Fiancée)

I had this opportunity only because of their determination, efforts, love, and support.

Acknowledgements

I want to sincerely express my gratitude to my advisor, Dr. Mark A. Williams, whose support has been pivotal to my Ph.D. I want to thank him particularly for going above and beyond to help me in my personal and professional life. I am fortunate to have learned from an intellect, Dr. Williams, who is an equally humble person and genuinely cares for others. I honestly enjoyed working with him.

I would like to express my deep appreciation to Dr. David Bevan and Dr. Roger Harris for the support and mentoring. I want to thank my committee members, Dr. Bingyu Zhao and Dr. Roderick Jensen, for their guidance and thought-provoking suggestions.

I want to thank my lab members, especially Rosana Pineda, for providing suggestions when I was choosing an advisor and showing me the ropes when I was new to the lab. I want to thank all the undergraduates in our lab for their help. I also want to thank Kerri Mills and Dennie Munson for their technical assistance.

I am grateful to Virginia Tech's Department of Horticulture, Graduate Student Assembly, and Genetics, Bioinformatics, & Computational Biology Program for providing a nurturing environment and helping me grow as a person.

Last but not least, I want to thank my family for their support in my pursuit of a Ph.D.

Attribution

Several collaborators have helped in the research in the different chapters presented as part of this dissertation. Their contributions are described below.

Chapter 2: Plant Invasions Associated with Change in Root-Zone Microbial Community Structure and Diversity, 2015, PLoS ONE 10(10): e0141424.

Authors and affiliations:
Richard R. Rodrigues[1], Rosana P. Pineda[2], Jacob N. Barney[3], Erik T. Nilsen[4], John E. Barrett[4], Mark A. Williams[1,2]
[1]Interdisciplinary Ph.D. Program in Genetics, Bioinformatics, and Computational Biology; [2]Department of Horticulture; [3]Department of Plant Pathology, Physiology, and Weed Science; [4]Department of Biological Sciences; Virginia Tech, Blacksburg, VA.

Conceived and designed the experiments: RRR RPP JNB ETN JEB MAW. Performed the experiments: RRR RPP. Analyzed the data: RRR RPP MAW ETN. Contributed reagents/materials/analysis tools: RRR RPP JNB ETN JEB MAW. Wrote the paper: RRR MAW JNB RPP ETN.

Chapter 3: Microbial Communities and Diazotrophic Activity Differ in the Root-Zone of Alamo and Dacotah Cultivars of Switchgrass.

Authors and affiliations:
Richard R. Rodrigues[1], Jinyoung Moon[2], Bingyu Zhao[2], Mark A. Williams[1,2]
[1]Interdisciplinary Ph.D. Program in Genetics, Bioinformatics, and Computational Biology; [2]Department of Horticulture; Virginia Tech, Blacksburg, VA.

Conceived and designed the experiments: RRR JM BZ MAW. Performed the experiments: RRR JM. Analyzed the data: RRR JM MAW. Contributed reagents/materials/analysis tools: RRR BZ MAW. Wrote the paper: RRR JM BZ MAW.

Chapter 4: The Search for the Core Microbiome in Switchgrass Rhizosphere.

Authors and affiliations:
Richard R. Rodrigues[1], Mark A. Williams[1,2]
[1]Interdisciplinary Ph.D. Program in Genetics, Bioinformatics, and Computational Biology; [2]Department of Horticulture; Virginia Tech, Blacksburg, VA.

Conceived and designed the experiments: RRR MAW. Performed the experiments: RRR. Analyzed the data: RRR. Wrote the paper: RRR MAW.

Table of Contents

CHAPTER 1
Introduction

Importance of nitrogen in agriculture

Nitrogen (N), *so much, yet so little*. N is the most abundant element in the atmosphere as di-nitrogen gas (N_2). It is a key component of chlorophyll, proteins, and nucleic acids, hence, essential in the growth and maintenance of living organisms. The "Nitrogen cycle" is the process by which N is interconverted between organic (*e.g.,* urea, proteins, nucleic acids) and inorganic forms (*e.g.,* ammonia, nitrous oxide, nitric oxide, nitrite, nitrate), and moves between the atmosphere, biosphere, and geosphere. Plant uptake of environmental and soil nitrogen is typically assimilated via inorganic forms, however small organic nitrogen molecules such as amino acids are also important. The uptake of N by the plants along with N export through crop-removal, denitrification, volatilization, soil erosion, and leaching end up depleting bio-available N reserves (Galloway *et al.*, 2003, Magill *et al.*, 2000). Thus, nitrogen supply is important to meet the nitrogen demand of high productivity cropping systems. Deficiency of bioavailable N limits plant productivity and has major implications in agriculture and feeding the world.

Negative effects of reactive N on environment and ecosystem's vegetative structure

Over the last century, plants have been supplied with increasing amounts of land-applied synthetic chemical fertilizers. Over the past 10 to 15 years, moreover, global agricultural consumption of nitrogen has increased from 83 million metric tons (MMT) per year to close to 104 MMT per year (http://www.eurochem.ru/ar/2010/4-importance-nitrogen). In agro-ecosystem and managed grasslands, N-fertilizers provide a majority of the reactive nitrogen (N_r) species, e.g. N ions and gases (Galloway *et al.*, 2003, Magill *et al.*,

2000). The increased use of nitrogen fertilizers has made a large contribution to increasing crop production to meet rising global demand.

The massive and increasing amounts of fertilizer application to the land base have also created numerous environmental problems. Annually, crops in agroecosystems merely use ~30-50% of the existing (*e.g.,* crop residue, animal manure, etc.) and newly added (*e.g.,* fertilizer and BNF) N_r. A majority of the remaining N is lost to the soil, water and air via denitrification and leaching (Blair *et al.*, 1998, Fowler *et al.*, 1997, Galloway *et al.*, 2003, Galloway *et al.*, 2008, Oenema *et al.*, 2009, Smil, 1999, Smil, 2002, Sommer & Hutchings, 1997), negatively affecting environment and human health (Galloway *et al.*, 2008, Wolfe & Patz, 2002). The only study, to my knowledge, to place an economic value on the threats of N pollution was conducted for the European Union (Sutton *et al.*, 2011a, Sutton *et al.*, 2011b), estimating societal costs of U.S. $100–400 billion per year. This is more than double the value that N fertilizers add to European farm income (Sutton *et al.*, 2011a). Much of the U.S. and Europe share many similar problems related to land application of N fertilizer. Some of the largest problems are fertilizer movement into local and regional waterways resulting in hypoxic dead-zones with significant environmental and economic implications on coastal ecosystems (Diaz & Rosenberg, 2011), *e.g.,* the Gulf of Mexico (Rebich *et al.*, 2011) and the Chesapeake Bay (Moore *et al.*, 2011).

The ill effects of excessive chemical fertilizers also extend beyond water pollution. Excessive N fertilizer applications and changes in nutrient cycling have been shown to be associated with increased invasion by exotic plants (Castro-Diez *et al.*, 2014, Chen *et al.*, 2013, Hawkes *et al.*, 2005, Parepa *et al.*, 2013, Parker & Schimel, 2010), which is a problem in both natural and managed plant systems. Invasive plants alter the plant community, yield, and soil nutrient cycling (Ehrenfeld, 2003, Liao *et al.*, 2008,

Vilà *et al.*, 2011) and greatly affect remnant and native ecosystems. While soil microbial communities have been implicated in the success of invaders (Jeschke *et al.*, 2012a, Tamura & Tharayil, 2014, Wolfe & Klironomos, 2005), there is a lack of data explicitly checking belowground community structure and function (Kourtev *et al.*, 2002). There is thus a need to use an individualistic approach to understand the plant-soil-microbe interactions associated with native and invasive plants (Jeschke *et al.*, 2012b).

Overall, the ability to continuously increase agricultural production to feed more people is not easily reconciled against the backdrop of reduced ecosystem functions and services caused due to N fertilizers, and the dichotomy of this predicament may be best summarized as having too much of a good thing (Sutton *et al.*, 2011b). It is clear cut that alternative management strategies and solutions are needed to confront the excessive use of synthetic nitrogen fertilizers.

Biological nitrogen fixation as an alternative to N fertilizers

If global crop nitrogen demand continues to increase, as expected, over the coming decades, environmental friendly alternative solutions will be needed to meet this demand. The conversion of atmospheric nitrogen to ammonia, "nitrogen fixation," is an important first step in the availability of N to plants. It is relatively easier to manage the input and efficient use of bio-available N. In addition, providing bio-available N in the right amounts is an important factor in agriculture, hence, further discussion and the major focus of the dissertation was on the study of nitrogen fixation.

There are two types of natural nitrogen fixation: (a) atmospheric fixation and (b) biological N fixation. The energy from lightning causes the atmospheric nitrogen to combine with oxygen to form nitrogen oxides which, when mixed with rain, form nitrates that can enter the soil. This process of fixing nitrogen is called "atmospheric fixation,"

which forms a minor source (~5–8%) of the total fixed nitrogen. Biological nitrogen fixation (BNF), on the other hand, is carried out by special groups of prokaryotes, mainly bacteria and archea, called "diazotrophs" and is a major source to fix atmospheric nitrogen (Lynch & Hobbie, 1988). Diazotrophs use energy in the presence of an oxygen sensitive enzyme, nitrogenase, to convert atmospheric nitrogen to ammonia (Hubbell & Kidder, 2009). Microbes play an important role in the N cycle, namely nitrogen-fixation, nitrification, and denitrification. However, in this thesis we mainly focus on the contribution from the nitrogen-fixing bacteria.

Symbiotic versus associative nitrogen fixers

Diazotrophs can be broadly separated into two major groups, namely symbiotic and associative. Symbiotic microbes have a *quid-pro-quo* relationship with host plants. The microbes receive carbon from plants, in the form of sugars, as a source of energy. In return, microbes provide the plants with fixed nitrogen. Sometimes, such a symbiotic relationship and the oxygen sensitivity of the nitrogenase enzyme cause the formation of specialized "zones" for nitrogen fixation *e.g.,* heterocysts, nodules, etc. Examples of symbiotic diazotroph-plant relationships include: the water fern *Azolla*'s with the cyanobacterium *Anabaena azollae*; actinorhizal trees and shrubs, such as Alder (*Alnus sp.*), with the actinomycete *Frankia*; and legumes, such as alfalfa and soybean, with *Rhizobium* (Wagner, 2011).

Associative microbes are thought to utilize plant exudates for energy, carbon, and nutrients, but do not form any specialized complex structures. Such microbes can be endophytes or present on the root surface or in the rhizosphere soil. Associative nitrogen fixation (ANF) provides significant amounts of nitrogen (>30-40 kg N ha^{-1} yr^{-1}) to agriculturally important non-legumes, *e.g.,* sugar cane (*Saccharum spp.*) and forage

grasses, such as *Panicum maximum, Brachiaria sp.* and kallar grass (*Leptochloa fusca*), when grown in uninoculated, N-deficient soils (Chalk 1991). Recent evidence indicate that numerous types of feedstocks (Miscanthus, energy-cane, and switchgrass (*cv.* 'Alamo')) obtaining 20-50% of their N per year through root-zone associated nitrogen fixation (Wewalwela, 2014). The contributions of associative nitrogen fixation to other C4 grasses have been estimated too (Weier, 1980). There is thus potential for ANF to make biologically significant contributions towards the plant's N demand.

Switchgrass as a system to study ANF

Switchgrass (*Panicum virgatum* L.) is a warm-season, C4, perennial tall grass, native to the United States. It represents a model system to study the associative nitrogen fixing microbial communities in feedstock grasses. The information gleaned from the study of switchgrass is expected to have application to world food crops, which are primarily domesticated grasses (e.g., rice, wheat, maize).

Switchgrass has two ecotypes: *lowland*–which are thick stemmed and better adapted to the moist and warmer southern U.S., and *upland*–which are thin stemmed and better adapted to the drier and cooler conditions of the mid and northern U.S. (Casler *et al.*, 2004, Wright & Perlack, 2011). Alamo, 'Miami,' and 'Kanlow' are lowland ecotypes, while 'Shawnee,' 'Trailblazer,' and 'Cave-in-Rock' are examples of upland ecotypes. Lowland cultivars of switchgrass are mostly tetraploid (2n = 4x = 36 chromosomes), while upland cultivars are mainly tetraploid and octoploid (2n = 8x = 72 chromosomes) (Zalapa *et al.*, 2011). Intercrossing between cultivars with same ploidy level is possible, irrespective of the ecotypes (McLaughlin *et al.*, 1999, Vogel & Mitchell, 2008). Hence, different cultivars of switchgrass are grown throughout the US.

Switchgrass can be grown from seeds and rhizomes, and stands can grow up to 10

feet tall (Weaver, 1968) and persist for more than 10 years after planting. Its ability to efficiently obtain water, store and reutilize nutrients (K, P, N, and carbohydrates) from rhizomes allows it to survive a variety of conditions (Moser & Vogel, 1995) and regrow after harvest or grazing (McLaughlin *et al.*, 1999, McLaughlin & Adams Kszos, 2005). Switchgrass biomass is primarily used as forage and feedstock for biofuel (Moser & Vogel, 1995). It also provides environmental benefits such as reducing atmospheric carbon dioxide by increased carbon sequestration (Follett *et al.*, 2012, Ma, 1999, Skinner, 2009), biological nitrogen-fixation (Tjepkema, 1975), preventing soil erosion by binding loose soils, and providing wildlife habitat (Wolf & Fiske, 1995). Switchgrass, thus, has great potential to produce sustainable feedstock and forage grasses.

Sustainable production of Switchgrass

Switchgrass, like other grass crops (e.g., maize, rice, etc.), tend to be vigorously fertilized. The cultivar's ecotype, temperature, moisture, harvest frequency and N-fertilizer application are important factors that affect switchgrass yield (Heaton & Long, 2004, Thomason *et al.*, 2005, Wullschleger *et al.*, 2010). It is suggested that N-fertilizer be applied at 50 kg ha^{-1} for the first year of switchgrass plantation and then around 90 kg ha^{-1} for the following years (Wolf & Fiske, 1995). The N uptake, accumulation, and switchgrass yield increases with fertilizer application (Madakadze *et al.*, 1999, Staley *et al.*, 1991, Thomason *et al.*, 2005). Overall, however, very little N from fertilizer is used (~15-40%) and decreases with an increase in pre-existing soil N (Staley *et al.*, 1991, Stout & Jung, 1995). Moreover, ANF may be suppressed in already N-fertile soils. These data indicate that excessive application of fertilizer, especially on high fertility land, would be costly and less productive. Hence, a more feedback-based strategy is needed, which would increase the amount of fixed nitrogen as per the requirement of switchgrass.

Despite indications in the literature that the rates of nitrogen fixation differ among plant species, ecotypes, and varieties, there is a lack of understanding about the nitrogen-fixing bacteria associated with switchgrass and other grasses in general. Previous works in cereals and biofuel grasses have studied the rhizosphere bacteria (Chaudhary *et al.*, 2012, Hargreaves *et al.*, 2015, Jesus *et al.*, 2010, Jesus *et al.*, 2015, Li *et al.*, 2014, Liang *et al.*, 2012, Mao *et al.*, 2014, Mao *et al.*, 2013, Mao *et al.*, 2011, Nautiyal *et al.*, 2013, Turner *et al.*, 2013). Such studies have shown the presence of bacterial phyla, of which some members conduct nitrogen fixation. Few studies, however, attempted to determine the overall functioning and nitrogen fixing potential of the microbial communities in the root-zone of switchgrass, or how these microbial processes may vary between switchgrass cultivars. Overall, our work mainly focuses on sustainable production of switchgrass; however, the knowledge obtained from this project can benefit production of other native bioenergy crops and grasses.

Choice of cultivars

Alamo and 'Dacotah,' as shown in Figure 1, are tetraploid cultivars of switchgrass, adapted to the southern and northern U.S. ecotypes, respectively. The cultivars differ in biomass yield, growth rate, disease susceptibility, and drought tolerance (Table 1). *We hypothesize that Alamo, the high biomass yielding and taller growing cultivar, has more nitrogen demand, and that these plant traits support bacterial communities with high rates of nitrogen fixation in the root-zone compared to Dacotah.* Using these contrasting cultivars will be useful for determining the bacterial communities associated with high and low productivity systems that are expected to support different types and amounts of diazotrophs.

Figure 1: Alamo and Dacotah. Images courtesy of Dr. Zhao.

Table 1: Phenotypic characteristics of Alamo and Dacotah

Alamo	Dacotah
Lowland ecotype	Upland ecotype
High productivity	Low productivity
Faster, taller growth	Slower, shorter growth
Drought tolerant	Less drought tolerant
Disease resistant	Disease susceptible
Cold susceptible	Cold tolerant
Thicker stem	Thinner stem
Flowers later	Flowers early

Identifying core microbes in Switchgrass in natural and managed ecosystems

It is important to identify the set of bacterial communities that are present in root-zones of switchgrass obtained form different locations and management patterns. There is a tradeoff in the ability to perform an experiment and the accuracy of the greenhouse setting to its natural field environment. Greenhouse studies offer the ability to remove external noise and focus on the factor of interest. However, the conditions in the controlled greenhouse differ as compared to the varying field environment. Many a times this results in differences in observations made by the greenhouse and field studies. Lab-based studies are often unable to replicate or extrapolate their results on a field scale and hence, there is a need to pair field and greenhouse studies (Jacobs & Latimer, 2012). This dissertation tries to address potential disparities by using field locations or field-soil wherever possible. This offers a new direction to the study of switchgrass growth-promoting bacteria by identifying universally present, switchgrass-associated bacteria. Ability to cultivate these bacteria, modifying these bacteria or their symbiotic bacteria can allow natural bacterial interactions for sustainable switchgrass production.

Broad overview of methods

Bioinformatics analyses of the root-zone microbial sequencing data are a major component of this dissertation. The following section provides a general overview of the workflows used in the following chapters (Figure 2). The exact details of the methods are provided in the actual chapters. Briefly, the root-zone microbial DNA is extracted, amplicons of interest (regions of the ribosomal DNA) are amplified and then sequenced. After removing the barcodes and trimming to obtain high quality sequences, the reads are clustered into Operational Taxonomic Units (OTUs) and the microbes are identified. The abundance of the microbes is used to perform alpha and beta diversity studies. Multivariate data analyses

methods are used to identify factors of interest (e.g. location, age, cultivar, etc.) or their interactions that affect the microbial composition and abundance. Functional and network analyses of microbes are used to identify the pathways and interactions that help to explain changes in the root-zone.

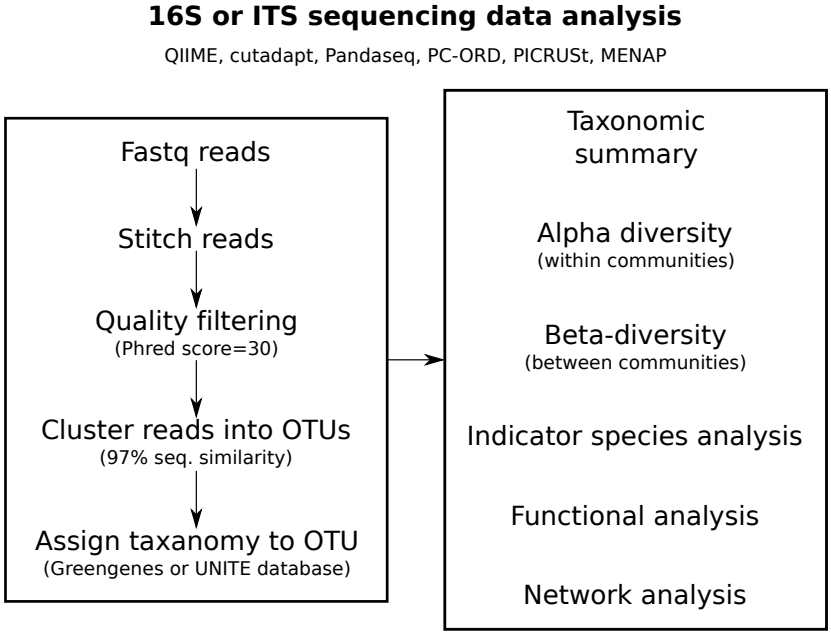


Figure 2: Analyses of amplicon sequencing data

Project Overview

The following dissertation will investigate the microbial communities associated with the root-zones of the natural and managed plants using a multifarious collection of methodologies with the eventual long-term use in sustainable production of native plants.

Chapter 2 investigates the hypothesis that there are associations between the microbial structure and function with the invasive ability of plants in natural ecosystem. This chapter performs a field study to identify the differences in the root-zone microbiome of invasive and native plants "in the real world" and unravels the effects of location and invasion status (or invasiveness). This chapter also provides a glimpse into the natural variation in root-zone microbial communities across plant species.

Chapter 3 studies the rhizosphere-associated microbes in a native, bioenergy-crop, switchgrass (*Panicum virgatum* L.), a prime example of managed ecosystem. The two switchgrass ecotypes, Alamo and Dacotah have different productivity and we hypothesize that the cultivars have different nitrogen demands and can be used to study the diazotrophic communities associated with cultivars of contrasting nitrogen requirements. The chapter uses field soil with the partially controlled conditions of the greenhouse to study the effects of plant growth stage and cultivar on the rhizosphere microbial communities in switchgrass. The research methods are also applicable and beneficial to other switchgrass cultivars and grasses.

Chapter 4 seeks to determine whether there is a core set of microbial communities in switchgrass that can be identified across natural and managed plant systems. We use two datasets with differences in switchgrass production (e.g., location, plant-age, etc.) and search for taxa that are always present in the root-zones of switchgrass. Identifying such communities provides new avenues for research to utilize the microbes universally associated with switchgrass and manipulate the interactions with known partners to move towards a community level growth promoting microbial inoculants.

Chapter 5 offers conclusions based on the previous work and offers directions where this work could be extended to provide deeper understanding of microbial communities that can aid in the sustainable production of switchgrass and other agriculturally important grasses. This work also helps in determining best management practices to support the long-term function and sustainability of native, remnant, and restored ecosystems.

CHAPTER 2

Plant Invasions Associated with Change in Root-Zone Microbial Community Structure and Diversity

Richard R. Rodrigues[1], Rosana P. Pineda[2], Jacob N. Barney[3], Erik T. Nilsen[4], John E. Barrett[4], Mark A. Williams[1,2*]

[1]Interdisciplinary Ph.D. Program in Genetics, Bioinformatics, and Computational Biology, Virginia Tech, Blacksburg, Virginia, United States of America.

[2]Department of Horticulture, Virginia Tech, Blacksburg, Virginia, United States of America.

[3]Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Blacksburg, Virginia, United States of America.

[4]Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, United States of America.

* Corresponding author

E-mail: markwill@vt.edu (MW)

This chapter has been published in an open-access journal, PLOS ONE. DOI: 10.1371/journal.pone.0141424

Abstract

The importance of plant-microbe associations for the invasion of plant species have not been often tested under field conditions. The research sought to determine patterns of change in microbial communities associated with the establishment of invasive plants with different taxonomic and phenetic traits. Three independent locations in Virginia, USA were selected. One site was invaded by a grass (*Microstegium vimineum*), another by a shrub (*Rhamnus davurica*), and the third by a tree (*Ailanthus altissima*). The native vegetation from these sites was used as reference. 16S rRNA and ITS regions were sequenced to study root-zone bacterial and fungal communities, respectively, in invaded and non-invaded samples and analyzed using Quantitative Insights Into Microbial Ecology (QIIME). Though root-zone microbial community structure initially differed across locations, plant invasion shifted communities in similar ways. Indicator species analysis revealed that Operational Taxonomic Units (OTUs) closely related to *Proteobacteria, Acidobacteria, Actinobacteria*, and *Ascomycota* increased in abundance due to plant invasions. The Hyphomonadaceae family in the Rhodobacterales order and ammonia-oxidizing *Nitrospirae* phylum showed greater relative abundance in the invaded root-zone soils. Hyphomicrobiaceae, another bacterial family within the phyla *Proteobacteria* increased as a result of plant invasion, but the effect associated most strongly with root-zones of *M. vimineum* and *R. davurica*. Functional analysis using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) showed bacteria responsible for nitrogen cycling in soil increased in relative abundance in association with plant invasion. In agreement with phylogenetic and functional analyses, greater turnover of ammonium and nitrate was associated with plant invasion. Overall, bacterial and fungal communities changed congruently

across plant invaders, and support the hypothesis that nitrogen cycling bacteria and functions are important factors in plant invasions. Whether the changes in microbial communities are driven by direct plant microbial interactions or a result of plant-driven changes in soil properties remains to be determined.

Keywords: plant invasion, microbial community, root-zone soil, function, N turnover, 16S rRNA, ITS, Microstegium, Rhamnus, Ailanthus

Introduction

Invasive plants are implicated in altering plant community dynamics, disturbance regimes, net primary productivity, and nutrient cycles [1-3], which threaten ecosystem functioning and stability. The soil microbial community plays a central role in ecosystem functioning, including serving as plant symbionts, mediating plant nutrient acquisition, nutrient cycles, and soil formation [4]. These belowground communities have been implicated in invasive species success, but only a few studies have assessed how belowground microbial taxa change with plant invasions into ecosystems [5].

Important feedbacks between plants and the soil biotic community have begun to shed new light on plant rarity and invasiveness. High density of native species, such as *Rhododendron maximum,* reduced soil nutrient availability and mycorrhizae abundance associated with surrounding plants [6-9]. *Alliaria petiolata* in contrast, an invasive plant, reduced arbuscular mycorrhizal fungi (AMF) colonization of native trees and overall tree growth [10]. It was thought that the reduction in AMF occurred as a result of the plant releasing glucosinolate containing root exudates [5]. Relatively uncommon native plants were also shown to be more negatively affected by

pathogens while invaders, in contrast, showed evidence of more positive plant-microbial feedbacks [11, 12]. These results have been further corroborated using reciprocal transplant studies of plant-soil-microbial feedbacks associated with invaded and native ranges of *Triadica sebifera* [13] and *Pinus contorta* [14]. Still, other effects related to soil nutrient cycling indicated that a mixture of the exotic grasses *Avena barbata* and *Bromus hordeaceous* had elevated levels of nitrate, ammonia oxidizers, microbial N, and gross nitrification rates compared to the native grass *Nasella sp*. [15]. Overall, these results show that microbial communities and their processes are altered due to the invasion of exotic plants, and provide evidence that invader and plants native to an ecosystem have underlying differences in their interactions with belowground microbial communities. Meta-analysis have concluded, specifically, that nitrogen turnover is greatly altered and often greater following exotic plant invasion of ecosystems dominated by native plants [16, 17].

Most of the microbial studies conducted have either been based on greenhouse plantings or field establishment of plants rather than observing changes that occur due to natural invasion in the landscape. There are also few studies that have measured microbial communities in the root-zones of native and invaded soil-ecosystems to determine the structure and composition of microbial communities and whether these field observations corroborate the multitude of different litter-based and experimental observations [18]. A recent meta-analysis suggested the importance of invader-ecosystem interactions and the lack of studies across taxonomic groups and habitats [19]. Meta-analyses help to unify ideas and hypotheses, but can mask the relationship between invasive plant species and their influence on soil nutrient pools and microbial dynamics, which are thought to be quite species specific [20, 21]. Studies that are inclusive of multiple invasive plants and their effects on root-zone microbial

community structure and function can thus help to inform whether belowground changes are specific, or broadly associated with plant invasion.

Our overall objective was to understand the effects of plant invasions on soil microbial community structure and its potential linkages to plant-ecosystem function. Specifically, we had two main questions: (1) Do invading species with different taxonomy and phenetic traits have similar or unique effects on microbial communities in root-zone soils?; and (2) are changes in root-zone communities consistent with changes associated with microbial function and soil processes?

Materials and Methods

Species and site descriptions

Study sites were selected that met the following criteria: (1) each site must have invaded and non-invaded (reference) areas, the latter of which represents the site pre-invasion; and (2) one invasive species dominates its strata in the invaded plot—no more than 10% cover of other invasive species are located in the invaded plot. Based on these criteria, three sites were selected in the Ridge and Valley Province of the central Appalachian Mountains in Virginia, USA (Table 1). One site (M) was invaded by a C_4 subcanopy grass (*Microstegium vimineum* [Trin] A. Camus; Japanese Stiltgrass) (Mv), another (R) was invaded by a shrub (*Rhamnus davurica* ssp *davurica* Pall.; Dahurian Buckthorn) (Rd), and the third (A) was invaded by a tree (*Ailanthus altissima* (Mill.) Swingle.; Tree of Heaven) (Aa). All three populations were chosen at locations where a nearby non-invaded reference site was available that was similar in plant community composition, slope, and aspect as the invasion. The native vegetation from these non-invaded sites was used as reference (MvR, AaR, RdR). In all cases it was concluded that the reference site was capable of being invaded, and

did not have overarching preexisting difference from the invaded site (Table 1). The term “invasion” is used to differentiate between invaded and non-invaded effects. Two sites were in use for another research grant funded by the USDA Joint Venture program (11-1480-01, 2011-2015). David Carr at the Blandy Experimental Farm provided permission to sample soils in the *Rhamnus* and reference sites. William McShea provided permission to sample soils at the Smithsonian Conservation Biology Institute forest site in *Ailanthus* and reference locations. We obtained permission from Eastern Divide District to sample soils at the Jefferson National Forest site in *Microstegium* and reference locations. The lands were public and no protected species were sampled.

Microstegium vimineum is a shade-tolerant C4 annual grass common to much of the Eastern US where it has been implicated in reducing tree recruitment (e.g., [20]), decreasing microarthropod diversity [24], and changing soil chemistry and soil microbial communities [25]. This *M. vimineum* invasion is located near an old homestead upslope from the site, but the exact date of establishment is unknown. The reference site was selected across an ephemeral stream likely acting as a barrier to dispersal to the *M. vimineum* population.

Table 1: Details of Sampling Locations.

Location	Invasive Species	Soil Type	Native Species
A: Smithsonian Conservation Biology Institute, Front Royal at an elevation of 378m. Latitude = 38.88553N Longitude = -78.13844W	*Ailanthus altissima* (**Aa**)	Montalto loam. *Taxonomic class:* Fine, mixed, semiactive, mesic Ultic Hapludalfs.	**AaR**: Red oak species (*Quercus species*), tulip poplar (*Liriodendron tulipifera*), and common hackberry (*Celtis occidentalis*). The understory had an abundance of spice bush (*Lindera benzoin*) and infrequent dunal pawpaw (*Asimina triloba*) and bush honeysuckle *(Lonicera maackii*).
M: Jefferson National Forest, Montgomery County at an elevation of 2280m. Latitude = 37.28108N Longitude = -80.47523W	*Microstegium vimineum* (**Mv**)	Berks-Weikert composition on slopes from 15 to 25 percent [22]. *Taxonomic class:* Loamy-skeletal, mixed, active, mesic Typic Dystrudepts.	**MvR**: The forest canopy is primarily red maple (*Acer rubrum*), white oak (*Quercus alba*), and red oak (*Quercus rubra*). The understory community composition is typical of Appalachian forests of Virginia with total site richness of 78 species [23].
R: Blandy Experimental Farm, Boyce at an elevation of 183m. Latitude = 39.05923N Longitude = -78.05428W	*Rhamnus davurica* (**Rd**)	Timberville silt loam. *Taxonomic class:* Fine, mixed, active, mesic Typic Hapludults Poplimento-Rock outcrop complex. *Taxonomic class:* Fine, mixed, subactive, mesic Ultic Hapludalfs.	**RdR**: Perennial grasses (e.g., *Panicum virgatum*) and infrequent annual and perennial herbaceous weeds

The following experimental groups were studied: (i) location (A, M, and R); (ii) invasion status (Invasive plants (I) and Native plants (N)); and (iii) interaction of location and invasion status (Aa, AaR, Mv, MvR, Rd, and RdR).

Rhamnus davurica ssp. Davurica is a deciduous short-lived shrub native to China, North Korea, Mongolia, eastern Siberia and Japan. It was commonly planted in the Northwestern US plains for windbreaks in the 1930's. Both *R. davurica* and *Rhamnus cathartica* L. (Common Buckthorn) were incorporated into the Virginia Arboretum in 1939, but only *R. davurica* has invaded into the Blandy Experimental Farm in Boyce, Virginia, USA. The site invaded by *R. davurica* has been unmanaged for over 3 decades and has not for the Blandy Experimental Farm. The *R. davurica* invasion into the grassland is well documented at this farm and has occurred over a 25-year period.

Ailanthus altissima is a common urban, roadside, and natural area invasive tree capable of growing in a variety of non-managed and disturbed systems worldwide; spreading both sexually and clonally [20, 25, 26]. This fast growing tree has putative allelopathic effects [23], though the ecological impacts of *A. altissima* are largely unknown [27]. The *A. altissima* invasion occurred at this site over the last 40 years following a clear cut on one side of a logging road. The other side of the road was not logged and is a non-invaded reference area. While logging removed overstory vegetation, the impacts on soil were relatively small.

Soil sampling and analyses

Soil sampling locations were selected by a stratified random technique. A 50 m transect was established along one edge of each plot (same for both invaded and non-invaded plots). The transect was divided into five replicate 10 m reaches. A random number generator was used to pick a meter mark within each 10 m reach for establishing a perpendicular transect. Once the position of the transect was identified, the random number generator was used to select a distance along the perpendicular transect for the soil sample. At this location, a coin was flipped to choose the right or left side of the perpendicular transect to sample. The soil sample was taken 1 m away from the perpendicular transect. If

the final location was occupied by a rock or tree, the closest location where a soil sample could be taken was used. Soils were sampled at each location using a standard 7-cm soil corer (Model # 402.25, AMS Inc., American Falls, ID, USA).

At each sample location, the litter and humus layers were removed. The soil corer was washed with 95% alcohol before sampling and between each soil sample. The soil sampler was then driven in to a depth of 10 cm using a professional slide hammer (Model 57780, AMS Inc., American Falls, ID, USA). Leaf litter, roots, and large debris were removed from each sample (100 cm^3) and the soil samples were placed in a sterile zip-top bag and refrigerated in a cooler until the samples could be stored at -5°C in the lab at Virginia Tech. This resulted in ten randomly selected soil samples at each site, five of which were from the invaded and five from the adjacent non-invaded reference. Each soil sample was sieved through an alcohol washed #20 soil sieve (Model H-3903, M & L Testing equipment, Calgary, Alberta, Canada), and individually mixed and homogenized. All precautions against contamination were taken. Subsamples of the sieved soil were analyzed for several nutrient cations and anions, extractable nitrogen, and microbial diversity. The subsamples for nutrient cation analysis were extracted with 1*M* KCL and analyzed using ICP. Soil parameters measured were: pH, cation exchange capacity, and concentrations of P, K, Ca, Mg, Zn, Mg, Cu, Fe, and B.

A separate subsample was incubated for seven days at field moisture water potential. Directly before and following the seven days of incubation, samples were extracted with 1M KCl to determine extractable inorganic nitrogen content. Based on water content and particle size analysis, it was estimated that water potential for all soil samples ranged between -100 to -500 KPa. Sampling in May ensured that each sample was near saturation and similarly moist. Total nitrate and ammonium ions ($\mu g\ g^{-1}$) were measured with a Lachet autoanalyzer (Quikchem 8500 Series 2) and turnover $(T_1\text{-}T_0)$ x $(100/\ T_0)$ was calculated following a one

week incubation of soil (25°C). Wilcoxon (rank-sums) test with a normal approximation to the two-sample test was performed in JMP® Pro, Version 11 (SAS Institute Inc., Cary, NC, 1989-2007) to check whether the turnover was different between invaded and non-invaded samples. Microbial community structure and diversity were determined on another subsample of soil DNA (see below).

Univariate statistical analysis on soil nutrients

A two-way analysis of variance was used to determine significant effects of location, invasion status (invaded or non-invaded), and their interaction on soil nutrition. Means were separated using Tukey HSD at alpha=0.05. All ANOVAs were performed with JMP statistical software (SAS Institute Inc., Cary, North Carolina).

DNA extraction and amplification

For both the 16S rRNA gene analyses and the ITS analyses, 0.5 g of freeze-dried homogenized soil was weighed and DNA was extracted from each soil sample using PowerSoil® DNA Isolation Kit (MoBio) according to the manufacture's protocol. DNA quality was checked on a 0.8% (w/v) agarose gel. DNA concentrations were determined by fluorometric quantification using the Qubit® 2.0 platform with Qubit dsDNA HS Assay Kit (Life Technologies). DNA was diluted to 50 ng μL^{-1} and stored in a -20°C freezer. It was used for the PCR-based protocol described in [28], using the PCR bacteria/archaeal primers 515F/806R targeting the V4 region of the 16S rRNA. ITS1FI2/ ITS2R were used to amplify the spacer ITS1 of the internal transcribed spacer (ITS) rDNA region [29, 30]. The reverse amplification primer also contained a twelve base barcode sequence. Both PCR primers contain sequencer adapter regions. The enzyme used in the PCR reaction was KAPA2G Robust (5 U/μL) from Kapa Biosystem. For 16S rRNA assay the 25 μL reaction mixture

contained 0.5 μL of dNTPs (10 mM), 0.5 μL of each primer (10 μM), 50 ng of the DNA template, 1 μL of DMSO (100%), 0.2 μL of the enzyme (5U/μL) and 5 μL of Buffer GC (Kapa Biosystem). For the ITS assay, the PCR reaction final volume was 25 μL, containing 0.5 μL of dNTPs (10 mM), 0.625 μL of each primer (10uM), 50 ng of the DNA template, 1.25 μL of DMSO (100%), 0.2 μL of the enzyme (5 U/μL) and 5 μL of Buffer A (Kapa Biosystem). The PCR conditions used were as follows: for the 16S assay, there was a denaturation step at 94°C for 3 minutes, 35 cycles of 94°C for 45 seconds, an annealing step at 60°C for 60 seconds, an extension step at 72°C for 90 seconds, and a final extension at 72°C for 10 minutes. For the ITS assay, there was a denaturation step at 95°C for 15 seconds, 35 cycles of 95°C for 30 seconds, an annealing step at 55°C for 30 seconds, an extension step at 72°C for 30 seconds, and a final extension at 72°C for 5 minutes. The specificity of the PCR was further evaluated by running on a 1.2% (w/v) agarose gel. The concentration of the DNA was obtained by Fluorometric Quantitation (Qubit® 2.0 Life Technologies) before sending samples to sequencing. From the bacterial experiments, two out of the 30 samples did not show 16S rRNA gene amplification. Hence, 28 samples were sent for 16S rRNA gene sequencing, whereas all 30 samples were sent for ITS rDNA sequencing. Sequencing on the Illumina MiSeq platform was conducted by the Virginia Bioinformatics Institute core facility.

Sequence data analyses

In the bacterial data, an 'Rd' sample (F8) was removed from further analysis due to contamination on the sequencing plate. The paired end reads were stitched using *Pandaseq* [31]. For the fungal data, only read-2s with a quality threshold of 30 were used for further analyses. The bacterial and fungal sequencing data were analyzed using *QIIME* [32]. Briefly, reads were clustered into OTUs based on 97% sequence similarity using *uclust* [33] and *usearch61* [33], for bacteria and fungi respectively, using an open reference OTU-picking

strategy. The representative sequence of an OTU was used to assign it a taxonomy, using *uclust* against the Greengenes reference database version 13_8 [34, 35] for bacteria, and *RDP classifier* [36] against the UNITE reference database version 12_11 [37] for fungi.

Comparison and statistics on groups

A sampling depth threshold of 80,000 and 3,200 sequences per sample, for bacteria and fungi respectively, was used for the diversity and taxonomic summary analyses. The beta diversity was calculated using weighted and unweighted Unifrac [38] (for bacteria), and Bray-Curtis [39](for fungi) distance metrics. To identify group differences, the distances were used for Principle Coordinate Analysis [40] and visualized in 3D-plots using EMPeror [41]. The chao1 [42] and observed species metrics were used to plot alpha rarefaction curves. The alpha diversity was calculated using PD whole tree (for bacteria only), chao1, observed species, and Shannon and Simpson indices for bacteria and fungi. The bar graphs with standard error bars were used to visualize microbial taxonomic summaries of the interaction between location and invasion at different levels and generated using custom python scripts. Multivariate data analysis methods of adonis [43] and Analysis of Similarity (ANOSIM) [44] were used to identify whether groups were significantly different. Indicator species analysis (ISA) [45] in PC-ORD Version 6 [46] was used to identify taxa that were significantly (indicator value > 70 and p-value < 0.01) associated with invasion when blocked by geographic sites/location. A seed of 16 and 18 with 5000 runs was used for the bacteria and fungi, respectively.

Functional analyses

The actual abundance (counts) of the OTUs belonging to the significant genera from ISA was used for functional analyses using *PICRUSt* [47]. OTUs not part of the closed

reference OTU picking were filtered out. Using default parameters, the filtered OTU table was normalized by the 16S rRNA copy number abundance to identify true abundance followed by metagenome functional prediction for each sample. The metagenomes were collapsed into KEGG pathways. Using STAMP [48], two-sided Welch's t-test [49] with Benjamini-Hochberg [50] and Storey [51] multiple testing corrections were performed to identify KEGG pathways that were significantly different (q-value < 0.05) between invaded and non-invaded samples.

Results

Soil nutrients change associated with invasion

Many soil parameters, particularly pH, P, K, Mg, Zn, and B varied among locations (Table 2). Four soil parameters varied between invaded and non-invaded plots across locations (Table 2). Interestingly, 7 of the 11 soil parameters varied between invaded and non-invaded plots among species, including pH, P, and CEC (Table 2).

In most cases, nutrient parameters were higher in the invaded patch compared to the non-invaded patch (Table 2). For example, *Microstegium vimineum* increased pH, K, and Ca, *Rhamnus davurica* increased K and Mn, while *Ailanthus altissima* lowered pH, Ca, Mn, Fe, and B (Table 2).

Concentrations of nitrate in soil ranged from 1.5 to 18.3 and ammonium from 9 to 29 mo g^{-1} soil. Following one week of incubation (22^{o}C), the concentrations increased, on average, ranging from non-detectable to 24 for nitrate and 33 to 51 mo g^{-1} soil for ammonium. Wilcoxon (rank-sums) test with a normal approximation to the two-sample test showed that turnover of nitrate during the one week incubation was observed to be significantly greater in association with invasion (p-value = 0.014), averaging 137 and 61 percent per week of incubation in invasive and non-invasive factors, respectively (Table 3)

On the other hand, turnover of ammonium during the one week incubation was observed to be greater, but not significant, in association with invasion, averaging 154 and 123 percent per week of incubation in invasive and non-invasive factors, respectively. These results suggest that invasion increased the rate of N cycling and availability of nitrogen for plant uptake from soil. The results also agree with the phylogenetic and functional analyses, which showed greater N cycling genes, and greater relative abundance of nitrifying and putative nitrogen-fixing bacteria in the invasive compared to non-invasive soil.

Table 2: Mean Values (St. Dev.) and Two-Way Analysis of Variance on Soil Nutrition Parameters from Three Sites in Central Appalachian Mountains with Invaded (I) and Non-Invaded (NI) Locations.

Location	**M**		**R**		**A**		**Location**	**Invasion Status**	**Location x invasion status**
Invader	***Microstegium vimineum***		***Rhamnus davurica***		***Ailanthus altissima***				
Invasion status	**I**	**NI**	**I**	**NI**	**I**	**NI**			
pH	5.36**	4.9	6.69*	6.66	6.29*	6.67	**<0.001**	0.608	**0.001**
	(0.27)	(0.15)	(0.2)	(0.12)	(0.12)	(0.31)			
P	2.4	2.2	11.8	4.4	2.6	2.0	**<0.001**	**0.015**	**0.016**
	(0.5)	(0.5)	(6.8)	(1.1)	(0.8)	(0.00)			
K	106.8**	52.4	104.0*	72.4	150.4	126.4	**0.002**	**0.002**	0.442
	(28.3)	(5.9)	(23.8)	(18.5)	(35.8)	(37.6)			
Ca	553.2**	156.2	1151.6	1123.6	1174.0*	1634.8	**<0.001**	0.872	**0.000**
	(208.1)	(26.3)	(175.9)	(117.4)	(285.4)	(265.6)			
Mg	65.0	32.8	97.0	88.4	164.0	208.4	**<0.001**	0.887	**0.003**
	(11.2)	(2.6)	(7.3)	(9.2)	(46.8)	(25.7)			
Zn	2.18	1.88	1.36	1.28	4.72	5.12	**<0.001**	0.980	0.538
	(0.37)	(0.29)	(0.31)	(0.25)	(1.18)	(1.11)			
Mn	12.62***	15.4	11.44***	7.48	14.96**	32.3	0.478	**0.004**	**<0.001**
	(1.12)	(7.61)	(5.7)	(0.64)	(2.58)	(7.74)			
Cu	1.4	1.48	0.60	0.82	1.24	0.78	**0.003**	0.677	0.091
	(0.22)	(0.50)	(0.23)	(0.18)	(0.55)	(0.19)			
Fe	18.5	22.4	16.4	18.5	3.98*	2.46	**0.027**	0.118	0.185
	(4.93)	(5.37)	(18.1)	(4.93)	(1.08)	(0.67)			
B	0.3**	0.2	0.5	0.5	0.8**	1.4	**0.000**	**0.017**	**<0.001**
	(0.1)	(0.0)	(0.1)	(0.1)	(0.2)	(0.3)			
CEC	6.4	6.1	6.9	6.5	8.14	10.26	0.052	0.222	**0.034**
	(0.42)	(1.16)	(0.86)	(0.62)	(1.60)	(1.50)			

Parameter = soil nutrition trait; Location = the three locations where each species was sampled; Invasion Status = invaded (I) and non-invaded (NI) plots. Bolded values indicate significant ($p \leq 0.05$) effects. All nutrient units are µg element g^{-1} soil. The statistical test (Tukey HSD means separation) is between invaded and non-invaded within site. * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$.

Table 3: Turnover (Percentage) of Inorganic Nitrogen (Mean, SE[a]) in Non-Invaded and Invaded Locations at Three Sites in Central Appalachian Mountains.

Location	M		R		A			
Invader	*Microstegium vimineum*		*Rhamnus davurica*		*Ailanthus altissima*		*All plant species*	
Invasion status	Invaded	Non-invaded	Invaded	Non-invaded	Invaded	Non-invaded	Invaded	Non-invaded
NO_3	42 (8)	-20 (20)	236 (106)	196 (70)	108 (24)	33 (5)	137 (45)	61 (31)
NH_4	247 (41)	347 (45)	6 (19)	-61 (20)	209 (48)	83 (17)	154 (35)	123 (48)

[a] The standard error (SE) of the mean is in given in parenthesis.

Alpha diversity of microbial communities associated with invasion

(a) Bacteria

A total of ~17.8 million high quality 16S rRNA gene sequence reads were obtained from the invaded and non-invaded plots. The sequences from 27 samples possessed a 254-bp average length and will be submitted to the NCBI Sequence Read Archive according to MIMS standard. There were a total of 210,007 distinct OTUs (observations) across samples with a total of 4,444,765 sequences (counts) that were assigned to these OTUs. The observation refers to the number of distinct OTUs; whereas the count refers to the abundance of bacteria belonging to these OTUs in samples. The mean and median counts per sample were 164,621 and 158,958, respectively. A sampling depth threshold of 80,000 counts per sample removed one sample from further analyses. The average Good's coverage for the bacterial data across 26 samples was 96.1%.

Chao1, observed species, Shannon, Simpson, and PD whole tree metrics were used to calculate alpha diversity (species diversity within the community). A non-parametric test with the default 999 Monte Carlo permutations with an FDR correction showed significant differences ($\alpha<0.05$) between locations and between location x invasion for alpha diversity but not between invaded and non-invaded samples (Shannon and Simpson metrics were not used) (Data not shown). However, the rarefaction curves, which are sample size independent,

showed trends that non-invaded samples have lower alpha diversity (S1 Fig). Without the sampling depth threshold on the 26 samples, a one-tail Mann-Whitney test showed that the alpha diversity of invasive samples was significantly greater ($\alpha<0.05$) than that in non-invaded samples for all five diversity metrics (Table 4). Kruskal Wallis test with a Chi-Square approximation of one-way test in JMP® Pro, Version 11 (SAS Institute Inc., Cary, NC, 1989-2007) suggested that the diversity metrics (except Simpson index) were significantly different ($\alpha<0.05$) between samples as per location and interaction of location and invasion status. Since the sample size variation can affect the diversity metrics, the sampling depth threshold was utilized for further analyses by taking a random subsample of 80,000.

Table 4: Alpha Diversity Metrics for Invasion, Location, and Location x Invasion in Bacteria.

	Chao1	Observed Species	Shannon	Simpson	PD Whole Tree
Invasion status					
I (n=11)	24,563	15,024	10.83	0.998	604
N (n=15)	20,566	12,328	10.54	0.997	512
p-value (one-tail)	**0.012**	**0.004**	**0.006**	**0.007**	**0.007**
Locations					
A	25,687	15,326	10.82	0.997	616
M	17,512	11,000	10.43	0.998	448
R	23,460	13,987	10.75	0.997	591
p-value (two-tail)	**0.001**	**0.002**	**0.003**	0.817	**0.002**
Location x Invasion status					
Aa	27,684	16,806	10.98	0.998	666
AaR	23,691	13,845	10.65	0.997	566
Mv	19,507	12,108	10.61	0.998	494
MvR	15,915	10,114	10.28	0.997	410
Rd	26,875	16,398	10.87	0.998	668
RdR	22,093	13,023	10.70	0.997	559
p-value (two-tail)	**0.003**	**0.002**	**0.002**	0.103	**0.002**

Bolded values indicate significant ($\alpha<0.05$) effects.

(b) Fungi

The read 1s were not used for the analysis due to the poor quality of sequences. A total of 204,835 high quality read 2s of the ITS gene sequence were obtained from the invaded and non-invaded plots. The sequences from 30 samples possessed a 230-bp average length and will be submitted to the NCBI Sequence Read Archive according to the MIMS standard. There were a total of 4,419 distinct OTUs (observations) across samples with a total of 182,009 sequences (counts) that were assigned to these OTUs. The mean and median counts per sample were 6,067 and 4,927 respectively. A sampling depth threshold of 3,200 counts per sample did not remove any sample from further analyses. The average Good's coverage for the fungal data across 30 samples was 95.5%.

Chao1, observed species, Shannon, and Simpson metrics were used to calculate alpha diversity. A non-parametric test with the default 999 Monte Carlo permutations with FDR correction showed significant differences ($\alpha<0.05$) between locations, invasion status, and their interaction (location x invasion status) for alpha diversity (Shannon and Simpson metrics were not used) (Data not shown). Similarly to the bacterial data, the rarefaction curves showed trends that non-invaded samples have lower alpha diversity (S2 Fig). Without the sampling depth threshold, a one-tail Mann-Whitney test showed that the alpha diversity of invasive samples is significantly higher ($\alpha<0.05$) than that in non-invaded samples for chao1 and observed species metrics (Table 5). Kruskal Wallis test with a Chi-Square approximation of one-way test in JMP® Pro, Version 11 (SAS Institute Inc., Cary, NC, 1989-2007) suggested that the diversity metrics were significantly different ($\alpha<0.05$) between samples as per locations and interaction of locations and invasion status. Since the sample size variation can affect the diversity metrics, the sampling depth threshold was utilized for further analyses by taking a random subsample of 3,200.

Table 5: Alpha Diversity Metrics for Invasion, Location, and Location x Invasion in Fungi.

	Chao1	Observed species	Shannon	Simpson
Invasion status				
I (n=15)	814	537	6.30	0.947
N (n=15)	728	483	5.87	0.935
p-value (one-tail)	**0.039**	**0.023**	0.076	0.221
Locations				
A	863	600	6.60	0.962
M	800	512	6.29	0.959
R	650	420	5.36	0.902
p-value (two-tail)	**0.015**	**0.022**	**0.006**	**0.006**
Location x invasion status				
Aa	844	570	6.65	0.965
AaR	883	629	6.55	0.959
Mv	935	604	6.56	0.961
MvR	666	421	6.02	0.958
Rd	664	438	5.68	0.916
RdR	636	401	5.03	0.889
p-value (two-tail)	**0.011**	**0.014**	**0.020**	**0.030**

Bolded values indicate significant ($\alpha<0.05$) effects.

Beta diversity of microbial communities associated with invasion

(a) Bacteria

Multivariate data analyses using adonis, ANOSIM, and MRPP on weighted and unweighted Unifrac distances showed significant differences ($\alpha<0.01$) in the beta diversity of the location and the interaction of location and invasion status.

(b) Fungi

The beta diversity of location, invasion status, and their interaction were significantly different ($\alpha<0.01$) as shown by adonis, ANOSIM, and MRPP on Bray-Curtis distances, with an exception of ANOSIM indicating a p-value of 0.014 for invasion.

The PCoA analysis of the weighted and unweighted Unifrac (for bacteria), and Bray-Curtis (for fungi) distances showed that the samples clustered as per the location and invasion

(Fig. 1), with location explaining the maximum variation (PC1). For the unweighted Unifrac and Bray-Curtis distances, invasion status (across all locations) consistently accounted for the second most variation (6% for bacteria and 17% for fungi on PC2). There was a lot of variation associated with the Rd samples as shown in Axis 2 of Fig 1. Overall, these results indicated the effects of invasion and location x invasion status. There were, thus clear patterns of change in soil microbial communities following the invasion of each species across geographically separated ecosystems.

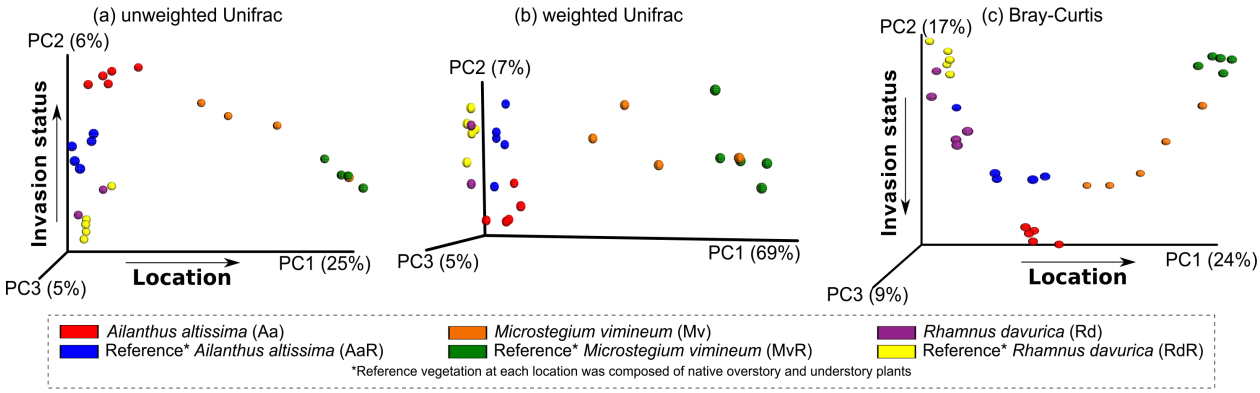


Fig. 1: PCoA plot describing (a) un-weighted and (b) weighted Unifrac for bacteria and (c) Bray-Curtis distances for fungi in the invaded and non-invaded sites. Each circle indicates a sample. Multivariate data analysis methods of adonis and ANOSIM were used to identify whether groups were significantly different.

Taxonomic summary and identification of microbial communities associated with invasion

Taxonomic summaries showed that *Acidobacteria* (~30%) and *Proteobacteria* (~22%), and *Ascomycot*a (~47%) and *Zygomycota* (~13%) were the most dominant phyla of bacteria and fungi, respectively (Fig. 2). A major proportion of taxa could not be assigned (~34%) to known taxa for the fungal data; however, they were a very minor portion for bacteria. The genus level taxonomic summaries were used for indicator species analysis (ISA) to identify taxa that were more abundantly associated with invaded or non-invaded samples (Table 6). Overall, the results suggested numerous types of taxa associated with invasion, whereas only one taxa was associated with non-invasion.

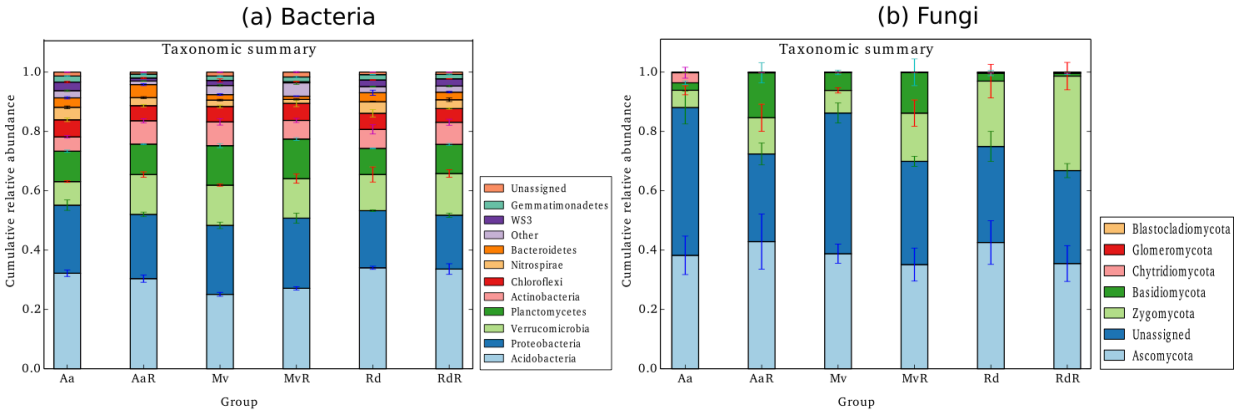


Fig. 2: Taxonomic summary of the relative abundance of (a) bacterial and (b) fungal phyla in the invaded and non-invaded sites. The taxa are arranged as per total relative abundance across all samples, with the most abundant phyla at the bottom and the least abundant phyla at the top of the y-axis. Similarly, the phylum names in the legend are arranged from the least abundant at the top to the most abundant at the bottom.

Table 6: Genera with a Greater Relative Abundance Associated with Invasion and Determined to have a Significant Effect Based on Indicator Species Analysis (IV > 70 and p-value < 0.01).

BACTERIA						
Phylum	**Class**	**Order**	**Family**	**Genus**	**I (%)**	**N (%)**
Acidobacteria	-	-	-	-	0.30	0.17
Acidobacteria	Holophagae	Holophagales	Holophagaceae	Geothrix	0.01	0.00
Acidobacteria	iii1-8	SJA-36	-	-	0.03	0.01
Acidobacteria	RB25	-	-	-	0.25	0.12
Acidobacteria	S035	-	-	-	0.08	0.05
Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter	0.02	0.01
Actinobacteria	Actinobacteria	Actinomycetales	Williamsiaceae	Williamsia	0.01	0.00
Chloroflexi	TK10	-	-	-	0.02	0.01
Gemmatimon-adetes	Gemmatimonadetes	-	-	-	0.03	0.01
Nitrospirae	Nitrospira	Nitrospirales	-	-	0.02	0.00
Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	0.52	0.10
OD1	SM2F11	-	-	-	0.01	0.00
OP3	koll11	-	-	-	0.01	0.00
OP3	PBS-25	-	-	-	0.01	0.00
Proteobacteria	Alphaproteobacteria	Rhodobacterales	Hyphomonadaceae	-	0.21	0.07
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobium	0.06	0.02
Proteobacteria	Betaproteobacteria	Methylophilales	Methylophilaceae	-	0.01	0.00
Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	Dechloromonas	0.03	0.00
Proteobacteria	Deltaproteobacteria	NB1-j	MND4	-	0.17	0.05
Proteobacteria	Deltaproteobacteria	Desulfuromonadales	Geobacteraceae	Geobacter	0.05	0.02
WS2	SHA-109	-	-	-	0.06	0.03
*Actinobacteria	Actinobacteria	Actinomycetales	Thermomonosporaceae	Actinomadura	0.00	0.01

FUNGI						
Phylum	**Class**	**Order**	**Family**	**Genus**	**I (%)**	**N (%)**
Ascomycota	-	-	-	-	1.34	0.34
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	Cladosporium	0.05	0.01
Ascomycota	Leotiomycetes	-	-	-	0.37	0.11
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	-	6.52	2.10
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Cylindrocarpon	0.95	0.45
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	0.83	0.14
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Neonectria	0.15	0.02
Ascomycota	Sordariomycetes	Incertae sedis	Plectosphaerellaceae	Plectosphaerella	0.24	0.03
Ascomycota	Sordariomycetes	Sordariales	-	-	1.07	0.53

The hyphen (-) indicates that no taxonomic information was available for that OTU at that level. The bacterial OTU indicated with asterisk (*) was the only OTU associated with non-invaded samples in the ISA. The last two columns indicate the percentage of relative abundance of taxa in the invaded and non-invaded samples, respectively.

(a) Bacteria

After removing OTUs assigned to archeal and unassigned taxa, OTUs with a total relative abundance of less than 0.1% across all samples were removed. The remaining 416 taxa were re-relativized and used for ISA blocked using soil/geographic locations. Out of 22 OTUs (Table 6) that showed significantly different abundance in invaded and non-invaded samples, 21 OTUs were associated with invasion. OTUs within *Proteobacteria* (6 OTUs), *Acidobacteria* (5 OTUs), and *Actinobacteria* (3 OTUs) had greater sequence abundance due to invasion as revealed by ISA and blocked across soil/geographic locations. Bacterial taxa responsible for nitrogen cycling in soil were increased in abundance in association with plant invasion. Taxa belonging to the ammonia-oxidizer *Nitrospirae* (phylum) and Nitrospira (class) were among the bacteria each with 1.5 times greater abundance in the invaded (3.5% compared to 2.4% in non-invaded) root-zone soils. Nitrifying bacteria appear to be a major result and perhaps driver of invasive plant species change in ecosystems.

The nitrogen-fixing bacterial community was also an important potential indicator of change noted in plant invasions. Several bacterial groups which are known to contain taxa involved in nitrogen-fixation were shown to increase as a result of plant invasion in our data. Rhodobacterales are commonly identified as nitrogen-fixing bacteria [52], and found to collectively contribute to (2.7 times) greater abundance in the invaded root-zone soils in our data (0.22% compared to 0.08% in non-invaded) and previous literature [53]. Hyphomicrobiaceae, another bacterial family within the phyla *Proteobacteria* were also greater as a result of plant invasion, but the effect was most strongly associated with the root-zones of *M. vimineum* (1.4 times abundant, 3.7% compared to 2.6% in non-invaded) and *R. davurica* (1.2 times abundant, 1.7% compared to 1.4% in non-invaded). Though nitrogen-fixation symbiosis are not widely considered key traits among the invasive plant species in this research study, the greater relative abundance of these putative diazotrophic taxa support

the idea that these traits may be important associations for many plant invader types.

(b) Fungi

After removing OTUs assigned to unassigned taxa, OTUs with a total relative abundance of less than 0.1% across all samples were removed. The remaining 226 taxa were re-relativized and used for ISA blocked across soil/geographic locations. All of the 9 OTUs (Table 6) that showed significantly different abundance in invaded and non-invaded samples were associated with invasive samples. OTUs within Ascomycota (9 OTUs) had a greater sequence abundance due to invasion as revealed by ISA blocked across soil/geographic locations. Taxa belonging to the Sordariomycetes were among the fungi with 1.2 times greater abundance in the invaded (21.3% compared to 17.5% in non-invaded) root-zone soils.

Predicting microbial functions in non-invaded and invasive samples

Currently, PICRUSt can only be used for functional analysis of bacterial taxa. To the best of our knowledge, we could not find a program for functional analysis of fungi, analogous to PICRUSt for bacteria. The fungal data resources AFTOL (http://aftol.org/) and FunSecKB [54] provide relevant but incomplete data for our purpose.

The actual counts from the OTU table were obtained for the bacterial species belonging to the genera that were significant from the ISA. OTUs not part of the closed reference OTU picking method were filtered out from the 3,385 OTUs belonging to the 22 significant genera and the remaining 365 OTUs (~11%) were used for functional analyses using PICRUSt. The 16S rRNA copy number normalized abundance was used to predict metagenome and collapse into KEGG pathways. Two-sided Welch's t-test with multiple testing corrections in STAMP was performed to identify KEGG pathways at different levels that are significantly different (q-value < 0.05) between invaded and non-invaded samples. At

Level 2 of KEGG, BH and Storey corrections found 9 and 27 pathways, respectively, to be significantly different between root-zone bacteria of invaded and non-invaded samples (S1 and S2 Tables). At Level 3 of KEGG, BH correction did not detect pathways to be significantly different between invaded and non-invaded root-zone bacteria. However, for the same level, Storey FDR detected 60 pathways to be different (S3 Fig). The significant processes were descending sorted as per the average of mean relative frequency (%) in non-invaded and invaded samples. The top 20 abundant processes were categorized as belonging to non-invaded (N) or invaded (I) samples depending on the difference of mean relative frequency (%) (S3 Table).

As expected from the taxonomies of bacteria from the ISA, nitrogen metabolism was also observed to be higher in the root-zone bacterial communities of invasive plants as compared to that of the non-invaded plants (S3 Table). The increase in nitrogen metabolism by invasive plants and the associated benefits to invasion are well known [55-57].

Discussion

Plant invasion theory has developed a broad number of hypotheses to explain the success of invasive plants [58]. Despite their likely importance, however, there is a dearth of research into aboveground-belowground linkages across landscape scales that have determined the effects of plant invasion on soil or root-zone microbial communities [59, 60]. Here we show that at three independent locations, three invasive plants are associated with uniform shifts in belowground root-zone soil microbial communities. This is important, further, because each of the invasive plants has a distinct phylogeny and life form. Our results are broadly relevant because belowground interactions between soil microbes and plants provide an important linkage to support plant invasions.

Bacterial community shifts due to plant invasion

Compared to adjacent non-invaded patches, fungal and bacterial communities were described by consistent ordinal shifts associated with invasion. Nitrospira sp. and *Nitrospirae* were among the bacteria with greater abundance in the invaded soils. Overall *Nitrospirae* was very abundant, and greater in the invaded (3.5% compared to 2.4% in non-invaded) root-zone soils. Previous studies have shown that *Nitrospirae*, which are most often found to be chemolithic autotrophs, and include taxa that are drivers of nitrification, tend to account for 0.2 to 0.7 % of OTUs in grasslands, agricultural systems, and forests [61, 62]. However, 2% or more have been observed in remnant deciduous forests [63], which is consistent with the forests described herein. Furthermore, our results corroborate that plant invasions are associated with major changes in the nitrogen cycle [18, 60] by showing greater rates of root-zone soil N turnover due to invasion.

Importantly, the results of our experiments support a major mechanism of plant invasion success, and link microbial phylogeny with functional measurements of nitrogen turnover. The greater rates of nitrogen turnover and estimates of metagenome composition and function using PICRUSt are in agreement that N cycling processes are important components of invader success. Nitrogen-fixing bacterial communities are also an important indicator of change previously documented in plant invasions [64]. The link between nitrogen-fixation and bacterial phylogeny, however, is not as strong as that with nitrification. Several bacterial groups which are known to contain taxa well known for nitrogen-fixation were observed to increase in our study as a result of plant invasion. Nitrogen-fixers can be free-living, and their abundance in soil tends to be low ($2.4x10^5$ copies g^{-1}); however, associative diazotrophs are generally more common ($1.3x10^7$ copies g^{-1}) in the root-zones of numerous types of plants if carbon is available to drive the energetically expensive process of N_2 reduction to ammonium [65]. Since these bacteria are closely linked to plant roots, their

greater abundance, and the confirmation that nitrogen fixation genomes are available to support greater nitrogen fixation (PICRUSt) associated with invaded soils, are in support of the argument that the result is not due to *a priori* soil habitat differences, but rather the impact of the root-zones of plant invaders. If greater N-fixation is the result of increasing abundance of diazotrophs, then greater supplies of N could help to foster greater nitrogen availability for plants and nitrifiers alike. These types of interactions have the potential to act as a positive feedback to support the habitat needs of the invader. Negative consequences of increased nitrogen-fixation and nitrification could also come from the leaching of nitrate to groundwater and gaseous losses through denitrification (N_2O).

Connections between plant traits and root-zone associated microbial communities have been considered [19, 60]. Less work, however, has been conducted to determine how root-zone soil microbes directly benefit and support the longer-term spread of invasive plants [66]. Though the work presented here does not directly address the long-term nature of invasion, they are representative of fairly mature invasions (>5y) and the potential consequences of changing microbial communities and alterations in ecosystem nutrient cycles.

The field results presented help to fill a major gap in understanding plant invaders and mechanisms of invasion success. The evidence provided in the research reported here are consistent with the idea that plant invaders shape belowground communities, and positively feedback to support the success of the plant invader. In addition, the research has shown that plant invaders are associated with change in soil properties, which might be driven by the plant invader and facilitated by positive feedbacks resulting from microbial community processes. Alterations in nutrient cycling have previously been described as potential drivers that feedback to support plant invasion. Often these results are tied to changes in plant tissue chemistry and the decomposition [21, 25] but less attention has been paid to the potential

effects that plant roots might have more directly on soil nutrient bioavailability. Plant root systems have the capacity to alter soil pH and therefore chemical equilibria and pH sensitive biological processes. Nitrification, for example, has been described as limited by pH below 5.5- 6.0 [67]. Chemical equilibrium of soil nutrients, such as phosphorus, potassium, and iron, furthermore, are strongly impacted by soil pH. The significant changes in bioavailable soil nutrient pools suggest further attention is needed to understand their role in sustaining plant invasions.

Fungal community shifts due to plant invasion

It was expected that invaded soils would tend to be less diverse and support greater dominance if invasive plants stimulated the activity of specific microbes that feedback to support invader growth. Invasion, however, was associated with greater diversity and richness of fungi (and bacteria). The importance and contribution of this microbial diversity to the success of the invaders is an open question, however, and despite attempts to link microbial diversity to function, diversity in soils is large and generally difficult to interpret. It is clear, though, that certain microbial types were associated with greater abundance in invaded soil and have the potential to feedback and support the growth and reproduction of invaders. The large changes in microbial diversity, though not straightforward to interpret, require further research and consideration of how it impacts plant invader success.

Unlike the structural and functional linkages that were made associated with bacterial community change and plant invasion, fungal communities in the current study were not as clearly demarcated phylogenetically nor linked with specific processes. There were, however, very similar directional shifts in fungal community structure that help to support the findings observed for bacterial communities. Indeed, shifts in fungal community structure accounted for up to 17% of the variation in the PCoA plot (Fig. 1). Fungi play critical ecosystem roles

as saprotrophs, mutualists, and pathogens and though pinpointing the exact nature of the effects are not possible in the current study, the patterns of community change support the idea that plant invaders drive and are driven by a positive plant-microbial feedback model that fuel their success.

The Ascomycota showed greater abundances associated with invasion, and as the compositionally largest phylum of fungi with 64,000 species and a range of traits that include saprobe, pathogens, and mutualists, the effects of the change are likely to be functionally important [68, 69]. It is important to recognize that fungi, like bacteria, can have multiple ecological roles; for example, many mycorrhiza are also saprotrophs. Using their methodology to sort orders into an ecological context, however, Sordariales were overwhelmingly characterized as Saprobes, and the Hypocreales and Capnodiales form a mix of saprobes, plant associates and plant pathogens. So although the primary ecological changes that were observed using these methodologies are still broad, they show the potential that phylogeny has for predicting fungal ecology and the effects of plant invasion.

It is notable that a considerable amount of study has been given to the pathogenic roles played by many of the fungal taxa in our surveys. Dothideomycetes and Nectriaceae, for example, are found to play multiple antagonistic roles to plants and plant growth. It cannot be known, however, if these fungi actually play this type of role or are perhaps recruited to support plant invasion through antagonization of non-invaded plant species [70]. Whether serving as a loose plant affiliate or a plant-microbial interaction, there would be opportunity for invasive plants to disrupt plant communities if invaders themselves were less prone to the antagonistic effects of the pathogens. Research is needed to understand the nature of the changes in fungal community structure and their consequences for plant invader success.

Conclusion

It is well known that invasive species have direct and indirect effects on the surrounding non-invaded plant community, especially through root exudates: *Centaurea* spp. [71]; *Ailanthus altissima* [72]; and *Artemisia vulgaris* [73]. Our study offers insights into microbial communities and plant invasions by showing a link between invasion and belowground community change. Functional predictions based on the phylogeny of bacteria agreed with field measurements of N turnover rates and suggest that changes in N cycling bacteria, which include nitrifiers and diazotrophs, may be a significant cog in the success of invasive plant encroachment and success into non-invaded/remnant ecosystems. If these results are further confirmed, management scenarios may soon be utilized to change the soil properties and outcome of plant-driven changes in microbial communities to help favor non-invaded plants and restore native ecosystem functions.

Acknowledgements

We would like to thank Zhe Bao, Austin Jackson, Daniel Tekiela for helping with soil collection and preparation. Sincere thanks to Bronte A. Lantin, Yoonji Ha, and Nolan Hodge for help with sample preparation for microbial analysis. We appreciate the help of Amanda Karstetter for laboratory help and edits to the final draft of the manuscript and Kevin Bamber in inorganic nitrogen analysis. We thank VTOASF for providing funds to cover the cost of publication.

References

1. Ehrenfeld JG. Effects of exotic plant invasions on soil nutrient cycling processes. Ecosystems. 2003;6(6):503-23. PubMed PMID: 200300559441.

2. Liao C, Peng R, Luo Y, Zhou X, Wu X, Fang C, et al. Altered ecosystem carbon and nitrogen cycles by plant invasion: a meta-analysis. New Phytologist. 2008;177:706-14.

3. Vilà M, Espinar JL, Hejda M, Hulme PE, Jarošík V, Maron JL, et al. Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. Ecology Letters. 2011;14:702-8.

4. Van Der Heijden MG, Bardgett RD, Van Straalen NM. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecology Letters. 2008;11(3):296-310.

5. Wolfe BE, Klironomos JN. Breaking new ground: soil communities and exotic plant invasion. Bioscience. 2005;55(6):477-87.

6. Hendrick R, N W. Rhododendron thickets alter N cycling and soil extracellular enzyme activity in southern Appalachian hardwood forests. . Pedobiologia. 2007;50:563-76.

7. Horton J, Clinton B, Walker J, Beier C, Nilsen E. Variation in soil and forest floor characteristics along gradients of ericaceous, evergreen shrub cover in the southern Appalachians. Castanea. 2009;74(4):340-52.

8. Nilsen E, Lei T, Semones S, Walker J, Miller O, Clinton B. Does *Rhododendron maximum* L. (Ericaceae) reduce the availability of resources above and belowground for canopy tree seedlings. . American Midland Naturalist. 2001;145:324-43.

9. Walker J, Miller OJ, Horton J, Beier C, Clinton B, Nilsen E. Distribution of ectomycorrhizal fungi on tree seedlings grown in gradients of ericaceous shrubs in the southern Appalachian Mountains. Inoculum. 2002;53(3):58.

10. Stinson KA, Campbell SA, Powell JR, Wolfe BE, Callaway RM, Thelen GC, et al. Invasive plant suppresses the growth of native tree seedlings by disrupting belowground mutualisms. PLoS Biology. 2006;4(5):727-31.

11. Klironomos JN. Variation in plant response to native and exotic arbuscular mycorrhizal fungi. Ecology. 2003;84(9):2292-301.

12. Klironomos JN. Feedback with soil biota contributes to plant rarity and invasiveness

in communities. Nature. 2002;417(6884):67-70. PubMed PMID: 200200430140.

13. Yang Q, Carrillo J, Jin H, Shang L, Hovick SM, Nijjer S, et al. Plant–soil biota interactions of an invasive species in its native and introduced ranges: Implications for invasion success. Soil Biology and Biochemistry. 2013;65:78-85.

14. Gundale MJ, Kardol P, Nilsson M-C, Nilsson U, Lucas RW, Wardle DA. Interactions with soil biota shift from negative to positive when a tree species is moved outside its native range. New Phytologist. 2014;202(2):415-21. doi: 10.1111/nph.12699.

15. Hawkes CV, Wren IF, Herman DJ, Firestone MK. Plant invasion alters nitrogen cycling by modifying the soil nitrifying community. Ecology Letters. 2005;8(9):976-85.

16. Castro-Diez P, Godoy O, Alonso A, Gallardo A, Saldana A. What explains variation in the impacts of exotic plant invasions on the nitrogen cycle? A meta-analysis. Ecol Lett. 2014;17(1):1-12. doi: 10.1111/ele.12197. PubMed PMID: 24134461.

17. Parker SS, Schimel JP. Invasive Grasses Increase Nitrogen Availability in California Grassland Soils. Invasive Plant Science and Management. 2010;3(1):40-7. doi: 10.1614/ipsm-09-046.1.

18. Hobbie SE. Plant species effects on nutrient cycling: revisiting litter feedbacks. Trends Ecol Evol. 2015. doi: 10.1016/j.tree.2015.03.015. PubMed PMID: 25900044.

19. Jeschke J, Aparicio LG, Haider S, Heger T, Lortie C, Pyšek P, et al. Support for major hypotheses in invasion biology is uneven and declining. NeoBiota. 2012;14:1-20.

20. Kourtev PS, Ehrenfeld JG, Haggblom M. Exotic plant species alter the microbial community structure and function in the soil. Ecology. 2002;83(11):3152-66. PubMed PMID: 200300067736.

21. Tamura M, Tharayil N. Plant litter chemistry and microbial priming regulate the accrual, composition and stability of soil carbon in invaded ecosystems. New Phytologist. 2014;203(1):110-24. doi: 10.1111/nph.12795. PubMed PMID: 24720813.

22. Brown R, Peet R. Diversity and invasibility of Southern Appalachian plant communities. Ecology. 2003;84:32-9.

23. Kowarik I, Säumel I. Biological flora of Central Europe: *Ailanthus altissima* (Mill.) Swingle. Perspectives in Plant Ecology, Evolution and Systematics. 2007;8(4):207-37. doi: 10.1016/j.ppees.2007.03.002.

24. Flory S, Clay K. Non-native grass invasion suppresses forest succession. Oecologia. 2010;164(4):1029-38. doi: 10.1007/s00442-010-1697-y.

25. McGrath D, Binkley M. Microstegium vimineum invasion changes soil chemistry and microarthropod communities in Cumberland Plateau Forests. Southeastern Naturalist. 2009;8(1):141-56.

26. Kourtev PS, Ehrenfeld JG, Huang WZ. Effects of exotic plant species on soil properties in hardwood forests of New Jersey. In: Wieder RK, Novak M, Cerny JE, editors. Biogeochemical investigations at watershed, landscape, and regional scales 1998; : 493 501: Kluwer Academic Publishers, PO Box 989, 3300 AZ Dordrecht, The Netherlands; Kluwer Academic Publishers, 101 Phillip Drive, Norwell, Massachusetts 02061, USA; 1998.

27. Gómez-Aparicio L, Canham CD. Neighbourhood analyses of the allelopathic effects of the invasive tree *Ailanthus altissima* in temperate forests. Journal of Ecology. 2008;96(3):447-58. doi: 10.1111/j.1365-2745.2007.01352.x.

28. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012;6(8):1621-4. doi: 10.1038/ismej.2012.8. PubMed PMID: 22402401; PubMed Central PMCID: PMC3400413.

29. Schmidt P-A, Bálint M, Greshake B, Bandow C, Römbke J, Schmitt I. Illumina metabarcoding of a soil fungal community. Soil Biology and Biochemistry. 2013;65:128-32. doi: 10.1016/j.soilbio.2013.05.014.

30. White T, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Innis M, Gelfand D, Shinsky J, White T, editors. San Diego: Academic Press; 1990.

31. Masella A, Bartram A, Truszkowski J, Brown D, Neufeld J. PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics. 2012;13(1):31. doi: 10.1186/1471-2105-13-31.

32. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335-6. doi: 10.1038/nmeth.f.303. PubMed PMID: 20383131; PubMed Central PMCID: PMC3156573.

33. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26(19):2460-1. doi: 10.1093/bioinformatics/btq461. PubMed PMID: 20709691.

34. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006;72(7):5069-72. doi: 10.1128/AEM.03006-05. PubMed PMID: 16820507; PubMed Central PMCID: PMC1489311.

35. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 2012;6(3):610-8. doi: 10.1038/ismej.2011.139. PubMed PMID: 22134646; PubMed Central PMCID: PMC3280142.

36. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007;73(16):5261-7. doi: 10.1128/AEM.00062-07. PubMed PMID: 17586664; PubMed Central PMCID: PMC1950982.

37. Abarenkov K, Henrik Nilsson R, Larsson KH, Alexander IJ, Eberhardt U, Erland S, et al. The UNITE database for molecular identification of fungi--recent updates and future perspectives. New Phytol. 2010;186(2):281-5. doi: 10.1111/j.1469-8137.2009.03160.x. PubMed PMID: 20409185.

38. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol. 2005;71(12):8228-35. doi: 10.1128/AEM.71.12.8228-8235.2005. PubMed PMID: 16332807; PubMed Central PMCID: PMC1317376.

39. Beals EW. Bray-Curtis Ordination: An Effective Strategy for Analysis of Multivariate Ecological Data. 1984;14:1-55. doi: 10.1016/s0065-2504(08)60168-3.

40. Gower JC. Principal Coordinates Analysis. Encyclopedia of Biostatistics. 2 ed. The Open University, Milton Keynes, UK: John Wiley and Sons, Ltd; 2005.

41. Vazquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. EMPeror: a tool for visualizing high-throughput microbial community data. Gigascience. 2013;2(1):16. doi: 10.1186/2047-217X-2-16. PubMed PMID: 24280061; PubMed Central PMCID: PMC4076506.

42. Hill TC, Walsh KA, Harris JA, Moffett BF. Using ecological diversity measures with bacterial communities. FEMS Microbiol Ecol. 2003;43(1):1-11. doi: 10.1111/j.1574-6941.2003.tb01040.x. PubMed PMID: 19719691.

43. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecology. 2001;26(1):32-46. doi: 10.1111/j.1442-9993.2001.01070.pp.x.

44. Clarke KR. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology. 1993;18(1):117-43. doi: 10.1111/j.1442-9993.1993.tb00438.x.

45. Dufrene M, Legendre P. Species Assemblages and Indicator Species:The Need for a Flexible Asymmetrical Approach. Ecological Monographs. 1997;67(3):345-66. doi: 10.1890/0012-9615(1997)067[0345:saaist]2.0.co;2.

46. McCune B, Mefford MJ. PC-ORD. Multivariate Analysis of Ecological Data. 6 ed.

MjM Software, Gleneden Beach, Oregon, U.S.A.2011.

47. Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. 2013;31(9):814-21. doi: 10.1038/nbt.2676. PubMed PMID: 23975157; PubMed Central PMCID: PMC3819121.

48. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics. 2014;30(21):3123-4. doi: 10.1093/bioinformatics/btu494. PubMed PMID: 25061070.

49. Welch BL. The Generalization of Students Problem When Several Different Population Varlances Are Involved. Biometrika. 1947;34(1-2):28-35. doi: 10.1093/biomet/34.1-2.28.

50. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300. doi: citeulike-article-id:1042553
doi: 10.2307/2346101.

51. Storey JD. A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2002;64(3):479-98. doi: 10.1111/1467-9868.00346.

52. Haselkorn R, Kapatral V. The nif Genes of Rhodobacter capsulatus, Rhodobacter sphaeroides and Rhodopseudomonas palustris. In: Palacios R, Newton WE, editors. Genomes and Genomics of Nitrogen-fixing Organisms. Nitrogen Fixation: Origins, Applications, and Research Progress. 3. Dordrecht, Netherlands: Springer; 2010. p. 71-82.

53. Wang G, Yoon SH, Lefait E. Microbial communities associated with the invasive Hawaiian sponge Mycale armata. ISME J. 2009;3(3):374-7. doi: 10.1038/ismej.2008.107. PubMed PMID: 18987678.

54. Lum G, Min XJ. FunSecKB: the Fungal Secretome KnowledgeBase. Database (Oxford). 2011;2011:bar001. doi: 10.1093/database/bar001. PubMed PMID: 21300622; PubMed Central PMCID: PMC3263735.

55. Thorpe AS, Callaway RM. Interactions between invasive plants and soil ecosystems: positive feedbacks and their potential to persist. In: Cadotte MW, McMahon SM, Fukami T, editors. Conceptual ecology and invasions biology. Great Britain: Springer; 2005. p. 333-51.

56. Wang C, Xiao H, Liu J, Wang L, Du D. Insights into Ecological Effects of Invasive Plants on Soil Nitrogen Cycles. American Journal of Plant Sciences. 2015;06(01):34-46. doi: 10.4236/ajps.2015.61005.

57. WeiHua L, ChongBang Z, JieYun L, ChaoJie Y. Characteristics of nitrogen metabolism and soil nitrogen of invasive plants. Journal of Tropical and Subtropical Botany. 2008;16(4):7.

58. Jeschke JM. General hypotheses in invasion ecology. Diversity and Distributions. 2014;20(11):1229-34. doi: 10.1111/ddi.12258.

59. Rout ME, Chrzanowski TH, Westlie TK, DeLuca TH, Callaway RM, Holben WE. Bacterial endophytes enhance competition by invasive plants. Am J Bot. 2013;100(9):1726-37. doi: 10.3732/ajb.1200577. PubMed PMID: 23935109.

60. Smith MS, Fridley JD, Goebel M, Bauerle TL. Links between belowground and aboveground resource-related traits reveal species growth strategies that promote invasive advantages. PLoS One. 2014;9(8):e104189. doi: 10.1371/journal.pone.0104189. PubMed PMID: 25105975; PubMed Central PMCID: PMC4126695.

61. Simonin M, Le Roux X, Poly F, Lerondelle C, Hungate BA, Nunan N, et al. Coupling Between and Among Ammonia Oxidizers and Nitrite Oxidizers in Grassland Mesocosms Submitted to Elevated CO2 and Nitrogen Supply. Microb Ecol. 2015;70(3):809-18. doi: 10.1007/s00248-015-0604-9. PubMed PMID: 25877793.

62. Placella SA, Firestone MK. Transcriptional response of nitrifying communities to wetting of dry soil. Appl Environ Microbiol. 2013;79(10):3294-302. doi: 10.1128/AEM.00404-13. PubMed PMID: 23524666; PubMed Central PMCID: PMC3685241.

63. Jangid K, Williams MA, Franzluebbers AJ, Schmidt TM, Coleman DC, Whitman WB. Land-use history has a stronger impact on soil microbial community composition than aboveground vegetation and soil properties. Soil Biology and Biochemistry. 2011;43(10):2184-93. doi: 10.1016/j.soilbio.2011.06.022.

64. Xu CW, Yang MZ, Chen YJ, Chen LM, Zhang DZ, Mei L, et al. Changes in non-symbiotic nitrogen-fixing bacteria inhabiting rhizosphere soils of an invasive plant Ageratina adenophora. Applied Soil Ecology. 2012;54:32-8. doi: Doi 10.1016/J.Apsoil.2011.10.021. PubMed PMID: WOS:000302110900005.

65. Coelho MRR, Marriel IE, Jenkins SN, Lanyon CV, Seldin L, O'Donnell AG. Molecular detection and quantification of nifH gene sequences in the rhizosphere of sorghum (Sorghum bicolor) sown with two levels of nitrogen fertilizer. Applied Soil Ecology. 2009;42(1):48-53. doi: 10.1016/j.apsoil.2009.01.010.

66. Elgersma KJ, Ehrenfeld JG, Yu S, Vor T. Legacy effects overwhelm the short-term effects of exotic plant invasion and restoration on soil microbial community structure, enzyme activities, and nitrogen cycling. Oecologia. 2011;167(3):733-45. doi: 10.1007/s00442-011-2022-0. PubMed PMID: 21618010.

67. Pang PC, Cho CM, Hedlin RA. EFFECTS OFpH AND NITRIFIER POPULATION ON NITRIFICATION OF BAND-APPLIED AND HOMOGENEOUSLY MIXED UREA NITROGEN IN SOILS. Canadian Journal of Soil Science. 1975;55(1):15-21. doi: 10.4141/cjss75-003.

68. James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, et al.

Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature. 2006;443(7113):818-22. doi: Doi 10.1038/Nature05110. PubMed PMID: WOS:000241362700042.

69. Schoch CL, Sung GH, Lopez-Giraldez F, Townsend JP, Miadlikowska J, Hofstetter V, et al. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Syst Biol. 2009;58(2):224-39. doi: 10.1093/sysbio/syp020. PubMed PMID: 20525580.

70. Eppinga MB, Rietkerk M, Dekker SC, De Ruiter PC, Van der Putten WH, Van der Putten WH. Accumulation of local pathogens: a new hypothesis to explain exotic plant invasions. Oikos. 2006;114(1):168-76. doi: 10.1111/j.2006.0030-1299.14625.x.

71. Callaway RM, Aschehoug ET. Invasive plants versus their new and old neighbors: a mechanism for exotic invasion. Science. 2000;290:521-3.

72. Heisey RM. Identification of an allelopathic compound from *Ailanthus altissima* (Simaroubaceae) and characterization of its herbicidal activity. American Journal of Botany. 1996;83(2):192-200.

73. Barney JN, Sparks JP, Greenberg J, Whitlow TH, Guenther A. Biogenic volatile organic compounds from an invasive species: impacts on plant-plant interactions. Plant Ecology. 2009;203:195-205.

Table S1: KEGG pathways (level 2) predicted by PICRUSt that were significantly different between root-zone bacteria of invaded and non-invaded samples using two-sided Welch's t-test with Benjamini Hochberg FDR for multiple testing corrections.

Level 2 KEGG pathways (increasing order BH corrected p-value)	**Group**
Energy Metabolism	I
Cell Motility	I
Amino Acid Metabolism	N
Immune System Diseases	N
Signaling Molecules and Interaction	N
Lipid Metabolism	N
Metabolism of Other Amino Acids	N
Transport and Catabolism	N
Transcription	N

I and N indicate pathway was abundant in root-zone bacteria of invaded and non-invaded samples, respectively.

Table S2: KEGG pathways (level 2) predicted by PICRUSt that were significantly different between root-zone bacteria of invaded and non-invaded samples using two-sided Welch's t-test with Storey FDR for multiple testing corrections.

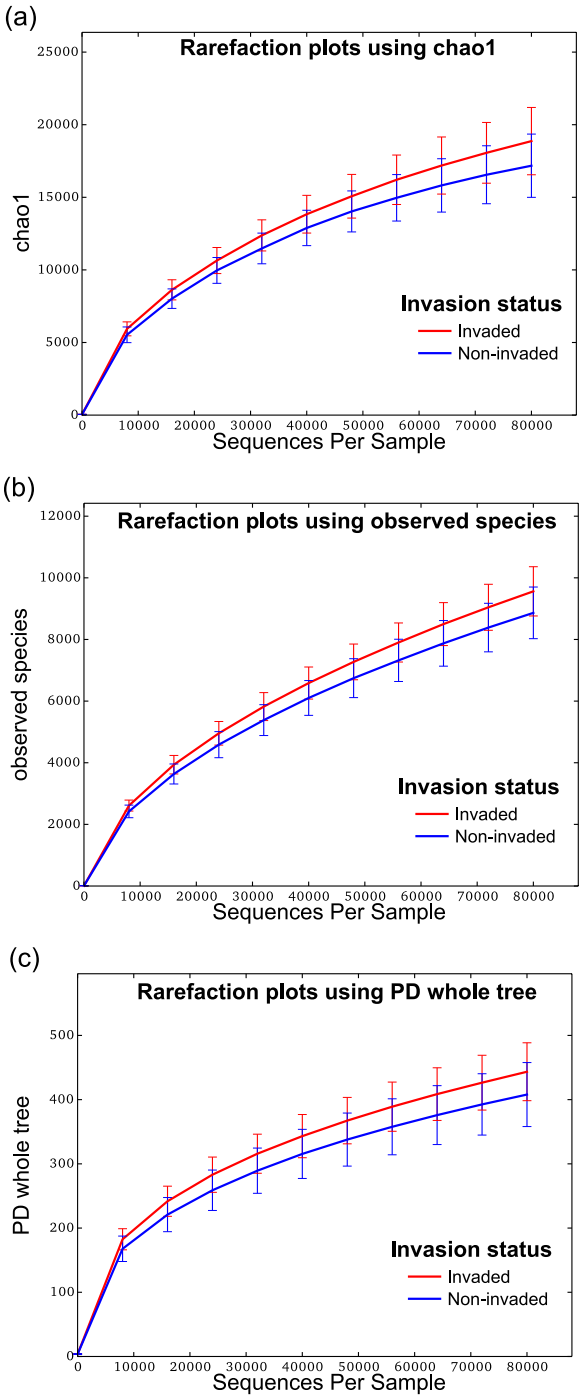
Level 2 KEGG pathways (increasing order Storey corrected p value)	Group
Cell Motility	I
Cellular Processes and Signaling	I
Poorly Characterized	I
Energy Metabolism	I
Translation	I
Genetic Information Processing	I
Glycan Biosynthesis and Metabolism	I
Signal Transduction	I
Folding, Sorting and Degradation	I
Metabolic Diseases	I
Infectious Diseases	I
Signaling Molecules and Interaction	N
Transcription	N
Immune System Diseases	N
Biosynthesis of Other Secondary Metabolites	N
Amino Acid Metabolism	N
Transport and Catabolism	N
Metabolism of Other Amino Acids	N
Lipid Metabolism	N
Metabolism of Terpenoids and Polyketides	N
Environmental Adaptation	N
Digestive System	N
Enzyme Families	N
Cancers	N
Endocrine System	N
Xenobiotics Biodegradation and Metabolism	N
Immune System	N

I and N indicate pathway was abundant in root-zone bacteria of invaded and non-invaded samples, respectively.

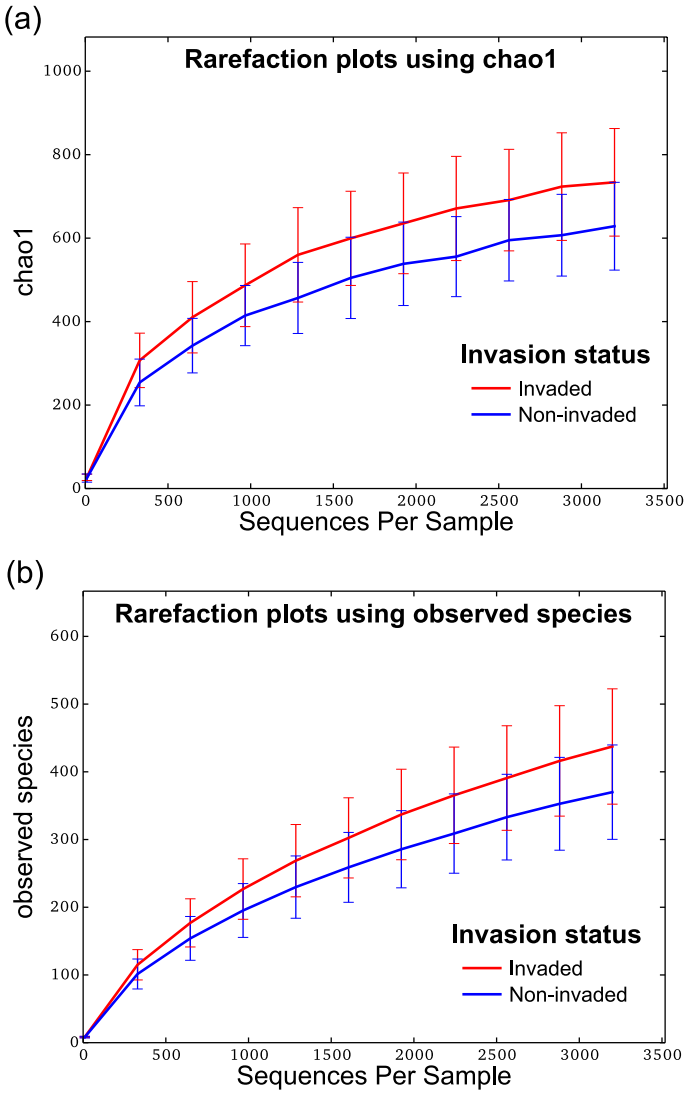
Table S3: Top 20 abundant and significant ($\alpha < 0.05$) level 3 KEGG processes by Storey FDR.

Level 3 processes significant by Storey FDR ($\alpha<0.05$)	Mean rel. freq. (%) I – Mean rel. freq. (%) N	Group
Bacterial motility proteins	0.195084675	I
Methane metabolism	0.148098067	I
Function unknown	0.120678949	I
Secretion system	0.103661933	I
Pyruvate metabolism	0.093091436	I
Flagellar assembly	0.07317117	I
Glycolysis / Gluconeogenesis	0.072181727	I
Citrate cycle (TCA cycle)	0.072041679	I
Carbon fixation pathways in prokaryotes	0.068991602	I
Nitrogen metabolism	0.051446033	I
Protein folding and associated processing	0.038192796	I
Chromosome	0.037170117	I
Phenylalanine, tyrosine and tryptophan biosynthesis	0.029974768	I
Chaperones and folding catalysts	0.018433534	I
DNA replication proteins	0.014744848	I
Ribosome Biogenesis	0.010685456	I
Lipid biosynthesis proteins	-0.02660903	N
Arginine and proline metabolism	-0.042705747	N
Peptidases	-0.067498986	N
Transcription factors	-0.070867248	N

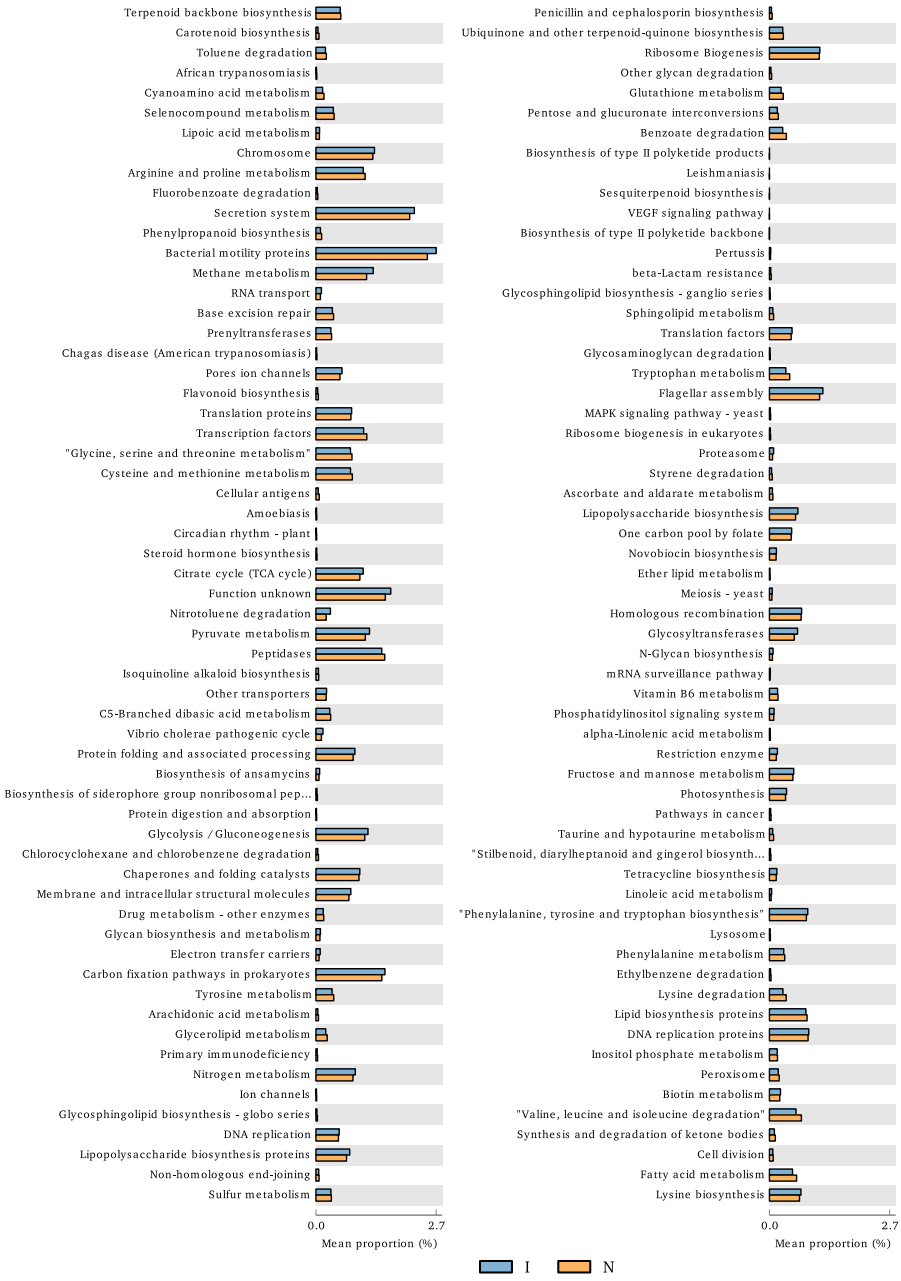
First, the significant processes were descending sorted as per the average of mean relative frequency (%) in native and invasive samples. The top 20 abundant processes were categorized as belonging to native (N) or invasive (I) samples depending on the difference of mean rel. freq. (%). Finally, in each category, the processes were descending sorted as per the difference in mean rel. freq. (%) between I and N.



S1 Fig: Rarefaction plots of bacterial alpha diversity for invaded and non-invaded samples using (a) chao1, (b) observed species, and (c) PD whole tree.



S2 Fig: Rarefaction plots of fungal alpha diversity for invaded and non-invaded samples using (a) chao1 and (b) observed species.



S3 Fig: KEGG pathways (level 3) predicted by PICRUSt that were significantly different between root-zone bacteria of invaded and non-invaded samples using two-sided Welch's t-test with Storey FDR for multiple testing corrections.

CHAPTER 3

Microbial Communities and Diazotrophic Activity Differ in the Root-Zone of Alamo and Dacotah Cultivars of Switchgrass

Richard R. Rodrigues[1], Jinyoung Moon[2], Bingyu Zhao[2], Mark A. Williams[1,2*]

[1]Interdisciplinary Ph.D. Program in Genetics, Bioinformatics, and Computational Biology, Virginia Tech, Blacksburg, Virginia, United States of America.

[2]Department of Horticulture, Virginia Tech, Blacksburg, Virginia, United States of America.

* Corresponding author

E-mail: markwill@vt.edu (MW)

Phone: 540-231-4887

Original Research Article

Keywords: Alamo, Dacotah, Switchgrass, structure, function, interactions, root-zone, diazotrophs, 16s rRNA, ITS

Abstract

Nitrogen (N) bioavailability is a primary limiting nutrient for crop and feedstock productivity. Associative nitrogen fixation (ANF) by diazotrophic bacteria has been shown to provide significant amounts of N in some tropical grasses, but the potential of Switchgrass, a warm-season, temperate, United States native, perennial tall-grass has not been widely studied to support ANF in the root-zone. “Alamo” and “Dacotah” are cultivars of switchgrass, adapted to the southern and northern regions of the US, respectively, and offer an opportunity to better describe this plant-bacterial association. The nitrogenase enzyme activity, microbial communities, and amino acid profiles in the root-zones of the two ecotypes were studied at three different plant growth stages. Differences in the nitrogenase enzyme activity and free soluble amino acid profiles indicated the potential for greater nitrogen fixation in the high productivity Alamo compared to the lower productivity Dacotah. Changes in the amino acid profiles of the root-zone suggest different plant-bacterial interactions can help to explain differences in nitrogenase activity. PICRUSt analysis revealed functional differences, especially nitrogen metabolism, that supported ecotype differences in root-zone nitrogenase enzyme activity. It is thought that the greater productivity of Alamo increased the belowground flow of carbon into roots and root-zone habitats, which in turn, support the high-energy demands needed to support nitrogen fixation. Further research is thus needed to understand plant ecotype and cultivar trait differences that can be used to breed or genetically modify crop plants to support root zone associations with diazotrophs.

Introduction

Nitrogen (N) bioavailability limits crop yields. Associative nitrogen fixation (ANF) by diazotrophic bacteria provides an alternative to chemical fertilizers. ANF is thought to be a loose form of *quid-pro-quo*, whereby energy and carbon fixed by the plant are exchanged for nitrogen fixed by bacterial diazotrophs. ANF has generally been considered a relatively small player in the annual nitrogen economy of grasses, however, in some cases it may provide up to 35% of nitrogen to agriculturally important non-legumes and forage grasses, such as sugarcane, Miscanthus, energy-cane, and switchgrass (Chalk, 1991, Weier, 1980, Wewalwela, 2014). ANF offers a natural solution to an immediate need for plant-available nitrogen while also reducing the footprint (Galloway *et al.*, 2003, Oenema *et al.*, 2009, Smil, 1999a, Smil, 1999b, Smil, 2002, Sutton *et al.*, 2011a, Sutton *et al.*, 2011b) of the land applied synthetic chemical fertilizers. Some gaps in knowledge, however, need to be filled before using ANF for sustainable production. The natural variation in ANF that occurs across plant cultivars, plant growth stages, and as a result of environmental conditions needs to be researched. This is an important step in the identification of plant-microbial traits and interactions that underlie potentially high rates of ANF, and can be used to eventually breed and develop cultivars with greater ANF capacities to support high and sustainable crop productivity. Research in this regard, carried out in model grasses and likely to be transferrable to other major worldwide grain crops (e.g. Maize, wheat) would thus support, more broadly, sustainable agriculture.

Switchgrass (*Panicum virgatum* L.) is a perennial, United States native grass. Its biomass is used for forage for livestock and for bioenergy. Switchgrass is useful in the prevention of soil erosion, provides wildlife habitat, and reduces atmospheric carbon dioxide (Follett *et al.*, 2012, Ma, 1999, Skinner, 2009). Due to the various environmental and monetary benefits of switchgrass, identifying microbes for its sustainable production has

been an active area of research (Chaudhary *et al.*, 2012, Hargreaves *et al.*, 2015, Jesus *et al.*, 2010, Jesus *et al.*, 2016, Mao *et al.*, 2014, Mao *et al.*, 2013, Mao *et al.*, 2011). These studies have shown presence of bacteria that are capable of nitrogen fixation, however, studies rarely confirm the occurrence and extent of nitrogen fixation. Taxonomic information without details about functions do not completely explain the role of this plant-bacterial interactions.

It is known that cultivars of plants often have differences in N requirements, abilities to support nitrogen fixation (Boddey & Dobereiner, 1988, Day *et al.*, 1975, Ledgard & Steele, 1992, Porter, 1966), unique rhizosphere microbiomes (Li *et al.*, 2014, Miller *et al.*, 1989), and thus are important sources of variation in plant-microbial traits. A recent study, for example, identified diverse sets of nitrogen-fixing bacteria associated with roots and shoots of switchgrass (Bahulikar *et al.*, 2014). Other studies focused on targeted approaches of using nitrogen-fixing bacterial endophytes to improve productivity across cultivars (Ker *et al.*, 2012, Lowman *et al.*, 2015, Xia *et al.*, 2013). Plant bacterial associations and how they differ among plant species and cultivars can help to understand the interplay of traits important for defining the drivers of a globally significant plant-bacterial function like that of ANF.

"Alamo" and "Dacotah," are tetraploid cultivars of switchgrass, adapted to the southern and northern U.S. ecotypes, respectively. Compared to Dacotah, Alamo shows higher biomass productivity, taller shoots, drought tolerance, and disease resistance. Nitrogen is required for the growth and maintenance of plants, so we hypothesized that Alamo has more nitrogen demand than Dacotah, and supports bacterial communities with higher rates of nitrogen fixation in the root-zone. **Here, we investigated the nitrogenase activity, amino acid composition, structure, function, and interactions of microbial communities in the root-zones of two switchgrass cultivars across multiple growth stages.** The holistic research offers novel insights into the dynamics of nitrogen fixing communities in the root-

zone of switchgrass that are useful for explaining the variation associated with this globally important plant-bacterial interaction.

Materials and Methods

Experimental setup

Root-zone soil was collected from the Virginia Tech Agronomy Farm/Urban Horticulture Center field growing Alamo, Dacotah (~7 years), and their F1 progeny lines (~2 years) to serve as an appropriate inoculum of potential microbes associated with switchgrass. Alamo and Dacotah seeds were imbibed and allowed to germinate at 28°C. After approximately 7 days, the germinated seeds were potted in sieved (4.75 mm) and homogenized field soil and grown in the greenhouse. The plants were kept moist, well aerated, and supplemented with light to maintain optimal growth. Sampling was performed in replicates at three time points from imbibition, indicative of important growth stages: Stages V0 (~2.5 weeks old) and V2 (~1.5 months old) from the vegetative phase, and Stage E3 (~3.5 months old) from the elongation phase (Moore *et al.*, 1991).

Nitrogenase enzyme activity

To estimate the rate of nitrogen fixation, nitrogenase activity was evaluated by measuring the quantity of acetylene reduced to ethylene using a modified protocol (S.Castle & Barger, 2010). Briefly, the day before the nitrogenase measurement the Alamo and Dacotah containing pots were wetted to saturation and allowed to drain gravimetrically to ~0.03 MPa. Pots were randomly sampled to obtain replicates (n=3 for V0, n=4 for V2, and n=5 for E3) that received acetylene ("treatment") or remained untreated ("controls"). Each replicate sample was sealed within a 500ml mason glass jar. Calcium Carbide (CaC_2) and water was mixed to produce acetylene gas, which was injected to create a ~10% (v/v)

headspace. Mason glass jars with sample but without acetylene served as controls, to assess the natural (background) production of ethylene. At four time points (immediately and then approximately every 30 minutes) headspace gas was mixed using a 10 ml syringe and needle to extract gas from the container and analyzed for acetylene and ethylene concentrations using a gas chromatography and column (Carbonplot, Agilent Inc.) in line with a flame ionization detector. The ethylene accumulation rate was used to determine nitrogenase enzyme activity.

Known ethylene standards were used to generate a linear calibration curve that was used to calculate ethylene concentrations in the sample and reported as the rate of ethylene production, $\mu g\ g_{system}^{-1}\ h^{-1}$. Following the confirmation of normality, equality of variances, and determination of outliers, the rates of ethylene production between Alamo and Dacotah were compared (p-value < 0.05) using two-sample t-test (V0 and E3) and Mann-Whitney test (V2). Only the 24 controls (3, 4, and 5 samples each of Alamo and Dacotah from V0, V2, and E3, respectively) samples were used for the following experiments.

Soluble amino acids

Shoots were removed, and the root and the rhizosphere soil were extracted in 10mL of 0.9% NaCl in a 50 mL falcon tube. The roots were washed to remove any attached soil and used to measure fresh and dry weight. Falcon tubes were vortexed at 200 RPM on electric orbital shaker for 20 minutes, allowed to sit for 5 minutes. Eight mL of solution was transferred to a new 15 mL falcon tube and centrifuged at 3500 G for 15 minutes. The amount of supernatant (approximately 5 mL) transferred to new smaller size falcon tubes was recorded and 8 µL of internal standard (2.5 mm α-Aminobutyric acid; AABA) was added to each sample. The volume for each sample was made to 10 mL by adding 0.9% NaCl. After shaking for 2 minutes, 1.5 mL of sample was transferred into a 2 mL microfuge tube through

0.22 µm polyvinylidene fluoride (PVDF; Thermo Scientific™ Target2™ Syringe Filters, Cat# 03377155) membrane syringe filter. 500 µL of filtrate from the 2 mL tube was transferred to a new 1.5 mL microfuge tube and dried in a vacufuge (speedvac) for 2.5 hours at 60°C (function 3). The dried pellet was derivatized using the AccQ FluorTM reagent kit (Fluorescent 6-Aminoquinoly-N-Hydroxysuccinimidyl Carbamate derivatizing reagent; Waters Co. Cat# WAT052880) following the standard protocol (Bosch *et al.*, 2006, Hou *et al.*, 2009). Chromatographic separation on the HPLC 1260 Infinity system (Agilent Technologies, USA) was carried out on a reversed phase column (Waters X-Terra MS C18, 3.5µm, 2.1 m x 150 mm). The mobile phase consisted of A: an aqueous solution containing 140 mM of sodium acetate, 17 mM of triethylamine (TEA; Fisher Chemical, Cat# O4884100), and 0.1% (g/L, w/v) disodium dihydrogen ethylenediaminetetraacetate dihydrate (EDTA-2Na·2H2O; Sigma, CAS# 6381-92-6), pH 5.05, adjusted with phosphoric acid solution, and B: Acetonitrile (ACN; HPLC grade, Fisher Chemical, Cat# A998-1): ultrapure water (60:40, v/v). The gradient conditions were 0 - 17 min 100 - 93% A, 17 - 21 min 93 - 90% A, 21 - 30 min 90 - 70% A, 30 - 35 min 70% A, 35 - 36 min 70 - 0% A, and then hold for 4 min before restoring to the initial composition at 40.5 min, with the final composition kept for 9 min. The column was thermostated at 50°C and operated at a flow rate of 0.35 ml/min. The sample injection volume was 5 µL. The analytes detection was carried out using a fluorescence detector (λex = 250 nm and λem = 395 nm) (Bosch *et al.*, 2006, Hou *et al.*, 2009). Soluble amino acids in the samples were qualified and quantified by comparison with amino acid standard solutions. Each amino acid standard solution contained 19 protein amino acids including alanine (Ala), arginine (Arg), aspartic acid (Asp), asparagine (Asn), cystine (Cys–Cys; more stable form in oxidative condition than monomer cysteine), glutamic acid (Glu), glutamine (Gln), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), serine (Ser), threonine (Thr), tyrosine (Tyr),

tryptophan (Trp), and valine (Val) and 1 non-protein amino acid Ornithine (Orn). Note that Asn and Ser were co-eluted and Cys-Cys and Tyr were co-eluted. A total of 20 peaks were quantified. The amino acid data was relativized per sample and used for multivariate data analyses using PC-ORD software version 6.0 (MjM Software, Gleneden Beach, OR, USA) and R.

DNA extraction

10 mL of 0.9% NaCl was added to the falcon tubes containing rhizosphere soil samples. The tubes were then turned upside down once and the soil-solution was homogenized by vortexing 1 min each (at level 3). A 3 mL sample of the re-suspended solution was taken immediately after vortexing in a 5 mL tube. After centrifugation of the 5 mL tubes at 10,000 RPM for 10 min at 20°C, the supernatant was removed and the tubes were turned upside-down on a Kim wipe (with cap open) to dry for 10 minutes. For each tube, the leftover soil (pellet) was mixed and 0.25g was used for the DNA extraction using MoBio's PowerSoil® DNA Isolation kit as per the manufacturer's protocol. The quality and concentration of the DNA was checked using 0.8% (w/v) agarose gel electrophoreses and Nanodrop 2000 spectrophotometry.

Real time PCR to detect nitrogenase reductase gene copy numbers

qPCR assay was performed on an ABI 7300 system (Applied Biosystems) to quantify the abundance of nitrogenase reductase gene (nifH). The 20 µl reaction volumes contained 10 µl of 2X PowerUp™ SYBR® Green Master Mix (Applied Biosystems), 2 µl template at 2 ng/µl, 4 µl each of PolF/PolR (10 µM) primers (Poly *et al.*, 2001) and standard conditions as per master mix's protocol. Triplicates of nuclease-free water and no-template controls were included. Technical triplicates of each biological replicate were averaged to get mean C_T

value. Dilutions of known concentrations of *Sinorhizobium meliloti* 1021 (provided by Dr. B. Scharf at Virginia Tech) were used to generate the C_T vs. log (N_0) standard curve. The standard curve was used to calculate the qPCR efficiency ($E = 10^{(-1/slope)}$) and nifH gene copy numbers in the samples (Brankatschk *et al.*, 2012). For each cultivar at each growth stage, biological replicates were checked for outliers using Grubb's test (Grubbs, 1969) in Graph Pad (http://graphpad.com/quickcalcs/Grubbs1.cfm) and boxplots were generated in BoxPlotR with Spear's criteria of whisker definition (Spitzer *et al.*, 2014). The cultivars at each growth stage were tested for significant (p-value < 0.05) differences in nifH gene copy numbers using Mann Whitney test.

Sequencing and data analyses

The 16S rRNA and ITS gene amplification were performed using bacterial primer pair (S-D-Bact-0341-b-S-17 and S-D-Bact-0785-a-A-21) (Klindworth *et al.*, 2013) and fungal primer pair (ITS1F and ITS2) (Smith & Peay, 2014), respectively. The library preparation was performed using Illumina 16S Metagenomic sequencing library preparation guide. DNA concentration of the pool was determined by fluorometric quantification using the Qubit® 2.0 platform with Qubit dsDNA HS Assay Kit (Life Technologies). Multiplexed 250 base pairs paired-end sequencing using Illumina MiSeq was performed at the Biocomplexity Institute (formerly Virginia Bioinformatics Institute) core facility at Virginia Tech.

Barcode adapters and primers were trimmed from each read using cutadapt v1.8.1 (Martin, 2011) with a quality-trimming threshold of 30 and minimum read length of 100. The paired-end reads were merged based on overlapping sequences into single reads using PANDAseq v2.8 (Masella *et al.*, 2012). The bacterial and fungal sequencing data were analyzed using QIIME v1.8.0 (Caporaso *et al.*, 2010) as previously described (Rodrigues *et*

al., 2015). Briefly, reads were clustered into OTUs based on 97% sequence similarity using uclust (Edgar, 2010) and usearch61 (Edgar, 2010) for bacteria and fungi, respectively, using an open reference OTU-picking strategy. The representative sequence of an OTU was used to assign it a taxonomy, using uclust against the Greengenes reference database version 13.8 (DeSantis *et al.*, 2006, McDonald *et al.*, 2012) for bacteria, and RDP classifier (Wang *et al.*, 2007) against the UNITE reference database version 12.11 (Abarenkov *et al.*, 2010) for fungi.

Post-processing included diversity and richness analyses, identifying and summarizing the most abundant taxons, and describing the taxons that are different and indicative of cultivars at different growth stages. Briefly, the alpha diversity was calculated on the OTU table for all samples using several different indices, including, PD whole tree (only for bacteria), chao1, observed species, Good's coverage, Shannon, and Simpson indices. After using a sequence threshold, the beta diversity for bacteria and fungi was calculated using weighted Unifrac (Lozupone *et al.*, 2011, Lozupone *et al.*, 2007) and Bray-Curtis (Beals, 1984) metrics, respectively, and were used for Principal Coordinate Analysis (Gower, 2005) and visualization. The genus level summary of communities was used for NMDS analysis using the Bray-Curtis dissimilarity in R (vegan). Using the collated alpha diversities from the rarefied OTU table, the PD whole tree (only for bacteria), chao1, and observed species indices were used to compare alpha diversity of groups (cultivar, stage, cultivar x stage) using a two-tail nonparametric t-test with FDR correction (q value < 0.05) using collated files. Multivariate data analysis methods of MRPP (Mielke, 1984), ADONIS (Anderson, 2001) and Analysis of Similarity (ANOSIM) (Clarke, 1993) were used to identify whether the cultivars and plant growth stage had an effect on the microbial communities.

Indicator species analysis (ISA) (Dufrene & Legendre, 1997) was performed in R (vegan) using the top 50% most abundant genera to identify taxa significantly (indicator value >50 and p-value <0.05) indicative of cultivars at the respective growth stage.

PICRUSt (Langille *et al.*, 2013) analysis using the bacterial OTU abundance was performed before and after ISA analysis to identify whether functions differed between cultivars at the different growth stages. OTUs not part of the closed reference OTU picking were filtered out and the metagenomes were collapsed into KEGG pathways. Using Statistical Analysis of Metagenomic Profiles (STAMP) (Parks et al. 2014), principal component analysis was used to visualize differences in KEGG pathways of bacteria between the cultivars at different growth stages. A two-sided t-test was performed to check whether N-metabolism pathway was significantly different (p-value < 0.05) between the cultivars.

Network analysis

The OTU table, at the Class level for bacteria and Order level for fungi, was summarized, relativized, and separated to create files as per samples belonging to cultivar irrespective of the plant growth stage, i.e. Alamo (V2, E3) and Dacotah (V2, E3). Phylogenetic ecological networks were generated for each group (cultivar x growth stage) using MENAP (Deng *et al.*, 2012) with an RMT threshold of 0.31 and module detection using leading eigenvector of the community matrix (Newman, 2006) and random walk methods (Pons & Latapy, 2005). A "module" is a group of OTUs that were highly connected (positively or negatively correlated via abundance) among themselves, but had much fewer connections with OTUs outside the group. For each network, connections between OTUs between different modules were discarded to help focus on interactions among OTUs within a module. We refer these sub-networks as "same-module networks" which show the connections between OTUs belonging to the same-module. For each cultivar, the networks

from leading eigenvector of the community matrix and random walk methods were compared to generate edges common between the two methods. These "COMmon Edge between Two methods" (COMET) networks help to remove any potential biases due to the module detection algorithms and represent high confidence edges. Finally, the COMET networks were compared between cultivars.

Cultivable, free-living diazotrophs

For samples from Stage V2, the NaCl solutions containing rhizosphere soil were (100x, 1000x and 5000x) serial diluted, spread on Yeast Mannitol (YM) Agar (YMA) petri-plates and incubated at 28°C for 3-5 days to selectively grow diazotrophs. Bacterial colonies were differentiated based on morphology and color. Colonies were counted and will be reported as colony forming unit per gram of root-soil. Representative colonies were sampled and isolation plated on new YMA plates to obtain pure cultures. After 3-5 days of incubation, the isolation-plated samples were grown in Yeast Mannitol broth and DNA was extracted using UltraClean® Microbial DNA Isolation Kit (Mo Bio). The quality of the DNA was estimated using 0.8% (w/v) agarose gel electrophoreses and Nanodrop 2000 spectrophotometry. After amplification of 16S rRNA genes using polymerase chain reaction using primers 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1492R (5'-TACGGYTACCTTGTTACGACTT-3') (Fredriksson *et al.*, 2013, Lane, 1991) and KAPA2G Robust PCR kit (Kapa Biosystem), and cleaned using Agencourt AMPure XP PCR product cleanup kit (Beckman Coulter). High quality, cleaned PCR-products, as determined by 1.2% (w/v) agarose gel electrophorese and Nanodrop 2000 spectrophotometry, were sent to Biocomplexity Institute (Virginia Bioinformatics Institute) for Sanger sequencing. The "ab1" files obtained from Sanger sequencing were converted to sequence files using 4Peaks and quality trimmed with a Phred score of 25. We used the taxonomy of the top hit from

nucleotide BLAST (megablast) on the sequences to identify the bacteria. Libshuff (Schloss *et al.*, 2004, Singleton *et al.*, 2001) analysis on the sequences and Pearson's Chi-squared test on the genus-level CFU abundance were performed to identify significant differences (p-value < 0.05) between the cultivable diazotrophic communities associated with Alamo and Dacotah.

Results

Alamo showed higher rates of biological nitrogen fixation

The rate of ethylene production, which indicates the nitrogenase enzyme activity or BNF rates, was highest at stage V2 than those at stages V0 and E3. At stages V2 and E3, Alamo had significantly greater nitrogenase activity than Dacotah (Mann Whitney test p-value = 0.017 at stage V2 and t-test with equal variance p-value = 0.030 at stage E3). Both Alamo and Dacotah showed (>22 µg ethylene per g system per hr) reduction in BNF rates at E3 as compared to those at V2. We quantified the nifH gene in the rhizosphere samples of switchgrass cultivars (Figure S9) using *Sinorhizobium meliloti* 1021 as the standard (C_T = 22.064 - 4.0825 log(N_0); R^2 = 0.985; E_{std} = 1.758), however, found no significant differences between cultivars.

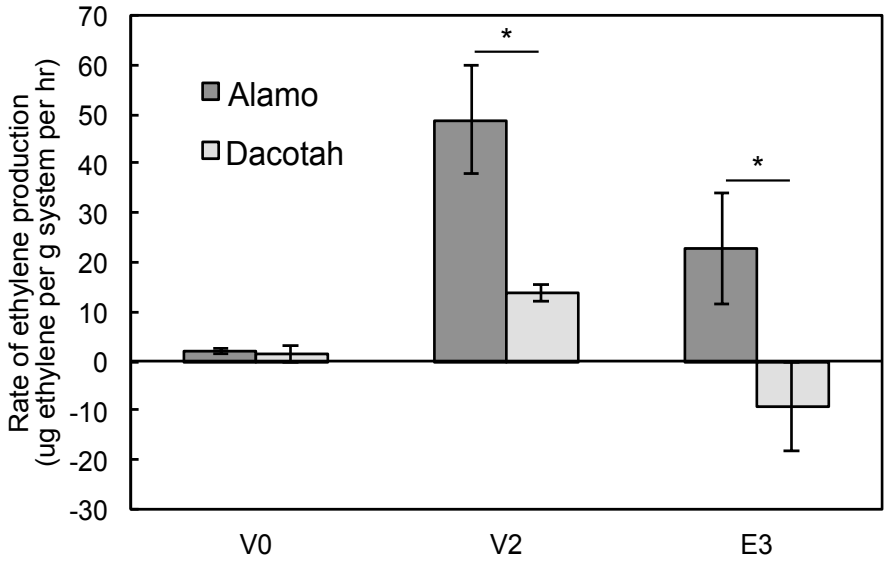


Figure 1: Rate of ethylene production (µg ethylene per g system hr^{-1}) in Alamo and Dacotah at three plant growth stages. The asterisk (*) indicates significant differences (p-value < 0.05) between the ethylene production rates of Alamo and Dacotah at the particular growth stage.

The cultivars showed differences in the free soluble amino acid profiles

Plant growth stages had a significant effect on the relative abundance of free soluble amino acids in the root-zone of the cultivars (adonis p-value = 0.001). Hence, further analysis was performed for each growth stage comparing the soluble amino acid profiles of Alamo and Dacotah. Adonis indicated that at stages V0 (p-value = 0.8) and E3 (p-value = 0.108) there were no differences in the relative content of root-zone soluble amino acids of cultivars. However, at stage V2, the relative abundance of soluble amino acids in the root-zones between Alamo and Dacotah were significantly different (adonis p-value = 0.04). The Mann-Whitney test showed that the cultivars had significantly (p-value < 0.05) different levels of the following amino acids in their root-zones: aspartic acid (Asp), glutamic acid (Glu), glutamine (Gln), glycine (Gly), and serine+aspargine (Ser+Asn) at V2; and arginine (Arg), histidine (His), and ornithine (Orn) at E3.

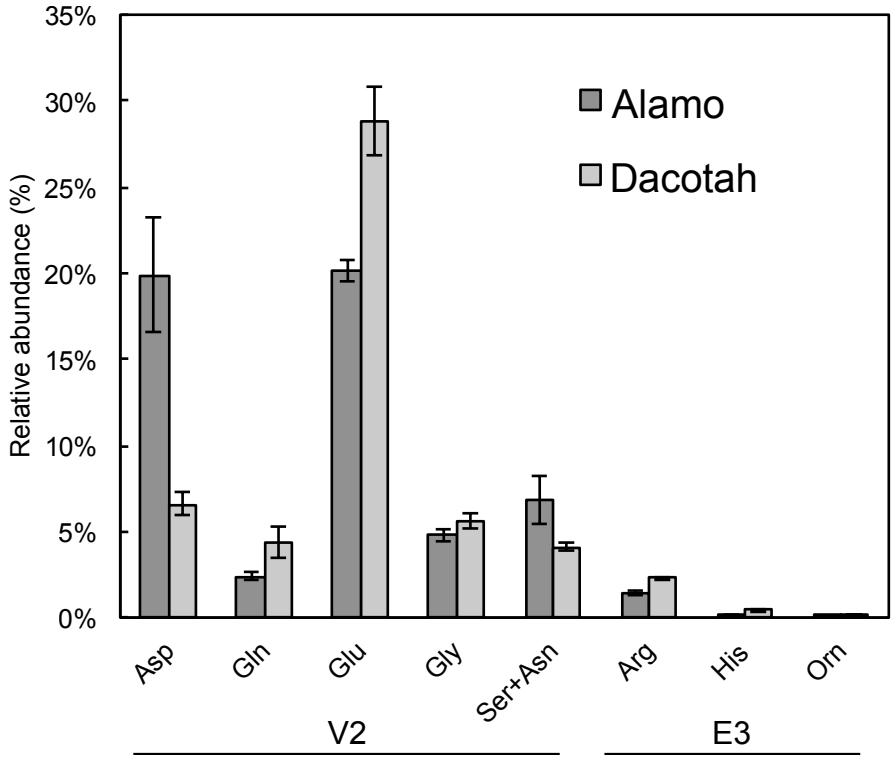


Figure 2: Relative abundance of amino acids that were significantly different (p-value < 0.05) between Alamo and Dacotah at stages V2 and E3.

Sequencing data

Removal of barcodes, adapters, and primers using cutadapt and stitching with PANDAseq gave ~ 1.9 million and ~ 1.7 million high quality 16S rRNA (bacterial) and ITS (fungal) gene sequence reads, respectively, for the 24 samples. The bacterial and fungal sequences, respectively, had average lengths of 333.2 and 254.3 bases with standard deviations of 57.7 and 48.7 bases. The raw data has been submitted to NCBI Sequence Read Archive according to MIMS standard. The bacterial dataset had total of 72,371 distinct OTUs with a total of 1,735,730 sequences (counts) assigned to these OTUs. The fungal dataset had total of 35,986 distinct OTUs with a total of 1,604,221 sequences (counts) assigned to these OTUs. The average Good's coverage of all samples was 92.5% and 97.0% for bacteria and fungi, respectively (Table S1). Table S1 contains the relevant statistics of the counts per sample and other alpha diversity metrics.

Alpha diversity

Using samples from all three growth stages and a sequence threshold (of 21,500 for bacteria and 16,000 for fungi) we compared the alpha diversity of groups with a two-tail nonparametric t-test with FDR correction using the collated alpha diversity file. The bacterial alpha diversity at stage E3 was greater than that at stage V2 (PD whole tree q-value = 0.003, chao1 q-value = 0.003). Furthermore, the bacterial alpha diversity in Dacotah at stage E3 was greater than that at stage V2 (observed species q-value = 0.045, chao1 q-value = 0.045).

The patterns of fungal alpha diversity resembled that of bacteria. The fungal alpha diversity at stage V2 was lesser than those at stages V0 (observed species q-value = 0.0135) and E3 (chao1 q-value = 0.003, observed species q-value = 0.003). Also, the fungal alpha diversity in Dacotah at stage E3 was greater than that at stage V2 (chao1 q-value = 0.03, observed species q-value = 0. 045). Certain patterns, however, were only present in the

fungal data. The alpha diversity in Alamo at stage E3 was greater than those at stage V2 (chao1 q-value = 0.045, observed species q-value = 0.0225) and stage V0 (chao1 q-value = 0.02).

Root-zone microbial communities structure (Beta diversity) different between stage and cultivars

Multivariate data analyses in QIIME using MRPP, ADONIS, and ANOSIM on the weighted UNIFRAC distance between samples, using all OTUs, detected significant differences (p-value < 0.05) in the bacterial communities as per stage and interaction of cultivar and stage (cultivar x stage), but not as per cultivar. These results were consistent when the analysis was performed on samples from (i) all three growth stages (V0, V2, E3), (ii) vegetative growth stages (V0 and V2), and (iii) vegetative (V2) and elongation (E3) stages (Table S2a).

PCoA (Fig. S2a) and NMDS (Fig. S2b) analyses of bacterial communities using weighted Unifrac and Bray-Curtis distances, respectively, showed that samples from Stages V2 and E3 clustered as per cultivars, however, samples from V0 did not. Hence, following analyses to compare cultivars were only performed at stages V2 and E3. *Proteobacteria, Actinobacteria,* and *Acidobacteria* were the most abundant phyla across samples (Fig. 3). Kruskal Wallis test with Benjamini Hochberg correction was performed on samples from stages V2 and E3 to identify significantly different (q-value < 0.05) phyla. For each growth stage, the significant phyla, except "Unassigned," were tested for differences (p-value < 0.05) in relative abundance between cultivars using a bootstrap Mann Whitney U test in QIIME. For easier illustration, the "Unassigned" phylum and phyla with total <1% abundance in samples across all growth stages were grouped as "Other" taxa and phyla showing differential relative abundance between cultivars from the above test are shown by stars (Fig.

3). Stage V2 had more phyla with differential relative abundance between cultivars. *Proteobacteria* was significantly different only at stage V2, however, it consistently showed higher relative abundance in Alamo across all plant growth stages.

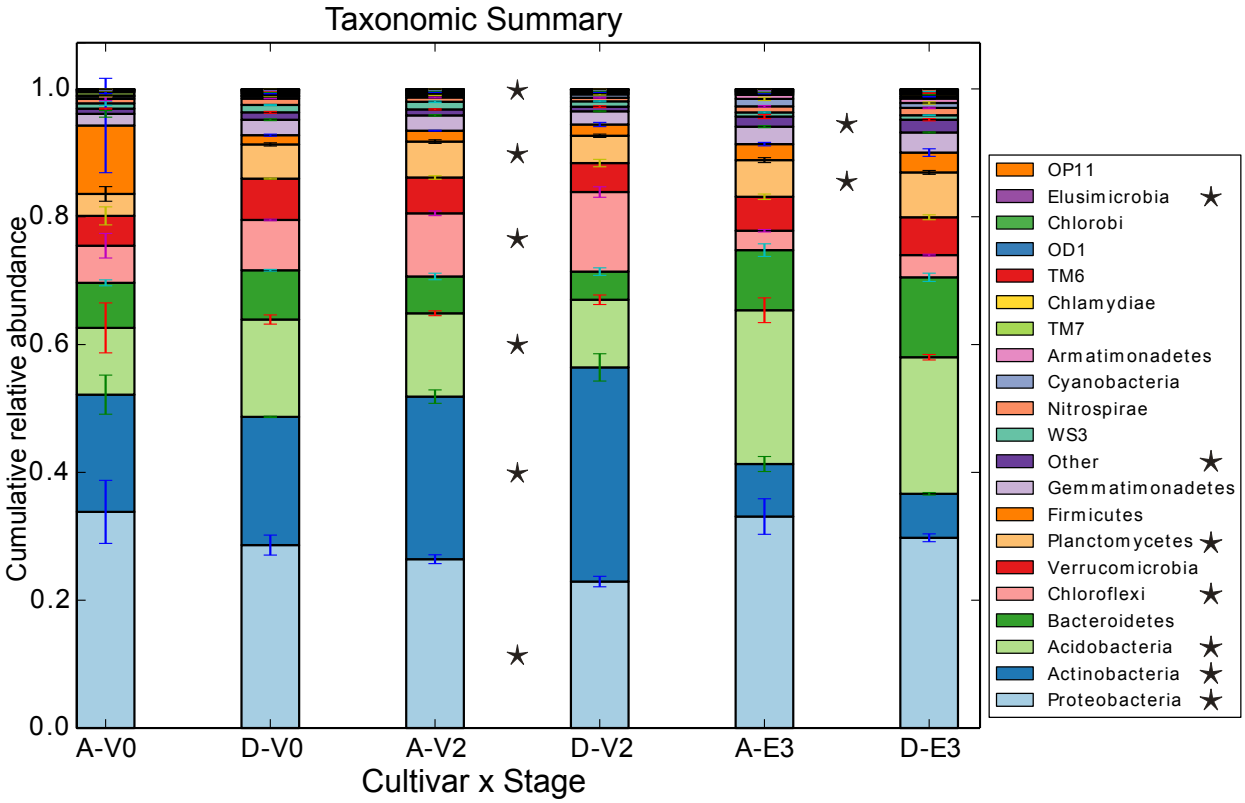


Figure 3: Phylum level taxonomic summaries of bacterial communities in cultivars at the different growth stages. Unassigned and less abundant taxa were grouped in "Other." Taxa in bars and legend from bottom to top are sorted as per decreasing abundance across all samples. Stars show phyla that are significantly different (q-value < 0.05) between cultivars at a specific growth stage. The significance was only calculated for stages V2 and E3.

Similar to the bacterial data, multivariate data analyses in QIIME using MRPP, ADONIS, and ANOSIM on the Bray-Curtis distance between samples, using all OTUs, detected significant differences (p-value < 0.05) in the fungal communities as per stage and interaction of cultivar and stage (Cultivar x Stage), but not as per cultivar (Table S2b). PCoA (Fig. S3) analyses of fungal communities using Bray-Curtis distances showed that samples from Stage V2 clustered as per cultivars, however, samples from V0 and E3 did not. *Ascomycota* was the most abundant phylum across all growth stages (Fig. S4).

Root-zone bacterial communities showed functional differences between cultivars

Indicator species analysis (ISA) was performed for the bacterial and fungal data to identify statistically significant indicator taxa (Table S3). Functional profiles were predicted using PICRUSt for the cultivars at the V2 and E3 stage before and after ISA. The predicted KEGG pathways in the metagenomes were different between cultivars at the corresponding growth stages for both before (Fig. S6) and after (Fig. S8) ISA. Before ISA, the KEGG pathway "N metabolism" showed significant differential abundance between cultivars at stage V2 (p-value = 0.039) but not at stage E3 (Fig S5). However, after ISA, the pathway was observed to have significant differential abundance between cultivars at stages V2 (p-value = 0.027) and E3 (p-value = 0.005) (Fig S7).

Cultivars differed in the root-zone bacterial interactions

For the bacterial data, V2 and E3 samples showed differences as per cultivar and hence, were used to generate bacterial COMET networks. In the bacterial COMET networks (Fig. 4), the number of classes (nodes) in Alamo and Dacotah were equal, whereas, the total number of bacterial interactions in Alamo (72) and Dacotah (129) were different, with 24 interactions present in both networks (colored in black). Alamo, however, had a higher number of modules (6) and percentage of positively correlated edges (56%) compared to the 3 and 46%, respectively, in Dacotah. For the fungal data, only samples from the V2 stage showed differences as per cultivar. Since MENAP software needs at least eight samples per group fungal COMET networks could not be generated.

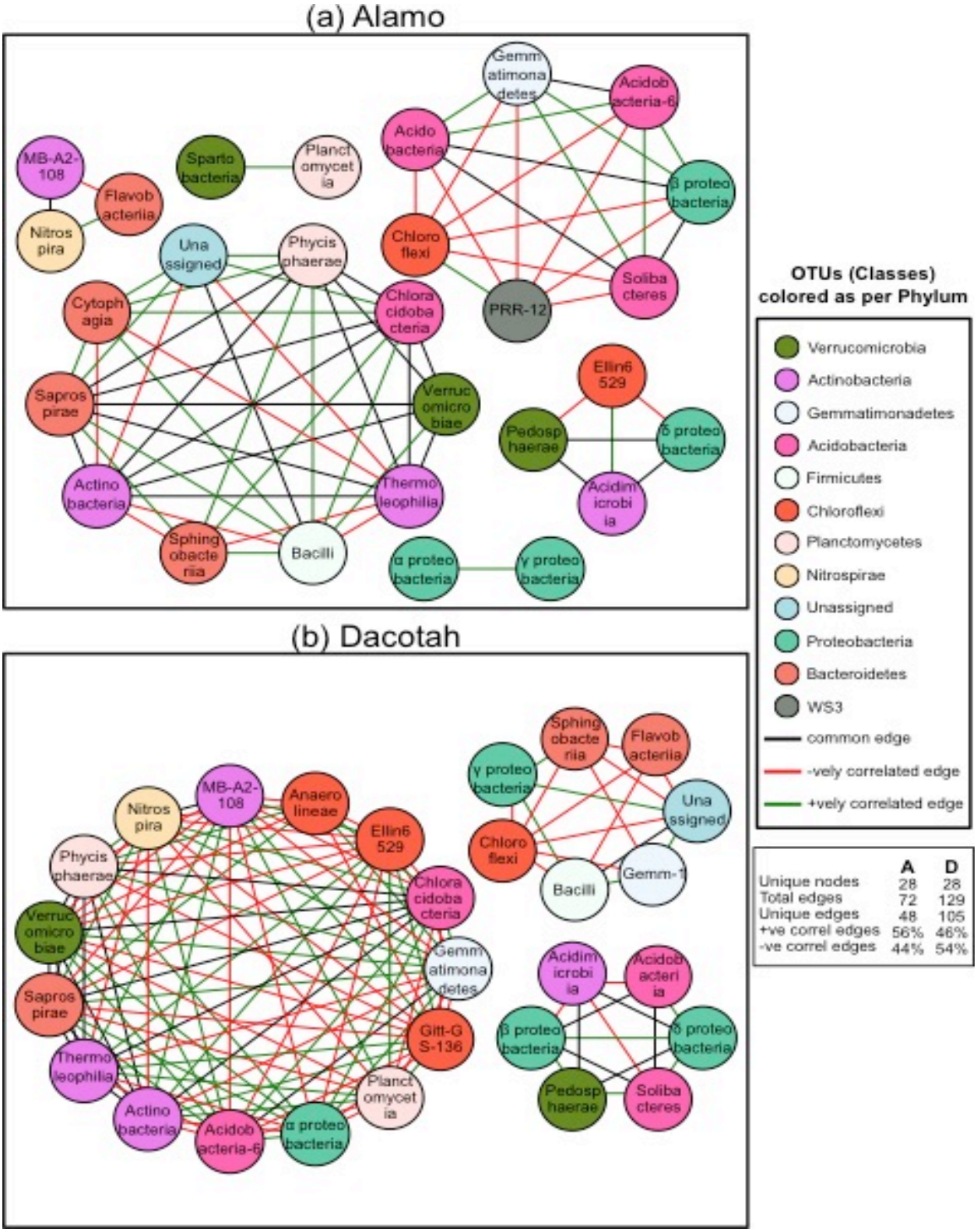


Figure 4: The COMET networks of bacterial interactions from root-zones of (a) Alamo and (b) Dacotah. The nodes represent classes and they are colored as per the phyla. The red and green colored edges represent negative, positive interactions respectively. A black colored edge represents that the interaction is present in both the Alamo and Dacotah networks.

Cultivable diazotrophs

After DNA extraction, 16S rRNA amplification, and Sanger sequencing of representative colonies cultivated on YM media we obtained 45 and 42 high quality reads for root-zone bacteria from Alamo and Dacotah, respectively. At stage V2, as per Libshuff analysis (Dacotah - Alamo p-value = 0.0054) and Pearson's Chi-squared test (p-value < $2.2e^{-16}$) (Fig. S10) we observed significant differences in the cultivable nitrogen-fixing bacteria (diazotrophs) between Alamo and Dacotah.

Discussion

The results of this study support the working quid pro quo model of plant-microbial interactions and the hypothesis that a higher productivity cultivar of switchgrass would be associated with a root-zone bacterial community with greater nitrogenase activity. The potential for greater carbon flow in exchange for bacterial fixed nitrogen provides a consistent framework to describe plant-bacterial interactions associated with switchgrass. The differences observed in the composition of amino acids, structure, interaction networks, and the general functional capacity (e.g., nitrogen metabolism) of the bacterial communities associated with the two cultivars, further support the above hypothesis. The results reinforce the idea that plant-diazotroph interactions and feedbacks (communications) are important descriptors of the greater nitrogen fixing potential of the high productivity Alamo, compared to the Dacotah ecotype. Root zone diazotroph communities, therefore, may help to meet the nitrogen demands of feedstocks and other grasses to support greater plant biomass production. The experimental outcomes also support the idea that molecular breeding of switchgrass can be used to manipulate plant-bacterial interaction.

Plant N demand a driver of switchgrass-diazotrophic interaction

BNF rates changed with growth stage and may reflect the temporal nitrogen demands of cultivars (Dong *et al.*, 2001, Scagel *et al.*, 2007). The increased BNF rates during the vegetative stages might be to meet the N demands of increasing biomass. On the other hand, the reduction of BNF rates during the elongation stage might be due the nearing of maturation and the reduction of plant nitrogen demand, especially with its ability to translocate nitrogen to its roots for conservation during winter and reuse during regrowth (Monti, 2012). While it cannot be unequivocally concluded that the higher activities of nitrogen fixing bacteria are a result of Alamo having greater productivity than Dacotah, the results are consistent with this idea and track expected growth stage changes in plant nitrogen demand and carbon flow to below ground roots (Qian *et al.*, 1997, Voisin *et al.*, 2003).

Plant-microbial amino acids in the root-zone

Plant root exudates, especially amino acids, are used as signals of communication between plant and the rhizosphere microbiome (Badri *et al.*, 2009, Morgan *et al.*, 2005). The amino acids glycine, glutamate, alanine, aspartate, glutamine, and asparagine were found to be the dominant soluble amino acids in relatively mature switchgrass root-zone soil habitats, (Figures 2 and S1), however, the contributions of these amino acids were altered with plant growth stage. The flux rates of these same molecules as exudates from plant roots have previously been shown to be relatively high (Carvalhais *et al.*, 2013, Carvalhais *et al.*, 2011, Lesuffleur *et al.*, 2007) but can vary due to fertilization (e.g. N, Fe), and plant maturation. The relative concentration of glycine, in particular was much higher in the root-zone of mature (E3) relative to young switchgrass plants (V0 and V2). There may thus be important changes in amino acid profiles that arise during plant maturation. The observations of amino acids in the root-zone of switchgrass represent a more complex habitat than studies that focus

on axenic, microbial-free root exudation. It is thus difficult to come to firm conclusions regarding the biological role that amino acid dynamics might play in the root-zone. The variation across cultivar and growth stage, nevertheless are consistent with the idea that root-zone amino acids reflect the contributions of an interactive plant-microbial system.

Asparagine and glutamine tend to be relatively low in the rhizosphere, but can vary between <1 and ~6% of the amino acid pool (Lesuffleur *et al.*, 2007, Paynel *et al.*, 2001). Intracellular concentrations within the roots of numerous plants (Lesuffleur *et al.*, 2007, Lodwig *et al.*, 2003, Paynel *et al.*, 2001) of glutamine and asparagine tend to be relatively high, especially within phloem and xylem. Because of the importance of nitrogen rich asparagine and the plant N shuttle glutamine, there is likely to be conservation and influx of these amino acids into organism biomass rather than efflux, particularly under N limiting growth conditions. Hence, the relatively low concentrations of these two amino acids in the root-zone of switchgrass, relative to their constituent acids, aspartic and glutamic acid, may be the result of N limiting growth conditions.

An amino acid feedback model has been proposed from the extensive study of legume (family Fabaceae) nodules (Dong *et al.*, 2001). However, the application of this model to plant roots and associative diazotrophic bacteria in the rhizosphere remains to be determined. The rhizosphere (associative BNF system) is a different habitat than that of a nodule (symbiotic BNF system), however, application of this model provides a useful starting point for understanding interactions and exchanges between plants and bacteria in the root zone. Glutamic- and aspartic- acid, and the associated bases of glutamine and asparagine appear to provide a feedback signal that helps to control nitrogen fixation in nodules of legumes (Lodwig *et al.*, 2003). These amino acids may thus be an important “amino-stat” that helps to regulate the exchange of plant-bacterial carbon and nitrogen.

We used this nitrogen quid pro quo model to describe the dominant and dynamic amino acid species in the root zone of the switchgrass ecotypes. We observed that Aspartic acid was ~3 fold greater in Alamo than Dacotah, and conversely, glutamic acid, was ~1.4 fold greater in Dacotah than Alamo. These differences could reflect different regulatory interactions between the plant and bacteria, perhaps related to differences in plant N demand and the availability of reduced nitrogen from diazotrophs.

Assuming that the different amino acid concentrations reflect differences in the flows of these molecules between plant and bacterial biomass, the high levels of aspartic acid in root-zone of Alamo could be interpreted as the result of higher rates of nitrogen fixation by diazotrophs. The greater flow of this bacterial-supplied N transporter could help meet the higher plant nitrogen demand of Alamo. In contrast, the relatively greater glutamic acid concentrations in the root-zone of Dacotah may reflect a build up of this amino acid as a consequence of relatively low rates of nitrogen fixation. The results are also concordant with the observation that glutamic acid is secreted at lower levels by nitrogen fixing bacteria under diazotrophic compared to adiazotrophic conditions (Gonzalez-Lopez *et al.*, 1995). Reductions in glutamine and asparagine, and elevated aspartate to asparagine ratios, on the other hand, have been shown to be a response to low N availability that is translated as deficiency or stress within the plant (Amarante & Sodek, 2006) and thus help to describe plant N demand (Atkins *et al.*, 1983, Lea *et al.*, 2007, Pate *et al.*, 1981, Pate *et al.*, 1984, Pate *et al.*, 1980). Though we did not measure concentrations of amino acids within plant tissues, differences in these ratios in the root-zone may help to provide information explaining the greater nitrogen fixation potential associated with Alamo relative to Dacotah.

Changes in the structure, functions, and interactions of rhizosphere microbiome

The broad differences in bacterial (Figure 3, Figure S2, and Libshuff results) and fungal communities (Figures S3) (Table S2) in the root-zone of Alamo and Dacotah may be linked to the above described changes in root exudates (e.g. amino acids), nitrogen fixation potential, and plant-microbial interactions (Bürgmann *et al.*, 2005, Mao *et al.*, 2014). Different plant cultivars and growth stages have previously been shown to impact microbial community structure in plant root zones (Berg & Smalla, 2009, Chaparro *et al.*, 2014, Chaudhary *et al.*, 2012, Mao *et al.*, 2011). The research reported in this manuscript build upon these differences to show specific changes in microbial communities, potentially mediated through plant-microbial signals (amino acids), and concomitant functional changes in the root-zone.

Bacterial nitrogen fixation potential tracked changes in the structure and metagenomic functional potential of the root-zone microbial community. The increased N metabolism potential in Alamo compared to Dacotah may reflect the overall differences in the nitrogen status of the ecotypes. The phylogenetic dominance of *Proteobacteria, Actinobacteria, Acidobacteria,* and *Firmicutes* observed in our results from culture dependent and independent methods across samples (Figure 3) have been observed previously in roots of switchgrass (Bahulikar *et al.*, 2014, Jesus *et al.*, 2010, Jesus *et al.*, 2016, Mao *et al.*, 2011, Plecha *et al.*, 2013). At stage V2, these phyla are among those that show differential abundances between ecotypes (Figure 3), suggesting the potential importance of phylogenetic changes for driving functional level (Figures S6 and S8) shifts, such as that observed for the relative increase in N-metabolism potential in Alamo relative to Dacotah. The relative abundances of *Proteobacteria* and *Actinobacteria,* for example, are indicators of Alamo and Dacotah (Table S3), respectively. The observations from our study (Table S3, Figures S5 and

S7) are consistent with the differential role that N-cycling bacteria play in the growth of switchgrass ecotypes (Mao *et al.*, 2013, Mao *et al.*, 2011).

The ecological niches of the two switchgrass cultivars also provide a framework for understanding the potential for different types of root-zone bacterial interactions in relation to the greater nitrogen fixation rates. The differences in bacterial interactions of COMET networks between cultivars point out significantly different root-zone community dynamics among the ecotypes (Figure 4). The common edges between Alamo and Dacotah COMET networks are interpreted as the core set of bacterial interactions that are associated with switchgrass, and with the unique edges being cultivar specific. The higher number of modules and greater proportion of positive interactions in Alamo hint at the potential of bacteria-bacteria and plant-bacteria interactions that could serve to help the plant achieve higher productivity. It is not known if these interactions reflect differences in the diazotroph root-zone communities of the ecotypes. Nevertheless, the model is useful for identifying positive feedbacks between and among plants and microorganisms.

Positive feedbacks between plants and microbes are often driven by environmental and nutrient limitations, and thus provide a working hypothesis to explain how two ecotype root-zone habitats with different energy and nutrient conditions may develop different microbial communities and plant-microbial interactions (Badri *et al.*, 2009, Berg, 2009, Morgan *et al.*, 2005). The *Proteobacteria*, for example, have diazotrophic bacterial members, and could help to explain their greater relative abundance and greater nitrogen fixation potential in Alamo relative to Dacotah. Similarly, the positive association of Bacilli with many other taxa within the Alamo bacterial network is consistent with a growth supportive role played by these bacteria. Presumably, the surrounding taxa that benefit would in return provide resources that support the growth of the Bacilli. While the diversity of interactions in root-zone and soil habitats is staggering, and description of them still in scientific infancy, the

networks described herein are useful for understanding, predicting and confirming possible positive interactions among organisms.

Conclusion

Though there is a scientific agreement regarding the importance of microbial interactions in the rhizosphere, the research presented is relatively unique in that it simultaneously studies the microbial structure, interactions, and functions. The research highlights that changes in the plant's nitrogen demands are associated with plant-microbial communication, which results in changes in the rhizosphere microbial communities. Understanding the interplay between the diazotroph communities with associated plants offer opportunities for genetic engineering and breeding of feedstock grasses and grain crops to select varieties that can better manage the microbial communities to satisfy their nitrogen demands.

Acknowledgements

We thank Virginia Tech's Genetics, Bioinformatics, and Computational Biology, Department of Horticulture, Multicultural Academic Opportunities, and Graduate Student Assembly for providing research, travel, and personnel funding. The authors sincerely thank Amanda Karstetter, Bronte Lantin, Brandi Edwards, Haley Feazel-Orr, Kelsey Weber, Deepak Poudel, Solmaz Eskandari, and Samantha Fenn for help with sample preparation for microbial analyses. Thanks to Dr. Li Ma for help with the soluble amino acid experiments. The authors acknowledge Virginia Polytechnic Institute and State University's Open Access Subvention Fund. Sincere thanks to Dr. Birgit Scharf for providing *S. melliloti* 1021.

References

Abarenkov K, Henrik Nilsson R, Larsson KH *et al.* (2010) The UNITE database for molecular identification of fungi--recent updates and future perspectives. New Phytol, **186**, 281-285.

Amarante L, Sodek L (2006) Waterlogging effect on xylem sap glutamine of nodulated soybean. Biologia Plantarum, **50**, 405-410.

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecology, **26**, 32-46.

Atkins CA, Pate JS, Peoples MB, Joy KW (1983) Amino Acid transport and metabolism in relation to the nitrogen economy of a legume leaf. Plant Physiol, **71**, 841-848.

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

Bahulikar RA, Torres-Jerez I, Worley E, Craven K, Udvardi MK (2014) Diversity of nitrogen-fixing bacteria associated with switchgrass in the native tallgrass prairie of northern Oklahoma. Appl Environ Microbiol, **80**, 5636-5643.

Beals EW (1984) Bray-Curtis Ordination: An Effective Strategy for Analysis of Multivariate Ecological Data. **14**, 1-55.

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

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

Boddey RM, Dobereiner J (1988) Nitrogen fixation associated with grasses and cereals: Recent results and perspectives for future research. Plant and Soil, **108**, 53-65.

Bosch L, Alegria A, Farre R (2006) Application of the 6-aminoquinolyl-N-hydroxysccinimidyl carbamate (AQC) reagent to the RP-HPLC determination of amino acids in infant foods. J Chromatogr B Analyt Technol Biomed Life Sci, **831**, 176-183.

Brankatschk R, Bodenhausen N, Zeyer J, Burgmann H (2012) Simple absolute quantification method correcting for quantitative PCR efficiency variations for microbial community samples. Appl Environ Microbiol, **78**, 4481-4489.

Bürgmann H, Meier S, Bunge M, Widmer F, Zeyer J (2005) Effects of model root exudates on structure and activity of a soil diazotroph community. Environmental Microbiology, **7**, 1711-1724.

Caporaso G, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Meth, **7**, 335-336.

Carvalhais LC, Dennis PG, Fan B *et al.* (2013) Linking plant nutritional status to plant-microbe interactions. PLoS ONE, **8**, e68555.

Carvalhais LC, Dennis PG, Fedoseyenko D, Hajirezaei M-R, Borriss R, Von Wirén N (2011) Root exudation of sugars, amino acids, and organic acids by maize as affected by nitrogen, phosphorus, potassium, and iron deficiency. Journal of Plant Nutrition and Soil Science, **174**, 3-11.

Chalk PM (1991) The contribution of associative and symbiotic nitrogen fixation to the nitrogen nutrition of non-legumes. Plant and Soil, **132**, 29-39.

Chaparro JM, Badri DV, Vivanco JM (2014) Rhizosphere microbiome assemblage is affected by plant development. ISME J, **8**, 790-803.

Chaudhary D, Saxena J, Lorenz N, Dick L, Dick R (2012) Microbial Profiles of Rhizosphere and Bulk Soil Microbial Communities of Biofuel Crops Switchgrass (Panicum

virgatum L.) and Jatropha (Jatropha curcas L.). Applied and Environmental Soil Science, **2012**, 1-6.

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology, **18**, 117-143.

Day JM, Neves MCP, Döbereiner J (1975) Nitrogenase activity on the roots of tropical forage grasses. Soil Biology and Biochemistry, **7**, 107-112.

Deng Y, Jiang YH, Yang Y, He Z, Luo F, Zhou J (2012) Molecular ecological network analyses. BMC Bioinformatics, **13**, 113.

Desantis TZ, Hugenholtz P, Larsen N *et al.* (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol, **72**, 5069-5072.

Dong S, Scagel CF, Cheng L, Fuchigami LH, Rygiewicz PT (2001) Soil temperature and plant growth stage influence nitrogen uptake and amino acid concentration of apple during early spring growth. Tree Physiology, **21**, 541-547.

Dufrene M, Legendre P (1997) Species Assemblages and Indicator Species:The Need for a Flexible Asymmetrical Approach. Ecological Monographs, **67**, 345-366.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics, **26**, 2460-2461.

Follett RF, Vogel KP, Varvel GE, Mitchell RB, Kimble J (2012) Soil Carbon Sequestration by Switchgrass and No-Till Maize Grown for Bioenergy. BioEnergy Research, **5**, 866-875.

Fredriksson NJ, Hermansson M, Wilen BM (2013) The choice of PCR primers has great impact on assessments of bacterial community diversity and dynamics in a wastewater treatment plant. PLoS ONE, **8**, e76431.

Galloway JN, Aber JD, Erisman JW, Seitzinger SP, Howarth RW, Cowling EB, Cosby BJ (2003) The Nitrogen Cascade. Bioscience, **53**, 341.

Gonzalez-Lopez J, Martinez-Toledo MV, Rodelas B, Pozo C, Salmeron V (1995) Production of amino acids by free-living heterotrophic nitrogen-fixing bacteria. Amino Acids, **8**, 15-21.

Gower JC (2005) Principal Coordinates Analysis. In: *Encyclopedia of Biostatistics.* pp Page. The Open University, Milton Keynes, UK, John Wiley and Sons, Ltd.

Grubbs FE (1969) Procedures for Detecting Outlying Observations in Samples. Technometrics, **11**, 1-21.

Hargreaves SK, Williams RJ, Hofmockel KS (2015) Environmental Filtering of Microbial Communities in Agricultural Soil Shifts with Crop Growth. PLoS ONE, **10**, e0134345.

Hou S, He H, Zhang W, Xie H, Zhang X (2009) Determination of soil amino acids by high performance liquid chromatography-electro spray ionization-mass spectrometry derivatized with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate. Talanta, **80**, 440-447.

Jesus, Susilawati E, Smith S *et al.* (2010) Bacterial Communities in the Rhizosphere of Biofuel Crops Grown on Marginal Lands as Evaluated by 16S rRNA Gene Pyrosequences. BioEnergy Research, **3**, 20-27.

Jesus EDC, Liang C, Quensen JF, Susilawati E, Jackson RD, Balser TC, Tiedje JM (2016) Influence of corn, switchgrass, and prairie cropping systems on soil microbial communities in the upper Midwest of the United States. GCB Bioenergy, **8**, 481-494.

Ker K, Seguin P, Driscoll BT, Fyles JW, Smith DL (2012) Switchgrass establishment and seeding year production can be improved by inoculation with rhizosphere endophytes. Biomass and Bioenergy, **47**, 295-301.

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glockner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res, **41**, e1.

Lane DJ (1991) 16S/23S rRNA sequencing. In: *Nucleic acid techniques in bacterial systematics*. (eds Stackebrandt E, Goodfellow M) pp Page. New York, NY, John Wiley and Sons.

Langille MG, Zaneveld J, Caporaso JG *et al.* (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol, **31**, 814-821.

Lea PJ, Sodek L, Parry MaJ, Shewry PR, Halford NG (2007) Asparagine in plants. Annals of Applied Biology, **150**, 1-26.

Ledgard SF, Steele KW (1992) Biological nitrogen fixation in mixed legume/grass pastures. Plant and Soil, **141**, 137-153.

Lesuffleur F, Paynel F, Bataillé M-P, Le Deunff E, Cliquet J-B (2007) Root amino acid exudation: measurement of high efflux rates of glycine and serine from six different plant species. Plant and Soil, **294**, 235-246.

Li X, Rui J, Mao Y, Yannarell A, Mackie R (2014) Dynamics of the bacterial community structure in the rhizosphere of a maize cultivar. Soil Biology and Biochemistry, **68**, 392-401.

Lodwig EM, Hosie AH, Bourdes A *et al.* (2003) Amino-acid cycling drives nitrogen fixation in the legume-Rhizobium symbiosis. Nature, **422**, 722-726.

Lowman S, Kim-Dura S, Mei C, Nowak J (2015) Strategies for enhancement of switchgrass (Panicum virgatum L.) performance under limited nitrogen supply based on utilization of N-fixing bacterial endophytes. Plant and Soil, 1-17.

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

Lozupone CA, Hamady M, Kelley ST, Knight R (2007) Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Appl Environ Microbiol*,* **73**, 1576-1585.

Ma Z (1999) Carbon sequestration by switchgrass. Unpublished PhD AUBURN UNIVERSITY, 107 pp.

Mao Y, Li X, Smyth E, Yannarell A, Mackie R (2014) Enrichment of specific bacterial and eukaryotic microbes in the rhizosphere of switchgrass (Panicum virgatum L.) through root exudates. Environmental Microbiology Reports*,* **6**, 13.

Mao Y, Yannarell A, Davis S, Mackie R (2013) Impact of different bioenergy crops on N-cycling bacterial and archaeal communities in soil. Environmental Microbiology*,* **15**, 928-942.

Mao Y, Yannarell A, Mackie R (2011) Changes in N-Transforming Archaea and Bacteria in Soil during the Establishment of Bioenergy Crops. PLoS ONE*,* **6**, e24750.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal*,* **17**, 10.

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

Mcdonald D, Price MN, Goodrich J *et al.* (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J*,* **6**, 610-618.

Mielke PW (1984) Meteorological applications of permutation techniques based on distance functions. In: *Handbook of statistics: Nonparametric methods.* (eds Krishnaiah PR, Sen PK) pp Page. Amsterdam: North-Holland.

Miller HJ, Henken G, Veen JaV (1989) Variation and composition of bacterial populations in the rhizospheres of maize, wheat, and grass cultivars. Canadian Journal of Microbiology, **35**, 656-660.

Monti A (2012) *Switchgrass: A Valuable Biomass Crop for Energy*.

Moore KJ, Moser LE, Vogel KP, Waller SS, Johnson BE, Pedersen JF (1991) Describing and Quantifying Growth-Stages of Perennial Forage Grasses. Agronomy Journal, **83**, 1073-1077.

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

Newman MEJ (2006) Finding community structure in networks using the eigenvectors of matrices. Physical Review E, **74**.

Oenema O, Witzke HP, Klimont Z, Lesschen JP, Velthof GL (2009) Integrated assessment of promising measures to decrease nitrogen losses from agriculture in EU-27. Agriculture, Ecosystems \& Environment, **133**, 280-288.

Pate JS, Atkins CA, Herridge DF, Layzell DB (1981) Synthesis, Storage, and Utilization of Amino Compounds in White Lupin (Lupinus albus L.). Plant Physiol, **67**, 37-42.

Pate JS, Atkins CA, Layzell DB, Shelp BJ (1984) Effects of N2 Deficiency on Transport and Partitioning of C and N in a Nodulated Legume. Plant Physiology, **76**, 6.

Pate JS, Atkins CA, White ST, Rainbird RM, Woo KC (1980) Nitrogen Nutrition and Xylem Transport of Nitrogen in Ureide-producing Grain Legumes. Plant Physiol, **65**, 961-965.

Paynel F, J Murray P, Bernard Cliquet J (2001) Root exudates: a pathway for short-term N transfer from clover and ryegrass. Plant and Soil, **229**, 235-243.

Plecha S, Hall D, Tiquia-Arashiro SM (2013) Screening for novel bacteria from the bioenergy feedstock switchgrass (Panicum virgatum L.). Environ Technol, **34**, 1895-1904.

Poly F, Monrozier LJ, Bally R (2001) Improvement in the RFLP procedure for studying the diversity of nifH genes in communities of nitrogen fixers in soil. Res Microbiol, **152**, 95-103.

Pons P, Latapy M (2005) Computing Communities in Large Networks Using Random Walks. **3733**, 284-293.

Porter CL (1966) An Analysis of Variation Between Upland and Lowland Switchgrass, Panicum Virgatum L., in Central Oklahoma. Ecology, **47**, 980.

Qian JH, Doran JW, Walters DT (1997) Maize plant contributions to root zone available carbon and microbial transformations of nitrogen. Soil Biology and Biochemistry, **29**, 1451-1462.

Rodrigues RR, Pineda RP, Barney JN, Nilsen ET, Barrett JE, Williams MA (2015) Plant Invasions Associated with Change in Root-Zone Microbial Community Structure and Diversity. PLoS ONE, **10**, e0141424.

S.Castle, Barger N (2010) Acetylene Reduction Assay (ARA): Measuring Nitrogenase Activity. In: *Lab Protocol.* pp Page, Department of ecology and evolutionary biology, University of Colorado at Boulder, Aridlands Ecology Lab.

Scagel CF, Bi G, Fuchigami LH, Regan RP (2007) Seasonal Variation in Growth, Nitrogen Uptake and Allocation by Container-grown Evergreen and Deciduous Rhododendron Cultivars. HortScience, **42**, 1440-1449.

Schloss PD, Larget BR, Handelsman J (2004) Integration of microbial ecology and statistics: a test to compare gene libraries. Appl Environ Microbiol, **70**, 5485-5492.

Singleton DR, Furlong MA, Rathbun SL, Whitman WB (2001) Quantitative Comparisons of 16S rRNA Gene Sequence Libraries from Environmental Samples. Applied and Environmental Microbiology, **67**, 4374-4376.

Skinner H (2009) Sequestration Potential of Switchgrass Managed for Bioenergy Production. pp Page, University Park, PA, ARS Pasture Systems and Watershed Management Research Unit.

Smil V (1999a) Detonator of the population explosion. Nature, **400**, 415-415.

Smil V (1999b) Nitrogen in crop production: An account of global flows. Global Biogeochemical Cycles, **13**, 647-662.

Smil V (2002) Nitrogen and food production: proteins for human diets. Ambio, **31**, 126-131.

Smith DP, Peay KG (2014) Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. PLoS ONE, **9**, e90234.

Spitzer M, Wildenhain J, Rappsilber J, Tyers M (2014) BoxPlotR: a web tool for generation of box plots. Nat Methods, **11**, 121-122.

Sutton MA, Howard CM, Erisman JW *et al.* (2011a) *The European Nitrogen Assessment: Sources, Effects and Policy Perspectives*, Cambridge University Press.

Sutton MA, Oenema O, Erisman JW, Leip A, Van Grinsven H, Winiwarter W (2011b) Too much of a good thing. Nature, **472**, 159-161.

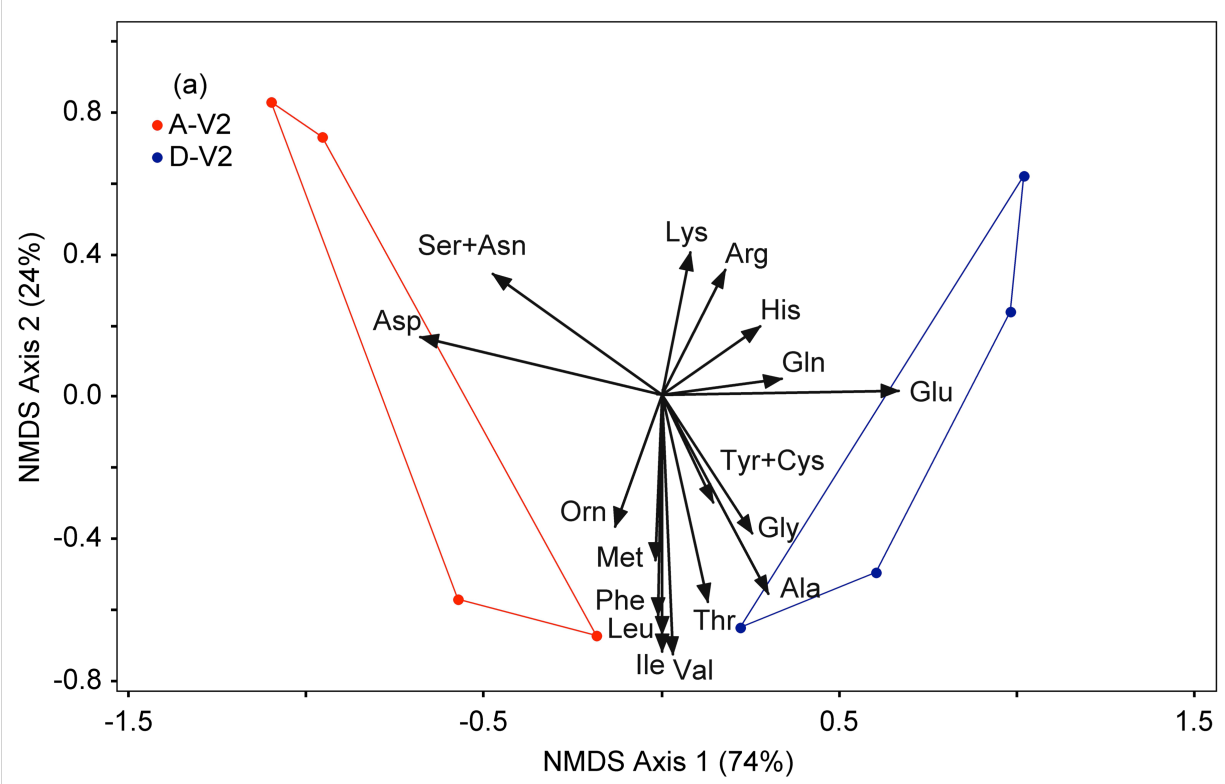
Voisin AS, Salon C, Jeudy C, Warembourg FR (2003) Root and nodule growth in Pisum sativum L. in relation to photosynthesis: analysis using 13C-labelling. Ann Bot, **92**, 557-563.

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

Weier KL (1980) Nitrogen fixation associated with grasses. Tropical Grasslands, **14**.

Wewalwela J (2014) Associative nitrogen-fixing bacteria and their potential to support the growth of bioenergy grasses on marginal lands. Unpublished Ph.D. Mississippi State University.

Xia Y, Greissworth E, Mucci C, Williams MA, De Bolt S (2013) Characterization of culturable bacterial endophytes of switchgrass (Panicum virgatumL.) and their capacity to influence plant growth. GCB Bioenergy, **5**, 674-682.



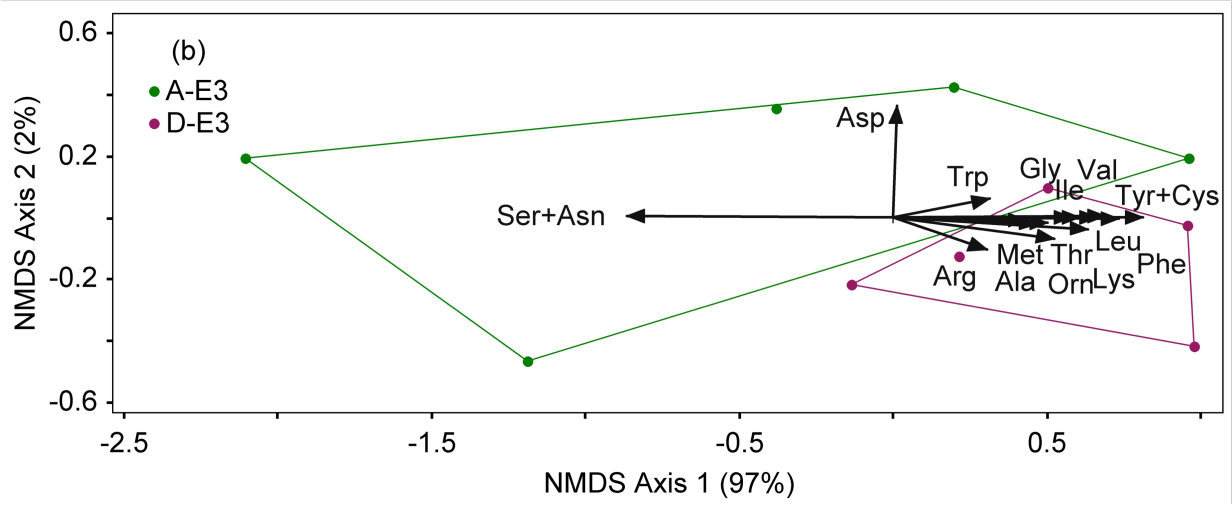


Figure S1: Non-metric multi-dimensional scaling plots of the free soluble amino acid profiles in the root-zones of the cultivars at stages (a) V2 and (b) E3.

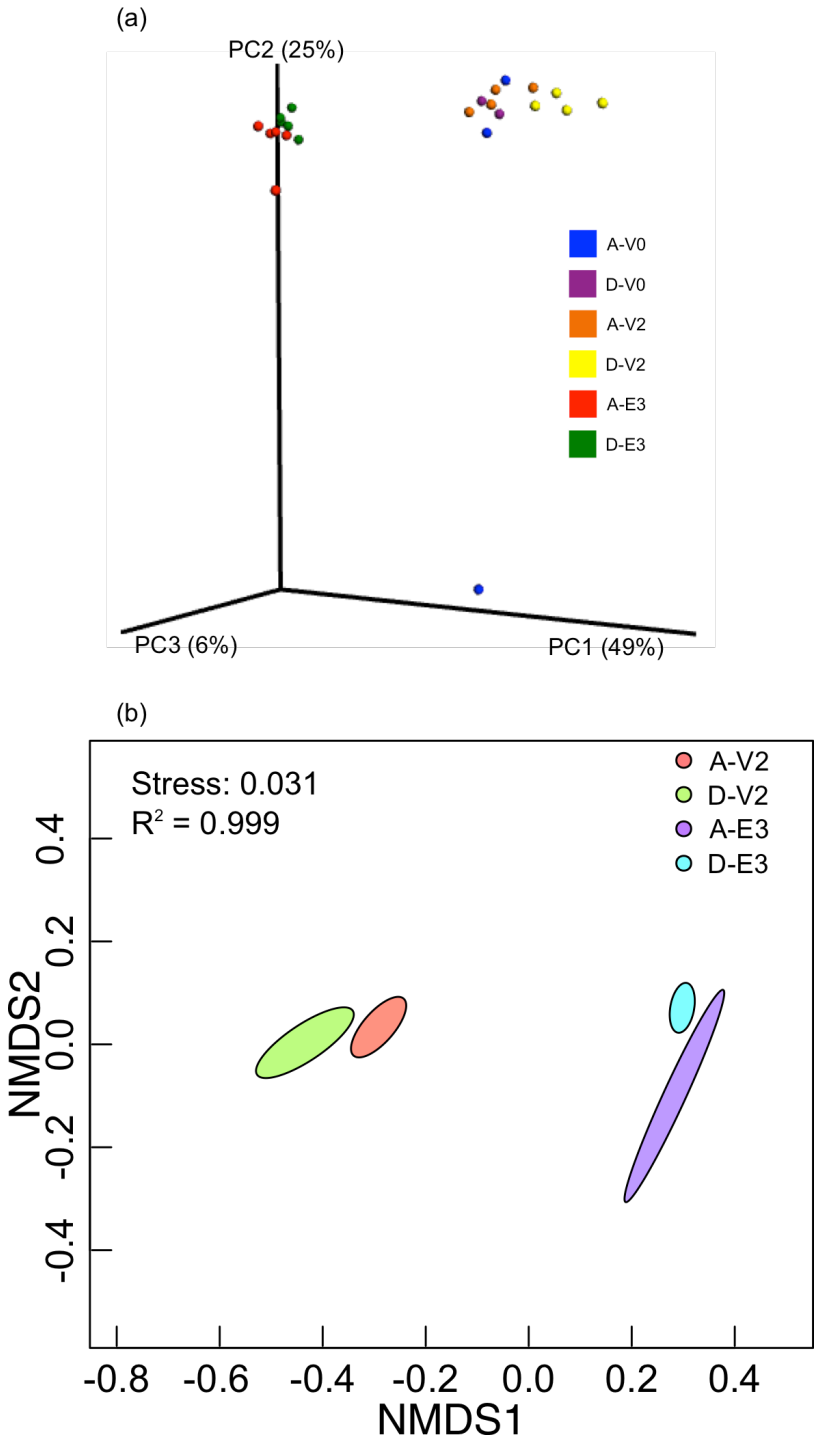


Figure S2: (a) PCoA and (b) Non-metric multi-dimensional scaling showing differences in bacterial communities as per plant growth stage and cultivars.

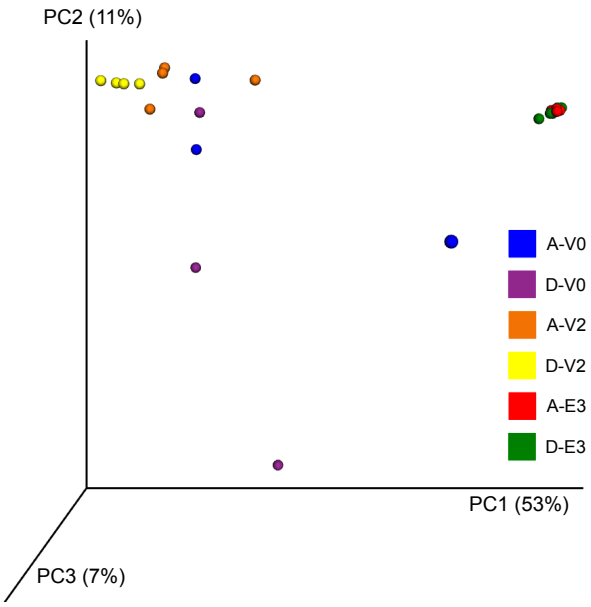


Figure S3: PCoA showing differences in fungal communities as per plant growth stage and cultivars.

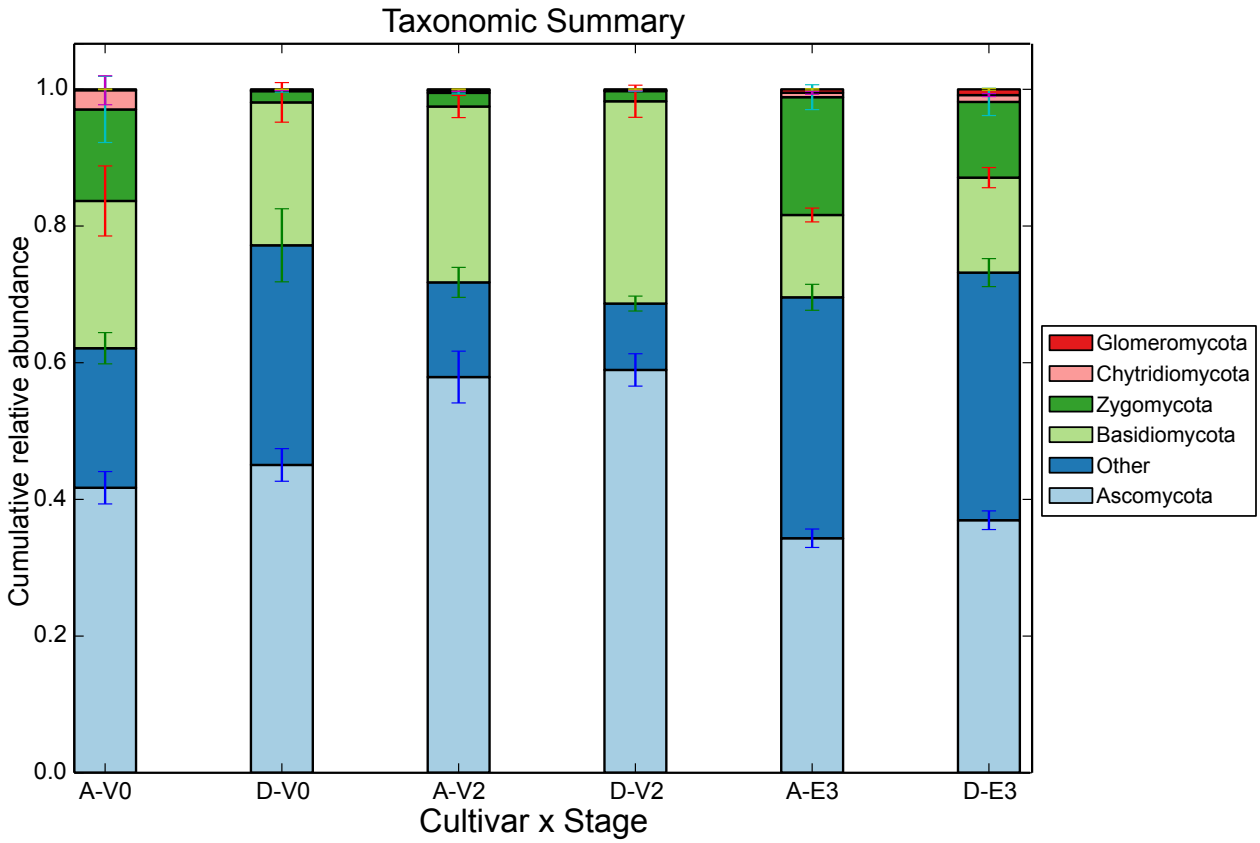


Figure S4: Phylum level taxonomic summaries of fungal communities in cultivars at the different growth stages. In the legend, "Unassigned" and less abundant taxa were grouped in "Other." Taxa in bars and legend from bottom to top are sorted as per decreasing abundance across all samples.

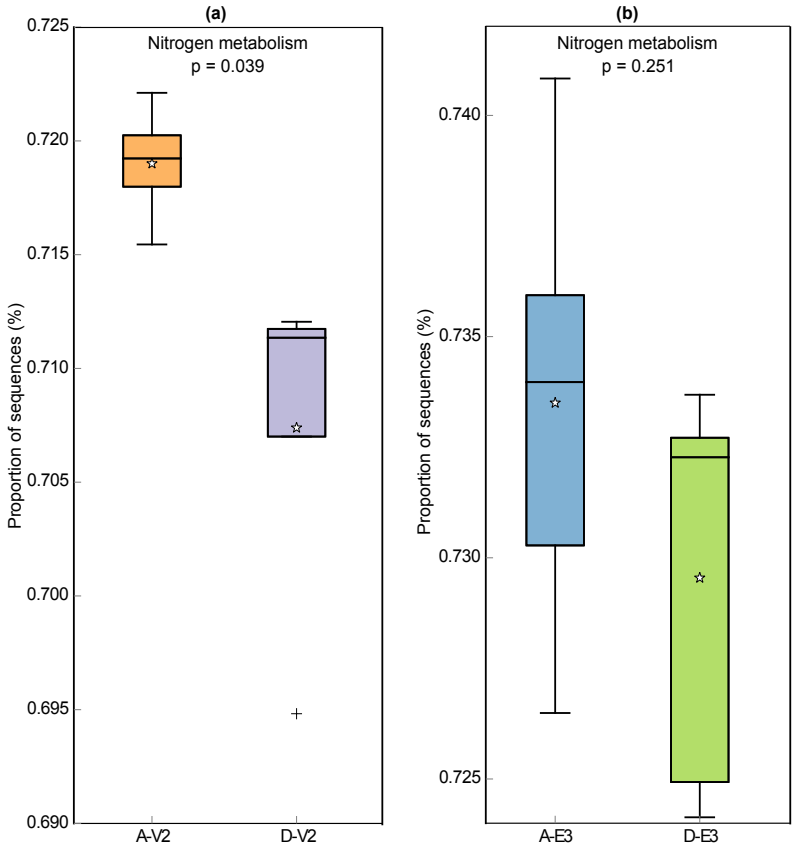


Figure S5: Pre-ISA abundance of the "N-metabolism" pathway in the cultivars at stages (a) V2 and (b) E3. Two-sided t-test (p-value < 0.05) was used to estimate differential abundance between cultivars.

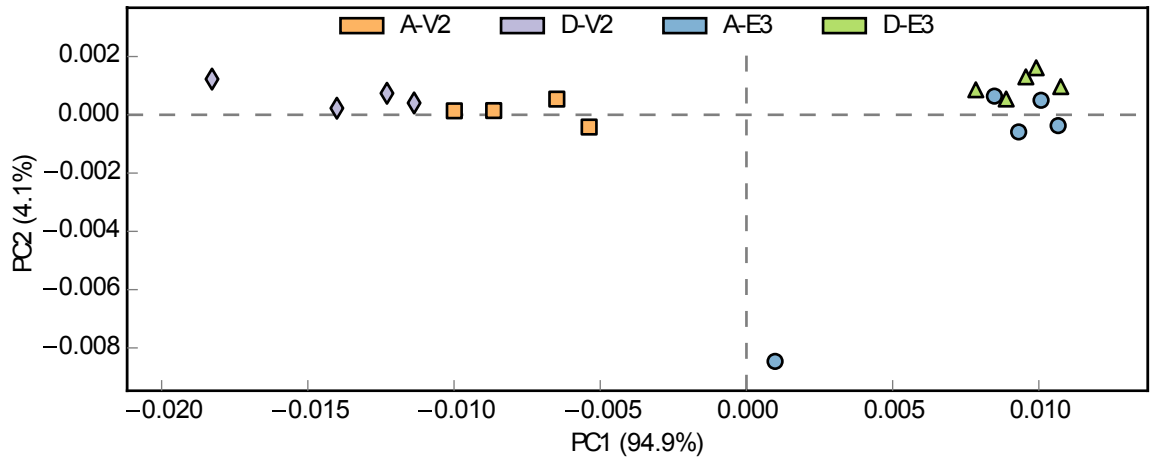


Figure S6: Principal component analysis of the KEGG pathways predicted in the metagenomes of the cultivars before ISA at stages V2 and E3.

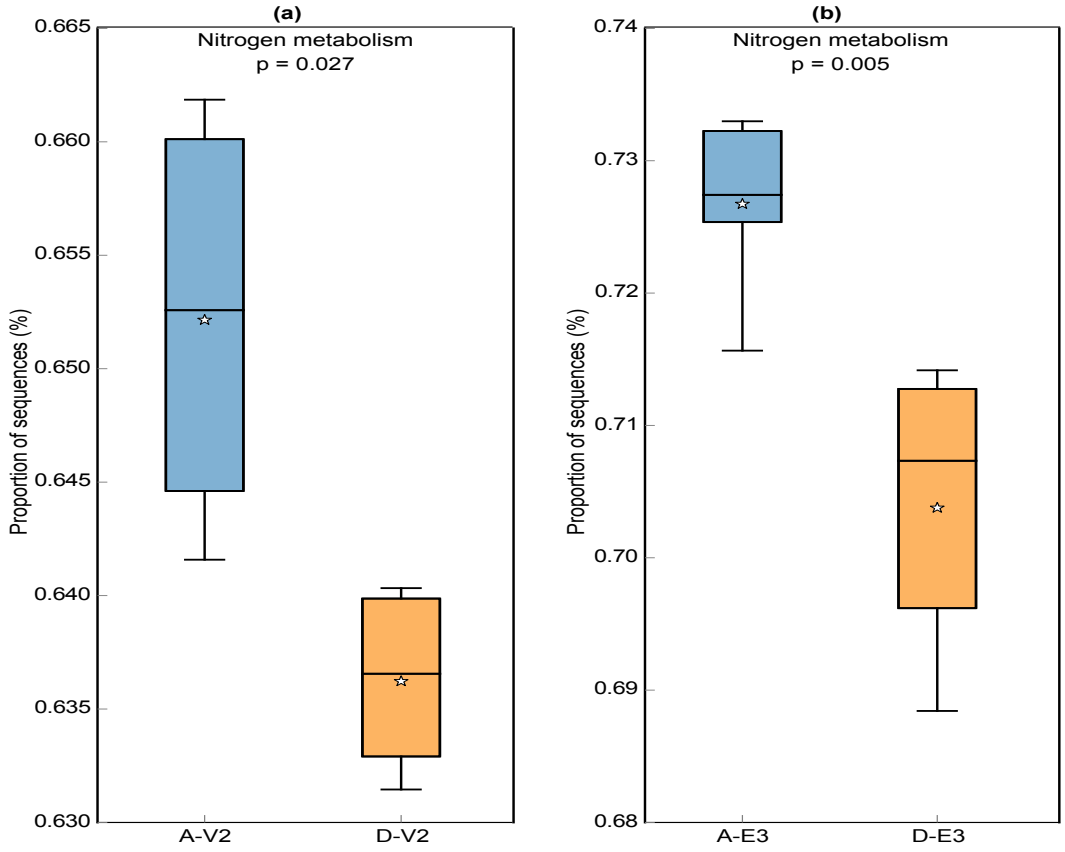


Figure S7: Post-ISA abundance of the "N-metabolism" pathway in the cultivars at stages (a) V2 and (b) E3. Two-sided t-test (p-value < 0.05) was used to estimate differential abundance between cultivars.

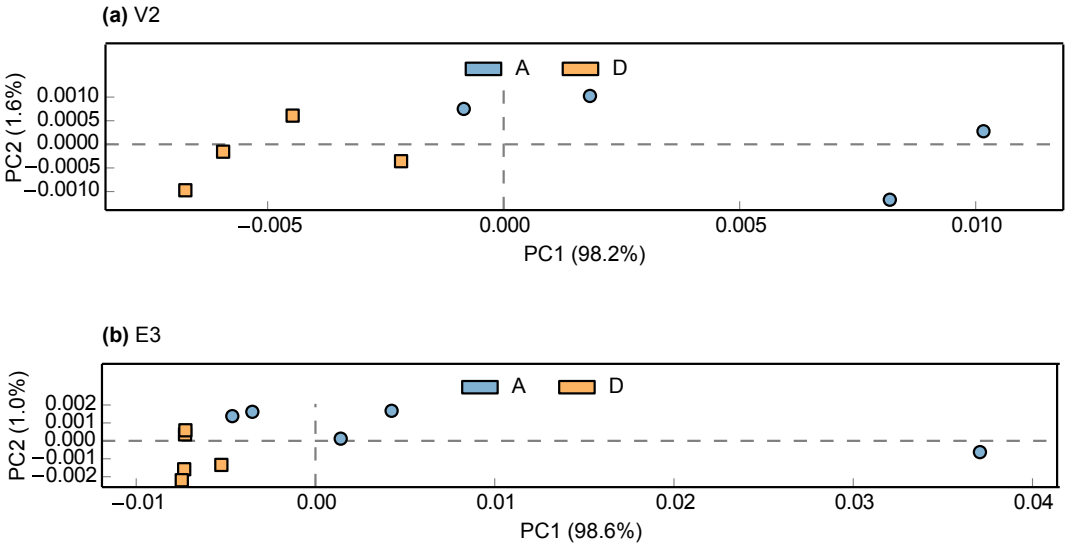


Figure S8: Principal component analysis of the KEGG pathways predicted in the metagenomes of the cultivars after ISA at stages (a) V2 and (b) E3.

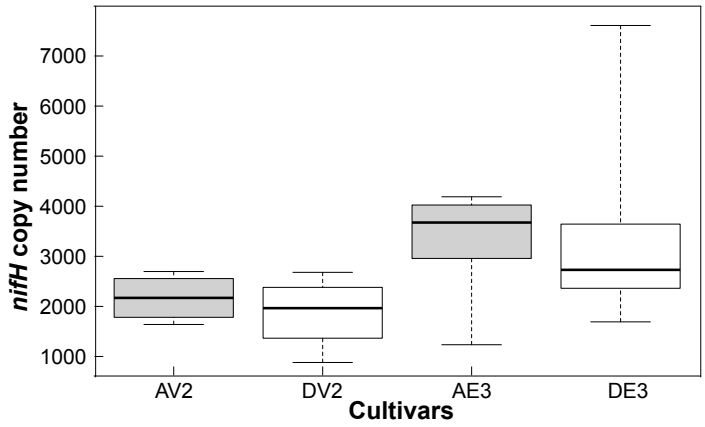


Figure S9: Boxplot of nifH copy numbers in cultivars at stages V2 and E3.

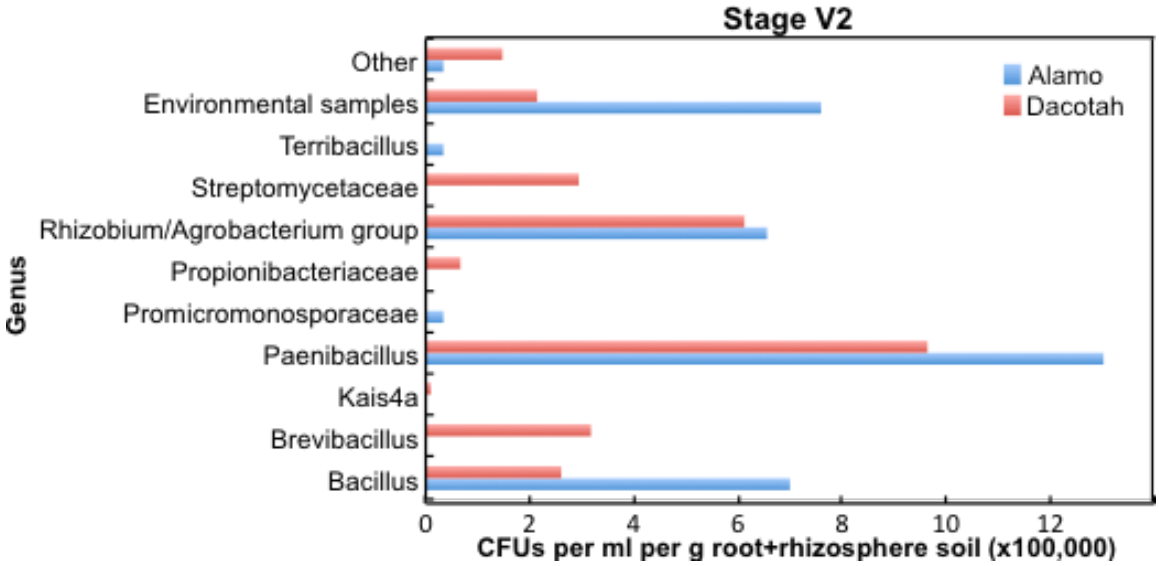


Figure S10: CFU abundance of cultivable root-zone diazotrophs from Alamo and Dacotah at Stage V2.

Table S1: (a) Alpha diversity and counts per sample of bacterial data.

#SampleID	Cultivar	Stage	chao1	Observed species	shannon	simpson	PD whole tree	ACE	goods_coverage	berger_parker_d	brillouin_d	dominance	equitability	Gini index	Simpson reciprocal	Counts/sample
1	A	V0	21,771	14,210	11.591	0.9989	549	23,802	0.948	0.012	7.865	0.001	0.84	0.9541	912	121,181
2	A	V0	9,303	5,377	11.156	0.9989	221	9,943	0.874	0.014	7.4	0.001	0.9	0.9712	879	21,613
3	A	V0	6,775	4,972	8.825	0.9906	208	7,359	0.971	0.054	6.012	0.009	0.719	0.9899	106	68,763
4	D	V0	13,639	8,422	11.414	0.9989	340	14,615	0.913	0.013	7.657	0.001	0.875	0.963	901	45,948
5	D	V0	13,066	8,106	11.449	0.999	336	13,951	0.906	0.012	7.663	0.001	0.882	0.9625	961	40,743
6	D	V0	4,855	2,716	10.645	0.9986	131	5,111	0.784	0.014	6.888	0.001	0.933	0.9804	736	6,910
7	A	V2	13,413	7,960	11.084	0.9984	317	14,023	0.92	0.015	7.452	0.002	0.855	0.9681	623	47,950
8	A	V2	18,426	11,270	11.096	0.9984	436	19,844	0.948	0.012	7.53	0.002	0.824	0.9657	617	100,490
9	A	V2	17,515	10,348	11.255	0.9985	409	18,707	0.931	0.012	7.599	0.002	0.844	0.9629	663	72,018
10	A	V2	12,418	7,210	11.117	0.9986	303	13,186	0.905	0.014	7.445	0.001	0.867	0.9686	700	37,762
11	D	V2	10,865	6,665	10.671	0.9978	255	11,139	0.941	0.023	7.207	0.002	0.84	0.976	453	50,433
12	D	V2	15,305	9,320	11.057	0.9983	363	16,350	0.938	0.017	7.475	0.002	0.839	0.9678	596	69,867
13	D	V2	15,341	9,231	11	0.9981	358	16,041	0.936	0.015	7.431	0.002	0.835	0.9677	515	67,437
14	D	V2	18,458	11,034	10.929	0.998	423	19,600	0.949	0.018	7.421	0.002	0.814	0.9678	501	102,799
23	D	E3	16,470	9,708	11.226	0.9987	408	17,891	0.926	0.011	7.572	0.001	0.848	0.9647	751	64,317
24	A	E3	18,244	11,639	11.343	0.9988	478	20,138	0.94	0.01	7.678	0.001	0.84	0.961	814	89,207
25	A	E3	21,798	14,916	11.714	0.9991	621	24,456	0.943	0.01	7.935	0.001	0.845	0.9497	1,067	114,813
26	A	E3	17,199	10,677	11.29	0.9988	450	19,121	0.931	0.01	7.627	0.001	0.844	0.9627	813	74,764
27	A	E3	15,939	9,711	10.49	0.9953	396	17,540	0.942	0.04	7.101	0.005	0.792	0.9709	213	80,122
28	D	E3	13,908	8,802	11.386	0.9987	380	15,101	0.915	0.019	7.642	0.001	0.869	0.9623	759	48,632
29	D	E3	27,498	18,843	11.405	0.9984	760	30,744	0.959	0.016	7.771	0.002	0.803	0.9516	630	201,251
30	D	E3	18,733	11,471	11.047	0.9979	476	20,467	0.94	0.019	7.48	0.002	0.819	0.9638	466	91,157
31	D	E3	14,500	8,586	11.382	0.9989	373	15,692	0.907	0.01	7.633	0.001	0.871	0.9632	944	45,698
32	A	E3	15,751	9,691	11.03	0.9983	402	17,319	0.936	0.016	7.456	0.002	0.833	0.9673	575	71,855

Table S1: (b) Alpha diversity and counts per sample of fungal data.

#SampleID	Cultivar	Stage	chao1	Observed species	shannon	simpson	ACE	Goods coverage	berger_parker_d	brillouin_d	dominance	equitability	Gini index	Simpson reciprocal	Counts/sample
1	A	V0	7,607	4,916	7.56	0.9731	9,038	0.976	0.122	5.173	0.027	0.616	0.9878	37.179	100,528
2	A	V0	4,338	3,250	7.142	0.9521	4,836	0.983	0.182	4.885	0.048	0.612	0.9906	20.87	75,907
3	A	V0	1,557	1,008	7.055	0.9802	1,738	0.972	0.066	4.8	0.02	0.707	0.9959	50.447	16,469
4	D	V0	5,701	5,414	8.164	0.9708	6,052	0.986	0.128	5.547	0.029	0.658	0.9742	34.197	73,573
5	D	V0	6,224	4,673	8.401	0.9834	7,208	0.965	0.081	5.706	0.017	0.689	0.9802	60.341	55,954
6	D	V0	6,643	3,992	7.77	0.9762	7,900	0.972	0.114	5.312	0.024	0.65	0.9886	42.1	75,223
7	A	V2	5,750	2,853	7.635	0.9782	6,821	0.965	0.106	5.208	0.022	0.665	0.991	45.909	46,459
8	A	V2	4,096	2,561	6.93	0.967	4,743	0.977	0.123	4.741	0.033	0.612	0.9937	30.271	56,940
9	A	V2	3,199	1,931	5.572	0.8878	3,814	0.981	0.305	3.81	0.112	0.51	0.9961	8.91	52,367
10	A	V2	6,527	3,933	7.001	0.9612	7,922	0.975	0.148	4.79	0.039	0.586	0.9911	25.8	85,152
11	D	V2	3,810	2,690	6.24	0.9409	4,480	0.982	0.191	4.271	0.059	0.548	0.9944	16.925	69,482
12	D	V2	3,443	2,088	6.931	0.9627	3,945	0.975	0.139	4.731	0.037	0.628	0.9938	26.776	41,364
13	D	V2	4,295	3,064	6.231	0.9555	5,053	0.987	0.12	4.279	0.044	0.538	0.9949	22.49	107,202
14	D	V2	4,283	3,005	6.065	0.9299	4,997	0.984	0.224	4.157	0.07	0.525	0.9946	14.267	89,322
15	D	E3	3,957	1,681	7.48	0.9727	4,246	0.94	0.122	5.046	0.027	0.698	0.9921	36.675	16,404
16	A	E3	6,640	2,974	7.636	0.9734	7,847	0.948	0.118	5.181	0.027	0.662	0.9891	37.594	35,065
17	A	E3	8,152	3,777	7.467	0.9672	9,695	0.963	0.151	5.092	0.033	0.628	0.9892	30.478	60,413
18	A	E3	8,023	4,049	7.131	0.9486	9,886	0.961	0.171	4.857	0.051	0.595	0.9885	19.474	62,162
19	A	E3	9,920	5,598	7.574	0.9687	11,758	0.967	0.119	5.172	0.031	0.608	0.9855	31.9	95,133
20	D	E3	10,737	6,188	7.481	0.9643	12,738	0.972	0.134	5.118	0.036	0.594	0.9857	28.009	123,565
21	D	E3	6,948	3,236	7.649	0.9777	7,947	0.958	0.087	5.205	0.022	0.656	0.9892	44.812	44,838
22	D	E3	6,861	3,214	7.831	0.9785	7,698	0.95	0.084	5.316	0.022	0.672	0.988	46.46	38,153
23	D	E3	9,771	4,742	7.605	0.9726	11,938	0.962	0.122	5.188	0.027	0.623	0.987	36.518	75,377
24	A	E3	10,623	5,855	7.859	0.9785	12,675	0.969	0.099	5.374	0.021	0.628	0.9853	46.577	107,169

Table S2 (a): Multivariate data analyses of samples using all bacterial OTUs to test for significant effects (p-value < 0.05) of stage, cultivar, and their interactions.

Dataset (Sequence threshold)	Categories	Adonis p-value	ANOSIM R statistic	p-value	MRPP p-value
V0, V2, and E3 (21,500)	Stage	0.001	0.7582	0.001	0.001
	Cultivar x Stage	0.001	0.6436	0.001	0.001
V0 and V2 (6,900)	Stage	0.003	0.5058	0.001	0.002
	Cultivar x Stage	0.001	0.4779	0.001	0.001
V2 and E3 (37,500)	Stage	0.001	0.9815	0.001	0.001
	Cultivar x Stage	0.001	0.7366	0.001	0.001

Table S2 (b): Multivariate data analyses of samples using all fungal OTUs to test for significant effects (p-value < 0.05) of stage, cultivar, and their interactions.

Dataset (Sequence threshold)	Categories	Adonis p-value	ANOSIM R statistic	p-value	MRPP p-value
V0, V2, and E3 (16,000)	Stage	0.001	0.8318	0.001	0.001
	Cultivar x Stage	0.001	0.7126	0.001	0.001

Table S3 (a): Indicator bacterial taxa for Alamo and Dacotah at stages V2 and E3. Alamo-V2: 21 taxa; Dacotah-V2: 11 taxa; Alamo-E3: 14 taxa; Dacotah-E3: 38 taxa.

Phylum	Class	Order	Family	Genus
Alamo (V2)				
Armatimonadetes	Chthonomonadetes	Chthonomonadales	Chthonomonadaceae	-
Chloroflexi	Anaerolineae	H39	-	-
Chloroflexi	Anaerolineae	S0208	-	-
Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus
Planctomycetes	Phycisphaerae	Cpla-3	-	-
Planctomycetes	Phycisphaerae	WD2101	-	-
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Phenylobacterium
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	-
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingobium
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	-
Proteobacteria	Betaproteobacteria	-	-	-
Proteobacteria	Betaproteobacteria	A21b	UD5	-
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Variovorax
Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Cupriavidus
Proteobacteria	Betaproteobacteria	-	-	-
TM7	TM7-3	EW055	-	-
TM7	TM7-3	-	-	-
Verrucomicrobia	[Spartobacteria]	[Chthoniobacterales]	[Chthoniobacteraceae]	OR-59
Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	Opitutus
WS6	B142	-	-	-
Dacotah (V2)				
Actinobacteria	Acidimicrobiia	Acidimicrobiales	C111	-
Actinobacteria	Acidimicrobiia	Acidimicrobiales	EB1017	-
Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	Phycicoccus
Q	Actinobacteria	Actinomycetales	Intrasporangiaceae	-
Actinobacteria	Actinobacteria	Actinomycetales	Thermomonosporacea e	Actinoallomurus
Actinobacteria	Thermoleophilia	Gaiellales	Gaiellaceae	-
Bacteroidetes	At12OctB3	-	-	-
Chlamydiae	Chlamydiia	Chlamydiales	Parachlamydiaceae	-
Chloroflexi	Ellin6529	-	-	-
Chloroflexi	Gitt-GS-136	-	-	-
Firmicutes	Clostridia	Clostridiales	Veillonellaceae	Pelosinus
Alamo (E3)				
[Thermi]	Deinococci	Deinococcales	Deinococcaceae	Deinococcus
Acidobacteria	Acidobacteriia	Acidobacteriales	Acidobacteriaceae	-
Acidobacteria	Solibacteres	Solibacterales	-	-
Acidobacteria	Solibacteres	Solibacterales	-	-
Actinobacteria	Actinobacteria	Actinomycetales	Cellulomonadaceae	Cellulomonas
Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Cryocola
Chloroflexi	-	-	-	-
Cyanobacteria	-	-	-	-
Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Coprococcus
Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	Propionivibrio
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfobulbus
Proteobacteria	Gammaproteobacteri a	Enterobacteriales	Enterobacteriaceae	-
Proteobacteria	Gammaproteobacteri a	Enterobacteriales	Enterobacteriaceae	Enterobacter
Proteobacteria	Gammaproteobacteri a	Enterobacteriales	Enterobacteriaceae	-

Dacotah (E3)				
Actinobacteria	Actinobacteria	Actinomycetales	Actinosynnemataceae	Lentzea
Actinobacteria	Actinobacteria	Actinomycetales	Micromonosporaceae	Catellatospora
Actinobacteria	Actinobacteria	Actinomycetales	Micromonosporaceae	-
Bacteroidetes	[Saprospirae]	[Saprospirales]	Chitinophagaceae	-
Chlamydiae	Chlamydiia	Chlamydiales	Rhabdochlamydiaceae	Candidatus-Rhabdochlamydia
Chloroflexi	Anaerolineae	envOPS12	-	-
Chloroflexi	Anaerolineae	S0208	-	-
Chloroflexi	Anaerolineae	SBR1031	SJA-101	-
Chloroflexi	Chloroflexi	[Roseiflexales]	[Kouleothrixaceae]	-
Chloroflexi	Chloroflexi	AKIW781	-	-
Chloroflexi	Gitt-GS-136	-	-	-
Chloroflexi	Ktedonobacteria	JG30-KF-AS9	-	-
Cyanobacteria	4C0d-2	SM1D11	-	-
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	-
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Clostridium
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	
Gemmatimonadetes	Gemm-1	-	-	-
Gemmatimonadetes	Gemm-2	-	-	-
Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Ellin5301	-
GN02	3BR-5F	-	-	-
OD1	Mb-NB09	-	-	-
Planctomycetes	-	-	-	-
Planctomycetes	Phycisphaerae	Phycisphaerales	-	-
Planctomycetes	Phycisphaerae	WD2101	-	-
Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	-
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Phenylobacterium
Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Cupriavidus
Proteobacteria	Deltaproteobacteria	-	-	-
Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrio
Proteobacteria	Gammaproteobacteria	Chromatiales	-	-
Proteobacteria	Gammaproteobacteria	Legionellales	Legionellaceae	
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Sinobacteraceae	Steroidobacter
TM7	SC3	-	-	-
Verrucomicrobia	[Methylacidiphilae]	S-BQ2-57	-	-
Verrucomicrobia	[Spartobacteria]	[Chthoniobacterales]	[Chthoniobacteraceae]	OR-59
Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Luteolibacter
WS6	SC72	WCHB1-15	-	-

Table S3 (b): Indicator fungal taxa for Alamo at stage V2. Alamo-V2: 4 taxa; No significant taxa were found for Dacotah-V2.

Phylum	Class	Order	Family	Genus
Alamo (V2)				
Ascomycota	Sordariomycetes	Hypocreales	Incertae_sedis	-
Basidiomycota	Tremellomycetes	Filobasidiales	-	-
Basidiomycota	Tremellomycetes	Tremellales	-	-
Zygomycota	Incertae_sedis	Mortierellales	Mortierellaceae	Mortierella

CHAPTER 4

The Search for the Core Microbiome in Switchgrass Rhizosphere

Richard R. Rodrigues[1] and Mark A. Williams[1,2*]

[1]Interdisciplinary Ph.D. Program in Genetics, Bioinformatics, and Computational Biology, Virginia Tech, Blacksburg, Virginia, United States of America.

[2]Department of Horticulture, Virginia Tech, Blacksburg, Virginia, United States of America.

* Corresponding author

E-mail: markwill@vt.edu (MW)

Phone: 540-231-4887

Original Research Article

Keywords: Switchgrass, core microbiome, root-zone, diazotrophs, 16s rRNA

Abstract

Improving yields of bioenergy crops is a current need to meet the world's ever-increasing demand of fuels. Sustainable production, however, of these crops is critical to avoid the toll of chemical fertilizers and traditional fuels on the environment. Previous studies have focused on the identification of root-zone microbes that promote plant growth. However, such studies are limited to plants and cultivars grown in specific locations or under specific conditions and assume that the root-zone microbes are mainly associated with the plant. In this paper, we aimed to identify whether any bacteria are always associated with the switchgrass rhizosphere, "core bacteria." We obtained two root-zone bacterial datasets of switchgrass grown under different conditions. The switchgrass from the two datasets had differences in the growing conditions (field vs. greenhouse), ages, cultivars, and locations, hence provided a good representation of the variety in which switchgrass can be produced. We used two different methods to identify bacteria associated with switchgrass rhizosphere: presence/absence and relative abundance. Our results suggest that there is a possibility of core bacteria in switchgrass rhizosphere. Using the presence/absence data, members of the Bacteroidetes, Firmicutes, and Proteobacteria were significantly present in the switchgrass root-zone. The genus *Lysobacter* (Proteobacteria) was observed to be associated with switchgrass rhizosphere using both presence/absence and relative abundance data. Previous studies indicate that the exoenzyme and antibiotic produced by the species from the *Lysobacter* genera protect the plant against bacterial and fungal diseases. Hence, *Lysobacter* genera is a top candidate in the search of the switchgrass associated root-zone core bacteria. Confirming this hypothesis, however, will need further experimentation.

Introduction

Efforts over the past decades have been directed towards achieving truly sustainable agriculture, however, it is still not a reality. The changing climate and the ever-growing population have rushed the need for economical and eco-friendly food crop production systems. On the other hand, the decreasing sources of traditional fuels and the increasing associated environmental costs (Galloway *et al.*, 2008, Sutton *et al.*, 2011) call for the production of bioenergy crops. Cereal and bioenergy crop yields, like any grass, are hindered by nitrogen limitations. Identifying the microbes that can serve as biofertilizers, especially the nitrogen-fixing bacteria, is an important first step in the quest for achieving sustainable grass crop production. These microbial genera should be able to associate and help meet the plant's nutrient demands to increase crop yield (Bhattacharjee *et al.*, 2008, Santi *et al.*, 2013).

Much interest and research is already under-way to develop cereals and other non-legumes that obtain fixed-nitrogen via symbiosis with diazotrophs to meet their N demands (Charpentier & Oldroyd, 2010, Santi *et al.*, 2013). However, understandably, much of the research is focused on human food crops (Santi *et al.*, 2013, Vessey, 2003) with minimal works studying plants used mainly for making biofuel and not as food for humans e.g. switchgrass, sorghum. Switchgrass is a US-native, perennial, bioenergy grass with many ecosystem benefits. Its ability to grow on marginal land provides little competition for land and nutrient resources with food crops. Few studies have identified the different bacteria found in the root-zones of switchgrass (Bahulikar *et al.*, 2014, Chaudhary *et al.*, 2012, Hargreaves *et al.*, 2015, Jesus *et al.*, 2010, Jesus *et al.*, 2016, Liang *et al.*, 2012, Mao *et al.*, 2014, Mao *et al.*, 2013, Mao *et al.*, 2011, Werling *et al.*, 2014), however, an integrative study of different datasets identifying the core microbiome in switchgrass rhizosphere is still lacking.

In this paper, we mined the root-zone bacteria obtained from switchgrass with differences in growth conditions (e.g. field vs. greenhouse), cultivars, locations, growth stages, and amplicon sequencing methods. Despite these numerous differences in the two datasets, we identified bacterial genera that are significantly associated with switchgrass rhizosphere. The root-zone bacteria from maize, restored prairie, and mixed grasses were used to test for the significance of a random occurrence of the core-bacteria (CB) from switchgrass. Based on traditional methods research can be focused to culture or modify members of the CB's genera. Alternatively, we can modify bacteria with well-known partnerships with members of the CB to employ the natural interactions between them to co-exist and improve plant yield. This would be an improvement from single species fertilizer and one step closer to a community level fertilizer, which is more reflective of the complexity of the rhizosphere. The research methodology used in this paper to identify bacteria is in contrast to the previously used methods of culturing root-zone bacteria and checking their plant growth promoting potential. We propose the use of omnipresent bacteria in switchgrass rhizosphere and suggest using directed efforts towards their modifications or of their natural partners. Additionally, this method could help to identify bacteria that may offer benefits other than nitrogen fixing ability, e.g. plant useful antibiotics, disease resistance, etc.

Materials and Methods

Dataset selection

The datasets used for this study were obtained from NCBI and were selected based on the availability of the raw (16S rRNA) sequence data of root-zone bacteria from switchgrass and comparable (native and/or bioenergy) grasses grown under different conditions. The dataset (Jesus *et al.*, 2016), hereby referred as "Jesus 2016," compared the rhizosphere soil microbial communities under restored prairie with three grass crops, namely corn,

switchgrass, and mixed prairie grasses. The grasses were grown in fields of Michigan and Wisconsin and were harvested after two and ten years. The V6-V8 region of the 16S rRNA gene was amplified and sequenced using the Roche 454 pyrosequencing. The dataset (Rodrigues et al., MS in prep.), hereby, referred as "Rodrigues 2016," compared the root-zone soil microbial communities under Alamo and Dacotah. The switchgrass were grown in the greenhouse using field soil obtained from Virginia and sampled at three time points. The V3-V4 region of the 16S rRNA gene was amplified and sequenced using the Illumina MiSeq sequencing. Overall, these two datasets served as valuable resources to compare the root-zone bacteria and identify the core-bacteria associated with switchgrass. The relevant differences have been summarized in Figure 1.

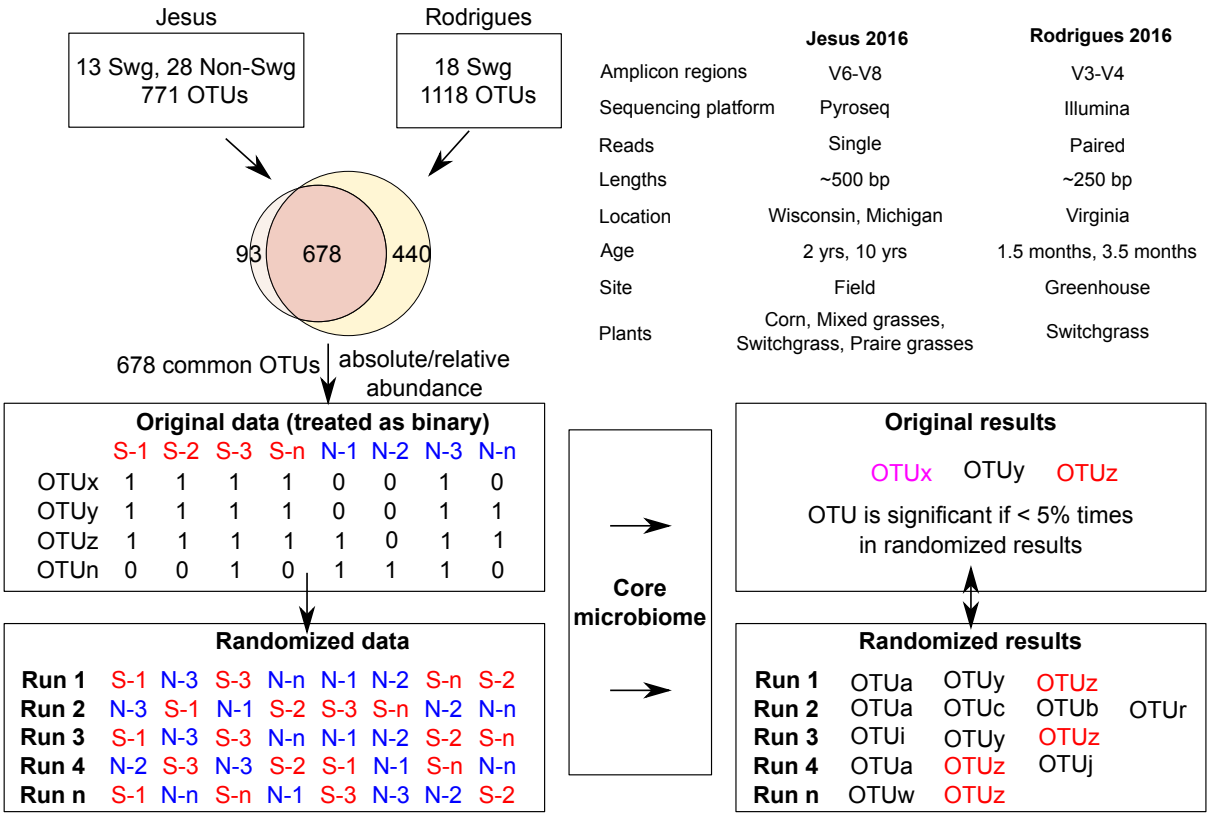


Figure 1: The workflow indicating the Jesus 2016 and Rodrigues 2016 datasets, differences between the datasets, and the methodology used to identify core microbiome. Switchgrass and other grasses are indicated by "Swg" and "Non-Swg," respectively.

Sequence data analysis and OTU picking

Rodrigues 2016: The rhizosphere OTUs for Alamo and Dacotah from stages V2 and E3 (4 A_{V2}, 4 D_{V2}, 5 A_{E3}, 5 D_{E3} = 18 switchgrass samples) were directly obtained from previously performed analysis (Rodrigues 2016). Briefly, barcode adapters and primers were removed from each read using cutadapt v1.8.1 (Martin, 2011) with a minimum quality (30) and read length (100) thresholds. The paired-end reads were merged using PANDAseq v2.8 (Masella *et al.*, 2012).

Jesus 2016: The 16S/18S reads for 43 root-zone samples (3 each from corn, switchgrass, mixed grasses (2 yrs. only), and restored prairie grown in Wisconsin and Michigan, and sampled after 2 years and 10 years; where, switchgrass grown in Michigan sampled after 10 years was an exception, with 4 samples) were obtained from PRJEB6704 (Jesus *et al.*, 2016). The quality score (25) and read lengths (150) thresholds were enforced using cutadapt (1.8.1) (Martin, 2011) to obtain high quality sequences. To allow comparison, this sequencing data analysis and the open reference OTU picking (enable_rev_strand_match True) were performed in QIIME v1.8.0 (Caporaso *et al.*, 2010) as previously described (Rodrigues *et al.*, 2015) (Rodrigues 2016). Mainly, uclust (Edgar, 2010) was used to cluster reads into OTUs (97% sequence similarity) and assign taxonomy against the Greengenes reference database version 13.8 (DeSantis *et al.*, 2006, McDonald *et al.*, 2012). Two samples were removed from downstream analysis due to very few counts (sequences assigned to OTUs).

Combining two datasets

Within each OTU table, sequences assigned to identical OTUs were summed to retain unique taxa. The common OTUs from the two datasets were selected, converted to biom format and used for further analyses (Figure 1). The data table was filtered and rarefied using

a sequence threshold, the beta diversity was calculated using Bray-Curtis (Beals, 1984) distance and used for Principal Coordinate Analysis (Gower, 2005) and visualization. Multivariate data analysis methods of MRPP (Mielke, 1984), ADONIS (Anderson, 2001) and Analysis of Similarity (ANOSIM) (Clarke, 1993) were used to identify whether the plant type (switchgrass versus non-switchgrass) has an effect on the bacterial communities. Group significance was calculated using Mann–Whitney U test on the rarefied data to identify OTUs that are differentially represented in switchgrass compared to other grasses. Taxons were summarized at the phylum level, those with less than 1% average abundance across all samples were grouped into "Other" and the abundance between plant types was compared using Mann–Whitney U test. Indicator species analysis (Dufrene & Legendre, 1997) was performed to utilize the relative abundance information and identify genera that are associated with switchgrass.

Core bacteria analysis

The OTUs that were present in all of the switchgrass samples were defined as the core bacteria and were identified using the "compute_core_microbiome" script from QIIME on the combined OTU table (original data). To test significance while preserving the OTU (data value) table, only the samples' group annotations were randomized a 1000 times and the core bacteria were calculated for each run. The core bacteria from the original data were considered to be statistically significant if they were observed less than 5% times in the randomized data.

Results

After quality filtering, a total of 319,821 reads were obtained from the Jesus 2016 dataset (mean 461.45 and std. dev. 69.34). Two samples with very few (48 and 75) counts

were removed; each of the remaining samples had more than 1150 sequences assigned to OTUs.

The number of OTUs in the Jesus 2016 and Rodrigues 2016 datasets was 771 and 1118, respectively. The combined dataset had 678 OTUs, 31 switchgrass samples and 28 non-switchgrass (other grasses) samples. Especially for the core microbiome analysis, absolute abundance is converted to presence/absence data.

The bacterial communities in switchgrass and grasses from the combined dataset were significantly different (Adonis, MRPP, and ANOSIM p values < 0.01) (Figure S1). Specifically at the phylum level, Mann Whitney test identified Bacteroidetes and Verrucomicrobia had significantly higher (p value < 0.05) relative abundance in switchgrass, whereas, Gemmatimonadetes were more abundant in other grasses (Figures 2). Some of the OTUs that were significantly different, as determined by Mann Whitney U test with FDR q value < 0.05, between switchgrass and other grasses are provided in Table S1.

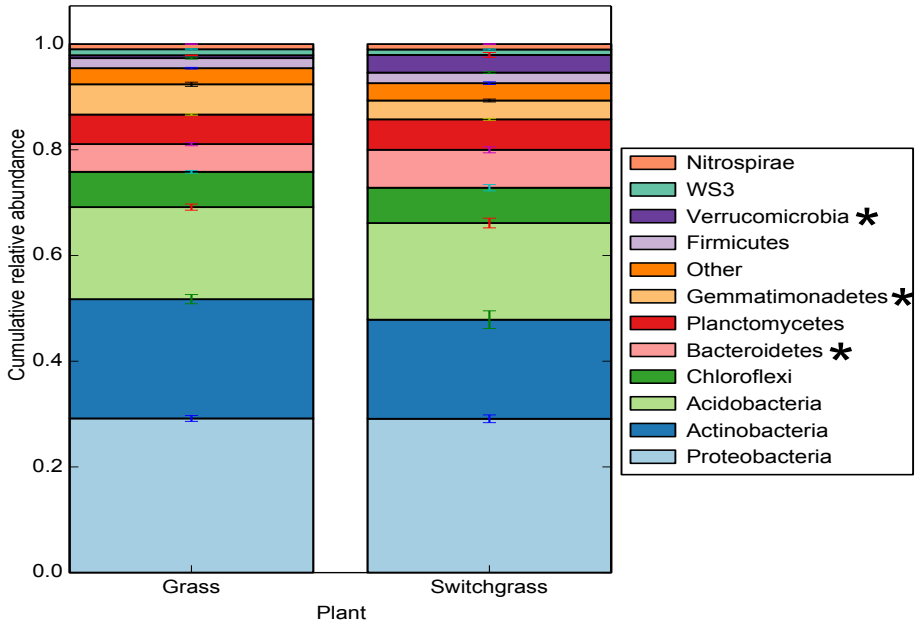


Figure 2: Taxonomic summary of the relative abundance of bacterial phyla in the combined dataset. The taxa and the labels are arranged as per total relative abundance across all samples, with the most abundant phyla at the bottom and the least abundant phyla at the top

of the y-axis. Mann Whitney test was used to identify phyla with significantly different (p value < 0.05) relative abundance.

We used two strategies to identify bacteria associated with switchgrass (Tables 1 and S2). The core microbiome method utilizes the presence absence data without taking into account the relative abundance. Also, except during significance calculation, it does not directly take into account the presence of the OTU in the other group of interest. To make it very conservative, we have used 100% presence criteria i.e., an OTU has to be present in all samples in the group of interest (switchgrass). The indicator species analysis takes into account the relative abundance and the consistency of occurrence of an OTU in the group of interest (switchgrass) compared to the other group (other grasses). This less conservative method offers a complementary analysis to the core microbiome analysis. Both these analyses are equally important since it is difficult to choose which criteria is most important: an OTU to be exclusively present, high/low abundance, or always present irrespective of consistent or increased abundance in switchgrass. Although, ISA identified many genera associated with switchgrass, the genus *Lysobacter* was observed to be significant by both strategies (Tables 1 and S2).

Table 1: The core bacterial OTUs those were significantly (FDR q value < 0.05) associated with switchgrass, calculated using presence/absence data and present in all switchgrass samples.

OTU	**q value**
p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae;g_Lysobacter;s_	0.036
p_Bacteroidetes;c_Flavobacteriia;o_Flavobacteriales;f_Cryomorphaceae;g_;s_	0.036
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Rhodobiaceae;g_Afifella;s_	0.046
p_Firmicutes;c_Bacilli;o_Bacillales;f_Alicyclobacillaceae;g_Alicyclobacillus;s_	0.046

Discussion

Identifying the microbes always associated with a plant offer a new direction for research. It can be studied to investigate the plant microbial communication and the exudates or extracts that are exchanged between them that favor such associations. This information can provide research avenues to modify these organisms and/or their microbial partners to enhance their associations with plants and improve plant growth.

The two datasets, Jesus 2016 and Rodrigues 2016, appropriately complement each other to identify core microbiome in switchgrass. Plots were located at the Kellogg Biological Station (Michigan), Arlington Agricultural Research Station (Wisconsin), greenhouse (Virginia), and other farms and reserves in Michigan and Wisconsin (Jesus *et al.*, 2016, Werling *et al.*, 2014). The locations used in this study were miles apart, hence can be treated as independent. Overall, the sampling covers (provides a good representation of) different grasses, conditions, soil types, and plant ages.

The mixed grass and prairie grass sites contained switchgrass along with many different forbs and other grasses (http://lter.kbs.msu.edu/datatables/269) (Werling *et al.*, 2014). This makes our analyses more stringent; so may be more bacterial OTUs that were associated with switchgrass but not found to be significant since they were observed in the switchgrass contaminated samples in "other grasses."

Currently, the core microbiome script only uses presence/absence data. For two groups (A and B) it checks whether (pre-determined percentage of) samples from group A have a non-zero value for the OTU, i.e., have at least one seq of the OTU. The script does not consider group B except for during randomizations. On the other hand, ISA uses an OTU's relative abundance (Specificity) and the proportion of samples in which it is present (Fidelity). ISA does not specifically enforce presence across all samples and could be the reason of observing higher number of OTUs from the ISA as compared to the core

microbiome analysis. Reducing the "present in percent samples" threshold (e.g. Table S3) of the core microbiome analysis could increase the number of OTUs observed in both analyses.

As an extension, we are currently implementing the core microbiome as a web-tool to allow multiple methods for randomizing data (Kallio *et al.*, 2011), e.g. randomize OTU labels, permutations of the all the entries in the OTU table, or only the entries within a row (OTU) and identifying core microbiome, e.g., using consistency in relative abundance across samples (Shade & Handelsman, 2012) or minimum number of sequences/relative abundance per sample.

There are many ways to expand on this work. Other switchgrass datasets could be incorporated to perform an integrative analysis (Hargreaves *et al.*, 2015, Jesus *et al.*, 2010, Mao *et al.*, 2013, Mao *et al.*, 2011). According to the research interest, this analysis could be used to identify core microbiome associated with field or greenhouse grown switchgrass.

Members of the *Lysobacter* genus are known to live in soil and are ecologically important due to their ability to produce exoenzymes and antibiotics (Reichenbach, 2006). Their antimicrobial activity against bacteria, fungi, unicellular algae, and nematodes are well studied (Islam *et al.*, 2005, Jochum *et al.*, 2006, Park *et al.*, 2008, Yin, 2010). For example, strains of this genus have been used for control of diseases caused by bacteria in rice (Ji *et al.*, 2008) and tall fescue (Kilic-Ekici & Yuen, 2004). Follow up analyses using PICRUSt and experiments could be performed to identify the functions that are more abundant and potentially beneficial to switchgrass.

In conclusion, the search for core microbiome is often ignored and underappreciated. Majority of the ecology-based studies merely focus on the highly abundant microbes, often present in more abundance in the group of interest. This results in ignoring the taxa that are rare and in lesser abundance in the group of interest (e.g., it is beneficial for the plant to have fewer pathogenic microbes). Research could then be directed to identify these core microbes

in individual or multiple bioenergy crops to identify common microbial targets that can be manipulated for sustainable agriculture. This combined with the ongoing efforts of plant breeding and genetic modifications can offer faster ways to improve crop yields without hurting the environment.

Acknowledgements

The authors would like to thank Dr. Roderick Jensen and James McClure for their suggestions in improving the manuscript and the web-tool, respectively.

References

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecology, **26**, 32-46.

Bahulikar RA, Torres-Jerez I, Worley E, Craven K, Udvardi MK (2014) Diversity of nitrogen-fixing bacteria associated with switchgrass in the native tallgrass prairie of northern Oklahoma. Appl Environ Microbiol, **80**, 5636-5643.

Beals EW (1984) Bray-Curtis Ordination: An Effective Strategy for Analysis of Multivariate Ecological Data. **14**, 1-55.

Bhattacharjee RB, Singh A, Mukhopadhyay SN (2008) Use of nitrogen-fixing bacteria as biofertiliser for non-legumes: prospects and challenges. Appl Microbiol Biotechnol, **80**, 199-209.

Caporaso G, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Meth, **7**, 335-336.

Charpentier M, Oldroyd G (2010) How close are we to nitrogen-fixing cereals? Curr Opin Plant Biol, **13**, 556-564.

Chaudhary D, Saxena J, Lorenz N, Dick L, Dick R (2012) Microbial Profiles of Rhizosphere and Bulk Soil Microbial Communities of Biofuel Crops Switchgrass (Panicum virgatum L.) and Jatropha (Jatropha curcas L.). Applied and Environmental Soil Science, **2012**, 1-6.

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology, **18**, 117-143.

Desantis TZ, Hugenholtz P, Larsen N *et al.* (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol, **72**, 5069-5072.

Dufrene M, Legendre P (1997) Species Assemblages and Indicator Species:The Need for a Flexible Asymmetrical Approach. Ecological Monographs, **67**, 345-366.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics, **26**, 2460-2461.

Galloway JN, Townsend AR, Erisman JW *et al.* (2008) Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. Science, **320**, 889-892.

Gower JC (2005) Principal Coordinates Analysis. In: *Encyclopedia of Biostatistics.* pp Page. The Open University, Milton Keynes, UK, John Wiley and Sons, Ltd.

Hargreaves SK, Williams RJ, Hofmockel KS (2015) Environmental Filtering of Microbial Communities in Agricultural Soil Shifts with Crop Growth. PLoS ONE, **10**, e0134345.

Islam MT, Hashidoko Y, Deora A, Ito T, Tahara S (2005) Suppression of damping-off disease in host plants by the rhizoplane bacterium Lysobacter sp. strain SB-K88 is linked to plant colonization and antibiosis against soilborne Peronosporomycetes. Appl Environ Microbiol, **71**, 3786-3796.

Jesus, Susilawati E, Smith S *et al.* (2010) Bacterial Communities in the Rhizosphere of Biofuel Crops Grown on Marginal Lands as Evaluated by 16S rRNA Gene Pyrosequences. BioEnergy Research, **3**, 20-27.

Jesus EDC, Liang C, Quensen JF, Susilawati E, Jackson RD, Balser TC, Tiedje JM (2016) Influence of corn, switchgrass, and prairie cropping systems on soil microbial communities in the upper Midwest of the United States. GCB Bioenergy, **8**, 481-494.

Ji G-H, Wei L-F, He Y-Q, Wu Y-P, Bai X-H (2008) Biological control of rice bacterial blight by Lysobacter antibioticus strain 13-1. Biological Control, **45**, 288-296.

Jochum CC, Osborne LE, Yuen GY (2006) Fusarium head blight biological control with Lysobacter enzymogenes strain C3. Biological Control, **39**, 336-344.

Kallio A, Vuokko N, Ojala M, Haiminen N, Mannila H (2011) Randomization techniques for assessing the significance of gene periodicity results. BMC Bioinformatics, **12**, 330.

Kilic-Ekici O, Yuen GY (2004) Comparison of strains of Lysobacter enzymogenes and PGPR for induction of resistance against Bipolaris sorokiniana in tall fescue. Biological Control, **30**, 446-455.

Liang C, Jesus E, Duncan D, Jackson R, Tiedje J, Balser T (2012) Soil microbial communities under model biofuel cropping systems in southern Wisconsin, USA: Impact of crop species and soil properties. Applied Soil Ecology, **54**, 24-31.

Mao Y, Li X, Smyth E, Yannarell A, Mackie R (2014) Enrichment of specific bacterial and eukaryotic microbes in the rhizosphere of switchgrass (Panicum virgatum L.) through root exudates. Environmental Microbiology Reports, **6**, 13.

Mao Y, Yannarell A, Davis S, Mackie R (2013) Impact of different bioenergy crops on N-cycling bacterial and archaeal communities in soil. Environmental Microbiology, **15**, 928-942.

Mao Y, Yannarell A, Mackie R (2011) Changes in N-Transforming Archaea and Bacteria in Soil during the Establishment of Bioenergy Crops. PLoS ONE, **6**, e24750.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, **17**, 10.

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

Mcdonald D, Price MN, Goodrich J *et al.* (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J, **6**, 610-618.

Mielke PW (1984) Meteorological applications of permutation techniques based on distance functions. In: *Handbook of statistics: Nonparametric methods.* (eds Krishnaiah PR, Sen PK) pp Page. Amsterdam: North-Holland.

Park JH, Kim R, Aslam Z, Jeon CO, Chung YR (2008) Lysobacter capsici sp. nov., with antimicrobial activity, isolated from the rhizosphere of pepper, and emended description of the genus Lysobacter. Int J Syst Evol Microbiol, **58**, 387-392.

Reichenbach H (2006) The Genus Lysobacter. 939-957.

Rodrigues RR, Pineda RP, Barney JN, Nilsen ET, Barrett JE, Williams MA (2015) Plant Invasions Associated with Change in Root-Zone Microbial Community Structure and Diversity. PLoS ONE, **10**, e0141424.

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

Shade A, Handelsman J (2012) Beyond the Venn diagram: the hunt for a core microbiome. Environ Microbiol, **14**, 4-12.

Sutton MA, Howard CM, Erisman JW *et al.* (2011) *The European Nitrogen Assessment: Sources, Effects and Policy Perspectives*, Cambridge University Press.

Vessey JK (2003) Plant and Soil, **255**, 571-586.

Werling BP, Dickson TL, Isaacs R *et al.* (2014) Perennial grasslands enhance biodiversity and multiple ecosystem services in bioenergy landscapes. Proc Natl Acad Sci U S A, **111**, 1652-1657.

Yin H (2010) Detection Methods for the Genus Lysobacter and the Species Lysobacter enzymogenes. Unpublished Master of Science University of Nebraska, Lincoln.

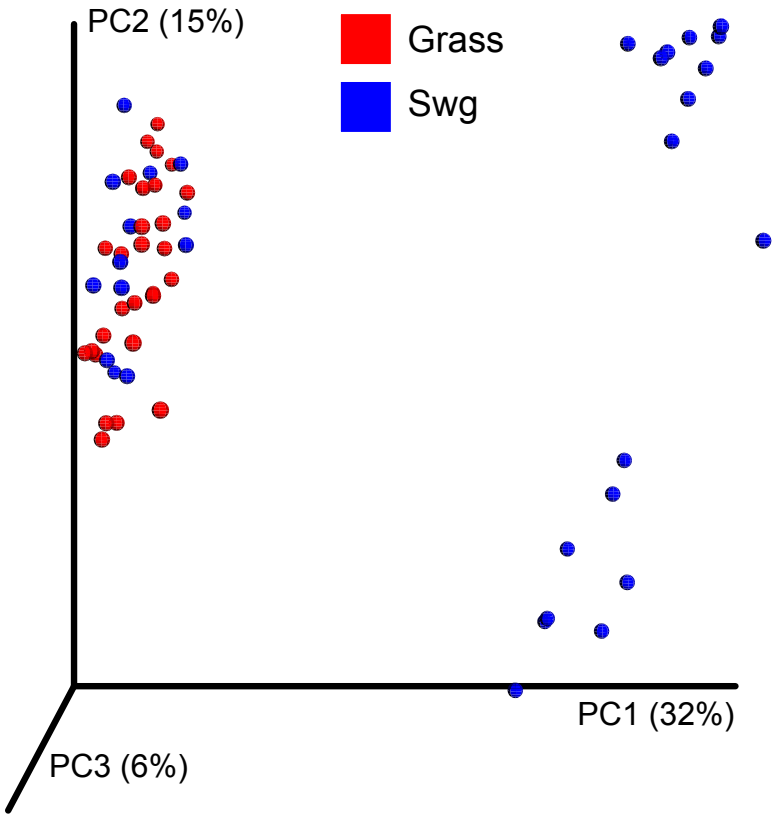


Figure S1: PCoA plot describing Bray-Curtis dissimilarities for bacterial communities at the OTU level in switchgrass and other grasses.

Table S1: The bacterial OTUs from rarified table those were significantly different (Mann Whitney U test with FDR q value < 0.05) between switchgrass and other grasses.

OTU	FDR_P (Mann Whit)	Sw mean abs abund	Gr mean abs abund
p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g_Sphingobium;s_	0.000	3.774	0.000
p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g_;s_	0.001	6.226	1.250
p_Verrucomicrobia;c_[Spartobacteria];o_[Chthoniobacterales];f_[Chthoniobacteraceae];g_;s_	0.002	3.258	0.071
p_Bacteroidetes;c_[Saprospirae];o_[Saprospirales];f_Chitinophagaceae;g_Flavisolibacter;s_	0.003	4.581	0.179
p_TM7;c_TM7-1;o_;f_;g_;s_	0.003	1.452	0.143
p_Actinobacteria;c_Thermoleophilia;o_Solirubrobacterales	0.003	1.581	3.964
p_Verrucomicrobia;c_Verrucomicrobiae;o_Verrucomicrobiales;f_Verrucomicrobiaceae;g_Luteolibacter;s_	0.003	3.226	0.036
p_Cyanobacteria;c_Nostocophycideae;o_Nostocales;f_Nostocaceae;g_;s_	0.003	1.484	0.036
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Geodermatophilaceae;g_;s_	0.003	1.774	4.429
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Pseudonocardiaceae;g_Pseudonocardia;s_	0.003	6.839	14.679
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micromonosporaceae [#]	0.004	2.516	5.250
p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g_Sphingomonas;s_	0.004	3.903	0.821
p_Verrucomicrobia;c_[Spartobacteria];o_[Chthoniobacterales];f_[Chthoniobacteraceae];g_DA101;s_	0.004	16.581	0.607
p_Proteobacteria;c_Deltaproteobacteria;o_Myxococcales;f_Polyangiaceae;g_Sorangium;s_	0.004	0.097	0.750
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Geodermatophilaceae;g_Geodermatophilus;s_	0.004	0.452	1.821
p_Proteobacteria;c_Deltaproteobacteria;o_Myxococcales;f_Haliangiaceae;g_;s_	0.004	4.548	8.714
p_Proteobacteria;c_Betaproteobacteria;o_MND1;f_;g_;s_	0.004	6.742	13.214
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Bradyrhizobiaceae;g_Bradyrhizobium;s_	0.004	1.452	3.857
p_Proteobacteria;c_Deltaproteobacteria;o_Bdellovibrionales;f_Bdellovibrionaceae;g_Bdellovibrio;s_	0.004	0.839	0.000
p_Verrucomicrobia;c_[Spartobacteria];o_[Chthoniobacterales];f_[Chthoniobacteraceae];g_CandidatusXiphinematobacter;s_	0.004	0.806	0.000
p_Verrucomicrobia;c_Verrucomicrobiae;o_Verrucomicrobiales;f_Verrucomicrobiaceae;g_;s_	0.004	0.935	0.000
p_Actinobacteria;c_Thermoleophilia;o_Solirubrobacterales;f_Solirubrobacteraceae	0.004	0.387	1.464
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micrococcaceae;g_Arthrobacter;s_	0.004	3.742	0.107
p_Armatimonadetes;c_0319-6E2;o_;f_;g_;s_	0.004	1.226	2.607
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micromonosporaceae;g_;s_	0.004	2.419	5.500
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_;g_;s_	0.005	7.129	13.393
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Bradyrhizobiaceae;g_;s_	0.005	3.581	1.000
p_TM6;c_SJA-4;o_;f_;g_;s_	0.005	1.000	0.107
p_Proteobacteria;c_Deltaproteobacteria;o_Myxococcales;f_;g_;s_	0.005	16.452	25.571
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micrococcaceae	0.005	0.903	1.857
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Kineosporiaceae	0.005	0.935	2.607
p_Acidobacteria;c_Acidobacteria-6;o_iii1-15;f_;g_;s_	0.005	79.258	55.500
p_Gemmatimonadetes;c_Gemmatimonadetes;o_Gemmatimonadales;f_;g_;s_	0.005	0.355	1.393
p_Proteobacteria;c_Betaproteobacteria;o_Nitrosomonadales;f_Nitrosomonadaceae	0.006	0.548	1.893
p_Actinobacteria;c_Thermoleophilia;o_Solirubrobacterales;f_Conexibacteraceae;g_;s_	0.006	1.968	4.286
p_Planctomycetes;c_Planctomycetia;o_Gemmatales;f_Isosphaeraceae;g_;s_	0.006	2.290	5.179
p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g_Novosphingobium;s_	0.008	2.452	0.393
p_Gemmatimonadetes;c_Gemm-5;o_;f_;g_;s_	0.009	0.484	2.000
p_Proteobacteria;c_Alphaproteobacteria;o_Rhodobacterales;f_Hyphomonadaceae;g_;s_	0.009	0.903	0.036
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Bradyrhizobiaceae;g_Bradyrhizobium	0.009	2.226	4.536
p_Verrucomicrobia;c_[Pedosphaerae];o_[Pedosphaerales];f_auto67_4W;g_;s_	0.009	0.935	0.000

p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Frankiaceae;g_;s_	0.009	1.032	3.000
p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g_Kaistobacter;s_	0.009	30.710	0.964
p_Fibrobacteres;c_Fibrobacteria;o_258ds10;f_;g_;s_	0.009	0.355	1.321
p_Bacteroidetes;c_[Saprospirae];o_[Saprospirales];f_Chitinophagaceae;g_;s_	0.010	42.968	19.964
p_Proteobacteria;c_Alphaproteobacteria;o_Rhodospirillales;f_Acetobacteraceae;g_;s_	0.010	2.419	4.714
p_Verrucomicrobia;c_[Spartobacteria];o_[Chthoniobacterales];f_[Chthoniobacteraceae];g_Chthoniobacter;s_	0.010	1.097	0.071
p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae	0.010	2.645	0.786
p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae	0.010	0.323	1.357
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Intrasporangiaceae	0.011	2.226	0.286
p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g_Sphingomonas;s_wittichi	0.011	1.258	0.143
p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae;g_Lysobacter;s_	0.013	2.903	0.857
p_Proteobacteria;c_Gammaproteobacteria;o_Legionellales;f_Coxiellaceae;g_Rickettsiella;s_	0.013	0.387	0.000
p_Proteobacteria;c_Deltaproteobacteria;o_Myxococcales;f_Myxococcaceae;g_Myxococcus;s_	0.014	0.129	0.750
p_Chloroflexi;c_TK10;o_B07_WMSP1;f_;g_;s_	0.014	1.419	2.929
p_Gemmatimonadetes;c_Gemm-3;o_;f_;g_;s_	0.015	1.032	3.107
p_Planctomycetes;c_Planctomycetia;o_Pirellulales;f_Pirellulaceae;g_;s_	0.017	5.871	9.179
p_Acidobacteria;c_iii1-8;o_32-20;f_;g_;s_	0.017	4.129	6.250
p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g_Methylibium;s_	0.017	4.226	7.214
p_Chlamydiae;c_Chlamydiia;o_Chlamydiales;f_;g_;s_	0.018	0.677	0.071
p_Verrucomicrobia;c_[Pedosphaerae];o_[Pedosphaerales];f_Ellin515;g_;s_	0.018	3.194	0.286
p_Proteobacteria;c_Alphaproteobacteria;o_Rhodobacterales;f_Rhodobacteraceae;g_Amaricoccus;s_	0.018	1.258	3.250
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales	0.020	0.935	2.107
p_Planctomycetes;c_Pla4;o_;f_;g_;s_	0.020	0.645	1.429
p_Armatimonadetes;c_Chthonomonadetes;o_Chthonomonadales;f_Chthonomonadaceae	0.020	0.774	0.036
p_Firmicutes;c_Bacilli;o_Bacillales;f_Bacillaceae	0.020	0.645	1.893
p_Verrucomicrobia;c_[Spartobacteria];o_[Chthoniobacterales];f_[Chthoniobacteraceae];g_OR-59;s_	0.020	0.452	0.000
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micrococcaceae;g_Arthrobacter	0.020	0.613	0.000
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Intrasporangiaceae;g_Phycicoccus;s_	0.020	0.903	0.036
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Propionibacteriaceae;g_;s_	0.024	2.129	3.429
p_Verrucomicrobia;c_[Pedosphaerae];o_[Pedosphaerales];f_;g_;s_	0.024	3.290	0.679
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Nocardioidaceae;g_;s_	0.024	10.452	17.786
p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae	0.025	3.742	5.750
p_Actinobacteria;c_Thermoleophilia;o_Gaiellales;f_;g_;s_	0.025	3.645	5.393
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Bradyrhizobiaceae	0.028	0.710	1.750
p_Planctomycetes;c_Phycisphaerae;o_WD2101;f_;g_;s_	0.030	28.355	16.071
p_Firmicutes;c_Clostridia;o_Clostridiales;f_Clostridiaceae;g_Clostridium;s_	0.037	1.032	0.393
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Nocardioidaceae	0.040	0.194	0.536
p_Bacteroidetes;c_Cytophagia;o_Cytophagales;f_Cytophagaceae;g_Adhaeribacter;s_	0.042	1.806	3.321
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Phyllobacteriaceae;g_Mesorhizobium;s_	0.043	1.000	0.357
p_Chloroflexi;c_Anaerolineae;o_envOPS12;f_;g_;s_	0.043	4.323	7.464
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Hyphomicrobiaceae;g_Hyphomicrobium;s_sulfonivorans	0.044	0.065	0.357
p_Proteobacteria;c_Betaproteobacteria;o_A21b;f_EB1003;g_;s_	0.044	2.258	4.893
p_Verrucomicrobia;c_Opitutae;o_Opitutales;f_Opitutaceae;g_Opitutus;s_	0.044	2.581	4.036
p_Acidobacteria;c_iii1-8;o_DS-18;f_;g_;s_	0.046	5.000	7.571
p_Planctomycetes;c_;o_;f_;g_;s_	0.049	0.484	1.429

[#] indicates the last OTU that is significant with a Bonferroni correction (adjusted p value < 0.05)

Table S2: Genera with a greater relative abundance associated with switchgrass and determined as important based on Indicator Species Analysis (IV > 70 and p-value < 0.01). Top 50% OTUs (Selected species: 152); IV=A*B*100; Association function: IndVal.g

Group: Grass	A	B	p.val	IV	Abund. Grass (%)	Abund. Swg (%)
p_Proteobacteria.c_Betaproteobacteria.o_Nitrosomonadales.f_Nitrosomonadaceae.Other	0.77	1	0	77.2	0.153	0.006
p_Actinobacteria.c_Actinobacteria.o_Actinomycetales.f_Geodermatophilaceae.g_Geodermatophilus	0.8	0.96	0	77	0.152	0.014
p_Gemmatimonadetes.c_Gemm.3.o_.f_.g_	0.77	0.96	0	73.8	0.32	0.038
p_Actinobacteria.c_Thermoleophilia.o_Solirubrobacterales.f_Solirubrobacteraceae.Other	0.76	0.96	0	73.2	0.108	0.006
p_Actinobacteria.c_Actinobacteria.o_Actinomycetales.f_Kineosporiaceae.Other	0.73	1	0.001	72.8	0.174	0.019
p_Actinobacteria.c_Thermoleophilia.o_Solirubrobacterales.Other.Other	0.73	1	0	72.5	0.382	0.025
p_Gemmatimonadetes.c_Gemm.5.o_.f_.g_	0.75	0.96	0.001	72.5	0.157	0.042
p_Gemmatimonadetes.c_Gemmatimonadetes.o_Gemmatimonadales.f_.g_	0.81	0.89	0	71.9	0.091	0.012
p_Proteobacteria.c_Gammaproteobacteria.o_Xanthomonadales.f_Xanthomonadaceae.Other	0.76	0.93	0.001	70.8	0.098	0.014
p_Fibrobacteres.c_Fibrobacteria.o_258ds10.f_.g_	0.7	1	0.001	70	0.087	0.015

Group: Switchgrass	A	B	p.val	IV	Abund. Grass (%)	Abund. Swg (%)
p_Proteobacteria.c_Alphaproteobacteria.o_Sphingomonadales.f_Sphingomonadaceae.g_Kaistobacter	0.97	0.94	0	90.9	0.079	4.321
p_Proteobacteria.c_Alphaproteobacteria.o_Sphingomonadales.f_Sphingomonadaceae.g_Sphingomonas	0.86	1	0	85.9	0.068	0.725
p_Verrucomicrobia.c_.Spartobacteria..o_.Chthoniobacterales..f_.Chthoniobacteraceae..g_DA101	0.97	0.87	0	84.3	0.046	2.144
p_Proteobacteria.c_Alphaproteobacteria.o_Sphingomonadales.f_Sphingomonadaceae.g_	0.81	1	0	80.5	0.111	0.699
p_Verrucomicrobia.c_.Pedosphaerae..o_.Pedosphaerales..f_.g_	0.88	0.9	0	79.2	0.042	0.497
p_Proteobacteria.c_Gammaproteobacteria.o_Xanthomonadales.f_Xanthomonadaceae.g_Lysobacter	0.79	1	0	79	0.082	0.433
p_Actinobacteria.c_Actinobacteria.o_Actinomycetales.f_Intrasporangiaceae.Other	0.89	0.87	0	77.5	0.025	0.29
p_Proteobacteria.c_Alphaproteobacteria.o_Sphingomonadales.f_Sphingomonadaceae.g_Sphingobium	0.95	0.81	0	76.9	0.017	0.535
p_Proteobacteria.c_Alphaproteobacteria.o_Sphingomonadales.f_Sphingomonadaceae.g_Novosphingobium	0.88	0.87	0	76.8	0.024	0.314
p_Verrucomicrobia.c_Verrucomicrobiae.o_Verrucomicrobiales.f_Verrucomicrobiaceae.g_	0.98	0.77	0	76	0.002	0.187
p_Firmicutes.c_Bacilli.o_Bacillales.f_Bacillaceae.g_	0.84	0.9	0.003	75.7	0.091	0.776
p_Proteobacteria.c_Alphaproteobacteria.o_Sphingomonadales.f_Sphingomonadaceae.Other	0.78	0.97	0	75.1	0.063	0.321
p_Verrucomicrobia.c_.Pedosphaerae..o_.Pedosphaerales..f_Ellin515.g_	0.96	0.77	0	74	0.016	0.551
p_Chloroflexi.c_Chloroflexi.o_Herpetosiphonales.f_.g_	0.89	0.81	0.001	72.1	0.032	0.384

p_Bacteroidetes.c_.Saprospirae..o_.Saprospirales..f_Chitinophagaceae.g_Flavisolibacter	0.97	0.74	0	71.7	0.009	0.67
p_Proteobacteria.c_Alphaproteobacteria.o_Rhizobiales.f_Phyllobacteriaceae.g_Mesorhizobium	0.79	0.9	0	71.1	0.021	0.123
p_Actinobacteria.c_Actinobacteria.o_Actinomycetales.f_Streptomycetaceae.Other	0.78	0.9	0.001	70.6	0.026	0.112
p_Proteobacteria.c_Alphaproteobacteria.o_Rhizobiales.f_Bradyrhizobiaceae.g_	0.78	0.9	0.001	70.4	0.099	0.5

Table S3: The core bacterial OTUs those were significantly (FDR q value < 0.05) associated with switchgrass, calculated using presence/absence data and present in 90% of switchgrass samples.

OTU	q val
p_Planctomycetes;c_Pla3;o_;f_;g_;s_	0.010
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Nocardiaceae;g_Rhodococcus;s_	0.010
p_Armatimonadetes;c_Armatimonadia;o_Armatimonadales;f_Armatimonadaceae;g_;s_	0.010
p_Planctomycetes;c_BD7-11;o_;f_;g_;s_	0.010
p_Actinobacteria;c_Acidimicrobiia;o_Acidimicrobiales;f_Microthrixaceae;g_;s_	0.013
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Phyllobacteriaceae;g_;s_	0.013
p_Verrucomicrobia;c_Opitutae;o_Opitutales;f_Opitutaceae;g_;s_	0.013
p_Acidobacteria;c_iii1-8;o_SJA-36;f_;g_;s_	0.013
p_Firmicutes;c_Bacilli;o_Bacillales;f_Bacillaceae;g_;s_	0.013
p_OP11;c_OP11-4;o_;f_;g_;s_	0.013
p_Chlorobi;c_;o_;f_;g_;s_	0.013
p_Verrucomicrobia;c_[Pedosphaerae];o_[Pedosphaerales];f_;g_;s_	0.013
p_Proteobacteria;c_Deltaproteobacteria;o_Myxococcales;f_Polyangiaceae	0.013
p_WS2;c_SHA-109;o_;f_;g_;s_	0.013
p_Planctomycetes;c_OM190;o_CL500-15;f_;g_;s_	0.013
p_Cyanobacteria;c_ML635J-21;o_;f_;g_;s_	0.013
p_Firmicutes;c_Bacilli;o_Bacillales;f_Planococcaceae;g_Sporosarcina;s_	0.013
p_Planctomycetes;c_vadinHA49;o_p04_C01;f_;g_;s_	0.015
p_Gemmatimonadetes;c_Gemm-2;o_;f_;g_;s_	0.016
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Bradyrhizobiaceae;g_;s_	0.021
p_Planctomycetes;c_C6;o_d113;f_;g_;s_	0.021
p_Chlorobi;c_OPB56;o_;f_;g_;s_	0.021
p_Chloroflexi;c_TK17;o_mle1-48;f_;g_;s_	0.021
p_Chloroflexi;c_Anaerolineae;o_H39;f_;g_;s_	0.023
p_Proteobacteria;c_Deltaproteobacteria;o_Spirobacillales;f_;g_;s_	0.023
p_Bacteroidetes;c_Flavobacteriia;o_Flavobacteriales;f_Cryomorphaceae;g_;s_	0.024
p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae;g_Lysobacter;s_	0.024
p_Proteobacteria;c_Betaproteobacteria;o_IS-44;f_;g_;s_	0.024
p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micromonosporaceae;g_Virgisporangium;s_ochraceum	0.026
p_Chloroflexi;c_Ktedonobacteria;o_JG30-KF-AS9;f_;g_;s_	0.038
p_Acidobacteria;c_S035;o_;f_;g_;s_	0.047
p_Chloroflexi;c_Anaerolineae;o_CFB-26;f_;g_;s_	0.047
p_Firmicutes;c_Bacilli;o_Bacillales;f_Bacillaceae;g_Bacillus	0.047
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Rhodobiaceae;g_Afifella;s_	0.047
p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g_Variovorax;s_paradoxus	0.047
p_Firmicutes;c_Bacilli;o_Bacillales;f_Alicyclobacillaceae;g_Alicyclobacillus;s_	0.047
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Beijerinckiaceae;g_;s_	0.047
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Bradyrhizobiaceae;g_Balneimonas;s_	0.049

CHAPTER 5

Conclusion

Outcomes and Future Directions

The overarching goal of this project was to determine the bacterial and fungal communities that support sustainable production of native plants, across natural and managed ecosystems, to develop low input, eco-friendly, economical, and sustainable cropping systems.

Chapter 2 focuses on the goal to identify the impact that other plants, especially, exotic invasive plants can have on the above- and below- ground communities of the native ecosystems. Network analyses identified the differences in the potential bacterial, fungal, and bacterio-fungal interactions in the rhizosphere of these native and invasive plants (Figs. S1-S3). Overall, the work highlights the shifts in microbial communities that are consistently observed across different plant types and warrants a *detailed study in plants from a variety of different taxonomic categories*. In addition, temporal dynamics within and between growing seasons, native (non-invasive), native (invasive), and (exotic) invasive plants spanning multiple orders could be used for this project. The use of both native- and exotic- invasive plants for comparison to non-invasive native plants could help to narrow down the importance of a plant's exotic nature and could be compared to plant traits associated with invasion.

Chapter 3 uses a diverse set of experiments with the goal of investigating the nitrogen fixing activity and microbes associated with the root-zones of switchgrass. It highlights the plant's temporal N-demand and dynamics in the composition, abundance, functions, and potential interactions of rhizosphere associated bacteria to support the plant N-requirements.

The next study could be to *determine whether plants can pass on to their progeny, the genetic traits that support and manage the growth of root-zone N_2-fixing bacteria.* Identification of these genes will provide molecular targets to decipher the plant-bacterial interaction and learning to efficiently manage the *rhizosphere microbiome* could have enormous potential for developing truly sustainable food and fiber agricultural ecosystems, e.g., production of a bioenergy crop, switchgrass, and major world food grass crops: maize, wheat, and rice.

Chapter 4 focuses on the goal of identifying whether and which root-zone bacteria are always associated with switchgrass, even when the plants are grown in differing conditions, such as natural and managed systems. An immediate extension of this chapter could be identifying fungi that are similarly omnipresent and associated with switchgrass. Identifying these microbes is the first step towards understanding the traits, interactions, and associations between the core microbes and switchgrass. So far studies have focused on genetic modifications of individual bacteria that promote plant growth. However, realizing the complexity of the belowground community research could be focused on modifying microbes (or their partners) that are universally associated with switchgrass. This can *utilize the inherent symbiosis certain microbes have to co-exist and promote plant growth* and move towards a microbial community level manipulation in switchgrass.

As a whole, this dissertation strongly supports sustainable agriculture through the understanding of the occurrence of natural plant-bacterial interactions and determine whether specialized diazotrophic bacteria can play a biologically significant role in plant N supply. The reduced need and applications of nitrogen fertilizers will lower direct costs to farmers and consumers. Environmental costs derived from excessive nitrogen losses into waterways and the atmosphere will also be reduced because BNF is regulated by a "builtin" feedback system, "nitrostat," that efficiently provides nitrogen by turning "on" to meet plant needs, and "off" when plant N demand is met (Menge & Hedin, 2009). In this way, nitrogen will not flow in

a huge pulse like that of N- fertilization, but rather, more like a drip to meet dynamic plant demand.

Over the long-term, similar studies will allow us to compare the genomics of C4 grasses that are associated with populations and activities of diazotrophs in their root-zones, and identify plant-microbial interactions and genes that can be used to predict and support high rates of nitrogen fixation. In particular, this offers the potential to breed for plant varieties that associate with specific bacterial communities to satisfy their temporal nitrogen demands. Such knowledge is crucial to the sustainable development of biofuel grass crops, which ubiquitously have high demand for nitrogen by providing them with bioavailable nitrogen without the costs of using chemical-based nitrogen fertilizers. This would satisfy human need for food and fiber while maintaining and improving the quality of soil, air and water.

References

Blair JM, Nellis MD, Turner CL, Henebry GM, Haiping S (1998) Terrestrial nutrient cycling in tallgrass prairie. In: *Grassland Dynamics: Long-Term Ecological Research in Tallgrass Prairie* (ed Alan K. Knapp JMB, David C. Hartnett, Scott L. Collins) pp Page. New York, Oxford University Press.

Casler MD, Vogel KP, Taliaferro CM, Wynia RL (2004) Latitudinal Adaptation of Switchgrass Populations. Crop Sci., **44**, 293-303.

Castro-Diez P, Godoy O, Alonso A, Gallardo A, Saldana A (2014) What explains variation in the impacts of exotic plant invasions on the nitrogen cycle? A meta-analysis. Ecol Lett, **17**, 1-12.

Chaudhary D, Saxena J, Lorenz N, Dick L, Dick R (2012) Microbial Profiles of Rhizosphere and Bulk Soil Microbial Communities of Biofuel Crops Switchgrass (Panicum virgatum L.) and Jatropha (Jatropha curcas L.). Applied and Environmental Soil Science, **2012**, 1-6.

Chen GQ, He YH, Qiang S (2013) Increasing seriousness of plant invasions in croplands of eastern china in relation to changing farming practices: a case study. PLoS ONE, **8**, e74136.

Diaz RJ, Rosenberg R (2011) Introduction to Environmental and Economic Consequences of Hypoxia. International Journal of Water Resources Development, **27**, 71-82.

Ehrenfeld JG (2003) Effects of exotic plant invasions on soil nutrient cycling processes. Ecosystems, **6**, 503-523.

Follett RF, Vogel KP, Varvel GE, Mitchell RB, Kimble J (2012) Soil Carbon Sequestration by Switchgrass and No-Till Maize Grown for Bioenergy. BioEnergy Research, **5**, 866-875.

Fowler D, Skiba U, Hargreaves KJ (1997) Emissions of nitrous oxide from grasslands. In:

Gaseous Nitrogen Emissions from Grasslands. (eds Sc J, Bf P) pp Page. North Wyke (United Kingdom): Institute of Grassland and Environmental Research.

Galloway JN, Aber JD, Erisman JW, Seitzinger SP, Howarth RW, Cowling EB, Cosby BJ (2003) The Nitrogen Cascade. Bioscience, **53**, 341.

Galloway JN, Townsend AR, Erisman JW *et al.* (2008) Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. Science, **320**, 889-892.

Hargreaves SK, Williams RJ, Hofmockel KS (2015) Environmental Filtering of Microbial Communities in Agricultural Soil Shifts with Crop Growth. PLoS ONE, **10**, e0134345.

Hawkes CV, Wren IF, Herman DJ, Firestone MK (2005) Plant invasion alters nitrogen cycling by modifying the soil nitrifying community. Ecology Letters, **8**, 976-985.

Heaton E, Long TVSP (2004) A quantitative review comparing the yields of two candidate C4 perennial biomass crops in relation to nitrogen, temperature and water. Biomass and Bioenergy, **27**, 21-30.

Hubbell DH, Kidder G (2009) Biological Nitrogen Fixation. pp Page.

Jacobs BS, Latimer AM (2012) Analyzing reaction norm variation in the field vs. greenhouse: Comparing studies of plasticity and its adaptive value in two species of Erodium. Perspectives in Plant Ecology, Evolution and Systematics, **14**, 325-334.

Jeschke J, Aparicio LG, Haider S, Heger T, Lortie C, Pyšek P, Strayer D (2012a) Support for major hypotheses in invasion biology is uneven and declining. NeoBiota, **14**, 1-20.

Jeschke JM, Aparicio LG, Haider S, Heger T, Lortie CJ, Pyšek P, Strayer DL (2012b) Taxonomic bias and lack of cross-taxonomic studies in invasion biology. Frontiers in Ecology and the Environment, **10**, 349-350.

Jesus, Susilawati E, Smith S *et al.* (2010) Bacterial Communities in the Rhizosphere of Biofuel Crops Grown on Marginal Lands as Evaluated by 16S rRNA Gene

Pyrosequences. BioEnergy Research, **3**, 20-27.

Jesus EDC, Liang C, Quensen JF, Susilawati E, Jackson RD, Balser TC, Tiedje JM (2015) Influence of corn, switchgrass, and prairie cropping systems on soil microbial communities in the upper Midwest of the United States. GCB Bioenergy, n/a-n/a.

Kourtev PS, Ehrenfeld JG, Haggblom M (2002) Exotic plant species alter the microbial community structure and function in the soil. Ecology, **83**, 3152-3166.

Li X, Rui J, Mao Y, Yannarell A, Mackie R (2014) Dynamics of the bacterial community structure in the rhizosphere of a maize cultivar. Soil Biology and Biochemistry, **68**, 392-401.

Liang C, Jesus E, Duncan D, Jackson R, Tiedje J, Balser T (2012) Soil microbial communities under model biofuel cropping systems in southern Wisconsin, USA: Impact of crop species and soil properties. Applied Soil Ecology, **54**, 24-31.

Liao C, Peng R, Luo Y *et al.* (2008) Altered ecosystem carbon and nitrogen cycles by plant invasion: a meta-analysis. New Phytologist, **177**, 706-714.

Lynch J, Hobbie J (1988) *Micro-organisms in action : concepts and applications in microbial ecology*, Oxford [England] ; Boston : Blackwell Scientific Publications.

Ma Z (1999) Carbon sequestration by switchgrass. Unpublished PhD AUBURN UNIVERSITY, 107 pp.

Madakadze IC, Stewart KA, Peterson PR, Coulman BE, Smith DL (1999) Cutting Frequency and Nitrogen Fertilization Effects on Yield and Nitrogen Concentration of Switchgrass in a Short Season Area. Crop Science, **39**, 552.

Magill AH, Aber JD, Berntson GM, Mcdowell WH, Nadelhoffer KJ, Melillo JM, Steudler P (2000) Long-Term Nitrogen Additions and Nitrogen Saturation in Two Temperate Forests. Ecosystems, **3**, 238-253.

Mao Y, Li X, Smyth E, Yannarell A, Mackie R (2014) Enrichment of specific bacterial and

eukaryotic microbes in the rhizosphere of switchgrass (Panicum virgatum L.) through root exudates. Environmental Microbiology Reports, **6**, 13.

Mao Y, Yannarell A, Davis S, Mackie R (2013) Impact of different bioenergy crops on N-cycling bacterial and archaeal communities in soil. Environmental Microbiology, **15**, 928-942.

Mao Y, Yannarell A, Mackie R (2011) Changes in N-Transforming Archaea and Bacteria in Soil during the Establishment of Bioenergy Crops. PLoS ONE, **6**, e24750.

Mclaughlin S, Bouton J, Bransby D *et al.* (1999) Developing switchgrass as a bioenergy crop. In: *Perspectives on new crops and new uses.* (ed Janick J) pp Page, Alexandria, VA.

Mclaughlin SB, Adams Kszos L (2005) Development of switchgrass (Panicum virgatum) as a bioenergy feedstock in the United States. Biomass and Bioenergy, **28**, 515-535.

Menge DN, Hedin LO (2009) Nitrogen fixation in different biogeochemical niches along a 120 000-year chronosequence in New Zealand. Ecology, **90**, 2190-2201.

Moore RB, Johnston CM, Smith RA, Milstead B (2011) Source and Delivery of Nutrients to Receiving Waters in the Northeastern and Mid-Atlantic Regions of the United States. J Am Water Resour Assoc, **47**, 965-990.

Moser LE, Vogel KP (1995) Switchgrass, big bluestem, and indiangrass. In: *Forages: An introduction to grassland agriculture.* (eds Barnes RF, Miller DA, Nelson CJ) pp Page. Ames, IA., Iowa State Univ. Press.

Nautiyal CS, Srivastava S, Chauhan PS, Seem K, Mishra A, Sopory SK (2013) Plant growth-promoting bacteria Bacillus amyloliquefaciens NBRISN13 modulates gene expression profile of leaf and rhizosphere community in rice during salt stress. Plant Physiol Biochem, **66**, 1-9.

Oenema O, Witzke HP, Klimont Z, Lesschen JP, Velthof GL (2009) Integrated assessment of

promising measures to decrease nitrogen losses from agriculture in EU-27. Agriculture, Ecosystems \& Environment, **133**, 280-288.

Parepa M, Fischer M, Bossdorf O (2013) Environmental variability promotes plant invasion. Nat Commun, **4**, 1604.

Parker SS, Schimel JP (2010) Invasive Grasses Increase Nitrogen Availability in California Grassland Soils. Invasive Plant Science and Management, **3**, 40-47.

Rebich RA, Houston NA, Mize SV, Pearson DK, Ging PB, Evan Hornig C (2011) Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico from Streams in the South-Central United States. J Am Water Resour Assoc, **47**, 1061-1086.

Skinner H (2009) Sequestration Potential of Switchgrass Managed for Bioenergy Production. pp Page, University Park, PA, ARS Pasture Systems and Watershed Management Research Unit.

Smil V (1999) Detonator of the population explosion. Nature, **400**, 415-415.

Smil V (2002) Nitrogen and food production: proteins for human diets. Ambio, **31**, 126-131.

Sommer SG, Hutchings NJ (1997) Components of ammonia volatilization from cattle and sheep production. In: *Gaseous nitrogen emissions from grasslands.* pp Page., CAB International, Wallingford, Oxon, UK.

Staley TE, Stout WL, Jung GA (1991) Nitrogen Use by Tall Fescue and Switchgrass on Acidic Soils of Varying Water Holding Capacity. Agronomy Journal, **83**, 732.

Stout WL, Jung GA (1995) Biomass and Nitrogen Accumulation in Switchgrass: Effects of Soil and Environment. Agronomy Journal, **87**, 663.

Sutton MA, Howard CM, Erisman JW *et al.* (2011a) *The European Nitrogen Assessment: Sources, Effects and Policy Perspectives*, Cambridge University Press.

Sutton MA, Oenema O, Erisman JW, Leip A, Van Grinsven H, Winiwarter W (2011b) Too much of a good thing. Nature, **472**, 159-161.

Tamura M, Tharayil N (2014) Plant litter chemistry and microbial priming regulate the accrual, composition and stability of soil carbon in invaded ecosystems. New Phytologist, **203**, 110-124.

Thomason WE, Raun WR, Johnson GV, Taliaferro CM, Freeman KW, Wynn KJ, Mullen RW (2005) Switchgrass Response to Harvest Frequency and Time and Rate of Applied Nitrogen. Journal of Plant Nutrition, **27**, 1199-1226.

Tjepkema J (1975) Nitrogenase activity in the rhizosphere of Panicum virgatum. Soil Biology and Biochemistry, **7**, 179-180.

Turner T, Ramakrishnan K, Walshaw J *et al.* (2013) Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. The ISME Journal, **7**, 2248-2258.

Vilà M, Espinar JL, Hejda M *et al.* (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. Ecology Letters, **14**, 702-708.

Vogel KP, Mitchell RB (2008) Heterosis in Switchgrass: Biomass Yield in Swards. Crop Science, **48**, 2159.

Wagner S (2011) Biological Nitrogen Fixation. Nature Education Knowledge, **3**, 15.

Weaver JE (1968) *Prairie Plants and Their Environment: A Fifty Year Study of the Midwest,* Lincoln, NE, University of Nebraska Press.

Weier KL (1980) Nitrogen fixation associated with grasses. Tropical Grasslands, **14**.

Wewalwela J (2014) Associative nitrogen-fixing bacteria and their potential to support the growth of bioenergy grasses on marginal lands. Unpublished Ph.D. Mississippi State University.

Wolf DD, Fiske DA (1995) Planting and Managing Switchgrass for Forage, Wildlife, and Conservation. pp Page, Virginia Polytechnic Institute and State Univ., Blacksburg.,

Virginia Cooperative Extension Publ. 418–013.

Wolfe AH, Patz JA (2002) Reactive Nitrogen and Human Health:Acute and Long-term Implications. AMBIO: A Journal of the Human Environment, **31**, 120-125.

Wolfe BE, Klironomos JN (2005) Breaking new ground: soil communities and exotic plant invasion. Bioscience, **55**, 477-487.

Wright LL, Perlack RD (2011) Switchgrass Production in the USA. (ed Berndes G) pp Page.

Wullschleger SD, Davis EB, Borsuk ME, Gunderson CA, Lynd LR (2010) Biomass Production in Switchgrass across the United States: Database Description and Determinants of Yield. Agronomy Journal, **102**, 1158.

Zalapa JE, Price DL, Kaeppler SM, Tobias CM, Okada M, Casler MD (2011) Hierarchical classification of switchgrass genotypes using SSR and chloroplast sequences: ecotypes, ploidies, gene pools, and cultivars. Theor Appl Genet, **122**, 805-817.

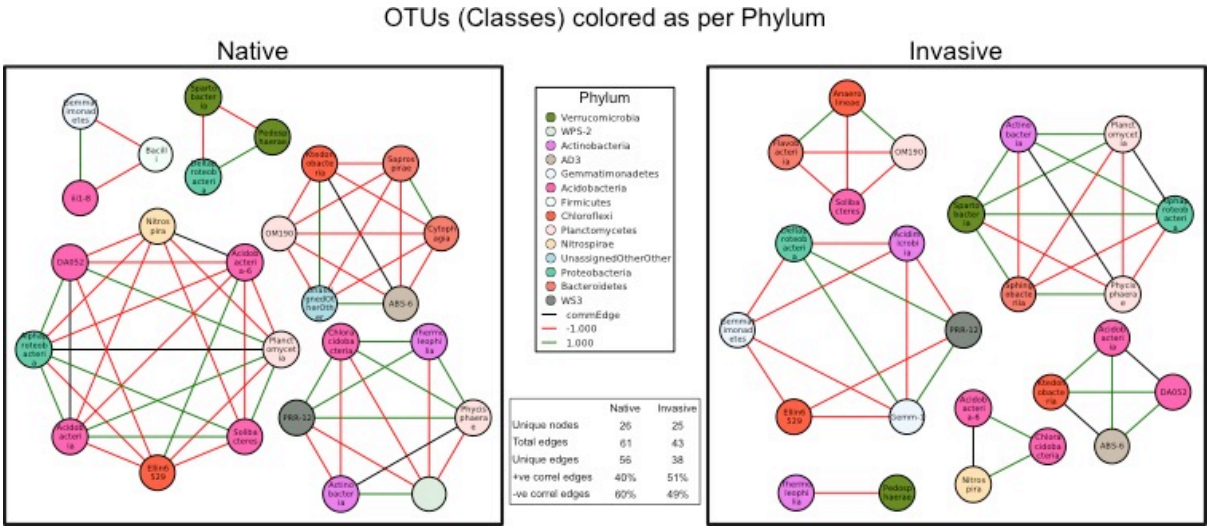


Fig S1: Potential bacterial interactions in the rhizosphere of native and invasive plants from the MENAP analysis.

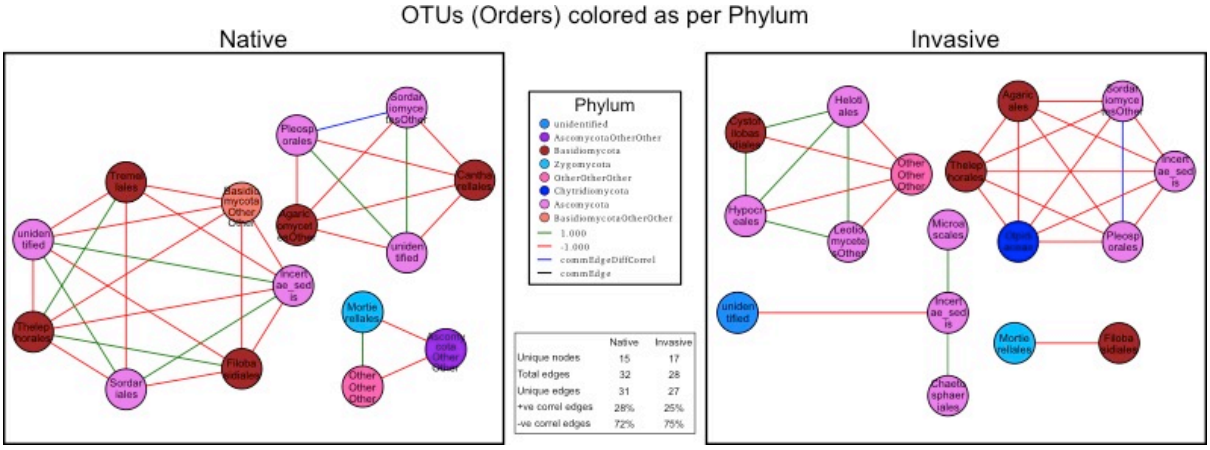


Fig S2: Potential fungal interactions in the rhizosphere of native and invasive plants from the MENAP analysis.

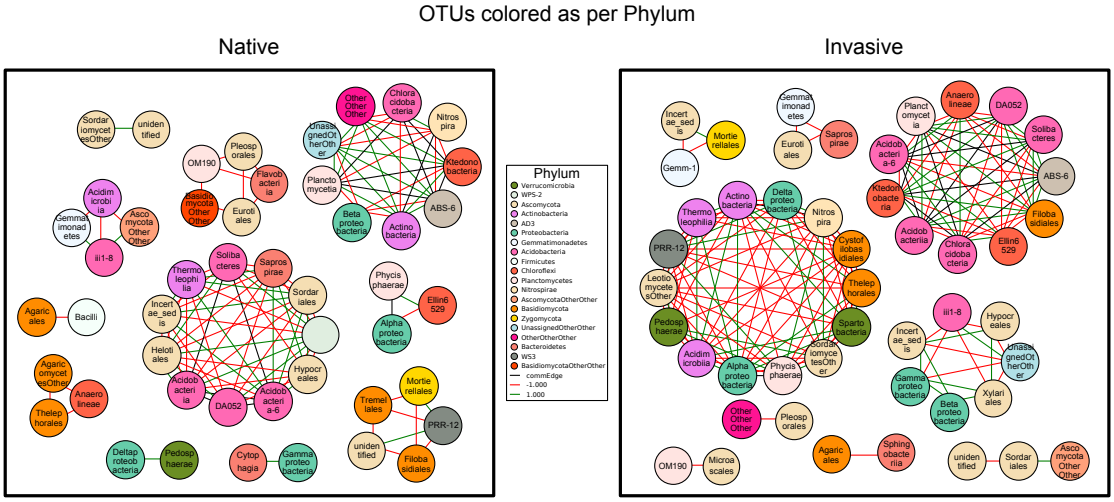


Fig S3: Potential bacterial-fungal interactions in the rhizosphere of native and invasive plants from the MENAP analysis.