DIVERSITY OF ARBUSCULAR MYCORRHIZAL FUNGI ASSOCIATED WITH SWITCHGRASS IN THE NEW JERSEY PINE BARRENS ECOSYSTEM

By

MOLLY R. BINDELL

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Ecology and Evolution

Written under the direction of

Ning Zhang

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January, 2019

ABSTRACT OF THE DISSERTATION

DIVERSITY OF ARBUSCULAR MYCORRHIZAL FUNGI ASSOCIATED WITH SWITCHGRASS IN THE NEW JERSEY PINE BARRENS ECOSYSTEM

by MOLLY R. BINDELL

Dissertation Director:

Dr. Ning Zhang

Arbuscular mycorrhizal fungi (AM fungi) are a group of mutualistic, root-colonizing, microorganisms that have enormous importance in ecosystem functioning, agricultural production, and habitat conservation. Despite their relatively small taxonomic diversity, they associate with over 80% of terrestrial plants. AM fungal species are often described through morphological differences, such as, spore wall features. However, next-generation sequencing (NGS) has become an increasingly popular method to expand our understanding of AM fungal global diversity patterns. The drivers of AM fungal diversity patterns remain poorly understood. This dissertation, therefore, combines Illumina MiSeq sequencing and microscopic observations to uncover the AM fungal communities inhabiting switchgrass roots in the Pine Barrens ecosystem, and compare those communities with those from agroecosystems. The AM fungal communities of the rare and unique Pine Barrens ecosystem were previously unknown. Illumina sequencing results uncovered several clades unique to the Pine Barrens

ecosystem, as well as, clades unique to agricultural fields, with many potentially novel species. This dissertation also developed a fully annotated, bioinformatic workflow for the study of AM fungal diversity via Illumina sequencing. Developing this AM fungal Illumina workflow showed that certain bioinformatic decisions can alter downstream AM fungal diversity results dramatically. Reference database selection was found to be a key decision in the workflow process. A pot experiment was also conducted in order to explore whether or not the acidic soils of the Pine Barrens influence the extent of switchgrass root colonization by and diversity of AM fungi. Native Pine Barrens AM fungi and soil were either left alone or amended with calcium carbonate (lime) in order to test whether soil pH plays a role in shaping the AM fungal communities, as previous studies have shown mixed results. Although the experiment didn't yield much significant data, it found that plants grown without AM fungi and those with acidic soil grew bigger than those with AM fungi and those with more neutral soil. Better control over inoculation procedures and abiotic factors may shed more light on these findings. While this dissertation research helps us better understand AM fungal global diversity patterns, lingering questions remain on how these fungi function in different environments, why species exist in certain places, and how these fungi can be realistically implemented in sustainable agricultural practices.

Acknowledgements

First and foremost, I must acknowledge my outstanding dissertation advisor, Dr. Ning Zhang. Not only did she have unwavering faith in my ability to work hard and succeed, but she gave me excellent research advice and guidance whenever it was needed. She afforded me the ability to make my own mistakes at times, in order to learn from them, and allowed me to use my creativity when designing experiments, all while sharing helpful tips along the way. I thank her for being supremely patient with me and for always believing in me. I would also like to thank my dissertation committee, Dr. James White, Dr. John Dighton, and Dr. Thomas Gianfagna, for giving me much necessary feedback on my experimental design and approaches. These extra perspectives were very much needed to successfully complete my research.

Second, I must thank my fellow laboratory members. In particular, Emily Walsh and Dr. Jing Luo were both extraordinarily helpful and instrumental in helping me figure out how to get through these last 7+ years of research and writing. I could not have done it without their insights, direction, and words of encouragement.

Third, I would like to thank all my sources of financial support throughout my time at Rutgers. Those include: Rutgers University awards, the National Military Family Association awards, Garden Club of America award, my GA support from Dr. Zhang, and my TA support from Rutgers University. In particular, I am very grateful to have had the opportunity to teach as a TA while a graduate student. I must thank Dr. Gregg Transue for giving me 5 years of amazing teaching and leadership opportunities within General Biology. Teaching was not what I anticipated it to be. I was surprised to discover that I could learn a thing or two from my students each semester. I was equally

surprised to find that I loved teaching Biology. I hope I have the chance to dive into teaching more in the days and years to come.

Last, I wish to give a massive thank you to all of my family and friends who have helped support me through these graduate school years. In particular, my parents, Anna and Alvin Silverstein, provided me with infinite babysitting help for my three kids, as well as, lots of love and emotional support along the way. My dad especially helped take care of my baby, Oren, in the last few months of my dissertation writing, making it possible for me to complete the writing phase. My in-laws, Paul and Laura Bindell, also helped care for my children when I needed to get some research or writing done. A special thank you, from the bottom of my heart, to my dedicated and patient husband, Yaakov (Jack). I could not have completed this huge undertaking without Jack's constant support. He guided me through the toughest times even when I couldn't see the light at the end of the tunnel.

I certainly cannot thank everyone enough for all that they have done to guide me through this challenging obstacle course called graduate school.

DEDICATION

This dissertation is dedicated to my precious children, Eliyahu, Chaim, and Oren, who, unbeknownst to them, put up with me and my late nights working on research, teaching, and writing for the last 7+ years. I love you boys to the moon and back!

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

Introduction

In this introduction to the dissertation, an overview of arbuscular mycorrhizal fungal (AM fungi) biodiversity is described in terms of this group's taxonomy, phylogeny, morphology, and ecology. The importance of AM fungi, both agriculturally and ecologically, is emphasized, as well as, impacts of global climate change on AM fungi. A brief history and significance of the AM fungal fossil records is reviewed. Problems and challenges along the path of discovery of global AM fungal biodiversity are discussed in depth, specifically, in the realm of AM fungal methodologies. Abiotic and other effects on AM fungal abundance and diversity are discussed. Host specificity of AM fungi is debated; confounding factors leading to conflicting results is a common theme throughout the literature. The great advantages and potential pitfalls of metagenomic methods for AM fungal community analysis are detailed. Background on switchgrass as an important, natural grass species and potential biofuel source is discussed, along with, the limited research on switchgrass associated AM fungi. Lastly, I address the necessity of studying AM fungi in understudied ecosystems and the unique qualities and significance of the Pine Barrens ecosystem, in particular. Thus far, there has been much progress in AM fungal diversity research, due, in part, to the advent of fungal metagenomics. No next-generation sequencing has yet been done on the AM fungal communities in Pine Barrens ecosystems in the United States, nor on native switchgrass AM communities. This dissertation aims to fill these knowledge gaps.

Arbuscular mycorrhizal fungi overview

Fungi are the second largest eukaryotic kingdom and one of the most ubiquitous organisms on the planet (Blackwell 2011; Mora et al. 2011). Mora et al. (2011) estimated the total number of eukaryotic land dwelling species to be ~8.7 million. The diversity of fungi is estimated to be anywhere from 1.5 million species (Hawksworth and Rossman 1997), to 2.27 million (Hawksworth 2001), to 5.1 million (Blackwell 2011), to as vast as 165.6 million species (Larsen et al. 2017). The number 165.6 million is possibly a large overestimate, based on the algorithms used by Mora et al. (2011) to calculate Earth's total species diversity, and frankly by most other accounts (above). Contrastingly, the diversity estimates of Mora et al. (2011) are most likely overly conservative. Nevertheless, we currently only have around 100,000 described fungal species (Blackwell 2011), about 0.02% of the likely 5.1 million fungal species. Despite the fact that scientists are discovering and describing new fungal species all the time, much work clearly has to be done on new fungal species descriptions. Expanded sampling efforts are needed in order to further elucidate the fungal tree of life (Torres-Cruz et al. 2017).

Arbuscular mycorrhizal fungi (AM fungi) are a group of symbiotic, plant-root associated, endomycorrhizal fungi within the phylum Glomeromycota (Schüßler et al. 2001; Tedersoo et al. 2018), more recently placed within a monophyletic subphylum, the Glomeromycotina, according to genomic evidence from the study by Spatafora et al. (2016). Symbiotic refers to "dissimilarly named organisms that live together", as Anton de Bary first coined the now popularized term in his 1878 lecture at the meeting of the Cassel Natural Scientists in Germany (Oulhen et al. 2016). Arbuscular mycorrhizal fungi are so named because of the distinctive "arbuscule" structure that they form inside plant

cortical root cells (Figure 1-1). These tree-like, fine hyphal structures are thought to be the main point of nutrient exchange between the plant host and AM fungus (Smith and Read 2008). Smith and Read (2008) point out that although we now call this important group of fungi "arbuscular mycorrhizal fungi", not long ago, we called them "vesicular arbuscular mycorrhizae", due to the fact that many of these fungi produce vesicles and arbuscules; more recently, we have seen that not all AM fungi produce vesicles and other fungi produce similar structures. Similarly, we have recently seen that not all AM fungi may be able to produce true arbuscules and some non-AM fungi produce arbuscules (Smith and Read 2008; Orchard et al. 2017a; Orchard et al. 2017b). Smith and Read (2008), therefore, argue for a more objective naming concept for AM fungi. Other notable features of AM fungi include aseptate (coenocytic) hyphae that grow mostly within the plant cortical root tissue; AM hyphae, unlike ectomycorrhizal hyphae, grow between and within cells, hence the name *endo*mycorrhizae (Smith and Read 2008). Many AM fungi produce hyphal coils instead of, or in addition to, arbuscules. Extraradical hyphae and intra- and extra-radical spores are also commonly produced by AM fungi.

As Smith and Read (2008) and others argue, naming of AM fungi must not be solely based on morphological features (Öpik and Davison 2016). As alluded to earlier, recent genomic research places AM fungi within a new monophyletic subphylum, the Glomeromycotina, and within a new fungal phylum, the Mucoromycota (Spatafora et al. 2016). Though Tedersoo et al. (2018) provide evidence for keeping the older name of Glomeromycota. This AM fungal naming feud continues because the phylum-level and subphylum-level taxonomic names are more subjective than lower ranking names

(McNeill and Turland 2012). There are currently ~288 AM fungal species descriptions (Öpik and Davison 2016) according to Dr. Arthur Schüßler's webpage dedicated to AM fungal phylogeny and taxonomy (http://www.amf-phylogeny.com/). Despite the small size of this group of fungi, the arbuscular mycorrhizae are ecologically very important, as they associate with over 80% of terrestrial plants (Wang and Qiu 2006; Smith and Read 2008); the arbuscular mycorrhiza is the most widespread mycorrhiza on the planet (Wang and Qiu 2006; Opik et al. 2013). In fact, AM associations are so ubiquitous that we often say that 'it is easier to list the plant families in which it is not known to occur than to compile a list of families in which it has been found' (Gerdemann 1968; Smith and Read 2008). AM fungi have been shown to influence plant communities and play a large role in total carbon storage and movement. Global AM fungal biomass has been estimated to total ~1.4 pg of dry weight (Treseder and Cross 2006) and their intraradical hyphae may account for about 4% of the worldwide total microbial carbon pool (Paul and Clark 1996; Treseder and Cross 2006).

It's important to learn more about AM fungal diversity and function, particularly in understudied ecosystems, because they may play a large role in in ecosystem recovery during climate change through soil carbon dynamics (Bago et al. 2000; Mohan et al. 2014). Microbes control a large portion of the greenhouse gases (Conrad 1996). Historically, when we have incorporated more microbial data into our climate change models, our models have been much better predictors (Powell et al. 2015; Treseder 2016). AM fungi are plentiful and ubiquitous; they can improve carbon cycling by often increasing plant primary production and carbon storage capacity belowground (Treseder 2016).

Studying AM fungal diversity in order to better predict global C flux is also realistic and within our capability. AM fungi are currently placed phylogenetically within one subphylum, the Glomeromycotina (Spatafora et al. 2016), with only ~25 known genera, 11 families, and 4 orders (Redecker et al. 2013). However, this doesn't consider all the Virtual Taxa (VT) that have been collected from natural environments through DNA sequencing (Opik et al. 2010). Clearly there are still more AM fungi that need study and description. But compared to other, much vaster groups of microbes, the study of AM fungal diversity and how different species may interact with their environment is completely within our reach.

Ectomycorrhizal fungi (ECM), by comparison, are a much larger group of fungi, consisting of ~5,000-6,000 species in 250 genera and across at least 3 phyla that associate mostly with woody perennial plants and form specific fungal structures, such as a sheath of mycelium around the root, a Hartig net within the root tissue and extraradical hyphae (Smith and Read 2008). They have similar mutualistic function to AM fungi, providing nutrients (mostly N and P) to their plant partner in exchange for fixed carbon (Sanders and Tinker 1971; Smith and Read 2008). They associate commonly with members of the Pinaceae and Fagaceae plant families, as well as, many tropical trees, such as members of the Dipterocarpaceae, and even many AM host plants also host ECM (Taylor and Alexander 2005; Smith and Read 2008). Ectomycorrhizal fungi, unlike AM fungi, are not monophyletic. Rather, they are a polyphyletic group of organisms lumped together for the sake of discussing organisms with very similar functions and morphology. They are more recently evolved fungi, compared to AM fungi; the first ECM producing fungus evolved ~145-200 million years ago (Martin et al. 2017). Additional mycorrhizal

associations include ectendomycorrhizas, arbutoid mycorrhizas, ericoid mycorrhizas, orchid mycorrhizas, and mycoheterotrophic mycorrhizas, each considered a different "type" based on their unique structures and host associations.

AM fungi are historically very significant, compared to all other mycorrhizae. Fossil records (Pirozynski and Malloch 1975) and other evidence (Heckman et al. 2001; Wang et al. 2010) have pointed to the Glomeromycotan lineage as the pivotal symbioses that allowed for plant adaptation to land life. The first Glomalean-like organisms were found in fossils from both the early Devonian, >400 million years ago (Remy et al. 1994), and the Ordovician , ~460 million years ago, possibly before the earliest plants adapted to life on land (Redecker et al. 2000). It is hypothesized that the plant-fungal symbiosis is what originally made the move to land possible (Pirozynski and Malloch 1975). Although there is disagreement as to whether or not the first plant to adapt to terrestrial life had a symbiotic association with an ancient arbuscular mycorrhizal fungus (Glomeromycotina) or with an ancient member of the Mucoromycotina (Field et al. 2015), it is clear that 1) AM fungi and Mucoromycotina associations were both significant in the adaptation of plants to terrestrial life challenges and 2) either way, the AM fungal association has become much more ubiquitous and arguably a more "successful" symbiotic relationship.

Naming of AM fungi has been a particularly complicated and ever changing problem, so much so, that even taxonomists have to use special taxonomic translation tables to clarify which clades are being discussed in the literature ((Öpik and Davison 2016). A lot of work still needs to be done and is currently being done all over the world to describe more AM fungal species. Part of the problem lies in the fact that AM fungal

spores are multinucleate, meaning each asexual spore contains more than one nucleus (Smith and Read 2008; Marleau et al. 2011), and in the case of AM fungi, often thousands of nuclei, making the spores heterokaryotic. The same AM fungus may contain anywhere from very little to tremendous amounts of polymorphism, making phylogenetic analysis complicated at best (Lin et al. 2014; Wyss et al. 2016).

Despite the constant struggle to make robust phylogenetic conclusions on the AM group of fungi, scientists are looking for novel species worldwide in order to attempt to bridge the current gaps in knowledge. Brazil, Switzerland, Spain, and Poland are the current hot spots of newly described AM fungal species (Gamper et al. 2009; Blaszkowski et al. 2013; Oehl et al. 2014; Öpik and Davison 2016). However, data have shown that there are potentially many more AM fungal species yet to be uncovered in all sorts of environments. AM fungi have been shown not to be limited to stereotypically diverse hot spots like tropical or temperate forests and grassland ecosystems (Davison et al. 2012; Fajardo et al. 2015), but also in arctic, boreal, and polar climates (Opik et al. 2003; Opik et al. 2013; Varga et al. 2015), disturbed and anthropogenic sites, (Pendleton and Smith 1983; Read 1991; Herrmann et al. 2016), geothermal soils (Appoloni et al. 2008), and even in aquatic ecosystems (Moora et al. 2016). AM fungi can be found in nearly every ecosystem on the planet (Manoharachary et al. 2010). This dissertation study aims to further elucidate the AM fungal tree of life through the exploration of AM fungal communities in underexplored ecosystems.

The importance of research on AM fungal biodiversity and function cannot be understated. For one, AM fungi are extremely ubiquitous associations that account for a large portion of our crop success (Smith and Read 2008; Gianinazzi et al. 2010). AM

fungi have been shown to enhance a plant's absorption of many nutrients. Phosphorus is one of the most important nutrients for which AM fungi enhance plant uptake, but nitrogen and other nutrients are also brought into the plant's roots via AM fungal extraradical mycelium (Smith and Read 2008; Koltai and Kapulnik 2010). Additionally, AM fungi have been shown to improve water absorption, help stabilize soil aggregate structure through uniquely AM fungal glomalin production, and decrease soil erosion (Wright and Upadhyaya 1996; Auge 2001; Rillig et al. 2002; Smith and Read 2008; Wilson et al. 2009). Many agricultural crops associate with arbuscular mycorrhizal fungi and, therefore, this relationship has been and must continue to be prioritized when thinking about crop production.

AM fungi contribute to ecosystem services, as a whole (Dighton 2016). Ecosystem services are the benefits that different ecosystems (forests, agroecosystems, streams, grasslands, etc.) provide to humanity and that plants provide (Gianinazzi et al. 2010). Not only do AM fungi increase a plant's absorption of certain limited or otherwise unavailable nutrients, but they can also provide protection from pathogens. The AM association has been shown to protect against pathogenic nematodes, viruses, fungi, and bacteria (Harrier and Watson 2004; Liu et al. 2007; Gianinazzi et al. 2010; Newsham 2011; Liang et al. 2015; Pozo et al. 2015). Additionally, in an ever-changing global climate, it is important to understand what shapes the belowground biodiversity so that we can better predict the tightly-linked changes in plant community resilience.

Abiotic effects on arbuscular mycorrhizae

A large focus of mycology has been on mycorrhizal relationships due to the fact that they are ubiquitous in nature and, in the case of arbuscular mycorrhizae, can be

intertwined with crop production. Because the world's massive human population size relies on consistent and successful crop management and production, AM fungi are often studied in cooperation with crop plant hosts as inoculum. Many studies suggest improved crop production with inoculation of specific AM fungi (Gamalero et al. 2004). The studies conducted here have focused on natural ecosystems and landscapes with small amounts of management (i.e. some nitrogen fertilizers and/or mowing occasionally).

Soil properties are often shown to play a role in influencing AM colonization and diversity. AM fungi are great at helping plants retrieve immobile nutrients, such as soil P and N. In return for the donations of P and N to the plant host, AM fungi receive fixed, organic carbon (C) (Hetrick 1989; Bolan 1991; Smith and Read 2008). One might, therefore, predict that soil P would be a large determiner of plant-AM symbioses and, so too, plant success. In other words, with less access to P, a plant will rely on and therefore become more colonized by AM fungi. Smith and Read (2008) insist this is a gross oversimplification. Despite the fact that plants often rely on AM fungi to retrieve P for them, the extent which P affects colonization is greatly influenced by host plant identity, by amount of sunlight, and other context dependent factors (Son and Smith 1988; Smith and Gianinazzi-Pearson 1990; Baon et al. 1992). However, results of studies of interacting effects of P and light are mixed. Studies by Son and Smith (1988), Hayman (1974) and many others have shown clear decreases in mycorrhizal colonization under decreased light conditions, with more exaggerated effects in response to addition of P. The reasoning being that less light means less sugars made by the plant, and therefore, less available to the fungus. Contrastingly, Tester et al. (1986) found no effect of

decreased light on AM colonization with *Trifolium*. Others found that the impact of decreased light was magnified when P was added to the system in both *Allium* (Amijee et al. 1989) and *Cucumis* (Bruce et al. 1994). It is plain to see that research untangling the effects of P, light, and host species under large sample sizes will be needed. This dissertation considers these potentially confounding factors through large sample size and the selection of several sample sites to best measure the impacts on AM fungal colonization and diversity.

In the meta-analysis conducted by Hoeksema et al. (2010), host plant functional group and nitrogen fertilization were unexpectedly more important in predicting plant response to AM fungi than other variables (i.e. phosphorus fertilization, mycorrhizal species identity, soil biotic complexity). Similarly, other papers have shown greater effects of nitrogen fertilization on AM fungal abundance, compared with phosphorus fertilization; increases in nitrogen fertilization yielded greater decline in AM fungal abundance (Treseder 2004; Gerz et al. 2016) and changes to AM communities (Egerton-Warburton et al. 2007; Porras-Alfaro et al. 2007). Phosphorus fertilization tends to only slightly decrease AM fungal colonization, while the addition of carbon (in the form of CO_2) increases AM fungal abundance (Treseder 2004). However, nitrogen fertilization effects may not be as consistent as P and CO_2 effects, which tend to be consistent throughout the literature (Treseder 2004). One reason for this may be that inorganic nitrogen is more available to plants, compared to inorganic phosphorus. Also, AM fungi might just be better at providing inorganic phosphorus to plants (Smith and Read 2008). Another reason might be that some landscapes are simply more nitrogen limited than others and, therefore, are more likely to increase AM fungal abundance (Treseder and

Allen 2000). The limiting nutrients in varying ecosystems may play a large role in determining the effects of various nutrients on AM fungal abundance and diversity, particularly because AM fungi derive much of their carbon from their plant partner. Without adequate nutrition, the plant host will likely not choose to "share" its carbon with their AM fungal symbionts (Treseder and Allen 2000).

Soil pH has been shown to be an important factor impacting AM fungal colonization and diversity. However, studies have shown varying effects of soil pH (i.e. higher colonization by AM fungi at neutral pH versus higher colonization at more acidic pH, etc.); the impacts of soil pH seem to be context dependent. For instance, a study done in the United Kingdom, looking at soil pH effects on oat and potato root AM colonization, showed no significant difference in colonization at soil pH levels ranging from 4.5 to 7.5 (Wang et al. 1993). The extreme soil pH levels used in the study by Wang et al. (1993) seem to have decreased AM colonization slightly, with a 'happy medium' pH of 6.5. Other studies show increased AM colonization in more alkaline soils (>7.0). Specifically, Ouzounidou et al. (2015) showed that chia roots had increased AM colonization at soil pH of 8.1, compared to neutral and acidic soil pH.

It seems that the particular AM species or host plant species can alter the outcome of AM fungal studies. For instance, in Ouzounidou et al. (2015), *Glomus mosseae* was used as a single inoculum on chia plants, compared to a mixed, natural selection of AM fungi used for inoculum by Wang et al. (1993) on spring oat and potato plants. Specific AM fungal species may have more or less tolerance to shifting soil pH (Sieverding 1991). Some AM fungi have been found only in acidic conditions, while others have been

captured from both acidic and alkaline conditions (Robson and Abbott 1989; Abbott and Robson 1991).

The biogeography and local adaptations to soil features can make a huge impact on AM fungal diversity and function. Johnson et al. (2010) provide evidence for local adaptation of AM fungi in their full factorial pot experiment. Their study showed that AM fungi adapt to their native soil type over time and are more likely to infect and colonize plants grown in their local soil compared to foreign soil (Johnson et al. 2010). Therefore, altering soil pH alone, may not capture all the major influencers of AM fungal abundance and diversity in nature.

Global AM fungi studies have shown that biomes and regional geographic patterns may matter more, overall, than soil pH or phosphorus alone when it comes to AM fungal global diversity and macroecology (Opik et al. 2013; Partel et al. 2017), in a sense, agreeing with Johnson et al.'s (2010) conclusions on adaptation to local soil types. Macroecology is the branch of ecology that emphasizes large, continental effects on diversity and how species interact with their environment; it looks at how space, time, and resources affect biota (Brown and Maurer 1989; Keith et al. 2012). Though Brown and Maurer (1989) would probably argue that regional effects impact AM fungal community dynamics more than local effects, it is likely a combination of both local and regional effects that dictate AM fungal diversity and ecology. For instance, Partel et al. (2017) used global AM fungal diversity data to show that local diversity, overall, was influenced by both historical and local (contemporary) variables. Species pool size was shown to be largest in the tropical grass biome, evidencing that species pool size is largely affected by historical (regional) biome distribution (Parr et al. 2014; Partel et al.

2017). Similarly, in the global AM study by Öpik et al. (2010), two thirds of all virtual taxa (VT), i.e. DNA sequences, uncovered, had restricted continental distribution. In other words, the majority of the AM fungi captured in this study were affected by regionality. Confounded factors are the source of major issues in all scientific fields, microbiology being no exception.

Geographic and climatic factors may be confounded in many studies claiming more effects of regionality on AM fungal diversity. Öpik et al. (2013) found a definite relationship between AM fungal diversity and continental distribution. However, they mentioned that their findings could be the result of confounding factors, as they were not able to tease apart geographic effects from climate effects. Therefore, future studies testing both of these effects would be greatly important to our understanding of AM fungal distribution patterns.

Host specificity of arbuscular mycorrhizal fungi

Host specificity has been assumed to be very low with regards to plant host selection by fungus. They have generally been assumed to be generalists, for the most part. Early on, Stahl (1949), Magrou (1936), and Gerdemann (1955) rightfully concluded that there is likely little host specificity when it comes to AM fungi preferring specific plant taxa over others. After all, there are < 300 described species of AM fungi and hundreds of thousands of associated plants; therefore, they must be associating with many plant taxa (Smith and Read 2008). Additionally, AM fungi do not disperse well and therefore, would do best to adapt to many host taxa, to better their chances of receiving organic C at any given site (Smith and Read 2008).

Despite what logical evolutionary hypotheses and some studies might show, some experiments have shown some host specificity of certain AM fungal species. For instance, Graw et al. (Graw et al. 1979) found in their study of 19 different plant hosts, that *Glomus gerdemanni* only formed a mycorrhizal association with 1 of the hosts. Vandenkoornhuyse et al. (2003) similarly found differences in AM fungal communities between different grassland species. More typically, though, AM fungi are able to associate with many hosts. This is seen in the usage of pot cultures to maximize healthy spore populations of native AM fungi from the field. Pot cultures notably use different hosts from that which the AM fungus was originally derived (Gilmore 1968). While there may be some exceptional AM fungi that have evolved to associate with a narrow range of hosts, host characteristics may be more important in delineating host-fungus associations. For example, Koorem et al. (2017) found AM fungal community membership was highly depending on the shade tolerance of the plant hosts. We also know from pot culture capture of AM fungal spores, that plant host identity can play a large role in the amount of sporulation of AM fungi (Bever et al. 1996; Jansa et al. 2002). It is plain to see that some AM fungi likely adapted to their environmental pressures by becoming generalists and others adapted strictly to certain host characteristics or taxonomic identity. This dissertation recognizes the likely host generalization of many AM fungi, and instead, aims to uncover any biogeographic or edaphic effects on AM fungal communities.

Methods in AM fungal biodiversity and relevance of metagenomics

In order to measure the abundance, diversity, and ecological function of AM fungi, it is important to be aware of the current available methodologies for observation

and quantification of AM fungi. Microscopy has been shown to be a terrific way to quickly and economically observe and quantify AM. Many staining protocols have been used over the years, with varying results, depending on thickness of roots, tannin presence, amount of time soaked, and dilution of stain, to name just a few important variables (Gange et al. 1999; Pitet et al. 2009; Diagne et al. 2011; Rath et al. 2014). Embedding roots in gelatin (Smith and Dickson 1991) is an older, less common mode of observing AM fungal colonization.

There are several ways to calculate the extent of root tissue colonized by AM fungi. Popular methods include those by McGonigle et al. (1990) and Giovannetti and Mosse (1980). Staining and microscopy, however, do not provide much diversity information; AM structures within the plant roots cannot provide much in the way of diversity information due to the non-species-specific nature of their structures. In order to better describe AM fungal diversity patterns, scientists have historically relied on spore morphology (Öpik and Davison 2016). Spores often have unique characteristics and are large enough to be seen with either the naked eye or a simple dissecting microscope. However, some AM fungal species have similar morphological features and can easily be parasitized, making it quite difficult to pinpoint the correct AM fungal identification; extracting, cleaning, and describing spores to the genus or species level, therefore, requires years of training and can be extremely laborious.

Growing attention and enthusiasm has been given to AM fungal diversity over the last half century. Interest in the diversity of AM fungi has been growing in particular, in order to help increase the essential ecosystem services these fungi can provide for humankind. However, methodological constraints are often barriers to increasing our

knowledge base. AM fungi are obligate biotrophs that cannot be grown on artificial media, unlike some other fungal groups; they need to be living within and amongst their living plant host to grow and thrive. Indeed, one of the greatest contributions to the study of AM fungi would be to figure out how to grow these symbiotic fungi axenically (Stajich et al. 2009). Select studies have used artificial media to grow AM fungi, but these studies are typically done in cooperation with living host tissue (Karandashov et al. 1999; Douds 2002; Debiane et al. 2011). The cost and time associated with this technique also prove it impracticable as a tool for studying diversity of AM fungi in natural systems.

Because AM fungi cannot be grown on artificial media and sequenced using conventional Sanger sequencing methods, molecular cloning of AM DNA into bacterial vectors has been used to determine the diversity of AM fungi from environmental samples. However, this procedure involves multiple problems, including but not limited to: positive clone selection bias, cost constraints, laborious process, and limitations on completeness of the true diversity (significantly unsaturated rarefaction curves). Therefore, next- generation sequencing (NGS) has become an invaluable and revolutionary tool for uncovering the diversity of microorganisms in different environments (DeLong et al. 2006; Nelson et al. 2010; Qin et al. 2010). In particular, the study of microbial, and AM fungal metagenomics through NGS, in particular, has mushroomed.

The term ‘metagenomics’ was first coined by Handelsman et al. (1998); metagenomics is the study of the genetic material obtained directly from the environment. NGS refers to the replacement of conventional Sanger sequencing with

much faster, massively parallel sequencing of fragments of DNA (Behjati and Tarpey 2013). NGS technologies have revolutionized biodiversity research since the late 1990s. Roche 454's pyrosequencing approach was the first technology to conduct massively parallel sequencing, along with the Illumina/Solexa tag-based platform and the lesser used ABI SOLiD, and the newer Ion Torrent platform following the trend towards NGS-based research (Bennett 2004; Margulies et al. 2005; Shendure et al. 2005). Between the ever-decreasing price per read and the increasing access and availability of bioinformatic resources, NGS technologies have become the wave of the future.

Most of the NGS literature on AM fungal diversity has utilized Roche 454 sequencing. 454, however, has very high read error (Quince et al. 2009), and up to 15% of reads can be artefactual (Gomez-Alvarez et al. 2009). In a direct comparison of 454 and Illumina, utilizing microbial data, Luo et al. (2012) showed that the two technologies are both viable options for microbial diversity studies. Nevertheless, they found that Illumina consistently revealed a higher percentage of the reference genomes, compared to 454. Illumina also showed an overall lower sequence error rate and, surprisingly, generated longer, more accurate reads (Luo et al. 2012). Additionally, Luo et al. (2012) the Illumina reads were run at about a quarter of the cost of the 454 reads, providing substantial basis for Illumina sequencing for future microbial diversity studies. These results are similar to those found in studies by Quince et al. (2009) and Margulies et al. (2005).

Hart et al. (2015) break down all the parts of designing and implementing NGS-based AM fungal studies and point out ways in which bias can be avoided through thoughtful and strategic decision making. Despite NGS's new and exciting possibilities,

these technologies bring the potential for a tangled web of issues. For one, the AM specialists of just twenty years ago are better trained in ecology and morphology of AM fungi, and less so in the ways of PCR, sequencing, and bioinformatic analysis. Exceptional specialization in AM fungal taxonomy and bioinformatics, as well as superb collaboration on the parts of the different specialists, will be required in order to generate accurate and informative data (Grube et al. 2017). PCR primers need to be carefully considered, as the ITS universal fungal barcode region is too variable among most AM fungal species (Stockinger et al. 2010; Schoch et al. 2012; Hart et al. 2015). The large subunit (LSU) region is generally thought to be the best for delimiting phylogenetic relationships for AM fungi. Nevertheless, the small subunit (SSU) region is most prevalent in the AM fungal nucleotide database (Opik et al. 2010) and is frequently used in AM fungal ecological studies (Öpik et al. 2014). AM fungal primer selection, analyzed in several reviews (Kohout et al. 2014; Öpik et al. 2014; Van Geel et al. 2014), is just one portion of the workflow involved in NGS-based studies of AM fungi. Many decisions need to be made when using NGS technology for biodiversity research and this can potentially result in nonreplicable studies, bias, and necessitates collaboration with knowledgeable experts.

Due to the new and potentially complicated nature of NGS, Illumina has only recently been utilized for AM fungal community studies. A few of the only published AM fungal community studies utilizing Illumina technology are those by Cui et al. (2016), Liu et al. (2017), and Ban et al. (2017), Orchard et al. (2017a), Wang et al. (2017), and Johansen et al. (Johansen et al. 2016) all published within the last two years. None of these papers include ample detail as to the molecular methodological choices

and reasoning made by the authors (i.e. the exact scripts used to sort and denoise the data). Morgan and Egerton-Warburton (2017) develop a workflow using Illumina NGS for AM fungal diversity analysis, however, they test only one primer pair and do not test the results of using different reference databases. While there is some available literature on how to analyze Illumina reads, they were mostly developed for the ITS or 16s (prokaryote) regions and/or require substantial bioinformatic decision making and script writing ability on the part of the user (Caporaso et al. 2010; Kuczynski et al. 2011; Hart et al. 2015). Vasar et al. (2017) compared Illumina to 454 sequencing of natural plant-root samples and AM specific primers; they found similar AM richness, after careful quality filtering, using both techniques.

This dissertation includes a user-friendly, step-by-step guide to AM fungal Illumina read analysis from start (i.e. site and sample selection) to finish (i.e. comparison of sequence diversity) (Chapter 3), including useful scripts and potential pitfalls. This dissertation also formulates a basic pipeline for the utilization and bioinformatic analysis of AM fungal rDNA sequences, using Illumina NGS and Linux based commands. It is important to note that decisions along the workflow can make a huge impact on the downstream analysis and on any conclusions drawn. These decisions include everything from sampling scheme, pooling strategies, PCR primer selection, PCR annealing temperatures, bioinformatic tools used, database of known sequences used, etc. And this is just the tip of the iceberg.

Significance of switchgrass (*Panicum virgatum)*

Switchgrass (*Panicum virgatum* L.*)* is a perennial, warm-season C_4 grass in the Poaceae plant family. Switchgrass is native to much of the United States; its native

habitat (mostly unforested areas) stretches for about two-thirds of the eastern Unites States until the western front of the Rocky Mountains, with northern limits in southern Canada and southern limits all the way to the coasts of Texas (Figure 1-2) even prior to European colonization and cultivation (Hitchcock 1935; Parrish and Fike 2005). There are two genetically distinct varieties of switchgrass, the lowland and upland ecotypes. The lowland ecotype is typically associated with warmer climate and is known to produce larger amounts of biomass; the upland ecotype is more cold hardy and is, therefore, found in more northern climates (Parrish and Fike 2005). It was originally bred for use as a forage crop, as it is a hardy grass with wide geographic range and hardy lifestyle (Parrish and Fike 2005). Even before Europeans began breeding switchgrass for animal forage, it was likely used in its natural, native state. Its historical ecology has been strongly tied to animal usage via stomping and grazing (Eom et al. 2001), as well as, frequent burning (Rice and Parenti 1978; Cuomo et al. 1998).

Switchgrass is now used for many other purposes, including wildlife conservation, prairie restoration, as well as, a potential biofuel source (Paine et al. 1996; Sanderson et al. 1996; Vogel 2004). The Department of Energy has, in the past ~25 years, researched the usage of switchgrass for potential as a bioenergy crop (Sanderson et al. 1996). It has been considered a great candidate for biofuel production because of its large habitat range and broad adaptability to varying conditions. It has been shown to be adaptable to drought, heat, nutrient, and salt stresses (Ashworth et al. 2016). Switchgrass performs very well under tough conditions, forming large, dense tufts of grass (Agriculture).

Climate change data has shown decreases in rainfall by a whopping 7.7% from just 1997-2008 (Karl et al. 2009). With increasing average temperatures and sudden bursts of rainfall, particularly in the southern US, switchgrass seems to be an excellent candidate to replace fossil fuels; it has been shown to adapt well to high heat conditions (Ashworth et al. 2016) Additionally, switchgrass has been shown to be environmentally beneficial in soil and wildlife conservation (McLaughlin et al. 1994), with adaptation to soil pH ranging from 4.5 - 8.0 (Agriculture) and the ability to improve water runoff and soil erosion (Self-Davis et al. 2003).

Switchgrass does not have high nitrogen needs, with particular adaptability to soils with little available inorganic N (Ashworth et al. 2016). Nevertheless, controlled studies have shown obvious increases in switchgrass biomass as a result of N fertilization (Jung et al. 1988; Ashworth et al. 2015). Maximized access to nutrients will surely be needed to produce extreme amounts of biomass (Parrish and Fike 2005). Switchgrass is currently an excellent candidate for biofuel production, for the aforementioned reasons. However, in order to meet the growing needs of families in the US alone, a larger effort will need to be made to produce the essential quantities of biofuel to replace fossil fuel usage. Amending the soil with loads of fertilizers and/or pesticides is not ecologically nor economically sustainable (Adler et al. 2007; Ashworth et al. 2015), nor is it entirely predictable, due to varied effects of different fertilizers on plant growth in regions of diverse land-uses and microbial communities, for starters. It, therefore, may prove more worthwhile to invest in research on the microbial communities associated with this potential biofuel crop.

The ubiquity of perennial grass-associated AM fungi has only recently been explored (van der Heijden et al. 2006). Switchgrass AM fungal communities are still not well understood (Research for Sustainable Bioenergy: Linking Genomic and Ecosystem Sciences 2014). To date, most studies on AM fungi associated with switchgrass focus on inoculation with select AM fungi for enhanced switchgrass growth and biomass yield (Clark et al. 1999; Clark 2002; Clark et al. 2005). Clark et al. (1999) showed enhanced mineral uptake, as well as, decreased Al toxicity of switchgrass under acidic soil conditions when inoculated with specific AM fungal isolates. However, this study was done with artificially created microbial communities, which may not occur together or in isolation in natural landscapes. Little explanation as to the reasoning behind their selection is given, other than to represent three distinct generic lineages of AM fungi. Similarly, Ghimire et al. (2009) showed improved performance of switchgrass biomass under AM fungi inoculated conditions. Clark et al. (1997) summarized the impacts of acidic soil (pH <5.0) on AM fungal host response and on AM fungi in general.

Other studies have looked at the impacts of field collected AM fungi on switchgrass growth (Hartnett et al. 1994; Wilson and Hartnett 1998; Hartnett and Wilson 1999; Schroeder-Moreno et al. 2012). However, no AM community analysis was conducted in the experiments by Hartnett and Wilson (1999) and Hartnett et al. (1994). In the experiments of Schroeder-Moreno et al.(2012) and Wilson and Hartnett (1998), only spore morphology was used to confirm which native AM species were used to influence host response; no root-associated AM fungal community analysis was conducted. Mao et al. (2014) conducted a molecular study on the diversity of microorganisms of switchgrass rhizospheric soil in a farm setting. However, because the

primers used in their study were originally designed as universal eukaryotic primers, mostly non-AM fungi were recovered and a deep phylogenetic analysis of AM fungi was not possible (Diez et al. 2001; Bailly et al. 2007). Mostly *Glomus spp.* were recovered in the Mao et al. (2014) study, further evidencing the findings of Clark et al. (2002) and Clark et al. (Clark et al. 1999) that various *Glomus* species can enhance switchgrass's acid-soil tolerance and nutrient uptake. No other AM fungal genera were able to be captured in the Mao et al. (2014) study. No one, thus far, has sufficiently investigated the diversity of root-inhabiting AM fungi associated with natural switchgrass populations; particularly lacking is an in-depth molecular analysis of these native AM fungi. Therefore, this dissertation aims to fill this knowledge gap with a molecular-based study of switchgrass root associated AM fungi from different locations.

The Pine Barrens ecosystem

The Pine Barrens ecosystem is a distinctive and rare ecosystem dispersed throughout the northeast United States and elsewhere, in lesser amounts, throughout the world (Forman 1998). The largest continuous Pine Barrens ecosystem lies in central/southern New Jersey, covering ~1.4 million acres (550,000 hectares) (Forman 1998). Figure 1-3 illustrates the extent of New Jersey considered Pinelands (i.e. Pine Barrens) territory and distinguishes land use, as of 2016 (Commission 2015). The Pine Barrens' distinction lies in its Northeastern pine-oak dominated forest with sandy, oligotrophic soils (Gamble 1963). The soils in the Barrens are distinctly acidic and low in nutrients. They are so acidic because there is little clay present, little humus, and low cation exchange capacity, thereby, inhibiting the soil from removing acidic compounds and maintaining its acidity (Forman 1998). There is also a lot of aluminum (Al) present

in the Pine Barrens soils, further preventing the soils from neutralizing (Douglas and Trela 1979; Geller 2002). The Pinelands were perceived by the earliest European settlers in the 1600's to be unfitting for agricultural production and 'barren', hence the name 'Pine Barrens' (O'Callaghan 1853; Forman 1998).

Another defining feature of the Pine Barrens is its frequent fires. Frequent intentional, prescribed burning was and still is practiced in the Pinelands (Wacker 1971). Despite receiving ~100-120 cm of rain annually (Biel 1958), much of the water is leached due to the sandy texture of the soil (Forman 1998). The effects of frequent fire on the Pine Barrens' vegetation are not to be ignored. ~50-80% of the Pine Barrens ecosystems are dominated by pitch pine (*Pinus rigida* L.*)* (Robichaud and Buell 1973; Forman 1998). Fire adapted plant species dominate the uplands section of the Barrens (the most common forest type in the Pine Barrens ecosystem, which was chosen for this dissertation's samples in New Jersey), with pitch pine, shortleaf pine (*Pinus echinata*), and black oak (*Quercus velutina*), blackjack oak (*Quercus marilandica*), and other oak species as the dominant uplands forest species. Where understory species are able to survive frequent fire disturbance, blueberry (*Vaccinium vaccilans),* sedges, switchgrass, and other grasses are common (McCormick and Buell 1957).

While fungal communities have been recently explored in the New Jersey Pine Barrens (Luo et al. 2014a; Luo et al. 2014b; Walsh et al. 2014; Walsh et al. 2015; Luo et al. 2017), with some work on ectomycorrhizal communities (Tuininga and Dighton 2004; Jonsson et al. 2006), only one study looked at arbuscular mycorrhization in the Pine Barrens (Dighton et al. 2013). The study by Dighton et al. (2013), however, only looked at colonization of one sedge and did not delve into any community analyses. In a study

on the bacterial and archaeal communities of the Long Island Pine Barrens in New York, Shah et al. (2011) found great diversity of microorganisms. The same belowground biodiversity trend could be seen with the AM fungal communities of the Pine Barrens.

AM fungal communities from ecosystems similar to the Pine Barrens have been explored. Southworth et al. (2014) review the AM fungal communities in serpentine ecosystems compared to non-serpentine ecosystems. They found that, despite predictions of low abundance and diversity of AM fungi in these oligotrophic soil systems, the opposite was true: serpentine soils housed abundant and diverse communities of AM fungi. In fact, AM abundance in serpentine soils was found to be similar to non-serpentine soils. Serpentine landscapes also had rich AM fungal communities, with some particularly dominant species of *Glomus* and *Acaulospora*. Interestingly, a study by Schechter and Bruns (2012), pointed out in the review paper by Southworth et al. (2014), showed that, when comparing the AM communities of the same host plant at serpentine versus non-serpentine sites, the communities had no overlap. This suggests that edaphic factors may play a role in shaping AM communities among certain host plants and may play less of a role in other hosts (Southworth et al. 2014). Similarly, in a review of AM fungal communities of acidic soils, Aguilera (2015) found *Glomus*, *Acaulospora*, *Gigaspora,* and *Scutellospora* to be the dominant genera of acidic soils. Aguilera (2015) emphasize the importance of AM fungi to the alleviation of high Al levels in some acidic soils. This is particularly relevant to the comparable soils of the Pine Barrens. There is great value to the study of AM fungal communities in underexplored habitats and extreme or disturbed habitats (Öpik and Davison 2016) and so this dissertation's focus is

on the investigation of AM fungal communities of the little-known, acidic soils of the Pine Barrens.

Conclusions, main objectives, significance of study

Arbuscular mycorrhizal fungi are a phylogenetically small group of fungi with big impact on our planet. Not only do they associate with virtually all major plant lineages, they can be found in most every ecosystem and under extreme conditions. Their true diversity is likely understated and unknown; their taxonomy and phylogenetic positions are tangled in years of assorted and confounded AM fungal ecological and phylogenetic research. Due to their ecological, agricultural, and economical importance, it is imperative that we continue to untangle the effects on AM fungal communities and their ecological function. In order to better understand AM fungal diversity and ecology, this dissertation sampled the roots of switchgrass (*Panicum virgatum* L.), a ubiquitous, historically and economically important plant species, from several sites in the New Jersey and Long Island Pine Barrens and beyond. This dissertation had three main objectives. The first objective of this study was to 1) identify and compare the switchgrass-associated AM fungal colonization patterns and communities between different Pine Barrens ecosystems. Additionally, we aimed to uncover whether differences in AM communities and the extent of colonization differed between Pine Barrens ecosystems and other ecosystems of different edaphic qualities and anthropogenic histories. These were accomplished using microscopic quantification of AM colonization and molecular analysis through Illumina NGS. The second objective was to 2) develop a practical, detailed and well-tested workflow for the bioinformatic analysis of AM fungal Illumina sequence data. Thus far, there are few user-friendly,

published pipeline for processing and interpreting such data. The last objective was 3) to begin to uncover the ecological function of native, New Jersey Pine Barrens AM fungi under controlled conditions, using switchgrass as the host plant. With this final pot experiment, we explored the impacts of decreased soil acidity on Pine Barrens AM fungal communities, as well as, on general AM fungal abundance and host response. This dissertation will add to our knowledge of AM fungal biodiversity in underexplored ecosystems and will help answer the long-standing question of what factors influence AM fungal diversity patterns.

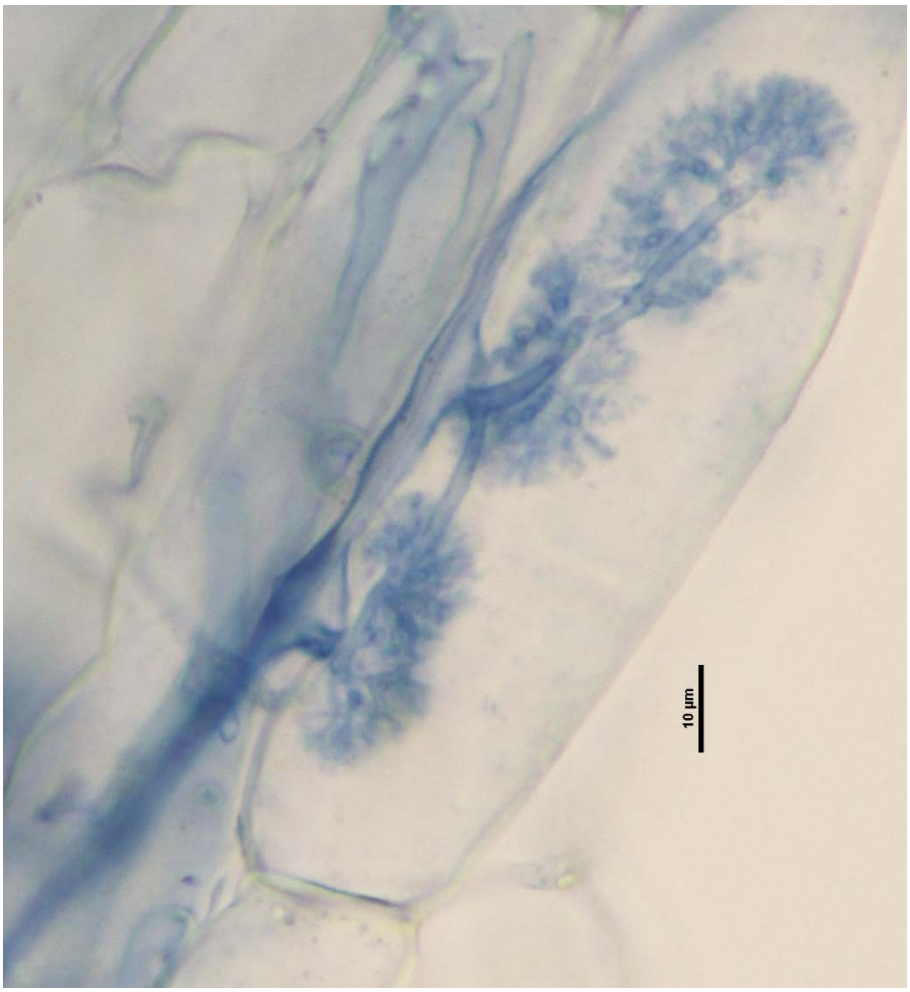

Fig. 1-1 Arbuscules colonizing the cortical root tissue of switchgrass (*Panicum virgatum* L.)

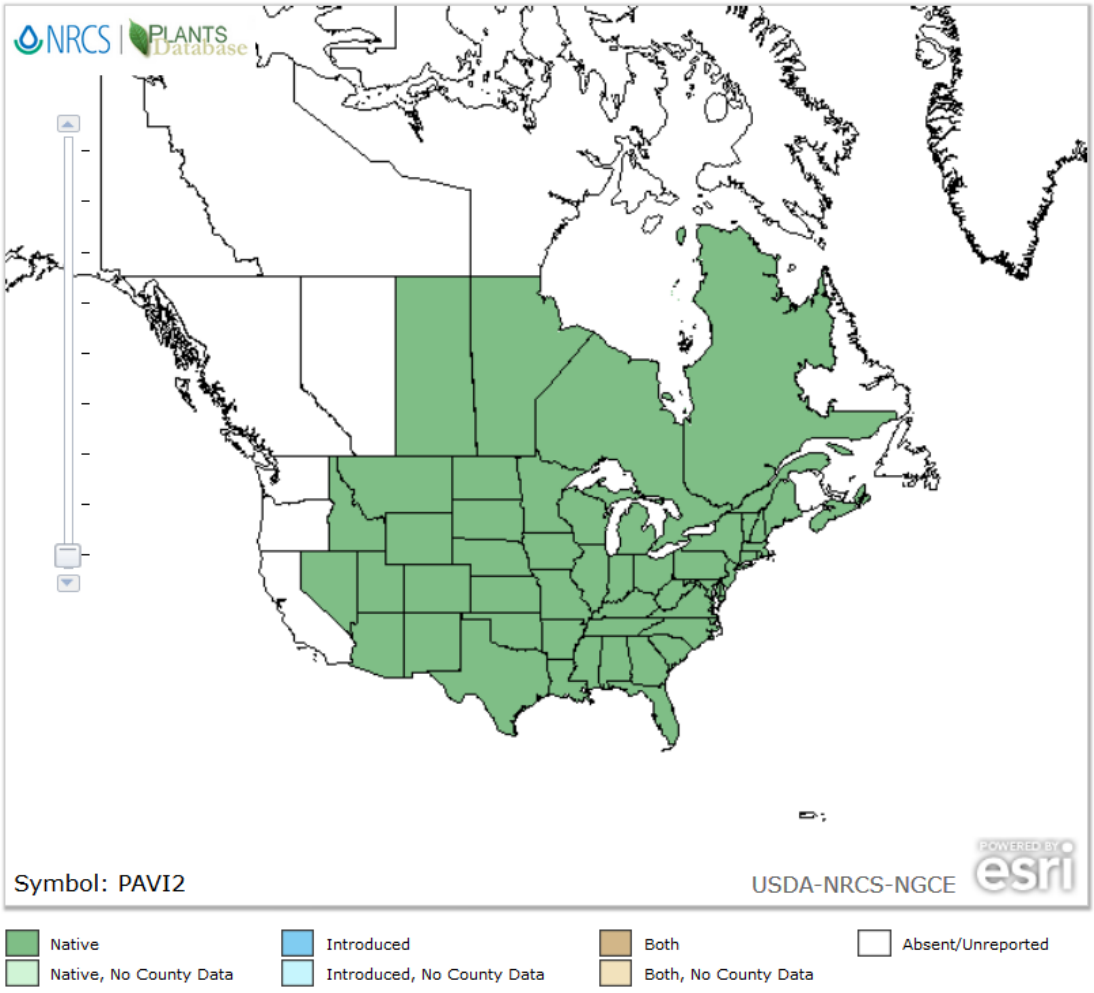

Fig. 1-2 Map of the distribution of native, North American switchgrass (*Panicum virgatum* L.), adapted from USDA-NRCS (USDA 2017).

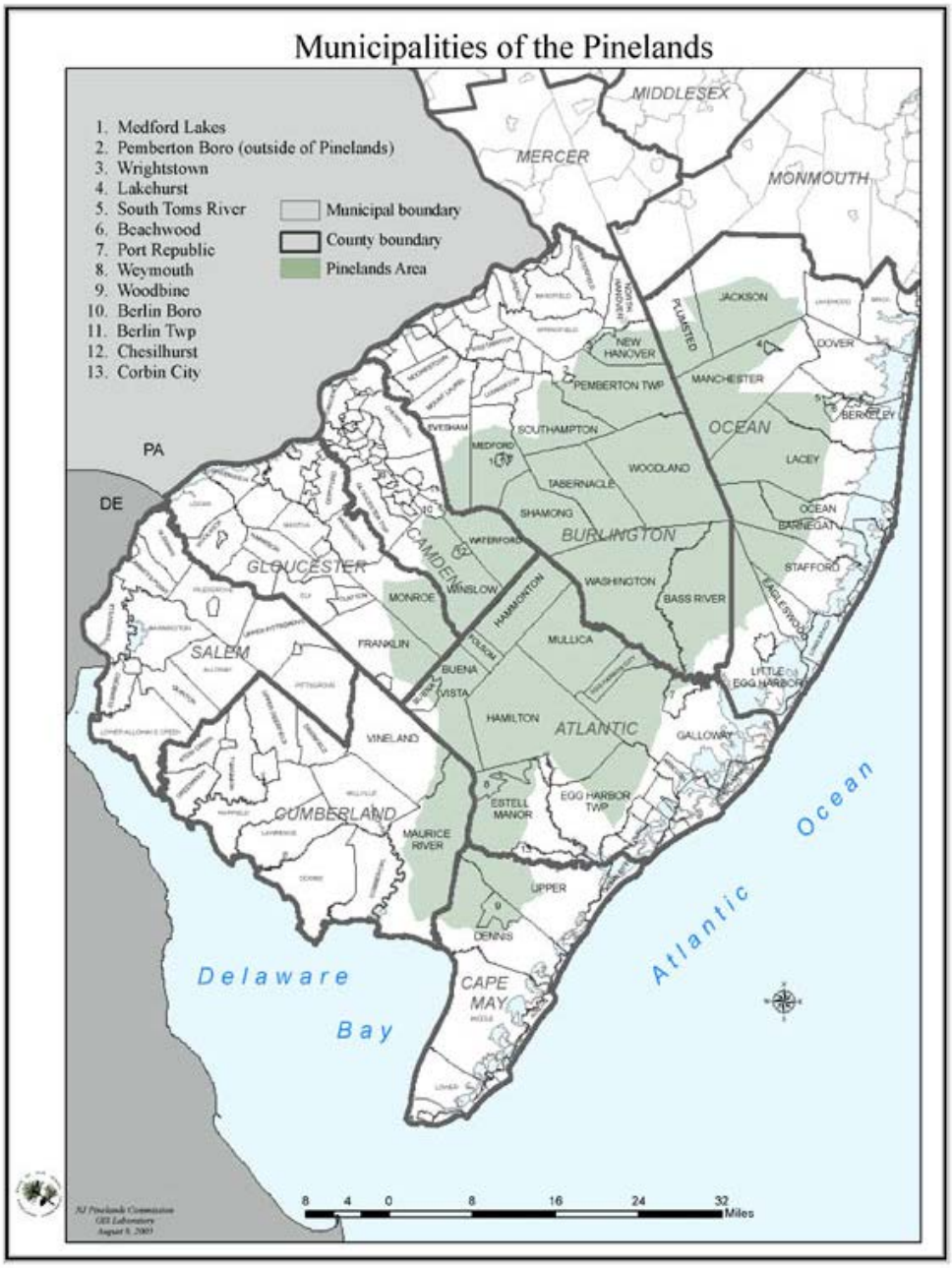

Fig. 1-3 Map of the distribution and municipalities of the New Jersey Pine Barrens ecosystem, adapted from the State of New Jersey Pinelands Commission (Commission 2015).

Chapter 2

Arbuscular mycorrhizal fungal communities associated with switchgrass roots are shaped by soil properties and land use[1]

Abstract

Arbuscular mycorrhizal (AM) fungi contribute globally to ecosystem services and play an important role in sustainable crop production in agricultural settings. However, it is unclear which factors contribute most to their colonization and community structure, particularly in understudied ecosystems. This study investigated how soil properties and varied land uses can influence the switchgrass associated AM fungal communities by comparing the understudied Pine Barrens ecosystem, an Iowa native prairie, as well as, more managed, agroecosystems. Both microscopy and sequencing with Illumina MiSeq technology were used. Results showed correlation between AM fungal communities and land use, soil pH, and exchangeable aluminum (Al). Soil pH was positively correlated with AM fungal diversity, while soil Al levels were negatively correlated with AM fungal diversity and abundance. *Glomus* was the most ubiquitous AM fungal genus recovered from all sites. *Acaulospora* and *Ambispora* were almost exclusively found in Pine Barrens sites. This study suggests differences in the AM fungal community structure under different soil conditions and land uses. This is the first sequence-based report of the AM fungal communities in the Pine Barrens ecosystem.

---

[1]This chapter was previously submitted for publication to Mycorrhiza as: Bindell, M, Luo, J, Walsh, E, Wagner, N, Miller, S, Zhang, N (2018) Arbuscular mycorrhizal fungal communities associated with switchgrass roots are shaped by soil properties and land use. The paper is reproduced here with few alterations other than formatting changes and minor edits for clarification purposes.

Introduction

Arbuscular mycorrhizal fungi (AM fungi) are root-associated, mutualistic fungi within the subphylum Glomeromycotina (Spatafora et al. 2016). Over 80% of all land plants associate with AM fungi, which are thought to have coevolved with their plant partners from the beginning of terrestrial plant life (Redecker et al. 2000; Smith and Read 2008). It was reported that more diverse AM fungal communities are associated with more diverse and productive plant communities (van der Heijden et al. 1998). Improving our knowledge of AM fungal species distribution will help us better understand why certain plants can survive under poor nutrient conditions and in heavily disturbed or extreme environments (Cumming and Ning 2003; Varga et al. 2015), particularly in a changing global climate (Johnson et al. 2013; Mohan et al. 2014).

However, our knowledge on AM fungal communities is still limited, particularly in many underexplored ecosystems. One understudied ecosystem is the Pine Barrens. The Pine Barrens ecosystem is scattered throughout the eastern United States and to a lesser extent, worldwide (Forman 1998). The Pine Barrens' distinction lies in its Northeastern pine-oak dominated forest with sandy, acidic, low nutrient soils and frequent fires (Gamble 1963; Forman 1998). They got their 'barren' name from the fact that the soils and frequent fire historically prohibited agricultural crop production. The Pine Barrens' soils are particularly acidic because there is little clay present in the soil, little humus, and low cation exchange capacity. These features inhibit soil from removing any naturally acidic compounds and, thereby, perpetuate its acidity (Forman 1998). There is also a large amount of soluble aluminum (Al) present in the Pine Barrens

soils, further preventing the soils from neutralizing (Douglas and Trela 1979; Geller 2002).

Hostile soil conditions (e.g., low nutrients, acidic and compacted soils) are common in both natural and anthropogenic lands. 30 – 40% of croplands are inhibited by acidic soils (below pH 5), making soil acidity one of the most crucial issues facing agroecosystems (von Uexküll and Mutert 1995; Iqbal 2012). Studies have shown varying degrees of impact of soil pH on AM fungal communities. Suzuki et al. (2014) and An et al. (2008) found soil pH to be major factors contributing to differences in AM fungal community composition under varying conditions. A biogeographic survey also found soil pH to have a major influence on AM fungal community composition, with some species found to be more prominent in more acidic soils (Stürmer et al. 2018). However, Johnson et al. (1991) showed little impact of soil acidity on overall AM fungal species richness or diversity. Different plants may have varied tolerance to acidic soil and to the accompanying mineral deficiencies (i.e. P and K) and metal stresses of these soils (i.e. Al and Mn toxicity) (Marschner and Dell 1994) (Marschner 1991).

Al is more soluble in low pH soils (Kochian 1995; Muthukumar et al. 2014) and may cause Al toxicity, which is a prime issue in crop production in acidic soils (Barinaga 1997; Kochian et al. 2005; Muthukumar et al. 2014). Elevated Al impacts root growth and, simultaneously, nutrient absorption (Kollmeier et al. 2000; Ma et al. 2001). Al^{3+} is the predominant form of Al in acidic soils, which is solubilized as the soil pH decreases (Kochian et al. 2005). In addition to more traditional lime and phosphorus amendments, AM fungi can be instrumental in decreasing Al phytotoxic effects (Mendoza and Borie 1998; Borie and Rubio 1999; Seguel et al. 2013). AM fungal taxonomic identity can play

a role in determining the effects these microbes have on Al remediation, as different species can be more or less adapted to acidic pH and aluminum toxicity (Siqueira et al. 1984; Kelly et al. 2005; Klugh and Cumming 2007). Therefore, gaining insight into the naturally occurring AM fungi in varying environments, including acidic soils and anthropogenic sites, may help determine which AM fungal species may be important in alleviating crop stresses.

This study looks at the AM fungal communities of switchgrass (*Panicum virgatum* L.), a C4 warm-season perennial grass that is native to most of North America and is a model bioenergy feedstock species (Parrish and Fike 2005). Switchgrass was an historically important forage grass but is now also used for wildlife conservation, prairie habitat restoration, and more recently, for biofuel production (Paine et al. 1996; Sanderson et al. 1996; Vogel 2004; Parrish and Fike 2005). Its great adaptability to many uses stems from switchgrass's adaptability to a wide range of soil and climatic conditions, including those with low water and nutrient availability (McLaughlin 1993). Switchgrass is one of the native and dominant grasses in the oligotrophic Pine Barrens in the eastern North America. Most switchgrass associated mycorrhizae studies on acidic soils have focused on the inoculation of the plant with specific AM fungal isolates to determine plant growth promotion (Clark et al. 1999; Clark 2002; Clark et al. 2005), while some have uncovered switchgrass associated AM fungal spore communities (Wilson and Hartnett 1998; Schroeder-Moreno et al. 2012). However, no study to date has comprehensively investigated native switchgrass associated AM fungal communities using a DNA sequencing approach.

While fungal communities and novel species have been recently explored in Pine Barrens forests (Luo et al. 2014a; Luo et al. 2014b; Walsh et al. 2014; Luo et al. 2015; Walsh et al. 2015; Luo et al. 2017), with some work on ectomycorrhizal communities (Tuininga and Dighton 2004; Jonsson et al. 2006), as well as, bacteria and archaea (Shah et al. 2011), only one study to date has investigated AM fungi in the Pine Barrens ecosystem (Dighton et al. 2013). However, this study by Dighton et al. (2013) only looked at the microscopic presence of AM fungi on a sedge species in the New Jersey Pine Barrens and did not identify or examine further on the AM fungal community. A comprehensive investigation into AM fungal diversity in the Pine Barrens is of great importance to understanding which AM fungi are present in the ecosystem with acidic, Al toxic soils.

A number of studies have been done to investigate the impact of land use, such as mowing and fertilizing, on AM colonization and community composition but the results were variable. Titus and Leps (2000) and Wang (2017) found decreased AM fungal colonization with the addition of fertilizers. Wang et al. (2017) and Johnson et al. (1993) showed that fertilization impacted AM fungal community composition, but Wang (2017) also found that increasing fertilization had no impact on the total number of AM fungal OTUs present in roots. Moora et al. (2014) and Van Diepen et al. (2011) similarly reported that intensive land usage can change community composition but Smilauer (2001) found little impact of mowing on AM fungal morphological diversity. Many studies focus on grasslands (Smilauer 2001; Santos et al. 2006; Valyi et al. 2015) or intensely farmed lands (Helgason et al. 1998), further widening the knowledge gap on rare ecosystems and low-input agricultural ecosystems; these lands may be the wave of

the future for agricultural production as climate change continues to alter our planet. A recent study by Sepp et al. (2018) detected differences in AM fungal community composition between different habitats with different land uses. However, they found pronounced differences between sites of the same habitat type, calling for further investigation into site differences (Sepp et al. 2018) and additional sample replication of study sites. The differences in results between the aforementioned studies could be due to differences in the N:P ratio in the soils (Johnson 2010), the specific habitat being studied (i.e. forest, grassland, or agricultural site) (Opik et al. 2006; Kivlin et al. 2011), and, notably, the method of measuring AM fungal abundance or community composition (Jumpponen et al. 2005).

Each method for AM fungal diversity studies has its drawbacks. Microscopy is one way in which we have learned about AM fungal presence in plant roots. However, this method is not able to distinguish between AM fungal species. Many have turned to next-generation sequencing (NGS) technology, such as Roche 454 sequencing platform, to gauge AM fungal community information from environmental samples (Opik et al. 2009; Moora et al. 2014; Klabi et al. 2015; Egan et al. 2017). Recently, more studies have used Illumina NGS technology for uncovering environmental AM fungal communities (Cui et al. 2016; Ban et al. 2017; Liu et al. 2017; Morgan and Egerton-Warburton 2017). A recent study by Vasar et al. (Vasar et al. 2017) showed that the Illumina platform can help uncover a level of AM fungal species richness comparable to that of 454, at a significant cost reduction. In this study, we utilized both microscopy and Illumina NGS to investigate AM fungal communities in the selected ecosystems, which allow us to gain information on both in vivo structural details, as well as, species-specific information.

The objectives of this study were: 1) to survey the diversity of AM fungal communities associated with switchgrass in the Pine Barrens ecosystem located in New Jersey and New York in the USA, 2) to determine if soil properties and land use play a role in shaping the AM fungal communities and colonization in the Pine Barrens ecosystems and ecosystems with more managed land use, and 3) to identify novel or Pine Barrens-specific AM fungal clades.

Materials and Methods

The Field Sites

In order to investigate the AM fungal community differences between 'natural' landscapes and 'managed' landscapes, as well as, determine if there are any unique Pine Barrens AM fungal species, we chose to collect switchgrass roots from several sites, including annually mowed switchgrass field plots, Pine Barrens forests, and an Iowa natural prairie land. We sampled from an Iowa prairie in order to have a natural switchgrass site that was not in a Pine Barrens forest. The Iowa prairie also supplied us with root samples from non-acidic soil so we could test whether soil pH plays a role in shaping the switchgrass associated AM fungal communities.

Switchgrass roots and soil were collected in the summer months (June, July, and August) of 2014, 2016, and 2017 from several 'natural' and 'managed' switchgrass populations, in New Jersey (NJ), New York (NY) and Iowa (IO) in the United States. The natural sampling sites include: the Doolittle Prairie State Preserve (IO), IO; and five Pine Barrens sites: Wharton State Forest (WSF), NJ; Colliers Mills (CM), NJ; Federal Aviation Administration (FAA), NJ; Rocky Point (RP), NY; and Long Island Pine

Barrens Preserve (LIPB), NY. The managed sites were: Somerset Research Field (SO), NJ; Adelphia Research Field (AF), NJ; and E.A.R.T.H. Center (EC), NJ. (For more information on sampling, see Table A1, Fig. A1). The sampled switchgrass variety in the managed sites was "Kanlow", a lowland ecotype. AF and LIPB were sampled two times, AF in 2014 and 2017 (AF14, AF17), and LIPB in 2014 and 2016 (LIPB14, LIPB16). Ten whole plant roots were collected from WSF, CM, FAA, RP, LIPB14, LIPB16, IO, and EC. Six whole plant roots were collected from SO, AF14, and AF17; this was because these research plots had limited supply of switchgrass available. Individual plant root samples were collected no less than 3 meters apart at each site, in order to avoid clonal ramets (Kleczewski et al. 2012). A total of 98 plant root samples were collected.

One pooled soil sample was collected at each site from 15-20 cm below soil surface, around the switchgrass roots. Samples were kept on ice before analysis. Pooled soil samples for each site were analyzed for chemical makeup, pH, and other soil features by Spectrum Analytic Inc. (Washington Court House, OH) (Table A1). Root samples were divided in two parts, one part for molecular work and one for microscopic observation. Roots for molecular analysis were initially rinsed under running water to rid the roots of excess soil. Then they were surface sterilized by washing with 95% ethanol for 30 seconds, 0.825% NaOCl solution for 2 minutes, and 70% ethanol for 2 minutes. Roots were then rinsed with sterile water 3-5 times and stored at -80°C. Roots used for microscopic observation were rinsed thoroughly under running water and stored at 4°C until staining. In order to confirm the plant host identity, DNA from the leaf sheath of a host plant sample from each site was confirmed to be *P. virgatum* (additional information available in Fig. A2).

Estimation of AM Fungal Colonization

Ten random root segments (1-2 cm each) from each site were stained with 0.05% aniline blue following a modified version of the procedures of Grace and Stribley (1991) (see Fig. A3 for detailed staining protocols). Random roots were subsampled in order to maximize AM fungal coverage (Oehl et al. 2005). Roots were examined with a crosshair reticle under a compound light microscope at 400X magnification in order to determine percent root colonization by AM fungi according to the magnified intersections method (1990). Root AM structures were tallied, (i.e. coenocytic hyphae, arbuscules, hyphal coils, vesicles, AM spores). In order to more accurately describe and compare AM colonization and because different stains have been shown to impact the clarity of certain structures (Gange et al. 1999), we chose to count the colonization by each fungal structure separately, similar to McGonigle et al. (1990). 'AM Colonization' refers to the colonization of roots by any aforementioned AM structures, combined into one category, whereas 'Arbuscular Colonization' refers to the colonization by arbuscules only. A random subsample of 10-15 root pieces per root sample collected were observed in this fashion. Approximately 100 intersections for each subsample of roots were observed for each of the 98 total root samples collected. Percent colonization was calculated by first counting how many of each AM structure was observed from the subsample of roots. Then we divided this number by the number of total intersections observed and multiplied by 100%. Then, average percent colonization for each site was calculated by taking the average of all colonization percentages per subsample. Percent colonization was also recorded for structures belonging to dark septate endophyte (DSE), a very diverse group of fungi that often co-inhabit plant root tissue with AM fungi. However,

because these fungi were not the emphasis of this dissertation and because no trends were found, this data is not included. Anecdotally, DSE were found in large amounts in nearly every site's switchgrass root samples, indicating that they may be important for plant health.

DNA Extraction, PCR, Illumina Sequencing

For DNA extraction, we first pooled the roots by site (i.e. all roots from the same site were pooled together). Then we randomly chopped 0.125 g from each set of pooled roots for extracting DNA. This sub-sampling was done three times to give us triplicate DNA samples for each of our 11 sites, totally 33 samples (11 X 3 = 33) (Dimitrov et al. 2017). Grinding of roots was done with liquid nitrogen. DNA was extracted from all 33 samples using a DNeasy PowerSoil DNA Isolation Kit (Qiagen, Germantown, MD), according to manufacturer's instructions. DNA concentrations were checked with a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA).

The primer pair AMV4.5NF/AMDGR (Sato et al. 2005) was used to amplify the variable region of the small subunit rRNA gene for AM fungi. Primers were designed for Illumina by attaching Nextera XT adapters (Illumina, San Diego, CA), designed to attach on one end to Illumina MiSeq adapters and on the other end, to our AM fungal primers. PCR was conducted with a mixture of 0.5 µl each of forward and reverse primers, 12.5 µl Taq 2X Master Mix (New England BioLabs, Maine), 1 µl template DNA, with PCR grade water added to a total volume of 25 µl. The PCR parameters were: 95°C for 2 min, then 35 cycles at 95°C for 45 s, 52/55/58°C for 45 s, and 72°C for 1.5 min, with a final extension of 72°C for 5 min. Three annealing temperatures were used to maximize the

amplification of AM fungal species (Schmidt et al. 2013). Each of the 33 samples were run at these three annealing temperatures, totaling 99 PCR reactions. PCR reactions from the three annealing temperatures were pooled together per site, leaving 33 samples in the end. Gel electrophoresis confirmed bands in the majority of the samples and no bands in the negative controls.

PCR products were cleaned up with an Agencourt AMPure XP kit (Beckman Coulter, Brea, CA) following manufacturer's instructions. A secondary PCR was run to attach Nextera indices and Illumina adapters to each sample, using a Nextera XT Index Kit (Illumina, San Diego, CA). Each reaction contained 25 µl NEBNext High Fidelity 2X PCR Master Mix, 5 µl clean, primary PCR amplicon, 5 µl Nextera XT index 1 primer, 5 µl Nextera XT index 2 primer, and 10 µl PCR grade water, yielding a 50 µl total reaction volume. The secondary PCR conditions were 95°C for 3 min, then 8 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s. The final extension was 72°C for 5 min. Secondary PCR products were cleaned again using the Agencourt AMPure XP kit. DNA concentration of each reaction was checked using a Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Inc., Waltham, MA). All samples were normalized via dilution to 4 nM (=1.33 ng/µl), pooled in equal volumes, and sequenced on Illumina MiSeq with 2 X 300 bp paired-end reads using the MiSeq Reagent Kit v3 (600 cycle).

Bioinformatic Analysis

Removal of sequencing adapters, PCR primers and low-quality bases was performed through the CLC Genomics Workbench v8.5.1 (CLC Genomics Workbench 2017). Then, using the same software, forward and reverse reads were merged and any

non-merged reads (with no overlap) were discarded since full overlap was expected on the 250 bp expected sequence size. Quality control parameters were set to reject any sequences <100 bp long. Sequences were de-replicated using the "fastx_uniques" command in USEARCH 9.0 (Edgar 2010). Sequences were sorted by size and singleton sequences (those with abundance of <2) were discarded from further analysis using the "sortbysize" and "minsize" commands, respectively, in USEARCH 9.0 (Edgar 2010). Singletons were removed because they were likely artifacts of the amplification process (Kunin et al. 2010; Tedersoo et al. 2010). Sequences were then pooled together (in order to allow for downstream statistical analyses) and clustered into operational taxonomic units (OTUs) at 97% similarity using the "cluster_otus" command in USEARCH 8.0 (Edgar 2010). Sequences were compared against the MaarjAM database (Opik et al. 2010) and chimeras detected via the "uchime2_ref" command in USEARCH 9.0 (Edgar et al. 2011). All remaining OTUs were subjected to a Basic Local Alignment Search Tool (BLAST) against the National Center for Biotechnology Information (NCBI) nucleotide database using the "query" and "db" commands in BLAST+ 2.5.0. All OTUs were observed in MEGAN Community Edition 6.6.7 (Huson et al. 2016) with the default lowest common ancestor (LCA) parameters (minimum score of 50.0, minimum support percent of 0.01, and with the minimum-complexity filter off). All OTUs with BLAST matches to Glomeromycotan fungi were kept for further analysis. All AM fungal OTUs with matches of "uncultured Glomeromycota", "uncultured Glomeromycetes", or "uncultured Glomerales" (i.e. all matches above family level) were subjected to manual queries against the MaarjAM database and NCBI database for further inspection, similar to the methods of Schlaeppi et al. (2016). Comparison across the family or genus level is

appropriate because AM fungal families are considered a phylogenetically significant level when it comes to ecosystem function (Powell et al. 2009; Yang et al. 2017). Sequences representing each AM fungal clade detected in this study, as well as, sequences that were endemic to the Pine Barrens sites were uploaded to GenBank under accession numbers MH908518-MH908579.

Statistical Analysis

For microscopic observations, presence of the aforementioned fungal structures was recorded. Percent colonization for each sample was recorded and mean percent colonization by each fungal structure was calculated. One-way analysis of variance (ANOVA) was used to test overall significant differences of percent AM and arbuscular colonization between sampling locations. Tukey's HSD post-hoc test was used to determine pairwise differences between locations. All analyses were done on RStudio 1.1.419 (RStudio 2016).

Illumina sequencing data was then analyzed. Because the sequencing process results in unequal sequencing depth among samples (Harris et al. 2010), OTU abundance data were resampled using the median number of reads from among all 33 samples (21,049), which was from SO, triplicate sample 1 (de Carcer et al. 2011). This was done with the 'rrarefy' function in vegan 2.4-4, using RStudio 1.1.419 (RStudio 2016; Oksanen et al. 2017). Table 1 shows the total number of reads per site. There is variation in the number of reads for different sites because samples that had less reads than the median (21,049) were not rarefied. This method was chosen because when the alternative method of rarefying to the minimum number of reads was utilized, many

OTUs (145 OTUs) were lost. In order to not lose too much diversity data, particularly those of rare OTUs, we chose to rarefy to the median number of reads. Raw OTU and read totals are shown in Table 1, but subsequent mention of OTU and read abundance use the rarefied data. In order to determine whether sufficient number of samples (reads) were obtained, rarefaction curves were drawn using the 'ggiNEXT' function in the iNEXT package (Chao et al. 2014) using RStudio 1.1.419 (RStudio 2016). To compare AM fungal diversity between samples, three diversity indices (Shannon, Simpson, and Fisher's α) were calculated for each site with the 'diversity' and 'fisher.alpha' functions in vegan 2.4-4, using RStudio 1.1.419 (RStudio 2016; Oksanen et al. 2017).

All soil properties were scaled prior to downstream analyses using the root-mean-square of each set of soil properties ('scale' function, without centering in RStudio 1.1.419) (RStudio 2016). Pearson's correlation coefficient (r) was used to test for correlations between fungal colonization, soil properties, AM fungal abundance (# reads), and AM fungal diversity (Shannon, Simpson, and Fisher's α diversity indices). Average root colonization for each sampled site was used for these correlations. Correlations and corresponding significance values were calculated using the rcorr() function in the Hmisc package using RStudio 1.1.419 (RStudio 2016; Harrell Jr. 2017).

Analysis of similarity (ANOSIM), permutational multivariate analysis of variance (PERMANOVA), and nonmetric multidimensional scaling (NMDS) were conducted to explore the effects of land use, soil pH, and exchangeable aluminum on AM fungal communities among the sites sampled in this study. Functions 'anosim', 'adonis', and 'metaMDS' were used for these analyses, respectively, in vegan 2.4-4 using RStudio 1.1.419 (RStudio 2016; Oksanen et al. 2017). Bray-Curtis dissimilarity measurements

were used to determine community similarity between groups. To explore the effects of land use, pH, and aluminum on AM fungal communities, we broke the sites down into categories. For comparing different land uses on AM fungal communities, we categorized our samples into three categories: 'Pine Barrens', 'Managed', and 'Iowa'. Iowa was separated for this comparison because its soil type (pH and micro- and macro-nutrient levels) and land use (native prairie) is dramatically different. Next, to explore the effects of soil pH, we separated our samples into two pH categories: samples with soil pH ≥ 6.0 (AF14, AF17, EC, IO) versus samples with soil pH < 6.0 (CM, WSF, LIPB14, LIPB16, FAA, RP, SO). Lastly, to explore the effects of aluminum, we separated our samples into two exchangeable aluminum categories: samples with exchangeable aluminum levels ≥ 10.0 ppm (AF14, EC, LIPB14, LIPB16, WSF, FAA, RP) versus samples with levels <10.0 ppm (AF17, IO, SO CM).

To further understand the phylogenetic diversity represented in this AM fungal survey, a phylogenetic tree was built to represent the overall AM fungal clades represented in this study. To generate the tree, 1-2 representative sequences from each clade obtained in this study were aligned with reference sequences from the MaarjAM and NCBI databases. Sequences were aligned with MUSCLE (Edgar 2004) in MEGA 6.06 (Tamura et al. 2013). The best model to fit our data was found to be the Tamura 3-parameter model with gamma distribution (Tamura 1992). A maximum likelihood tree was then built in MEGA 6.06 (Tamura et al. 2013) with 1000 bootstrap replications, using the tree with the highest log likelihood (-2482.7931). Bootstrap values >70% are shown. Branch lengths represent the number of substitutions per site (Fig. 1).

Results

AM Fungal Colonization Based on Microscopic Observations

Based on microscopy data, AM fungi were observed from all the sampled sites. However, the degree to which the plants were colonized by the AM fungi varied. Average AM fungal colonization ranged from 1.7-52.1%, while arbuscular colonization ranged from 0-15.3% (Fig. 2) (Table A2). WSF and CM had the highest average percent AM colonization (52.21% and 37.03% respectively), arbuscular colonization (15.29% and 11.44% respectively), and AM fungal hyphal colonization (30.62% and 18.84% respectively) from all the sites sampled. LIPB16 had the lowest average percent AM and arbuscular colonization (1.7% and 0% respectively) and RP had the lowest average AM fungal hyphal colonization (Fig. 2) (Table A2).

AM Fungal Diversity Based on Illumina Sequencing

AM fungal reads were clustered at 97% similarity, as this threshold has been deemed appropriate for separating AM fungal taxa to the morphospecies-level in prior studies (Santos-Gonzalez et al. 2007; Lumini et al. 2010; Sepp et al. 2018). After quality filtering and removal of singletons, a total of 1,822,463 raw AM fungal reads were suitable for downstream analysis. Based on 97% similarity, 497 AM fungal OTUs were detected from the 33 root samples. After rarefying the data to the median number of reads, 423 AM fungal OTUs remained, consisting of 537,522 total reads. The number of AM fungal reads (rarefied) ranged from 26,451 in LIPB16 to 63,147 in CM, RP, and AF17 (Table 1). The number of AM fungal OTUs ranged from 19 in EC to 111 in IO (Table 1).

This study uncovered a wide array of diversity, with all four AM fungal orders represented (Fig. 1, Table 2). Thirteen genera or genus-rank taxa were obtained through Illumina sequencing of switchgrass roots from all sampled sites (Table 2). OTUs that could not be identified to the genus level were labeled as 'Undescribed' taxa, representing either undescribed clades or discrepancies in the reference database. The vast majority of this study's total AM fungal OTUs (59%) and reads (78%) belonged to *Glomus* (251 OTUs, 419,908 reads) (Table 2). The second and third most diverse OTUs (clades with the next highest amount of OTUs) were *Claroideoglomus* and *Acaulospora*, with 42 and 38 OTUs, respectively. The top five most abundant OTUs (the OTUs with the most reads) all matched to *Glomus* species and these five OTUs accounted for 54% of all reads (288,355 reads). The second and third most abundant OTUs were *Gigaspora* and *Claroideoglomus*, with 26,731 and 26,213 reads, respectively (Table 2).

IO had the highest OTU richness. Additionally, all three diversity indices showed IO with the highest diversity index scores, indicating higher total AM fungal diversity in that location (Table 1). The second highest indices were FAA (Simpson and Shannon), and SO (Fisher's α). The lowest diversity indices were found in LIPB16 (Simpson and Shannon) and EC (Fisher's α) (Table 1). The rarefaction curves match, overwhelmingly, with the diversity index findings.

Overall, *Glomus* was, by far, the most common genus detected from three land use categories in this study (Managed, Pine Barrens, and Iowa). All OTUs identified in the Ambisporaceae and Acaulosporaceae families were found exclusively from the natural (non-managed) ecosystems, and almost entirely in the Pine Barrens sites. Only three OTUs in the Acaulosporaceae family were found in a non-Pine Barrens location

(IO), while all Ambisporaceae OTUs were found in the Pine Barrens exclusively (Table A3, Table 2). Some of these uniquely Pine Barrens derived OTUs had poor matches in both the NCBI and MaarjAM databases, indicating that these may represent new, undescribed AM fungal clades. These include CM1_1135, FAA1_959, WSF3_55871, and WSF3_204765. These isolates matched *Acaulospora* sp. in the MaarjAM database with 94%, 92%, 92%, and 93% identity matches, respectively. Such findings may necessitate additional investigation. Additionally, there were some taxa that were more abundant in managed field sites. *Diversispora* spp. were more than four times more prolific in managed sites (358 reads) than in the Pine Barrens (87 reads) (Table 2). The most abundant *Archaeospora* sequence in this study (AF172_1184) was found exclusively in managed sites (Fig. 1, Table 2).

Sampling Effort

After rarefying the AM fungal OTU and abundance data to the median number of reads (21,049 reads) as stated above, 537,522 AM fungal reads and 423 AM fungal OTUs remained. Rarefaction curves show that some sites had AM fungal species saturation with the sampling effort put forth. However, other sites, specifically FAA, CM, WSF, and AF17, did not appear to reach asymptotic species saturation (Fig. 3). More sampling effort in these sites may result in additional species detection.

Soil and Management Impacts on AM Fungal Communities

Pearson's correlation tests conducted between soil properties, diversity indices, AM fungal abundance, and colonization data provided evidence that AM fungal diversity was indeed correlated with certain soil properties (Table A4). All soil properties were

tested but those soil properties that had negligible *r* correlation values ($r < 0.3$) and high *P* values ($P > 0.05$) were omitted from Table A4. Exchangeable aluminum was negatively correlated with both observed and rarefied AM fungal OTU richness ($r = -0.600$, $P = 0$ and $r = -0.625$, $P = 0.0002$, respectively), AM fungal diversity (See Shannon, Simpson, and Fisher diversity indices' correlations), and rarefied AM fungal abundance ($r = -0.468$, $P = 0.006$) (Table A4). Soil pH was positively correlated with AM fungal diversity (Shannon, Simpson, and Fisher indices) ($r = 0.614$, $P = 0.0001$; $r = 0.488$, $P = 0.004$; $r = 0.511$, $P = 0.002$ respectively). With respect to trends in AM fungal colonization rate based on the microscopy data, nitrate was positively correlated with hyphal and arbuscular colonization ($r = 0.583$, $P = 0.0004$; $r = 0.513$, $P = 0.002$ respectively). Organic matter was positively correlated with AM vesicles and spores ($r = 0.764$, $P = 0$; $r = 0.487$, $P = 0.004$ respectively). Extractable calcium and cation exchange capacity (CEC) were also positively correlated with AM fungal diversity (Table A4).

ANOSIM and PERMANOVA analyses showed significantly different AM fungal communities among 'Pine Barrens', 'Managed', and 'Iowa' sites ($R = 0.357$, $P = 0.001$; pseudo-$F = 4.286$, $P = 0.001$, $R^2 = 0.222$ respectively). ANOSIM and PERMANOVA analyses showed significantly different AM fungal communities under acidic vs. non-acidic soil pH ($R = 0.234$, $P = 0.005$; pseudo-$F = 3.440$, $P = 0.006$, $R^2 = 0.100$ respectively). Additionally, ANOSIM and PERMANOVA analyses showed significantly different AM fungal communities under high vs. low exchangeable aluminum ($R = 0.231$, $P = 0.003$; pseudo-$F = 2.827$, $P = 0.011$, $R^2 = 0.084$ respectively). NMDS analysis (Fig. 4) showed that Pine Barrens, managed, and Iowa AM fungal communities grouped

separately. NMDS analysis also showed that triplicate samples grouped together, with the exception of EC.

Discussion

Our study found that switchgrass populations surveyed in this study are inhabited by a large diversity of AM fungi, representing all four known AM fungal orders. Additionally, different land uses and soil conditions harbored different communities of AM fungi. We found that managed switchgrass sites had a significantly different AM fungal community compared to natural sites. Soil properties were also correlated with AM fungal OTU richness and diversity. Interestingly, almost all OTUs detected from the Ambisporaceae and Acaulosporaceae families were derived from Pine Barrens sites, some of which may represent new AM fungal lineages.

Soil properties impact on AM fungi

Our Illumina sequencing results indicate that soil pH and exchangeable aluminum were correlated with AM community diversity, which is in line with several previous studies (Suzuki et al. 2014; Bouffaud et al. 2016; Kawahara et al. 2016). However, some metanalyses showed little impact of soil pH on regional differences in AM fungal communities (Kivlin et al. 2011) and, instead, found biogeographic history and AM fungal dispersal (or lack thereof) to play a larger role (Morton et al. 1995; Opik et al. 2013). The extent to which these fungi are specifically adapted to certain soils is an ongoing debate (Opik et al. 2009; Kawahara et al. 2016), with some studies showing more importance of host plant (Scheublin et al. 2004; Jansa et al. 2008; Martinez-Garcia et al. 2015) or climatic features (Dumbrell et al. 2011) on AM fungal community

structuring. However, this may be because studies like those of Jansa et al. (2008) and Scheublin et al. (2004) used only a few host species. While host specificity at the plant functional group level may play a role in fungal community establishment in some circumstances, environmental features are thought to be the underlying 'proxy indicators' of AM fungal communities in more recent studies (Chaudhary et al. 2018; Sepp et al. 2018).

Primer specificity

The primer pair used (AMV4.5NF-AMDGR) was not as AM fungal specific as we expected. Out of 1,964 total OTUs clustered in this study (5,734,024 total reads), 497 raw OTUs (1,822,463 raw reads) belonged to the Glomeromycotina, representing about 25% of our study's raw OTUs (32% of raw reads). Several other fungal clades were observed: 542 OTUs matched to Basidiomycota, 130 OTUs to Ascomycota, 129 OTUs to Chytridiomycota, 41 OTUs to Mucoromycotina, as well as, OTUs that were identified as other Eukaryotic lineages or not assigned any taxonomic name. Van Geel et al. (2014) and Lumini et al. (2010) found 72% and 76% AM fungal specificity with this primer pair, respectively. They used different plant hosts and PCR parameters though. Cui et al. (2016) found many other organisms represented in their AM fungal Illumina study, similar to our study, though they do not mention any particulars on the other taxa found. Cao et al. (2016) found only 24% AM fungal specificity with this primer pair, very similar to the specificity of our study. They also found large abundance of members of the Basidiomycota and Chytridiomycota (Cao et al. 2016). A longer and more Glomeromycotina-specific DNA barcode may improve the specificity and taxonomic resolution problems (Opik et al. 2016).

Microscopic observations

Based on the microscopy, the highest AM colonization rates of *Panicum virgatum* by AM fungi were from CM and WSF, the NJ Pine Barrens sites (Fig. 2). However, correlation tests (Table A4) showed no correlation of arbuscular colonization (based on microscopy data) to soil pH (r = -2.00), and Tukey's HSD tests showed differences in colonization between the acidic soils of, say, WSF and LIPB and similarities between sites with dissimilar soil pH levels (i.e. WSF, CM, and AF). Our microscopic observations of arbuscules and AM fungal hyphae correlated, instead, with nitrate levels, total aluminum, and iron. This apparent conflict of results in our microscopic observations of AM fungal colonization compared to our Illumina diversity data could be due to the relatively small number of sample size in this study. Another possible explanation for finding variable colonization patterns is that we cannot gauge species information from microscopy data. Different AM fungal species have different levels of tolerance to fluctuating pH (Sieverding 1991) or aluminum tolerance. And studies have shown that different AM fungal families colonize plant roots at varying rates, with Glomeraceae species colonizing the most extensive amounts of roots the quickest, while Acaulosporaceae species often colonize the least (Hart and Reader 2002; Jansa et al. 2005). Additionally, although not a focus of this survey, the stoichiometry of soil nutrients may also play a role in shaping the colonization and/or community patterns seen in this AM fungal study (Johnson 2010). Lastly, it is also important to keep in mind that varied sample quality, innate bias in the process of PCR and sequencing, and bioinformatic decisions could also have played a role in varied downstream results (Engelbrektson et al. 2010; Hart et al. 2015; Alberdi et al. 2018).

Ambispora, Acaulospora prominence in the Pine Barrens

Interestingly, *Ambispora* spp. were found exclusively in the Pine Barrens, and *Acaulospora* spp. were found almost exclusively in the Pine Barrens sites, with only three OTUs observed in non-Pine Barrens ecosystems (Iowa). This distribution pattern corroborates the finding that *Acaulospora* and *Ambispora* may be more adapted to acidic soils compared to other AM fungi (Young et al. 1985; Oehl et al. 2011; Palenzuela et al. 2013). In the study by Oehl et al. (2011), spores of *Ambispora* sp. were only observed from soil pH of 5.0, while Nicolson and Schenck (1979) found only *Acaulospora laevis* spores in soil pH of 4.0-4.5, and Young et al. (1985) found *A. laevis* to predominate in soils of pH 4.3-4.8. Castillo et al. (2006) found *Acaulospora* sp. to also be the most common AM fungal species found in the acidic soils of Southern Chile (pH 5.5). Oehl et al. (2004) reported *Acaulospora* sp. to be rare in conventional farmland and French et al. (2017) found *Acaulospora* sp. in only natural sites, just as our study found. This study found switchgrass associated *Ambispora* spp. in samples strictly from soils of pH 4.8 and 5.0 (RP and CM, respectively) and most *Acaulospora* spp. from pH 4.8-5.2 (Table A3). Clark (1997) summarizes the findings of other authors which found *Glomus, Acaulospora,* and *Gigaspora* to be the most predominant genera inhabiting acidic soils. This gives credence to our findings, since *Glomus* and *Acaulospora* were similarly found to be the dominant genera in this study and *Gigaspora* was much more dominant in the Pine Barrens forests (17,396 reads) compared to managed fields (9,335 reads) and the Iowa prairie (0 reads) (Table 2). It is possible that these species of AM fungi are more adapted to acidic soils. Additional experiments are needed to test whether the endemic Pine Barrens species are specialist or generalist fungi.

Impact of land usage on AM fungal communities

Our results corroborate the previous finding that anthropogenic land use impacts AM fungal communities (Jansa et al. 2003; Moora et al. 2014; Valyi et al. 2015; Ciccolini et al. 2016). While some studies find that intensive agricultural practices tend to yield lower AM fungal diversity compared to natural sites (Helgason et al. 1998; Daniell et al. 2001; van der Gast et al. 2011), others find surprisingly diverse in agricultural sites, particularly under low amounts of fertilizers and tillage (Jansa et al. 2003; Hijri et al. 2006). Differences in these results could stem from differences in methodology (primer bias, NGS bioinformatic decisions, cloning limitations) or the extent and type of field management. The 'managed' sites surveyed in this study had low levels of fertilizers added several years prior to sampling, were mowed only once annually, and were polycultures, which may explain the relatively high AM fungal diversity. Additionally, a recent study by Garcia de Leon et al. (Garcia de Leon et al. 2018) showed that anthropogenic impacts on AM fungal diversity are not always consistent in nature. They, instead, argue that anthropogenic land use causes AM fungal diversity to equalize over different sites, not simply increase or decrease in a consistent manner. This may help explain why some managed sites in our study (i.e. AF14) had a very diverse AM fungal community, while other managed sites (i.e. EC) had much less AM fungal diversity. Future studies must take into account that anthropogenic change to a site may not drive AM fungal diversity in a consistent fashion, but rather, it might may equalize the community diversity over a larger scale (Garcia de Leon et al. 2018).

Certain AM fungi were more common in the managed field sites in this study. *Diversispora* spp. were more than four times more prolific in managed sites (358 reads)

than in Pine Barrens sites (87 reads) (Table 2). Moora et al. (2014) found Archaeosporaceae and Diversisporaceae (as well as Claroideoglomeraceae) to be indicator taxa for disturbed, agricultural sites. Therefore, it seems that perhaps these two clades could be associated less with acidic soils and more with disturbed or managed sites. However, in our study, Archaeosporaceae was found in similar quantities among different land uses. Overall, *Glomus* was, by far, the most common genus uncovered from all habitats. This is similar to findings in other Illumina AM fungal soil studies (Cui et al. 2016; Liu et al. 2017; Zhao et al. 2017) and AM fungal root studies (Johansen et al. 2016; Ban et al. 2017), as well as 454 (Varela-Cervero et al. 2015) and spore morphology studies (Oehl et al. 2004), which found *Glomus* and Glomeraceae to be the most prolific AM fungal genus and family, respectively. In this study, the second-most abundant taxon differed among site type. Despite the fact that *Glomus* spp. are almost equally prolific at each site, the complementary AM fungi may be more adapted to local soil types or differentially adapted to land use intensity (i.e. mowing, fertilizers), as local adaptation of AM fungi has been shown to be an important factor determining AM fungal communities (Johnson et al. 2010; Ji et al. 2013).

Iowa's natural prairie's switchgrass roots harbored the most diverse AM fungal community according to rarefaction analysis and according to all three diversity indices calculated. Other studies have similarly shown that grasslands have diverse AM fungal communities compared to forest habitats (Öpik et al. 2006; Moora et al. 2014) and croplands (Oehl et al. 2017). However, diversity indices and OTU richness are not the only pieces of the puzzle. It is also noteworthy that Pine Barrens locations had 12 AM fungal genera represented, while managed sites had 10 and Iowa had 6. Some of the

managed switchgrass sites in this study contained high AM fungal diversity. AF17 and SO in particular had a large amount of AM fungal diversity, corroborating the findings of Jansa et al. (2003) and Hijri et al. (2006) that managed lands can have surprisingly great AM fungal diversity. Despite the fact that AM fungi are instrumental in providing phosphorus to plant roots (van der Heijden et al. 2006), phosphorus was not shown to be associated with switchgrass AM fungal colonization, abundance, or diversity in this survey study. Although some studies have shown slightly improved switchgrass growth under P fertilization (McKenna and Wolf 1990), switchgrass is typically thought to be frugal in its use of this nutrient (Hall et al. 1982; Brejda et al. 2000). Our finding that switchgrass AM fungal communities do not seem highly affected by P is not surprising, as this host plant does not need much P and host is often an important variable determining impacts of P and mycorrhizal fungi (Gosling et al. 2013).

Future directions

Despite our uncovering many potentially undescribed, new AM fungal lineages, the path to declaring new taxa must be tread on lightly because the phylogenetic positions of many AM fungi are still in flux (Orchard et al. 2017a; Walker et al. 2018). Furthermore, we need to understand that when doing ecological studies using solely or mostly molecular evidence (i.e. DNA sequences), our data are only as good as our reference databases. Much more work needs to be done to barcode herbarium specimens so that our databases have adequate supply of sequences for already described species (Brock et al. 2009; Rocha et al. 2014). Additionally, our AM fungal diversity and community composition findings rely on our limited and uneven sampling efforts, which can be problematic for downstream conclusions. We were able to collect 10 samples

from Iowa, 28 from managed sites, and 60 from Pine Barrens ecosystems. Further work should include additional sites from Iowa or other locations with similar soil type, including managed fields, as well as, more managed sites in New Jersey and New York for proper comparison of fungal communities between land uses. Additional Pine Barrens sites could help confirm that our *Acaulospora* and *Ambispora* spp. are indeed endemic to Pine Barrens ecosystems.

Conclusion

This survey of switchgrass root-inhabiting AM fungi uncovered a large amount of diversity from different soil types and land uses. This study is the first to detail the AM fungal communities associated with switchgrass roots in the Pine Barrens ecosystem using both microscopy and NGS methods. The NGS data showed that soil pH and aluminum impacted root AM fungal community composition and diversity. *Glomus* was the most prolific AM fungal genus inhabiting in all the sampled sites; however, different land use types were inhabited by different AM fungal communities. A significant finding of this study was that *Acaulospora* and *Ambispora* were almost exclusively associated with the Pine Barrens ecosystems. Through Illumina sequencing, this study further enhances the breadth of knowledge on the DNA sequence diversity of AM fungi and begins to uncover the AM fungal communities associated with Pine Barrens ecosystems. More research on switchgrass associated AM fungi across a larger geographic region would help gain insight into what makes switchgrass thrive in different regions and under different land uses.

Table 2-1. The number of AM fungal operational taxonomic units (OTUs) and reads detected through Illumina MiSeq sequencing.

Three diversity indices (Shannon, Simpson, and Fisher's α) are compared. 'Observed' refers to raw abundance and 'rarefied' refers to abundance after rarefying to the median number of reads. Site abbreviations are: AF14 – Adelphia Research Field 2014, AF17 – Adelphia Research Field 2017, CM – Colliers Mills, EC – EARTH Center, IO – Iowa, FAA – Federal Aviation Administration, LIPB14 – Long Island Pine Barrens 2014, LIPB16 – Long Island Pine Barrens 2016, RP – Rocky Point, SO – Somerset Research Field, WSF – Wharton State Forest.

Site	Samples	OTUs (observed)	OTUs (rarefied)	Reads (observed)	Reads (rarefied)	Shannon	Simpson	Fisher's α
AF14	3	74	74	41,819	41,819	2.366	0.869	8.732
AF17	3	153	89	710,459	63,147	1.736	0.675	10.973
CM	3	97	81	156,356	63,147	1.023	0.367	9.549
EC	3	19	19	35,821	35,245	0.952	0.4	1.937
FAA	3	83	76	169,083	54,286	2.661	0.877	8.826
IO	3	113	111	48,525	44,057	3.148	0.927	13.892
LIPB14	3	30	30	33,757	33,757	0.756	0.283	3.243
LIPB16	3	25	25	26,451	26,451	0.655	0.261	2.723
RP	3	39	33	367,234	63,147	1.911	0.8	3.352
SO	3	103	102	63,304	56,003	2.12	0.791	11.949
WSF	3	102	84	169,654	56,463	2.089	0.811	9.428

Table 2-2. Distribution of all AM fungal operational taxonomic units (OTUs) and reads (rarefied) obtained through Illumina MiSeq sequencing.

Taxa that clustered most closely to known isolates at the order or family level were named as "Undescribed" taxa. Colors represent a common order. Abundance of each taxon found under different land uses (Managed, Pine Barrens, Iowa) is included.

Order	Family	Genus	OTUs	Reads	Reads by Land Use		
					Managed	Pine Barrens	Iowa
	Ambisporaceae	Ambispora	4	65	0	65	0
Archaeosporales	Archaeosporaceae	Archaeospora	10	569	301	268	0
	N/A	Undescribed Archaeosporales	1	20	0	20	0
	Acaulosporaceae	Acaulospora	38	23,685	0	23,042	643
	Claroideoglomeraceae	Claroideoglomus	42	26,213	14,068	7,885	4,260
Diversisporales	Diversisporaceae	Diversispora	3	1,063	358	87	618
	Gigasporaceae	Undescribed Gigasporaceae	12	13,874	1,789	12,085	0
	Gigasporaceae	Gigaspora	15	26,731	9,335	17,396	0
	Glomeraceae	Glomus	251	419,908	156,470	235,695	27,743
Glomerales	Glomeraceae	Undescribed Glomeraceae	12	2,690	742	156	1,792
	N/A	Undescribed Glomerales	20	21,689	12,430	266	8,993
Paraglomerales	Paraglomeraceae	Paraglomus	14	1,006	712	286	8
N/A	N/A	Undescribed Glomeromycetes	1	9	9	0	0

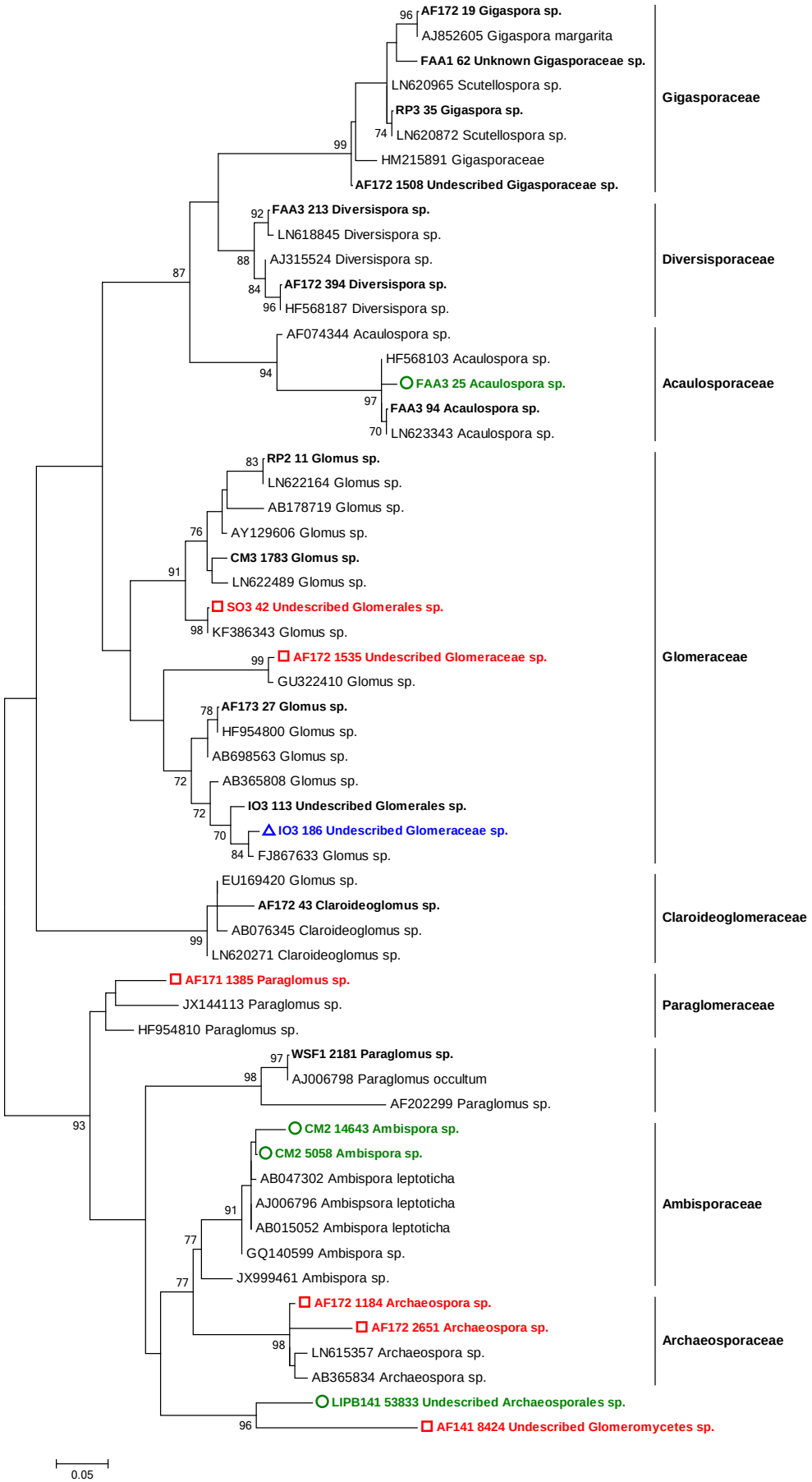
AF172 19 Gigaspora sp.
AJ852605 Gigaspora margarita
FAA1 62 Unknown Gigasporaceae sp.
LN620965 Scutellospora sp.
RP3 35 Gigaspora sp.
LN620872 Scutellospora sp.
HM215891 Gigasporaceae
AF172 1508 Undescribed Gigasporaceae sp.
Gigasporaceae
FAA3 213 Diversispora sp.
LN618845 Diversispora sp.
AJ315524 Diversispora sp.
AF172 394 Diversispora sp.
HF568187 Diversispora sp.
Diversisporaceae
AF074344 Acaulospora sp.
HF568103 Acaulospora sp.
FAA3 25 Acaulospora sp.
FAA3 94 Acaulospora sp.
LN623343 Acaulospora sp.
Acaulosporaceae
RP2 11 Glomus sp.
LN622164 Glomus sp.
AB178719 Glomus sp.
AY129606 Glomus sp.
CM3 1783 Glomus sp.
LN622489 Glomus sp.
SO3 42 Undescribed Glomerales sp.
KF386343 Glomus sp.
AF172 1535 Undescribed Glomeraceae sp.
GU322410 Glomus sp.
AF173 27 Glomus sp.
HF954800 Glomus sp.
AB698563 Glomus sp.
AB365808 Glomus sp.
IO3 113 Undescribed Glomerales sp.
IO3 186 Undescribed Glomeraceae sp.
FJ867633 Glomus sp.
Glomeraceae
EU169420 Glomus sp.
AF172 43 Claroideoglomus sp.
AB076345 Claroideoglomus sp.
LN620271 Claroideoglomus sp.
Claroideoglomeraceae
AF171 1385 Paraglomus sp.
JX144113 Paraglomus sp.
HF954810 Paraglomus sp.
Paraglomeraceae
WSF1 2181 Paraglomus sp.
AJ006798 Paraglomus occultum
AF202299 Paraglomus sp.
CM2 14643 Ambispora sp.
CM2 5058 Ambispora sp.
AB047302 Ambispora leptoticha
AJ006796 Ambispsora leptoticha
AB015052 Ambispora leptoticha
GQ140599 Ambispora sp.
JX999461 Ambispora sp.
Ambisporaceae
AF172 1184 Archaeospora sp.
AF172 2651 Archaeospora sp.
LN615357 Archaeospora sp.
AB365834 Archaeospora sp.
Archaeosporaceae
LIPB141 53833 Undescribed Archaeosporales sp.
AF141 8424 Undescribed Glomeromycetes sp.
0.05

Fig. 2-1 Maximum-likelihood phylogenetic tree of the arbuscular mycorrhizal (AM) fungi represented in this study, obtained from Illumina sequencing of switchgrass (*Panicum virgatum* L.) roots.

Numbers on branches indicate percent bootstrap support for 1000 replications. Bootstrap values >70% are shown. Bold names represent sequences obtained from this study. AM fungal families are aligned vertically. Red coloring/square symbol – sequences retrieved in this study exclusively from managed lands; Blue coloring/triangle symbol – sequences retrieved in this study exclusively from the Iowa prairie; Green coloring/circle symbol – sequences retrieved in this study exclusively from Pine Barrens forests.

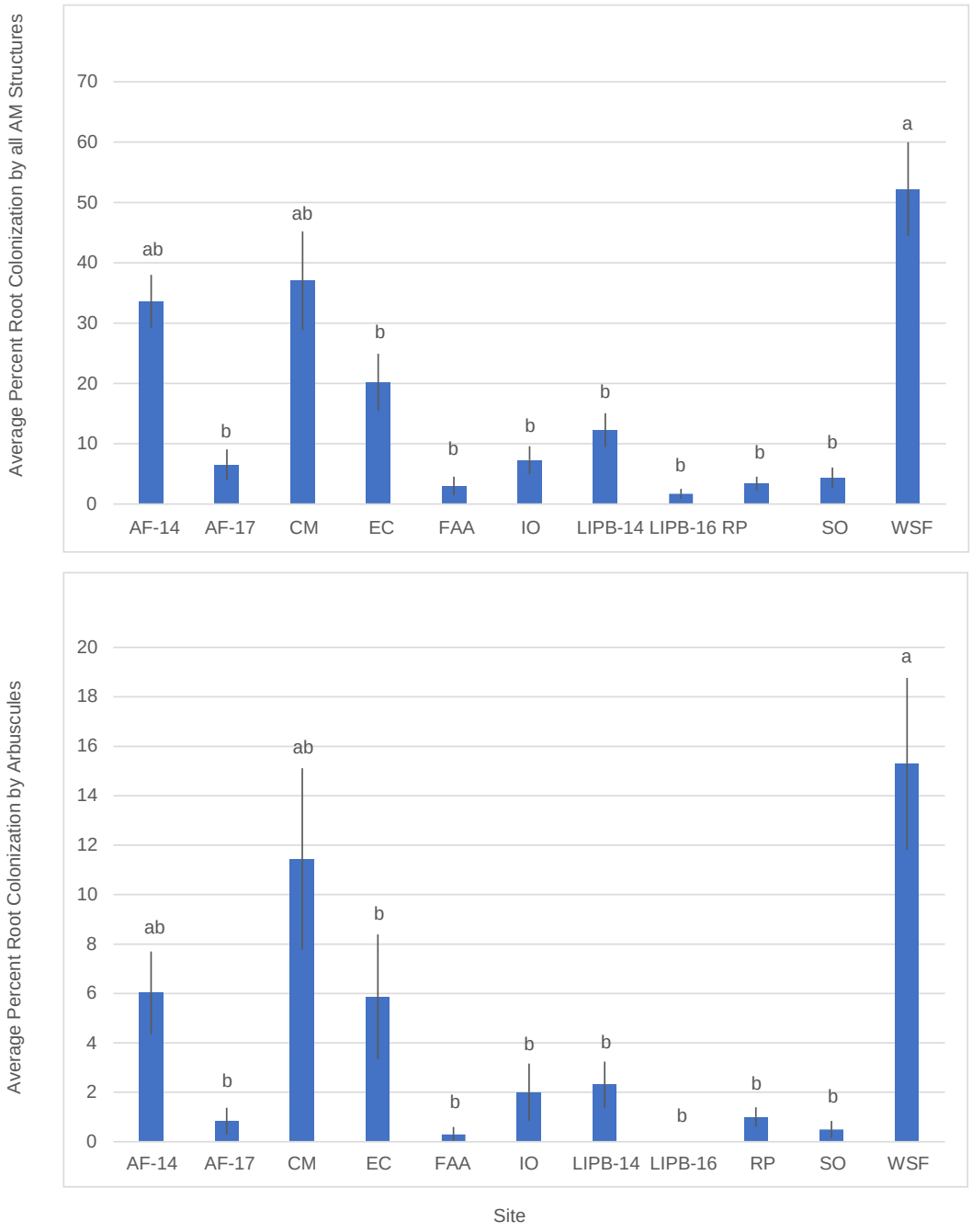

Fig. 2-2 Average percent root colonization of switchgrass by all AM structures and by arbuscules only.

Letters above bars represent significant differences in average fungal colonization between locations (Tukey HSD; $P < 0.05$). Error bars indicate ± 1 standard error. AF14 –

Adelphia Research Field 2014, AF17 – Adelphia Research Field 2017, CM – Colliers Mills, EC – EARTH Center, IO – Iowa, FAA – Federal Aviation Administration, LIPB14 – Long Island Pine Barrens 2014, LIPB16 – Long Island Pine Barrens 2016, RP – Rocky Point, SO – Somerset Research Field, WSF – Wharton State Forest.

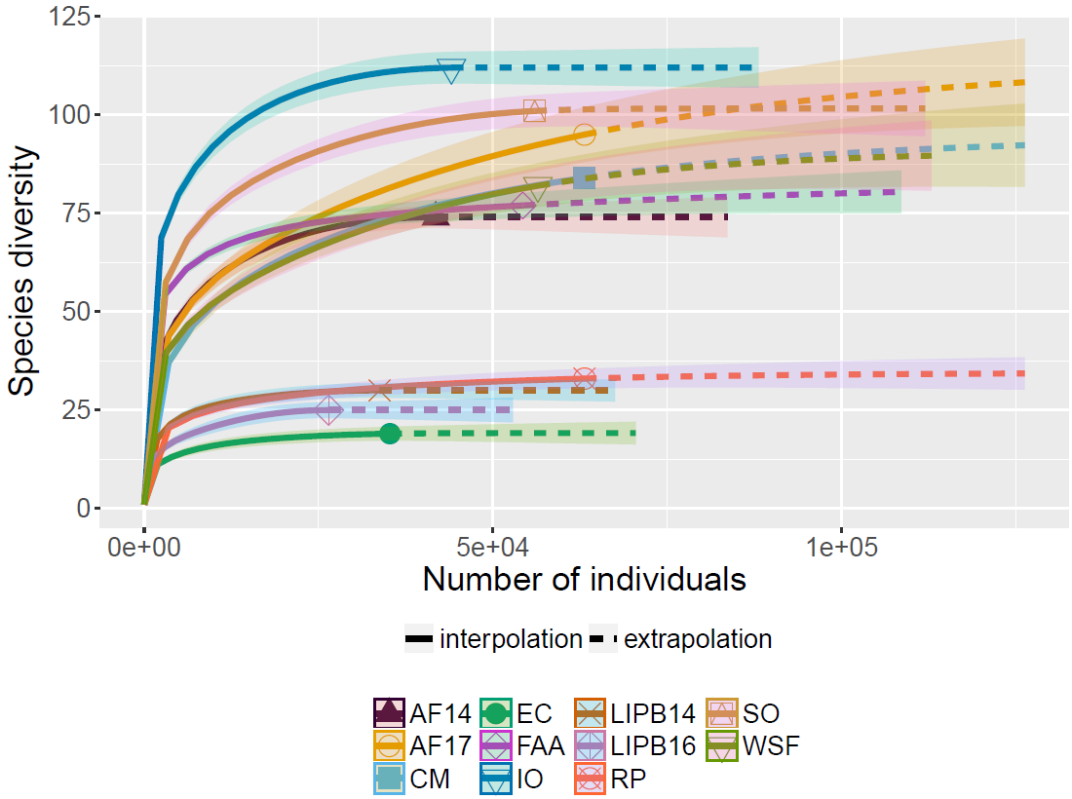

Fig. 2-3 Rarefaction analysis of AM fungal communities from all sampled sites.

AF14 – Adelphia Research Field 2014, AF17 – Adelphia Research Field 2017, CM – Colliers Mills, EC – EARTH Center, IO – Iowa, FAA – Federal Aviation Administration, LIPB14 – Long Island Pine Barrens 2014, LIPB16 – Long Island Pine Barrens 2016, RP – Rocky Point, SO – Somerset Research Field, WSF – Wharton State Forest.

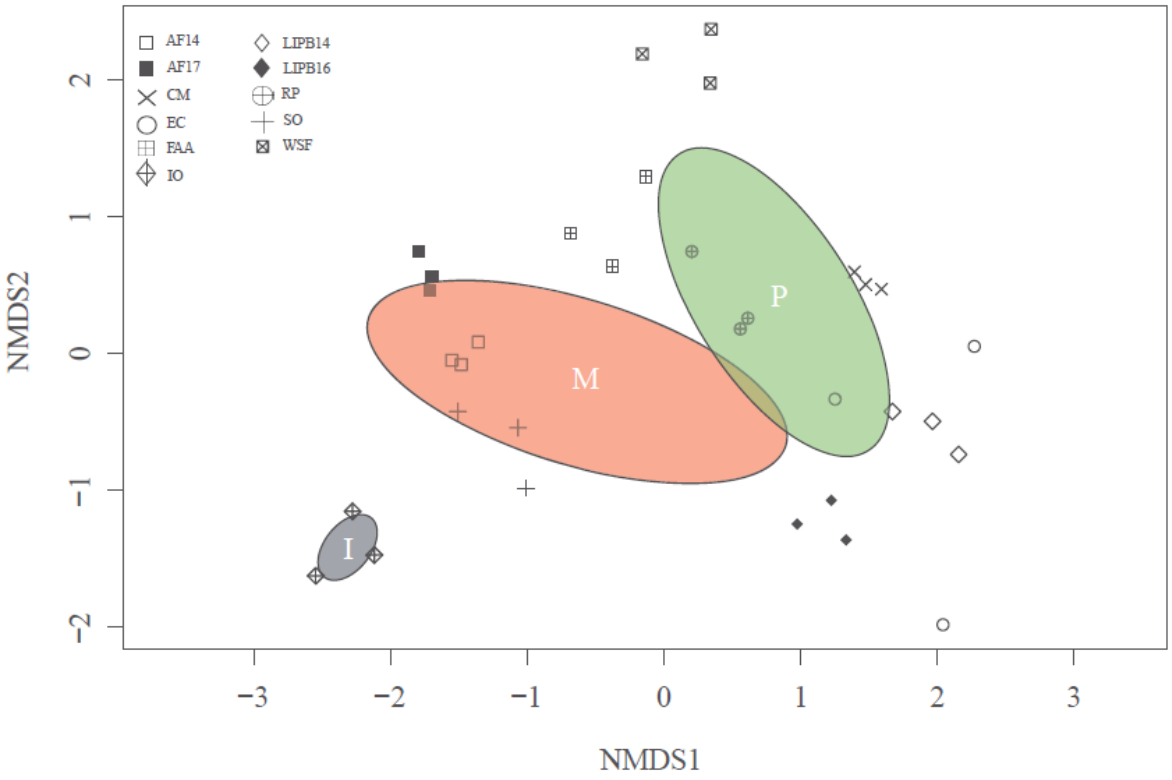

Fig. 2-4 Nonmetric multidimensional scaling (NMDS) plot of AM fungal communities from thirty-three switchgrass samples.

This analysis was based on Bray–Curtis dissimilarity measurements between samples (non-pooled) (stress = 0.147). A green ellipse with a 'P' indicates Pine Barrens AM fungal diversity, a red ellipse with an 'M' indicates managed AM fungal diversity, and a grey ellipse indicates Iowa AM fungal diversity. Pine Barrens sites included CM, FAA, LIPB14, LIPB16, RP, and WSF. Managed sites included AF14, AF17, EC, and SO. Iowa sites were exclusively IO. Different shapes represent these 11 sites.

Chapter 3

Impacts of bioinformatic workflow decisions on the results of arbuscular mycorrhizal fungi diversity and community composition: an 'Illumina'ting study

Abstract

Next-generation sequencing (NGS) has become a popular method for assessing arbuscular mycorrhizal fungi (AM fungi) fungal diversity. However, there exist few comparisons of central bioinformatic decisions with their respective downstream results. Therefore, this study seeks to assess AM fungal diversity from environmental samples using Illumina NGS technology and select computer programs for analyzing the AM fungal sequences**.** The main goals of this study were to [1-] develop an approach to assess AM fungal communities, using Nextera XT barcoded, AM fungal designed primers and Illumina NGS technology, [2-] provide computer scripts for others to use for future AM fungal Illumina studies, and [3-] determine any impacts of bioinformatic decisions on downstream results. To accomplish our goals, 30 switchgrass (*Panicum virgatum* L.) root samples were collected from three Pine Barrens forest sites. After genomic DNA extraction, PCR amplification of a fragment of 18s rDNA, and sequencing on an Illumina MiSeq, sequence data were subjected to various bioinformatic trajectories. All sequences were subjected to: [1-] clustering at 97% and 95% similarity, [2-] removal of singleton sequences, removal of singleton and doubleton sequences, and no rare sequence removal [3-] comparison against two DNA reference databases (NCBI and MaarjAM). Results showed that different workflow trajectories yielded dramatically different downstream

results, specifically, when different reference databases were use. Using the MaarjAM database as a reference database yielded higher AM OTU richness and overall diversity than the NCBI database. *Paraglomus* was the most prolific OTU found when using the MaarjAM database, whereas this genus was hardly detected at all when using the NCBI database. Removing low frequency AM fungal operational taxonomic units (OTUs) (i.e. rare sequences) and clustering at different thresholds made little impact on overall AM OTU richness and diversity. This Illumina NGS and dual indexing technique proved to be cost-efficient and provided informative data on AM fungi inhabiting an understudied ecosystem. However, careful attention to workflow decisions and reference database selection, in particular, are of utmost importance when analyzing AM fungal sequence diversity. Annotated computer code is included in a step-by-step fashion in order to aid future investigations into global AM fungal diversity.

Introduction

Next generation sequencing (NGS) technology has revolutionized the study of microbiology, helping scientists uncover unexplored and cryptic microbial diversity (DeLong et al. 2006; Nelson et al. 2010; Qin et al. 2010; Sun et al. 2014). NGS allows scientists to quickly and economically understand the genomic diversity in environmental samples with relative ease (Shokralla et al. 2012). NGS technologies have been used in many studies of arbuscular mycorrhizal fungi (AM fungi) ecology and diversity (Opik et al. 2009; Cui et al. 2016; Schlaeppi et al. 2016; Ban et al. 2017; Liu et al. 2017). Arbuscular mycorrhizal fungi (AM), recently placed within the Mucoromycota phylum and Glomeromycotina subphylum (Spatafora et al. 2016) are mutualistic, symbiotic microorganisms that associate with over 80% of terrestrial plants (Smith and Read 2008). These root and soil inhabiting fungi play a major role in ecosystem productivity and nutrient cycling (Bago et al. 2000; Rillig 2004; Mohan et al. 2014; van der Heijden et al. 2015). The amount of described AM fungal species has increased dramatically in a short time span, from 150 species in 1993 (Walker and Trappe 1993) to 1658 taxa estimated based on the internal transcribed spacer region 'sequence hypothesis' (Koljalg et al. 2013), ~300 morphospecies (Öpik and Davison 2016) and over 500 molecular or 'virtual' taxa (VT) based on SSU rDNA (Kivlin et al. 2011). Novel AM fungal taxa are being discovered at ever-increasing rates.

There has been a major methodology shift in AM fungal species identification. We have gone from a morphology-based (Morton et al. 1995) to a molecular-based approach (Helgason et al. 1998) to NGS-based approaches using platforms such as 454 and Illumina (Hiiesalu et al. 2014; Liu et al. 2017). Because of the varied methodologies,

a comparison of AM fungal taxa between studies is often challenging. There are many potential reasons for the lack of clarity in AM fungal taxonomy and nomenclature. Many AM fungi cannot be or have not been cultured and so there is often a disconnect between morphological AM species and VT species. Secondly, there is constant resorting and renaming of AM genera (Redecker et al. 2013) and even higher-order taxa (Spatafora et al. 2016).

In addition to these reasons that scientists struggle to interpret the true AM fungal diversity in nature, there is a scarcity of published, detailed pipelines for the bioinformatic analysis of AM fungal environmental sample sequencing. One recent publication by Morgan and Egerton-Warburton (2017) begins to address this issue, developing a pipeline for AM fungal community analysis via Illumina MiSeq sequencing. However, they use a particular primer set that could not successfully amplify the rDNA from most of the root samples in our study. This could be because those primers simply did not provide a good match to the target genomes from the AM fungal samples collected in this study. Therefore, this study set out for an alternative set of primers, in accordance with Nextera XT barcoding. Although few other AM fungal Illumina studies have used this approach, it has been shown to be a promising method for collecting high-throughput sequence data (up to 96 libraries) and needs more experimentation (Vasar et al. 2017). Other portions of the Illumina workflow were not included in the study by Morgan and Egerton-Warburton (2017), such as, the comparison of AM communities when using NCBI vs. the MaarjAM databases (Opik et al. 2010).

This lack of consistency and, often, transparency in workflow can lead to dramatically different results between authors. Few AM fungal researchers have the

necessary background in computer programming to properly dissect their data. All portions of the NGS research pipeline can incur biases and cause major downstream variation. For instance, database selection can bias the amount and identity of taxa collected in the study. The percent similarity used in clustering (95% or 97%) can also greatly impact the observed AM fungal species richness (Morgan and Egerton-Warburton 2017). Additionally, in order to attempt to rectify innate PCR bias (Bellemain et al. 2010), studies often remove rare OTUs (aka: rare sequences) from downstream analyses. The question becomes how to define "rare" OTUs. They are typically defined as OTUs found only once or a few times in a particular sample. There is large variation in NGS fungal diversity studies as to how we should define and how we should treat rare sequences. Some studies show that, for AM fungi, singletons, doubletons, tripletons, and even 4-tons do not impact downstream AM richness (Unterseher et al. 2011) or overall AM community patterns (Kivlin et al. 2011); however, these studies looked at 454 sequencing data, not Illumina data. Contrastingly, it is common practice to remove singleton sequences (Opik et al. 2013; Moora et al. 2014; Vasar et al. 2017) out of fear of including potential PCR artefacts (Tedersoo et al. 2010). With so much variation in rare sequence removal decisions, there is little research as to how removing rare sequences actually impacts the observed AM fungal communities in nature. A study by Morgan and Egerton-Warburton (2017) compares keeping all rare sequences, removing singletons, and those OTUs with 10 constituent sequences (10-tons) using Illumina sequencing. Morgan and Egerton-Warburton (2017) found that increasing the percent similarity for OTU clustering inflated the number of AM fungal OTUs per sample, as did keeping rare sequences (<1, 2, or 10 constituent sequences). And although they found that overall

diversity indices remained constant no matter how the data was treated, AM community composition was impacted. No study has compared these important bioinformatic decisions using the recently popular primer set AMV4.5NF/AMDGR (Sato et al. 2005). As Illumina and this primer set are becoming popular in the exploration of AM fungal communities, it is critical to see what the impacts are of varied OTU clustering and rare sequence cutoffs on AM communities.

Most AM fungi NGS studies don't publish their scrips or computer code for the bioinformatic analysis necessary to interpret their results. However, there is a great need for this "sharing" of computer code (Barnes 2010). For one, without shared code, there is lingering mystery as to what steps scientists took to determine their results. Secondly, sharing codes allows for consistency in methods between different studies, thereby allowing for comparison and reproducibility of results (Mesirov 2010). Sharing code provides important information about what specific parameters were used to obtain the results, which are too often left out of methods sections (LeVeque 2012). This study, therefore, shares all computer code used, in order to help fill some of the knowledge gaps in bioinformatic decision making and aid other AM fungal specialists in their search for more undescribed species.

Here, we test AM fungal primers (AMV4.5NF-AMDGR) with Nextera XT overhang adapters to sequence environmental switchgrass root samples on the Illumina MiSeq platform. Our main goals were to [1-] develop a replicable workflow for this high throughput AM fungal PCR primer set for Illumina sequencing and analysis [2-] compare workflow decisions to determine the impacts of bioinformatic choices on downstream

AM fungal diversity results and $^{3-}$ present annotated program code for others to incorporate into their respective AM fungal diversity studies.

Materials and Methods

Sampling Sites

Samples of switchgrass (*Panicum virgatum* L.) roots and surrounding soil were collected from the following 3 Pine Barrens, temperate forest locations in the Northeastern United States (US) during June-July, 2014 (summer season): Long Island Pine Barrens forest (David A. Sarnoff Pine Barrens Preserve) in Westhampton, New York (LIPB); Wildlife Management Area in Colliers Mills, New Jersey (CM); Wharton State Forest in Hammonton, NJ (WSF) (Table 1). Switchgrass plants were all of similar size and had similar surrounding plant communities consisting of *Pinus rigida* (Pitch Pine), *Quercus* sp. (oak), ericaceous shrubs, and other grasses. The Pine Barrens forests have distinctly acidic and sandy soils with low amounts of available nutrients and frequent fires (Forman 1998). Sites were chosen based on keen interest in this underexplored, rare habitat type and based on microscopic detection of AM fungal structures within the roots.

Sample Collection and Processing

Soil and *Panicum virgatum* (switchgrass) root samples were collected from 10 *Panicum virgatum* L. (switchgrass) plants per site and stored on ice until processing (within 24 hours). Plants and soil were collected after removing the first 15-20 cm of topsoil. Samples were collected at least 3 meters apart at each site to best represent the average AM fungal diversity of the site. Samples were stored on ice until processing.

Roots were surface sterilized with a modified version of the technique used by Arnold and Lutzoni (2007): immersion in 95% ethanol for 30 seconds, 0.825% NaOCl for 2 minutes, and 70% ethanol for 2 minutes. Roots were then rinsed three times with autoclaved, double distilled water, patted dry with autoclaved paper towels, and stored at -80°C. 100g of air dried, pooled soil from each site was analyzed for soil pH, heavy metals, available P, and other soil properties by Spectrum Analytic (Washington Court House, OH) in order to confirm similarity of soil type among all three sites (Table 1).

DNA Extraction

Stratified subsampling was conducted for DNA extraction purposes. For example, for the DNA samples for LIPB, pieces of root from each of ten plants collected from the site were combined into 0.125g of pooled roots. This was done to best represent the AM fungal community at each site. Subsampling was done three times per site to provide triplicates for each of the three sites. (3 sites X 3 replicate DNA extractions = 9 samples). Grinding with liquid nitrogen was performed to best pulverize the root pieces and expose fungal DNA inside. Total genomic DNA was then extracted from powdered roots using a DNeasy PowerSoil DNA Isolation Kit (Qiagen Germantown, MD) according to manufacturer's instructions. A NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA) was used to measure DNA concentration and quality.

Primer Design

Thoughtful primer selection is vital, as the specific AM primers and gene region of interest can bias downstream AM fungal community descriptions (Kohout et al. 2014;

Van Geel et al. 2014). Therefore, an AM specific primer pair was selected for this study based on several factors: [1-] Minimal non-specific amplification based on literature searches [2-] Ability to amplify all AM fungal clades, including rare taxa [3-] Ability to produce high quality DNA reads and [4-] Ability to amplify a region of DNA that is common in the literature, so that comparison could easily be done with known databases. Additionally, the primers had to amplify <600 bp, as the Illumina 600 cycle platform would not perform well with larger fragments. Based on these priorities and on several previous studies successfully using our choice of primer (Lumini et al. 2010; Dai et al. 2013; Van Geel et al. 2014; van Geel et al. 2015), the AMV4.5NF/AMDGR primer pair was chosen (Sato et al. 2005). This primer pair amplifies 300 bp in the center of the SSU region (Van Geel et al. 2014). Although there is variation in this primer pair's specificity for AM fungi, Lumini et al. (2010) and Van Geel et al. (2014) found it to have 76% and 72% AM fungal specificity, respectively. DNA amplification and species resolution of AM fungi with a single primer pair has been historically difficult, with many barcode regions available but imperfect (Simon et al. 1992; Lee et al. 2008; Van Geel et al. 2014; Lekberg et al. 2018). This primer pair (Sato et al. 2005) is able to resolve to the VT or morpho-species level.

Illumina's Nextera XT overhang adapters (Oligonucleotide sequences 2018 Illumina, Inc.) were incorporated into our AM specific primers for primary PCR. A major advantage of this protocol is that with the Nextera overhang adapters, you can multiplex up to 96 samples into one Illumina MiSeq run, making the sequencing endeavor more economical and higher throughput. The primary PCR primers used in this study were

Forward primer:

5’ **TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**aagctcgtagttgaatttcg

Reverse primer:

5’ **GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**cccaactatccctattaatcat

where boldfaced nucleotides represent the Nextera overhang adapter and non-boldfaced nucleotides represent the AM specific primers AMV4.5NF/AMDGR (Sato et al. 2005).

Secondary PCR was performed to attach the complementary Nextera XT indices (96 sample Nextera XT Index Kit FC-131-1002) and the Illumina sequencing primer to the PCR products. The secondary PCR primers used in this study were

Index 1 primer:

5’ <u>CAAGCAGAAGACGGCATACGAGAT</u> [i7] gtctcgtgggctcgg

Index 2 primer:

5’ <u>AATGATACGGCGACCACCGAGATCTACAC</u> [i5] tcgtcggcagcgtc

where underlined nucleotides represent the Illumina primer sequence for binding to the MiSeq flow cell, [i7] and [i5] are 8 nucleotide Nextera XT index sequences (Oligonucleotide sequences © 2018 Illumina, Inc.), and lowercase nucleotides represent the complementary nucleotides to the Nextera overhang adapters. This method utilizes 2 indices per sample (Index 1 on one end and Index 2 on the other) in order to allow for up to 96 different index combinations (8 Index 2 sequences X 12 Index 1 sequences = 96 barcode combinations) for multiplexing on a single Illumina MiSeq run.

PCR and Illumina Sequencing

A 300bp fragment of the SSU rDNA was amplified using our Nextera XT AM specific primers AMV4.5NF and AMDGR. The SSU region was targeted, as this region is the most widely used (Opik et al. 2009; Öpik et al. 2014) and because it is sufficiently variable for species delineation. The 25ul total reaction mixture contained 12.5ul Taq 2X Master Mix (New England BioLabs, Maine), 0.5ul of each primer (10uM concentration), and 1ul DNA template. The reaction was conducted in a thermal cycler 2720 (Applied Biosystems) for initial denaturation at 95°C for 2 min, 35 cycles of 95°C for 45 s, 52°C/55°C/58°C for 45 s, 72°C for 1.5 min, and a final elongation step of 72°C for 5 min. Three different annealing temperatures were used for optimal amplification of diverse clades of AM fungi (Schmidt et al. 2013).

For each of the 9 samples, there were three PCR products (from 3 rounds of PCR, as stated above). All reaction products were observed through gel electrophoresis. Despite the fact that some PCR products did not show bands on the gel for verification of band size and strength, all PCR products were used. PCR products from each of the 3 annealing temperatures were pooled, per sample, leaving us with 9 total samples. All PCR products were cleaned up with AMPure XP beads (Beckman Coulter).

A secondary Nextera indexing PCR step was performed in order to place Nextera indices on each library and also to attach the Illumina adapters, as mentioned earlier. PCR cleanup was performed and the library sizes were validated on a Bioanalyzer High Sensitivity DNA chip. This portion of our workflow follows the protocols set forth in the 16S Metagenomic Sequencing Library Preparation guide (Illumina, San Diego, CA, USA). Sequencing was done with the MiSeq Kit v3 600 cycle.

Bioinformatics

The following bioinformatic workflow used in this study is depicted in Figure 1. This figure shows the steps in our workflow in which major decisions were made. Additionally, the Bash commands and scripts used in this bioinformatic study can be found in Figs. S1 and S2. These were all run using PuTTY 0.67 and WinSCP user interfaces for Windows OS using in-house bioinformatic scripts and freely available code.

Sequences retrieved from the Illumina run were demultiplexed according to the Nextera barcodes, and sequences were oriented into the correct directions. Sequence adapters, PCR primers, and low-quality bases were filtered out using CLC Genomics Workbench v8.5.1. The same software was used to merge all forward and reverse reads. Any non-merged reads (reads with insufficient length, and therefore with no overlap region) were removed from analysis. The headers of the joined reads were edited using BBTools software suite (Bushnell 2016) so that the headers would reflect sample number and site information. Reads were then dereplicated using USEARCH 9.0 (Edgar 2010) in order to create individual sequence units (ISUs). Dereplicated ISUs from all samples in this study were concatenated into one sequence file. Dereplicated ISUs were then sorted by size using USEARCH 9.0 (Edgar 2010) in order for them to properly cluster in the following step. Decision making in the bioinformatic process occurred at the sorting step. ISUs with only one constitutive sequence (i.e. singleton) or only two constitutive sequences (i.e. doubleton) were either kept or removed from further analysis, depending on the trajectory of the workflow (Fig. 1). These sorted sequences were then clustered at either 95% or 97% similarity threshold, using USEARCH 8.0. We chose 97% because

that is the most commonly used clustering cutoff for rRNA comparison of AM fungi, and is thought to differentiate to the species level (Öpik et al. 2014). We chose 95% because the rRNA region is known to be variable for AM fungi (Sanders et al. 1995), particularly in certain genera and species (Clapp et al. 1999; Rodriguez et al. 2005), separating species too liberally. Clustered sequences were then checked for chimeric sequences against the MaarjAM database (Opik et al. 2010) using the UCHIME algorithm from USEARCH 9.0 (Edgar 2016) in order to remove potentially spurious sequences from the community analysis (Reeder and Knight 2009). Sequences were then queried against either the NCBI nucleotide database or the MaarjAM database (AM fungal type sequences) using BLAST+ 2.2.28 (Camacho et al. 2009). BLAST results were observed and organized in MEGAN Community Edition, v. 6.7 (Huson et al. 2016) with the default lowest common ancestor (LCA) parameters (minimum score of 50.0, minimum support percent of 0.01, and with minimum-complexity filter off).

Statistical Analysis

Because NGS resulted in unequal sequencing depth between samples (Harris et al. 2010), OTUs were resampled using the minimum number of reads per sample (site) in RStudio 1.1419 with the 'rrarefy' function in vegan 2.4.4 (RStudio 2016; Oksanen et al. 2017). Both rarefied (resampled) and raw read data are used for all analyses. Table A1 summarizes the raw and rarefied read and OTU data for each of the 9 samples used in this study. Tables S2 and S3 summarize the composition of AM fungi (OTUs and read abundance, respectively) found from each of the 12 workflow decisions using the raw read data. Tables S4 and S5 summarize the same information, except using the rarefied data.

AM fungal diversity between all 12 workflow decisions was calculated with the use of three diversity indices (Shannon, Simpson, and Fisher's α). These were calculated using the 'diversity' and 'fisher.alpha' functions in vegan 2.4-4, using RStudio 1.1.419 (RStudio 2016; Oksanen et al. 2017). In order to determine whether enough samples (sequences) were obtained in each of 12 workflow decisions, as well as, to compare AM fungal diversity among samples, rarefaction curves were drawn using the 'ggiNEXT' function in the iNEXT package (Chao et al. 2014) with RStudio 1.1.419 (RStudio 2016), using both raw and rarefied data. Subsequent discussion of OTUs and reads for comparison of AM fungal community composition, uses the raw data for simplicity's sake and because no major differences in AM community composition were found between the raw and rarefied data.

Results

Comparison of workflow decisions

After quality filtering, 1,555,515 reads remained for processing. After removal of potentially chimeric sequences, and removal of non-AM fungal reads, the data sets observed in this study contained a wide range of AM fungal diversity, depending on the bioinformatic decisions made. The largest amount of AM fungi (2,317 OTUs and 1,330,656 reads) were detected when the 97% clustering threshold, MaarjAM database, and no rare sequence removal was used (Table 2). However, much of this apparent diversity were likely artefacts (more on this in Discussion). The least amount of AM fungi (56 OTUs and 329,630 reads) were detected when the 95% clustering threshold, NCBI database, and singleton and doubleton removal were used (Table 2). Selection of a reference database had the largest noticeable impact on downstream results of AM fungal

OTU and read amounts. Using the MaarjAM database always supplied us with more reads than did the NCBI database, no matter what clustering threshold or rare sequence removal decisions were made (Table 2).

AM fungal community composition

AM fungal community composition varied dramatically based on the reference database selected. Fig. 2 summarizes the main differences in community composition detected when using the NCBI vs. the MaarjAM database. *Glomus* was the most common AM fungal genus detected when using the NCBI database, with 180,393 reads, accounting for 54% of all reads (97% clustering, singletons and doubletons removed) (Fig. 2, Table A3). The most common genus detected when using the MaarjAM database was *Paraglomus*, no matter whether raw data (Fig. 2, Table A3) or rarefied data was used (Fig. A3). Interestingly, when comparing against the NCBI database, only one or two *Paraglomus* spp. matches were detected (95% clustering and 97% clustering, respectively, both with singleton and doubleton removal) (Table A2). Additionally, the NCBI database captured no reads associated with *Ambispora* spp., while the MaarjAM database captured between 244 (97% clustering, singleton and doubleton removal) and 1,199 reads (95% clustering, no rare sequence removal) matching to this clade. Similarly, when using the NCBI database, we detected much fewer *Archaeospora* spp. and *Claroideoglomus* spp. than when we used the MaarjAM database (Table A3). No *Gigaspora* spp. were detected when using the MaarjAM database, whereas, they were detected in small amounts (8-18 reads) when the NCBI database was used. *Geosiphon* (a monotypic clade which associates with a cyanobacterium, not plant roots) was only detected as singleton sequences, and therefore, was likely an artefact.

Removing all singletons and clustering at different thresholds made only a small impact on the overall proportion of AM fungi represented in this study (Tables S2, S3). The AM fungi primarily amplified in this study were *Glomus*, *Acaulospora*, and *Paraglomus*; all other genera were detected in relatively small amounts (close to 0%), no matter which clustering or rare sequence removal decisions were made. Similarly, the top two most abundant OTUs when using the MaarjAM database were consistently *Paraglomus* spp. and *Glomus* spp., no matter which rare species removal cutoffs or clustering thresholds were utilized. These clades accounted for 43-45% and 32-33% of all reads, respectively. Also, the top two most abundant OTUs when using the NCBI database were consistently *Glomus* spp. and *Undescribed Glomeromycota*. These clades accounted for 54-65% and 16-21% of all reads, respectively.

Rarefaction curves and diversity indices

Rarefaction curves show saturation of OTUs when both singletons and doubletons are removed from the analyses. In all four workflow decisions in which singletons and doubletons are removed, all sites sampled (CM, LIPB, and WSF) show sufficient sampling to detect all AM fungal OTUs (Fig. 3a, 3d, 3g). When no rare sequences are removed, the rarefaction curves show no plateau, meaning more sampling effort is needed to reveal all AM fungal OTUs. Although not a focus of this study, it should be noted that a difference in rarefaction curves between sites showed that the removal of rare sequences impacted their differences in diversity. For example, when 95% clustering, no rare sequence removal, and the NCBI database were used, CM has the highest species diversity according to the rarefaction curve (Fig. 3l). However, when using the same clustering threshold and reference database but removing singletons (Fig. 3k) or

singletons and doubletons (Fig. 3j), WSF has the highest species diversity. Similar trends can be seen in the other workflow decisions (i.e. 97% clustering, MaarjAM database) (Fig. 3). Therefore, rare sequence removal impacted the resulting differences in diversity between sites.

AM fungal diversity, as evidenced by Shannon, Simpson, and Fisher's α index scores, was impacted by the bioinformatic workflow decisions. All three diversity indices decreased as rare sequences were removed from analyses (Tables S6 and S7). Using the NCBI database also decreased diversity, compared to using the MaarjAM database. Additionally, clustering at 95% decreased diversity, compared to using the 97% clustering threshold (Tables S6 and S7).

Unclear taxonomic assignment

As the data for this study was sorted, we recognized several important results. Firstly, when using either of the two reference databases, many OTUs matched above the genus level. For example, when using the MaarjAM database, 97% clustering, and with removal of singleton and doubleton sequences, 40% of OTUs (205 out of 509 total AM fungal OTUs) and 21% of reads (221,579 out of 1,038,345 total AM fungal reads) were identified above the genus level. Similarly, when using the NCBI database, 95% clustering, and with removal of singletons, 43% of OTUs (39 out of 91 total AM fungal OTUs) and 34% of reads (123,434 out of 364,373 total AM fungal reads) were identified above the genus level (Table A3). Secondly, many OTUs were named as one taxon when using the MaarjAM database and named as a different taxon when using the NCBI database, often with distant evolutionary similarity. For example, when the results for the trajectory with 95% clustering, singleton and doubleton removal was inspected, CM2_98

was identified by the MaarjAM BLAST as '*Paraglomus* sp.' with an 88% identity match. The same OTU was identified as 'Dikarya' by the NCBI BLAST with a 94% identity match. The reference sequence length for each of these matches was 520 bp (MaarjAM hit) and 1171 bp (NCBI hit). This particular OTU was relatively abundant (7,572 reads), but this mismatching of names occurred also with less abundant OTUs (Table A8). Because of the relatively low percentages of identity in this OTU hit and others, and because reference sequence size was often much larger than the template sequence, many reads were likely misidentified. This lack of resolution for many of the OTUs in this study is a major problem which is further examined in our discussion.

Discussion

One of the main goals of this study was to determine whether using Nextera XT barcoded, AM fungal primers could effectively uncover AM fungal communities from plant root samples. In order to test this question, we sampled 9 switchgrass plant roots from 3 Pine Barrens ecosystems, amplified a fragment of SSU rDNA using Nextera XT AM fungal primers and sequenced our PCR products on the Illumina MiSeq. We found that using this protocol, we were able to successfully amplify AM fungal DNA. Even though all samples came from harsh environmental conditions (i.e. acidic, sandy soils with low nutrients), an impressive amount of AM fungal diversity was observed. In particular, when we compared our results against the MaarjAM database, between 7-9 known AM fungal genera were detected, depending on the bioinformatic workflow decisions utilized. Using bioinformatic conditions often seen in the literature (MaarjAM, 97% clustering, singleton and doubleton removal), we uncovered 509 AM fungal OTUs. This amount of OTUs exceeds the total number of currently described AM fungal

morphospecies (c. 300). This excessive amount of OTUs may have been detected because AM fungi are known to have several to many nuclei in a single spore and a lot of genetic variation even within the same morphospecies (Hijri and Sanders 2005; Stockinger et al. 2009). This amount of OTUs also exceeds those found in some recent AM fungal ecological studies, which found 186 (Liu et al. 2017) or 130 total AM fungal OTUs (Schlaeppi et al. 2016) and, of course, greatly exceeds those in cloning based studies, in which one study found only 12 OTUs and only 2 genera (Schechter and Bruns 2012). Therefore, it seems we succeeded in developing a workflow that uncovers great AM fungal diversity.

With regards to whether our primers successfully targeted all major AM fungal groups, it should be noted that members of the *Diversispora* were hardly found in any samples. Only 2 OTUs belonging to *Diversispora* spp. were found, and they were both singletons. *Diversispora* spp. may have not been well detected in this study because the primers were not able to amplify this group of fungi. It has been shown that SSU rDNA is often insufficient in providing adequate resolution for the Diversisporales (Opik et al. 2013; Ohsowski et al. 2014). Another explanation for the lack of *Diversispora* fungi in this study is that members of this genus were simply not present in the roots collected from the Pine Barrens. Moora et al. (2014) found *Diversispora* spp. to be indicator species of the anthropogenic landscapes surveyed in their study. Similarly, results from a study looking at the AM fungal communities inhabiting switchgrass in Pine Barrens versus managed landscapes found *Diversispora* spp. to be prolific in anthropogenic landscapes (3,666 reads), but only minimally present in Pine Barrens sites (237 reads)

(Bindell et al. unpublished). Therefore, we concluded that this set of primers and our workflow should still work well for other environments and conditions.

A second major goal of this study was to determine whether our bioinformatic decisions had an impact on downstream AM fungal diversity results. We found that some bioinformatic decisions altered our downstream results substantially while others had negligible impacts. One of the major impacts on our results was caused by our reference database selection. It was surprising that despite the much larger size of the NCBI database compared to the MaarjAM database, the MaarjAM database always detected more diversity of AM fungi and was able to identify more sequences to the genus level. This is partly due to the fact that the queries against the NCBI database did not exclude environmental samples. This may have skewed our results to dubious species matches and higher order matches. For example, OTUs matching to the genus, *Paraglomus*, were found exclusively when using the MaarjAM database. Hardly any OTUs identified as '*Paraglomus*' were found when using the NCBI database, no matter which clustering threshold or rare sequence removal decisions were made. This major difference must not be taken lightly, as this was the second most prolific genus uncovered when using the MaarjAM database (26% of reads) (Table A3). Through further investigation into this peculiar finding, we found that many OTUs that were being matched to '*Paraglomus*' were identified as completely different fungi when using the NCBI database for querying. For instance, in the dataset which used 95% clustering and singleton and doubleton removal, one of the most abundant OTUs (6212 reads), WSF3_42, was identified as '*Paraglomus*' by the MaarjAM database, but as 'Dikarya' by the NCBI database. Under deeper inspection, we found this apparent misidentification by

the NCBI database may have been due to the fact that the top hit in NCBI was to 'Basidiomycota', while the other hits were to a large variety of other taxa (Table A8). When NCBI has hits to different taxa, the BLAST results will show the consensus name at the highest taxonomic level, 'Dikarya' in this case. Much of these issues of taxonomic assignment stem from the fact that the NCBI database is full of 'environmental samples' that have no corresponding, described fungal names attached to them (Melo et al. 2017) and, therefore, have no clear taxonomic name. Another problem is that we did not first filter out all non-AM fungal reads before aligning against the MaarjAM database. This would certainly decrease the number of false positives seen in the MaarjAM results. Part of the problem may also be the low threshold for similarity being used by the BLAST+ function in the MaarjAM database. The MaarjAM database only showed 84% similarity between WSF3_42 and the reference sequence. It is plausible that this is a new species within *Paraglomus*. Another possibility is that when we clustered sequences into phylotypes (i.e. molecular taxonomic units) (Melo et al. 2017), we may have erroneously grouped several species into the same OTU or separated one species into multiple OTUs, thereby giving us a faulty representation of AM fungal diversity in the environment (House et al. 2016), similarly proposed in Melo et al. (2017). Clearly, our taxonomic data is only as good as our databases. More work needs to be done to further expand our AM fungal databases to describe novel species both morphologically, as well as, molecularly. Sequencing of the SSU (and/or additional regions) of rDNA for currently non-sequenced specimens in the AM fungal collections, as well as, whole genome sequencing for important basal clades is necessary for the future of AM fungal taxonomy (Bruns et al. 2017).

Another potential pitfall of this workflow is that our data rely on a relatively small segment of the SSU region (~250 bp). The set of primers we used in this study have been used in a handful of studies (Cui et al. 2016; Ban et al. 2017; Liu et al. 2017), with mixed amount of specificity for AM fungi. Despite the fact that the SSU region is commonly used in AM fungal ecological studies and makes up the bulk of the MaarjAM reference database, there is little evidence that this locus is a barcode that can accurately distinguish AM fungal species from other fungi (interspecific variation) or between species within the Glomeromycotina (intraspecific variation) (Öpik et al. 2014). As it is well known, sequencing the DNA from several loci is the gold standard for taxonomic clarification. Taxonomic studies are increasingly using SSU, ITS, LSU, and protein coding genes for the proper phylogenetic placement of other fungal lineages (James et al. 2006). Using only one gene region makes it difficult, albeit impossible, to obtain species-level identifications on metagenomic data (Bruns et al. 2017). Even matches to the genus-level are often impossible to obtain. For instance, when using the NCBI database (95%, singleton and doubleton removal), no OTUs were identified to the species level, while 19 out of 56 OTUs were identified above genus level, giving limited insight on the diversity present in nature. And even when we used the AM specific, well curated, MaarjAM database (95%, singleton and doubleton removal), only 21 out of 350 of OTUs were identified to the species level, while 160 OTUs were identified above the genus level. Future ecological studies would certainly benefit from an improved barcode for AM fungal molecular species identification.

Despite these observations, we found little impact of clustering threshold or rare sequence removal on overall AM fungal diversity (Table 2). However, it must be noted

that we only compared 95% to 97% clustering thresholds. Comparing a wider range of thresholds may produce different results. Table 2 shows that the Shannon and Simpson's diversity indices didn't vary much with differing clustering thresholds or rare sequence removal. However, we did find that clustering at 97% greatly inflated our total number of AM fungal OTUs. As it has been shown before, the process of high throughput sequencing (including PCR biases and bioinformatic workflow biases) can often inflate the number of OTUs way beyond what is seen in the natural environment (Reeder and Knight 2009). For instance, 97% clustering detected more than 3-fold more OTUs (654 OTUs) compared to 95% clustering (201 OTUs) (NCBI, no rare sequence removal) in this study. Nevertheless, as rare sequences were removed, OTU inflation, occurring with the use of both reference databases, was substantially mitigated. For example, there were a whopping 2,317 OTUs detected when no rare sequences were removed (MaarjAM, 97% clustering). But when we removed both singleton and doubleton sequences from our analyses, we removed more than three quarters of these artefactual OTUs, with 509 OTUs remaining (Table 2). Similarly, decreasing the clustering threshold to 95% further removed artefactual sequences, bringing our total OTU count to 350, which is on par with the total number of morphologically described AM fungal species (Öpik and Davison 2016). Comparable trends were found when using both reference databases. Our results are similar to those of Morgan and Egerton-Warburton (2017), who detected a whopping 1,000-5,817 AM fungal OTUs (at two locations, respectively) at 97% clustering and no rare sequence removal. They, too, found that using 95% clustering and rare sequence removal decreased OTU inflation to between 285-1,178 OTUs. These are very important and pertinent findings because most AM fungal diversity studies use a 97% clustering

threshold (Kluber et al. 2012; Liu et al. 2017; Sepp et al. 2018). Using 95% clustering threshold seems to be a better choice for AM fungal ecological studies, as it detects AM fungal OTU richness that is most consistent with our current estimates of AM fungal, global OTU richness (Opik et al. 2010).

The debate on whether to remove rare sequences and how to define them will most certainly continue beyond our present study, as different groups of microbes may require different methodological tweaks. In this study, some clades were present only when rare sequences were not removed. *Geosiphon*, for instance, was found only when singleton sequences were kept (Table A4). We can be fairly certain that a *Geosiphon* species was not living in the root of our switchgrass sample because members of this clade have, thus far, been found exclusively in association with cyanobacteria, not with plant roots (Gehrig et al. 1996). Nevertheless, rare sequences may not always represent artefacts; rare sequences found similarly in multiple sample sets can sometimes indicate a truly rare species. The 'leave-one-out' rule may, at times, oversimplify the true diversity present and inappropriately leave out some significant rare species (Reeder and Knight 2009). As for AM fungi, though, it seems that keeping singleton sequences will unfortunately overinflate both total AM fungal OTU numbers as well as overall diversity estimates. Although not a focus in this study, future AM diversity studies may want to compare replicate samples for any shared rare sequences before blindly removing them from downstream analyses, as shared rare sequences may not be spurious. Once confirming that rare sequences are not shared among replicate samples, it would be wise to remove all singleton *and* doubleton sequences. This would help reach the level of

OTU richness estimated by our current databases and global studies (Opik et al. 2010; Opik et al. 2013)

Our study uncovered different AM fungal communities when using different reference databases, encouraging careful future selection of a reference database. Because many studies rely strictly on the MaarjAM database (Davison et al. 2012), with some exceptions (Schlaeppi et al. 2016; Melo et al. 2017), it is important to know how this database works with a variety of workflow decisions. Therefore, when using alternative workflow decisions, future studies must keep in mind that this may alter the way the reference database should be used.

Conclusion

This study developed a novel approach to studying AM fungal communities by using AM specific primers, Nextera XT indices, and dual-indexed Illumina MiSeq sequencing. This approach uncovered a wide array of microbial diversity in the Pine Barrens ecosystem. The bioinformatic tools and associated protocols were tested here to determine whether or not they would impact downstream AM fungal diversity and community composition results. In conclusion, we found that reference database selection made a large impact on both AM fungal diversity, as well as, OTU richness. Several AM fungal clades, such as, *Paraglomus*, were not observed when using the NCBI database, but found to be prolific when using the MaarjAM database. Varying the clustering threshold made little impact on AM fungal diversity and OTU richness, particularly when rare sequences were removed from the analyses. This study emphasizes the need for examining each bioinformatic decision in ecological AM fungal diversity studies. It is suggested that future studies using this primer set utilize the 95%

clustering threshold, as this threshold lowers the extent of OTU inflation, compared to the 97% threshold. Additionally, singletons and doubletons should be removed, as they are likely spurious species and often overinflate OTU richness. Lastly, although the NCBI database provided ample AM fungal taxonomic information, the MaarjAM database is a better choice for ecological studies, as this database has more updated taxonomic names, uses only published, verified sources, and is able to identify a more extensive amount of AM fungal taxa.

Table 3-1 Sampling site information.

Site	State	Root Samples	Site Coordinates	Soil Properties						
					pH	SOM [a] (%)	CEC [b]	P (ppm)	K (ppm)	Ca (ppm)
Colliers Mills	New Jersey	3	°40.0680667, °-74.4449333		5	0.2	0.6	11	911	1,129
Long Island Pine Barrens	New York	3	°40.8975333, °-72.6586500		4.9	0.4	0.5	8	138	8,913
Wharton State Forest	New Jersey	3	°39.7557667, °-74.6947333		5.2	2.6	7.3	41	4,341	48,443

[a] Soil organic matter content

[b] Cation exchange capacity

Table 3-2 Arbuscular mycorrhizal (AM) fungal operational taxonomic units (OTUs) and AM fungal reads detected in this study through different workflow decisions. Decisions include clustering at different similarity thresholds (95% and 97%), after different levels of rare species removal (removal of singleton, doubleton reads), and through comparison via the Basic Local Alignment Search Tool (BLAST) query against two DNA reference databases, the National Center for Biotechnology Information (NCBI) and the MaarjAM database for AM fungal sequences (Opik et al. 2010). Blue indicates all results from 95% clustering and red indicates all results from 97% clustering.

Clustering Threshold	Reference Database	Rare sequence removal	OTUs	Reads
95%	NCBI	N[a]	201	433,839
		NS[b]	91	364,373
		ND[c]	56	329,630
	MaarjAM	N	975	1,345,563
		NS	527	1,140,139
		ND	350	1,051,398
97%	NCBI	N	654	438,833
		NS	185	363,094
		ND	87	331,380
	MaarjAM	N	2,317	1,330,656
		NS	913	1,124,121
		ND	509	1,038,345

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

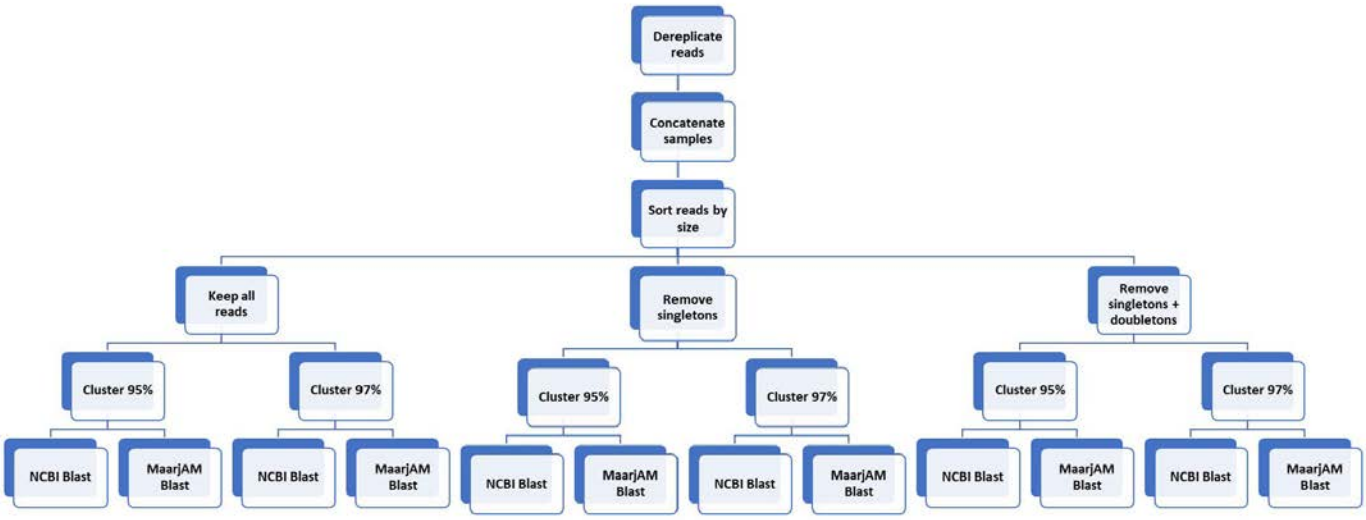

Fig. 3-1 The bioinformatic workflow.

Major workflow steps in this study include: dereplication of reads, concatenation of reads from all samples, sorting reads by size, removal of rare species (i.e. singleton or doubleton reads), clustering reads, and the Basic Local Alignment Search Tool (BLAST) query against reference databases. Bifurcations in the workflow represent bioinformatic decisions made in this study.

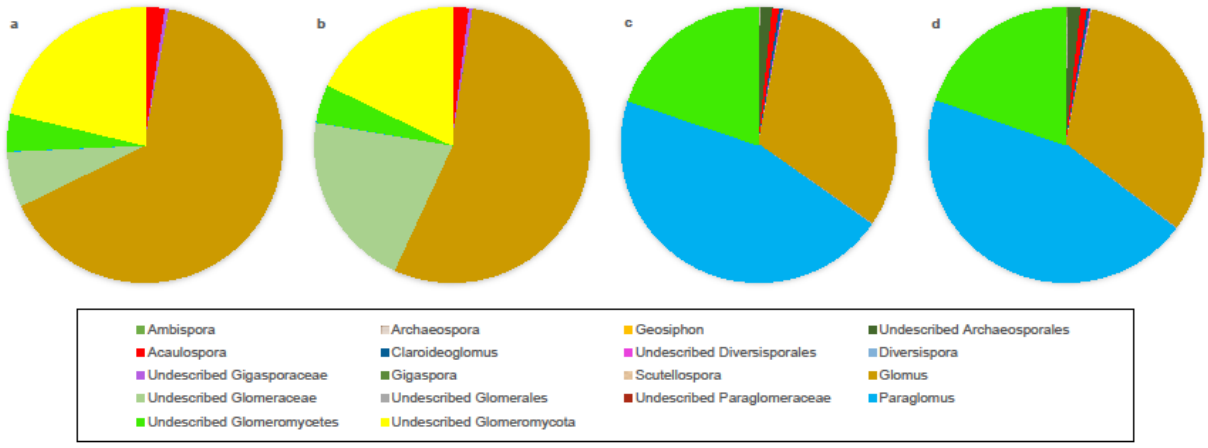

Fig. 3-2 Arbuscular mycorrhizal (AM) fungal diversity under different workflow decisions.

AM fungal diversity detected when using 95% clustering threshold, NCBI reference database, singleton and doubleton removal (a), 97% clustering threshold, NCBI reference database, singleton and doubleton removal (b), 95% clustering threshold, MaarjAM reference database, singleton and doubleton removal (c), and 97% clustering threshold, MaarjAM reference database, singleton and doubleton removal (d). Raw read data are used.

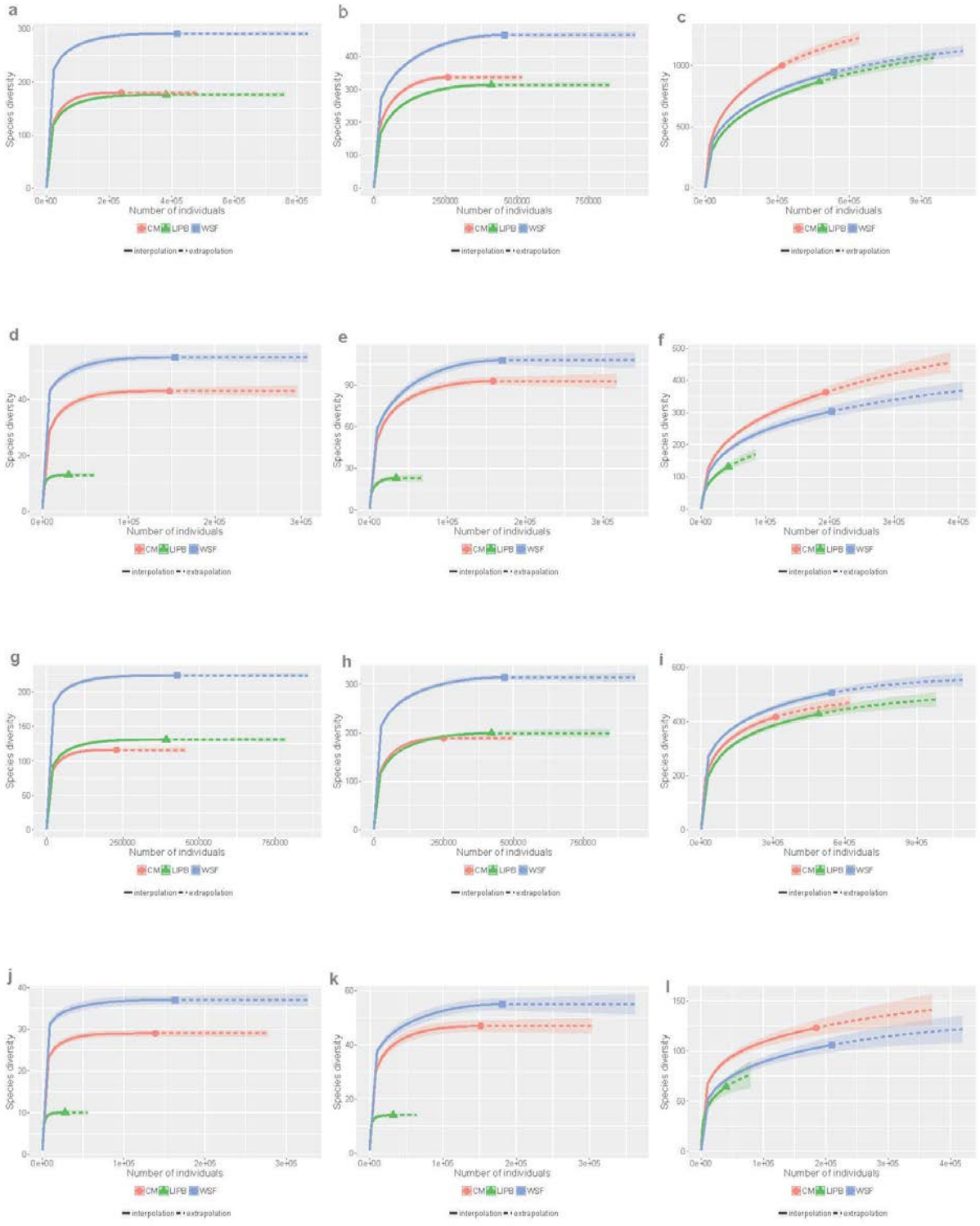

Fig. 3-3 Rarefaction curves of operational taxonomic units (OTUs) from all twelve workflow decisions.

Curves represent samples from Colliers Mills (CM) (red), Long Island Pine Barrens (LIPB) (green) and Wharton State Forest (WSF) (blue). Workflow decisions were: 97% clustering, singletons and doubletons removed, MaarjAM database (a), 97% clustering, singletons removed, MaarjAM database (b), 97% clustering, no rare sequences removed, MaarjAM database (c), 97% clustering, singletons and doubletons removed, NCBI database (d), 97% clustering, singletons removed, NCBI database (e), 97% clustering, no rare sequences removed, NCBI database (f), were 95% clustering, singletons and doubletons removed, MaarjAM database (g), 95% clustering, singletons removed, MaarjAM database (h), 95% clustering, no rare sequences removed, MaarjAM database (i), were 95% clustering, singletons and doubletons removed, NCBI database (j), 95% clustering, singletons removed, NCBI database (k), 95% clustering, no rare sequences removed, NCBI database (l). Raw read data are used.

Chapter 4

Soil pH impact on native Pine Barrens AM fungal colonization and on switchgrass plant growth: a pot experiment

Abstract

Acidic, oligotrophic soils are commonplace, with natural and anthropogenic soils often impacted by soil acidity and metal toxicity. Little is known about how naturally acidic soils impact arbuscular mycorrhizal (AM) fungal life and how this in turn can affect plant growth, particularly in rare ecosystems. This experiment sought to address whether the acidic soils of the New Jersey Pine Barrens forest promote AM fungal growth in plant roots, as well as, whether these acidic soils, in cooperation with naturally occurring AM fungi, produce larger, stronger plants. These questions were tested through a greenhouse pot experiment. Switchgrass seeds were planted in four sets of soil treatments, using soil collected directly from the New Jersey Pine Barrens forest. The treatments were as follows: sterile, whole soil (- AM control); non-sterile, whole soil (+ AM), sterile, whole soil with increased soil pH (- AM control), and non-sterile whole soil amended with increased soil pH (+ AM). Despite our efforts, plants desiccated and many died during the experiment. Additionally, little difference in plant growth was found between + AM and – AM treatments and between more and less acidic treatments. Lessons learned were: you can never over-plan, manageable size experiments are key, and never give up.

Introduction

Soil acidity is a common phenomenon. Acidic soils, however, can be a major constraint to plant growth. Because metal toxicity often goes hand in hand with low soil pH, acidic soils are a particular problem for agricultural ecosystems, which often need to utilize acidic soils for crop production (von Uexküll and Mutert 1995; Iqbal 2012). Crop production is often challenged by acidic and metal-toxic soils. This has forced farmers and scientists to try to grow crops naturally tolerant to acidic soils. One such crop is switchgrass (*Panicum virgatum* L.). This grass has been used for biofuel production for years (Parrish and Fike 2005). It is widely adaptable to different soil and climate conditions, as it is native and hardy within most of the United States and southern Canadian provinces (McLaughlin 1993). Additionally, this grass is known to associate with arbuscular mycorrhizal fungi (Bentivenga and Hetrick 1992; Schroeder-Moreno et al. 2012) (AM fungi). These fungi are extremely important because they help plants gain access to otherwise inaccessible but significant nutrients, like phosphorus, under acidic soil conditions (Smith and Read 2008).

In order to understand how AM fungi can best be utilized to improve plant production, it is important to know what influences their community structure in nature. Just like plants, arbuscular mycorrhizal (AM) fungi can be impacted by soil acidity. Wang et al. (1993) showed no difference in colonization at soil pH levels ranging from 4.5 to 7.5, a relatively wide range. However, Ouzounidou (2015) found that more alkaline soils increased AM fungal colonization. With regards to AM fungal communities, Kawahara et al. (2016) found that acidifying neutral soil shifted the AM fungal community, whereas, neutralizing acidic soil did not change that AM fungal

community. They found that the acidic-soil inhabiting AM fungi were a subset of the neutral-soil inhabiting AM fungal community. Johnson et al. (1991) showed minimal effects of soil acidity on AM fungal diversity. However, more recent and larger scale studies have shown a relatively large impact of soil acidity on AM community composition (Lekberg et al. 2011; Hazard et al. 2012; Suzuki et al. 2014; Stürmer et al. 2018).

Of course, the findings of these authors are often contradictory. Select or unusual ecosystems might have different pressures affecting their plant and fungal communities. One such ecosystem of interest is the Pine Barrens forest. The Pine Barrens is an unusual ecosystem found in only a handful of places on Earth. It's uniqueness stems from its acidic (pH ~ 5), sandy soils that are very low in nutrients and high in toxic metals like aluminum (Forman 1998). Frequent forest fire and a specific plant community associated with the Pine Barrens also lead to its intrigue. The plant species that populate these 'barren' soils are trees like pitch pine, shrubs like blueberry, and grasses such as switchgrass. In order to delve into why switchgrass may be able to survive so well in an otherwise unhospitable soil environment, we decided to look into switchgrass's native AM fungal communities from the Pine Barrens of New Jersey, USA. Our main goal was to see if the acidic soil influences AM fungal colonization and diversity and whether increased colonization under acidic soil conditions impacts plant growth.

Materials and Methods

First, it should be noted that this experiment was the second round of experiments in the greenhouse, for which the goal was to test whether native Pine Barrens-derived AM fungi impact switchgrass growth and if the soil acidity increases the AM fungal root

colonization. The first round of experiments did not result in AM fungal colonization and so the current study was employed to help further our knowledge. (More on this in 'Discussion').

Soil Collection

Soil was collected from around *Panicum virgatum* plant roots in a Pine Barrens forest in Colliers Mills, NJ, USA in November 2015. The Pine Barrens ecosystem consists of a pitch pine-scrub oak dominated forest with herbaceous and ericaceous understory and it is known for its acidic, oligotrophic soils (Forman 1998). Soil was collected in the fall season, so that photosynthesis would be slower and energy would, therefore, switch from AM hyphal growth to overwintering spore production. The idea being that more spores in the soil would provide more inoculum for inoculating the experimental plants.

Experimental Design

Seeds of *Panicum virgatum* v. Kanlow were purchased and used in order to provide genetic similarity between seeds and limit confounding variables. *P. virgatum* seeds were surface sterilized by soaking in 95% ethanol for 30 s, 0.06% NaOCl for 30 s, and 70% ethanol for 2 min. Seeds were then rinsed thoroughly with sterile, distilled water and air dried under a laboratory fume hood. 35 seeds per planting cell were pregerminated with daily misting on twice autoclaved, Colliers Mills soil for 15 days. Seedlings were then transplanted to 14 cm pots (Fig. 1). Soils used were a 1:1 ratio of native Colliers Mills field soil and trap culture soil. Trap culturing is a common technique for growing AM fungal inoculum in soil over a period of several months, in

this case, for the purposes of increasing the amount of Pine Barrens native-propagules in the collected soil. We followed the procedures with sudangrass used as the trap culture host. More details on how to produce trap cultures of AM fungi can be found on the website for the International Culture Collection of Arbuscular Mycorrhizal Fungi (INVAM): https://invam.wvu.edu/methods/cultures/trap-culture . All soil was sieved to remove large debris and weighed in order to ensure evenness and equal amounts of soil in each container. Plants were grown in a climate-controlled glasshouse (greenhouse) at Rutgers University, New Brunswick, NJ, USA from September 2016-May 2017.

The following were the four soil treatments used at transplanting: autoclaved whole soil (- AM control) (abbreviated AC), non-autoclaved whole soil (+ AM) (abbreviated NAC), autoclaved whole soil amended with CaO powder for increased pH (- AM control) (abbreviated AC-pH), and non-autoclaved whole soil amended with CaO powder for increased pH (+ AM) (abbreviated NAC-pH) (Fig. 1). CaO (lime) was added to the pH-increased treatments in order to observe any impacts increased soil pH would have on AM fungal colonization and diversity. Autoclaved treatments were autoclaved for 30 min, let to rest for 24 hours, and then re-autoclaved for 30 min in order to sufficiently kill all biota but still preserve abiotic properties. 10 g of CaO was added to 13.6 kg of once-autoclaved soil while still warm, in order to better mix the soil. Soil was then re-autoclaved the following day. The same was done to 13.6 kg of the non-autoclaved soil. Soil pH was tested prior to potting using a pH duplex indicator test kit (LaMotte Company, Chestertown, MD). Soil pH measurements were: 7.0, 6.0, 7.0, and 6.0 for AC-pH, AC, NAC-pH, and NAC treatments, respectively. Soil was subsequently tested at approximately 6 weeks after transplanting and the soil measurements were 6.5,

5.5, 7.0, and 6.0 for AC-pH, AC, NAC-pH, and NAC treatments, respectively. Some fluctuation was expected, but overall, our soils treatments stayed distinct from their non-pH-increased counterparts. Soil was also tested at the end of the experiment for final soil pH measurements and other soil properties by Spectrum Analytic (Washington Courte House, OH) (Table 1).

During transplantation, 75 ml of a soil slurry was poured onto the roots of all seedlings. This was done in order to make sure that all treatments started out with similar native microbial communities, excluding the AM fungi. The soil slurry was made by first soaking non-autoclaved Colliers Mills soil (mixed with pot culture soil) in water over night. The soil to water ratio was 1:3. The soil slurry was then passed through a 500 µM sieve and then a 38 µM sieve. Large soil particles and spores were caught on the sieves, allowing smaller microbiota and small soil particles through. In order to confirm that the soil slurry contained no AM fungal propagules, leek and sudangrass seeds were sown in twice autoclaved Colliers Mills soil that was soaked in the soil slurry. 3 sets of both leek and sudangrass seeds were sown (several seeds in each plant cell), along with negative controls (no soil slurry). Leek and sudangrass are considered highly mycorrhizal plants (https://invam.wvu.edu) and were, therefore, chosen as good indicators of whether AM fungi were erroneously present in our soil slurry. Total roots of both plants were stained with 0.05% Aniline Blue (Grace and Stribley 1991) at 4 weeks, ample time for colonization to take hold for highly mycorrhizal plants such as these. Non-AM fungal hyphae were observed in sudangrass and leek roots, but no AM fungal structures were observed, as expected. Therefore, the -AM treatments (AC and AC-pH) were deemed to be sufficiently free of AM fungi.

Each of the 4 soil treatments were replicated 10 times, for a total of 40 pots. One random pot from each treatment was removed during the experiment, though, to have a pot from which to harvest and observe root colonization by AM fungi, helping gauge when to harvest all pots. Seedlings were thinned to 9 plants per pot during the first two weeks after transplantation. Shoot length of each plant was measured at 2-, 3-, and 4-week intervals, as the experiment progressed, in order to measure differences in shoot growth rate between treatments and controls.

Root and Shoot Measurements

At 31 weeks, plants were harvested. Root length (rooting depth) and shoot lengths were measured for each plant. Root lengths were calculated by measuring the longest root from each plant. Shoot lengths were measured in the same fashion. Soil pH of each replicate pot was measured from the soil around the root zone using a pH duplex indicator test kit (LaMotte Company, Chestertown, MD). After root and shoot length measurements, roots were rinsed thoroughly in running water. A similar amount of randomly selected roots from each pot was set aside for microscopy and metagenomic analysis, respectively. Shoots were pooled together, as were, the remaining roots, and dried in a dehydrator until all moisture was removed. Dry weight was measured for both shoots and roots.

Root Microscopic Observation

Roots set aside for microscopy were pooled within treatments and 3 subsamples of pooled roots were stained with 0.05% Aniline Blue (Grace and Stribley 1991). These 3 subsamples of roots were observed under a Nikon light microscope and AM fungal

colonization was quantified (McGonigle et al. 1990). 100 intersections per subsample were observed. AM fungal structures counted included AM fungal hyphae, hyphal coils, arbuscules, vesicles, and spores.

DNA Extraction, PCR, and Sequencing

Roots set aside for metagenomic analysis were frozen at -80°C after rinsing. 3 replicate tubes of 0.25g pooled roots for each of 4 treatments were ground with liquid nitrogen (12 samples total). DNA was extracted from ground roots using a DNeasy Plant Mini Kit (Qiagen, CA, USA) following manufacturer's instructions. AM fungal specific primers, AMV4.5NF/AMDGR (Sato et al. 2005), were used in order to amplify a 300 bp region of the SSU rDNA, a commonly used gene region (Lee et al. 2008; Opik et al. 2010) that can delimit virtual taxa (VT) of AM fungi (slightly higher level than the morpho-species level) (Opik et al. 2016). Because the plants of the +AM treatments looked visually smaller than -AM treatments, fungal ITS primers, ITS1F KY01-OA/ITS4-OA (Toju et al. 2012), were used to detect any pathogenic or other notable general fungal differences between treatments. Nextera XT adapters (Oligonucleotide sequences © 2018 Illumina, Inc.), designed to attach on one end to Illumina MiSeq adapters and on the other end, to the fungal primers, were used in place of overly length Illumina adapters. This was done to overcome primer dimers, which consistently occurred when using Illumina adapters attached to our fungal-specific primers.

PCR was conducted with a mixture of 0.5 ul each of 10X forward and reverse primers, 12.5 ul Taq 2X Master Mix (New England BioLabs, Maine), and 1 ul template DNA. PCR parameters were adjusted from the Nextera XT protocols ((Illumina, part 15031942 rev. C, October 2012) because of inconsistent PCR product concentrations and

with few bands verified through gel electrophoresis. PCR parameters were therefore modified as follows: 95°C for 2 min, 35 cycles at 95°C for 45 s, 50/52/55°C for 45 s, and 72°C for 1.5 min. Final extension was 72°C for 5 min. Three annealing temperatures were used to best amplify all AM fungal species present in the samples (Schmidt et al. 2013). Gel electrophoresis confirmed bands in most of the +AM samples.

PCR products were cleaned up with an Agencourt AMPure XP kit (Beckman Coulter, Brea, CA) following manufacturer's instructions. Secondary PCR was then run to add Nextera indexes and Illumina adapters to each sample. Each reaction contained 25 µl NEBNext High Fidelity 2X PCR Master Mix, 5 µl clean, primary PCR amplicon, 5 µl Nextera XT index 1 primer, 5 µl Nextera XT index 2 primer, and 10 µl PCR grade water, yielding a 50 µl total reaction volume. Secondary PCR was: 95°C for 3 min, then 8 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s. Final extension was 72°C for 5 min. Secondary PCR products were cleaned again using the Agencourt AMPure XP kit. Paired end 300x300 bp sequencing was performed using the MiSeq Reagent Kit v3 (600 cycle). A positive control was included to confirm our protocol. This control was PCR product from *P. virgatum* root sample from Collier's Mills previously tested on the MiSeq and found to have positive presence of AM fungi. Two negative controls were also included to test whether contaminants made their way into the libraries and to further test our protocol. One negative control consisted of PCR grade water while the other consisted of all PCR reagents except template DNA. All samples were subjected to sequencing on the Illumina MiSeq™ platform.

Bioinformatic Analysis

Low quality bases, primer regions, and adapters were removed using CLC Genomics Workbench v 8.5.1 (CLC Genomics Workbench 2017). Forward and reverse reads were then merged together. Any non-merged reads (very few), i.e. those reads with no overlapping region, were discarded. This was because we expected full overlap for the 250 bp reads. Any reads <100 bp long were discarded, as well.

Sequences were dereplicated (combining reads that are identical) using the *fastx_uniques* command in USEARCH 9.0 (Edgar 2010). After dereplication, sequences were sorted by size using the *sortbysize* command in USEARCH 9.0. Singleton sequences (those with less than 2 constituent reads) were removed from our analyses using the *minsize* command in USEARCH 9.0. Reads from all treatments were then combined and clustered into operational taxonomic units (OTUs) at a 97% similarity threshold using the *cluster_otus* command in USEARCH 8.0 (Edgar 2010). Sequences were then queried against the MaarjAM database (AM fungal DNA reference database) (Opik et al. 2010) in order to detect and remove any potentially chimeric sequences. This was performed with the *uchime2_ref* command in USEARCH 9.0 (Edgar et al. 2011). All remaining OTUs were subjected to a Basic Local Alignment Search Tool (BLAST) against the National Center for Biotechnology Information (NCBI) nucleotide database using the *query* and *db* commands in BLAST+ 2.5.0. BLAST results (OTU identities) were observed in MEGAN Community Edition 6.6.7 (Huson et al. 2016) with the default lowest common ancestor (LCA) parameters (minimum score of 50.0, minimum support percent of 0.01, and with the minimum-complexity filter off). All matches to the "Glomeromycota" clade in MEGAN were kept for further inspection.

Statistical Analysis

Student's t tests were conducted to compare the dry weights of the shoots and roots of the four treatments, as well as, shoot lengths, and root lengths. One-way analysis of variance (ANOVA) was conducted on these data, as well, to account for the statistical problem of running multiple pairwise t tests. The student's t test was also conducted on the NAC and NAC-pH microscopic colonization data to test whether soil pH impacted the percent AM fungal colonization in these treatments. No further analyses were conducted because many plants died and colonization by AM fungi was not nearly as high as we expected it to be in any treatment, indicating problems with experimental design or implementation. Illumina sequences were not analyzed in depth because of issues with plant growth during the experiment. These issues are discussed below, under 'Plant Death' and in 'Discussion'.

Results

Plant death

Despite our best efforts and multiple resuscitation attempts in a previous greenhouse experiment with similar results, as well as pilot experiments in the laboratory, the switchgrass plants in this study became yellow and nutrient deprived. As can be seen in Fig. 2, many plants had spindly, thin growth, as well as, chlorosis and desiccation. Some of the 9 plants within each pot died during the experiment, but the amount of death within each pot varied even within the same treatment. Plants were harvested after it was clear that they were suffering from either lack of nutrients, disease, and/or stress from living in a small space with sandy soil and non-natural sunlight and watering (i.e. watering from drip irrigation instead of natural rainfall).

Shoot and root growth

Measurements of shoot and root length, as well as, dry weight biomass, were still taken, despite plant health issues. After drying, shoot and root biomass didn't amount to very much weight overall. Average shoot biomass, in grams, was 1.60, 0.38, 0.55, and 0.33 for AC, AC-pH, NAC, and NAC-pH treatments, respectively. Student's t test results showed a significant difference between AC and AC-pH shoot biomass, as well as, a significant difference between AC and NAC-pH treatments (P = 0.025, P = 0.021, respectively). Significant difference was found between AC and NAC treatments (P = 0.047) shoot biomass. Average root biomass, in grams, was 0.93, 0.78, 0.39, and 0.51 for AC, AC-pH, NAC, and NAC-pH treatments, respectively. Student's t tests showed a significant difference between AC and NAC treatments and borderline difference between AC and NAC-pH treatments (P = 0.016, P = 0.046, respectively). The number of individual plants per treatment that were harvestable at the end of the experiment varied, as many plants desiccated. The number of plants remaining were: 29, 35, 42, and 38 for treatments AC, AC-pH, NAC, and NAC-pH respectively. Average shoot length was highest in NAC-pH (26.62 cm) and lowest in AC-pH (18.64) (Fig. 3). Average root length was highest in AC and lowest in AC-pH (Fig. 4). ANOVA analysis showed no significant difference between treatments for root length (p = 0.50) or shoot length (p = 0.67).

Colonization

Minimal amounts of colonization by AM fungi were observed (Table 2). In NAC-pH, arbuscular colonization ranged from 0-4% and in NAC it ranged from 0-10%. Structures resembling arbuscules were observed one time in AC-1 and not at all in AC-

pH. Hyphae presumed to be from Dark Septate Endophyte (DSE) fungi were observed in small amounts in the AC and AC-pH treatments. Vesicles and AM fungal spores were only present in NAC-pH and not in any other treatment. Hyphae and hyphal coils were more prolific in the NAC treatment. T tests showed no difference in AM fungal hyphal or arbuscular colonization under acidic or more neutral soil conditions ($P = 0.17$ and $P = 0.34$, respectively). Despite lack of significant difference between NAC and NAC-pH with regards to AM fungal colonization, it is clear that the NAC-pH treatments had substantially more colonization by vesicles and spores than did NAC, whereas, NAC had more colonization by AM hyphae and arbuscules (Fig. 5).

Illumina sequencing

For several reasons, we decided not to pursue the AM and other fungal sequence-based communities within the roots of these plants. Firstly, many plants desiccated during the experiment (See Fig. 6 for photograph of pots just before harvest). Secondly, there were not large differences in shoot/root lengths found between + AM and – AM or pH increased and non-pH altered treatments. Lastly, because we saw only small amounts of colonization when observing the roots microscopically, we decided not to delve into potentially misleading or artefactual bioinformatic data retrieved from Illumina sequencing.

Discussion

This greenhouse experiment was the second large scale greenhouse experiment attempted during this dissertation, in addition to many rounds of smaller-scale pilot experiments within the laboratory. The discussion will focus on lessons learned.

In 2015, a full-factorial greenhouse pot experiment was set up. The main questions being addressed were similar to those in this study, namely, whether soil pH and naturally occurring AM fungi impact plant growth. We also were trying to test whether dark septate endophytes (in this case, *Acidomelania panicicola*) work in conjunction with or antagonize the AM fungi. Unfortunately, not enough pilot experimentation was performed prior to the large-scale study, and it had several flaws. The main flaw in the 2015 experiment was that the AM fungi did not infect the switchgrass roots essentially at all. AM fungal spores were picked out of soil slurries by hand, manually, and so there were likely not nearly enough spores to colonize the roots. Also, we didn't consider that many of the spores may have been parasitized or dead. Additionally, by not using whole soil on our 2015 greenhouse project, we unintentionally put a bias on the community of AM fungi we added to the mix because we selected only specific spores. Another issue with picking spores out is that the spores are not the only structure from which the fungus can replicate. Hyphae also make great inocula. By using whole soil for the second greenhouse experiment (the one in this study), we allowed for the entire AM fungal community to be present and have the potential to colonize the switchgrass roots. Also, whole soil includes hyphal fragments which could help increase colonization potential.

In the current experiment, we decreased the amount of replication to make the experiment much more manageable. We also decreased the number of variables, in order to better test more focused hypotheses. We got rid of the DSE portion of the experiment because it was not the focus of our questions. We also got rid of the greenhouse potting mix portion of the experiment. Instead we just altered the soil pH, in the current

experiment, with lime powder. This allowed us to directly test any impacts of soil pH and decreased the number of confounding variables.

Despite improvements made for our current study, the switchgrass plants still did not thrive like we expected them to. Some possible reasons why the switchgrass plants may have not done as well as expected could have been: limited nutrients in the Pine Barrens sandy soil, limited nutrients and space to grow in relatively small containers, overwatering, or disease. Even though greenhouse experiments have as many variables controlled as possible, the aforementioned issues are tough, if not impossible, to completely control. Some work arounds for future studies may include using PVC pipes instead of pots, as personal conversations with other scientists have shown this to be a useful container for growing switchgrass in Pine Barrens soil.

Other pitfalls of this study were possibly insufficient colonization microscopic observations. A lot of variation within the three replicate root samples existed. For instance, 10% colonization by arbuscules was observed in one NAC-pH sample, while 0% was found in a second replicate. Future work using microscopy should probably use more replicates. Secondly, starting the experiment in the natural growing season (spring) would be wise, in order to increase the amount of natural light available to the growing seedlings. The artificial light in the greenhouse may not have been sufficient to sustain photosynthetic activity. Future work should also expect some problems with using naturally sandy, oligotrophic soil. Using fields instead of greenhouse space may improve this experiment. This way, the plants will have access to more of their natural surroundings and, therefore, will be more likely to withstand non-ideal conditions if they

are well adapted to their environment. More AM fungi would be living in field soil in the Pine Barrens, which may increase the rate of colonization in future studies too.

Nevertheless, some interesting findings were seen by the close of this experiment. Even though there was not significant difference between treatments for any type of AM fungal colonization, some general trends were observed. Firstly, NAC plants had more colonization by AM hyphae and arbuscules than NAC-pH plants. This could mean that the increased soil pH in the NAC-pH pots decreased the amount of vegetative growth. This contrasts to other studies which found increased AM fungal colonization under more neutral pH conditions (Ouzounidou et al. 2015). However, different AM fungal species may be present in the Pine Barrens soils compared to those in the experiment of Ouzounidou et al. (2015) and it has been shown that different AM fungal species colonize plant roots at very different rates (Hart and Reader 2002; Jansa et al. 2005). Secondly, NAC-pH plants had more colonization by vesicles and spores, both of which are storage-type structures, compared with NAC plants. This could be because in the pots with more neutral pH, the plants were able to photosynthesize and gain access to nutrients more readily, thereby, permitting the AM fungi to share in the fixed carbon allotment. If more replication were to be done in a future experiment and/or different statistical tests used, statistical differences may begin to appear between treatments, with regards to hyphal, vesicular, and spore colonization of the roots.

Additionally, AC plants had significantly more shoot and root biomass compared to the other treatments, particularly compared to the + AM treatments. While this seems surprising that AM fungi would seem to be inhibiting growth in our study, it is possible that it wasn't the AM fungi that inhibited growth but maybe naturally occurring

pathogens in the whole soil inoculum. Because we used a limited amount of soil for the soil slurry, it is possible that the whole soil used for NAC and NAC-pH pots had much higher quantity of pathogens compared to AC and AC-pH pots, which only received 75 ml of slurry. Also, it is not uncommon for autoclaved soils to have higher levels of accessible nutrients, stemming from the autoclaving process. As seen in Table 1, many important nutrients and CEC were found to be higher in the autoclaved treatment. Although our theories on why AC plants grew larger is not proven, future studies should try to use a much larger quantity of soil slurry just in case. Future studies may also want to utilize microwaving of soil instead of autoclaving to see if this adequately kills off microbiota because this may keep nutrients in recalcitrant form.

Conclusion

Despite best efforts, plants in this experiment did not thrive. This made analysis of AM fungal colonization and community analysis difficult. However, some trends were seen. NAC plants had more colonization by AM hyphae and arbuscules, whereas, NAC-pH plants had more colonization by vesicles and spores. However, neither of these trends were shown to be statistically significant differences. Additional findings showed that AC plants surprisingly gained more shoot and root biomass than did other treatments. More replication and better control over environmental conditions and plant growth may prove beneficial to future experiments looking at impacts of soil pH on AM fungal colonization and diversity and how these factors influence plant growth. Some lessons learned were: you can never do enough pilot experimentation, keeping experiments to a manageable size is a must, and most importantly, never give up!

Table 4-1 Soil properties used for the four soil treatments in this study.

Soil was collected from the New Jersey Pine Barrens. The four treatments were: autoclaved whole soil, autoclaved whole soil with increased soil pH, non-autoclaved whole soil, and non-autoclaved whole soil with increased pH. Exchangeable Al (Exch. Al), P, K, Al, NO3, and Ca were extracted by Mehlich-3 (ICP) and are reported in ppm. Total N (N), Organic Matter are reported as percentages. Cation Exchange Capacity (CEC) is reported as meq/100 g soil. Soil measurements were taken by Spectrum Analytic (Washington Courte House, OH). Treatments were: AC = Autoclaved soil, AC-pH = Autoclaved soil + pH, NAC = non-autoclaved soil, NAC-pH = non-autoclaved soil + pH.

Treatment	Soil pH	Organic Matter	P	K	N	NO3	Exch. Al	CEC	Ca
AC	6.5	0.1	21	11	0.21	3	1.3	2.6	389
AC-pH	7.0	0.1	18	5	0.25	2	1.1	1.4	237
NAC	6.2	0.1	14	5	0.25	2	2.3	0.9	175
NAC-pH	6.9	0.1	21	7	0.32	3	1.4	1.5	278

Table 4-2 Percent arbuscular mycorrhizal (AM) fungal colonization of switchgrass roots, observed via microscopy.

Structures observed include: AM hyphae, hyphal coils, vesicles, AM fungal spores, and arbuscules. Treatments were: AC = Autoclaved soil, AC-pH = Autoclaved soil + pH, NAC = non-autoclaved soil, NAC-pH = non-autoclaved soil + pH.

Treatment	AM Hyphae	Coils	Vesicles	Spores	Arbuscules
AC-1	1	0	0	0	0
AC-2	3	0	0	0	1
AC-3	0	0	0	0	0
AC-pH-1	0	0	0	0	0
AC-pH-2	1	0	0	0	0
AC-pH-3	1	0	0	0	0
NAC-1	13	10	0	0	10
NAC-2	5	2	0	0	0
NAC-3	18	6	0	0	5
NAC-pH-1	6	6	3	2	4
NAC-pH-2	5	5	22	2	0
NAC-pH-3	2	3	9	0	0

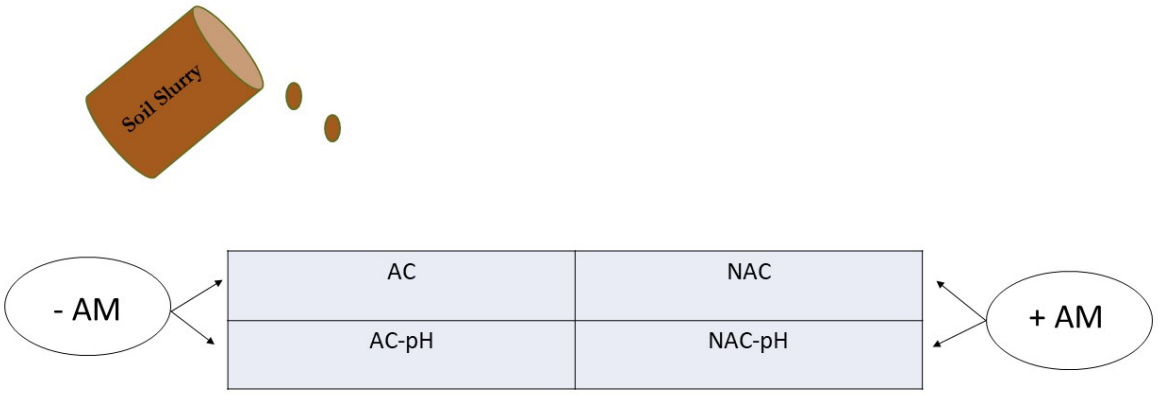

Fig. 4-1 Four soil + fungal treatments used in this study.

Treatments were: AC = Autoclaved soil, AC-pH = Autoclaved soil + pH, NAC = non-autoclaved soil, NAC-pH = non-autoclaved soil + pH.

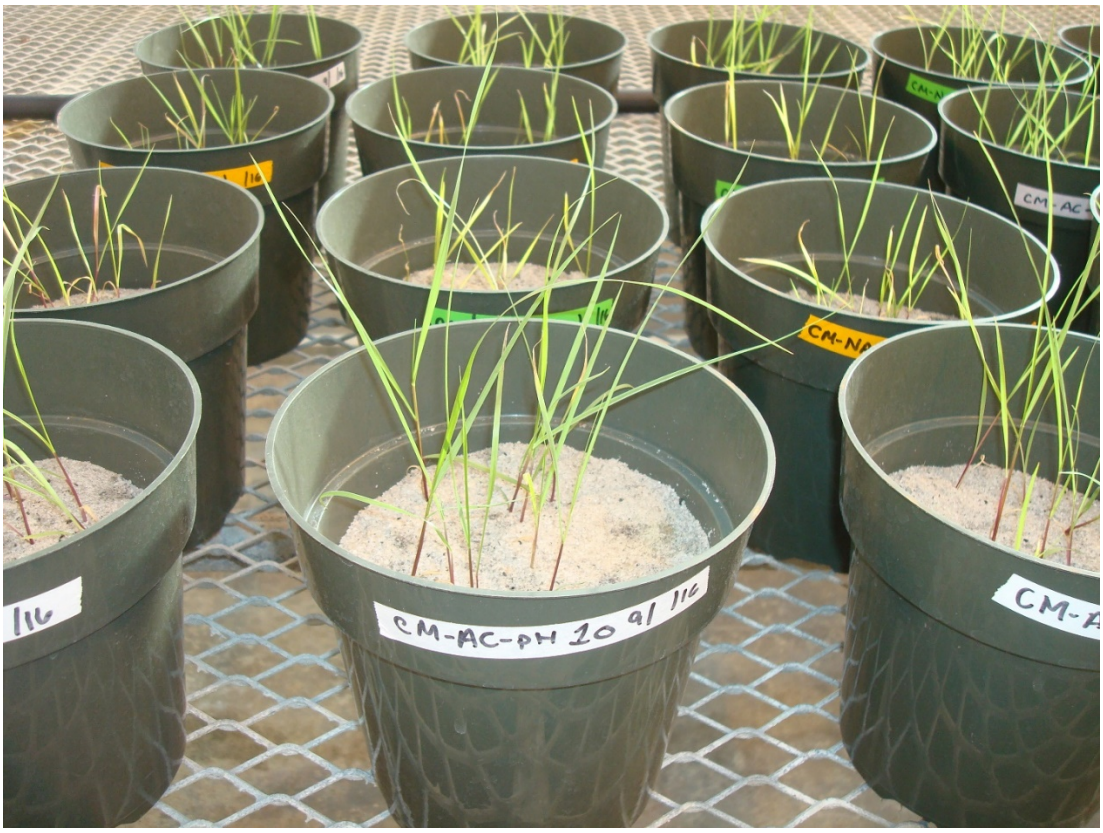

Fig. 4-2 Photograph of switchgrass pots at six weeks old.

Switchgrass shoots can be seen, along with the sandy, acidic soil from the New Jersey Pine Barrens, used in this experiment.

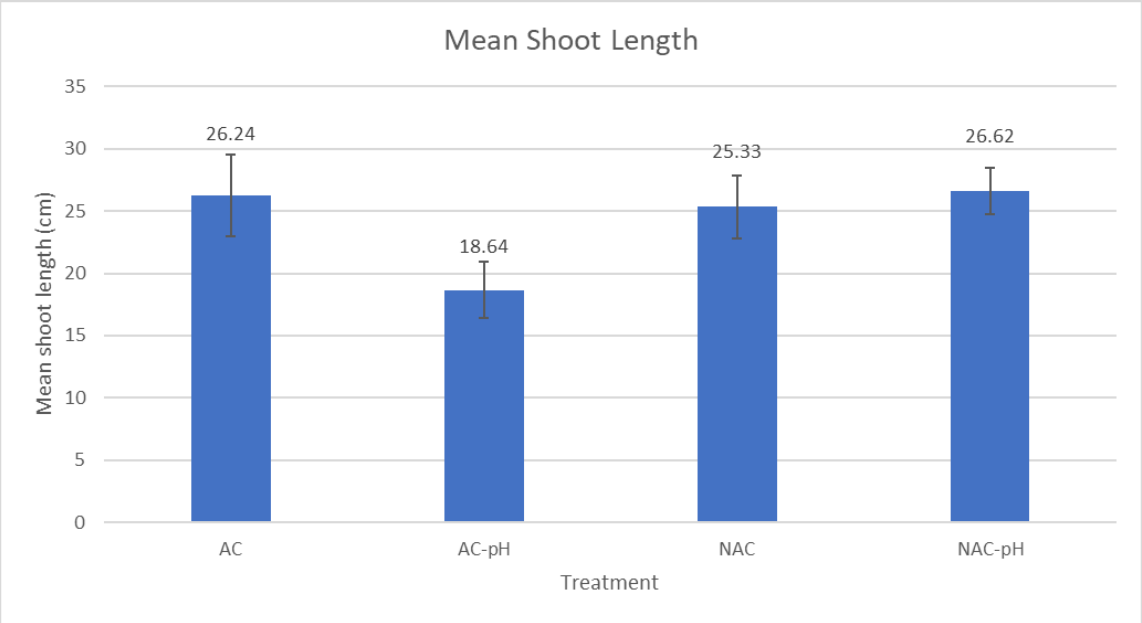

Fig. 4-3 Mean shoot length (cm) of the shoots from each of four experimental treatments.

Treatments were AC = Autoclaved soil, AC-pH = Autoclaved soil + pH, NAC = non-autoclaved soil, NAC-pH = non-autoclaved soil + pH. Error bars are 1 standard error above and below the mean shoot length. Numbers above bars are the mean shoot length values per treatment.

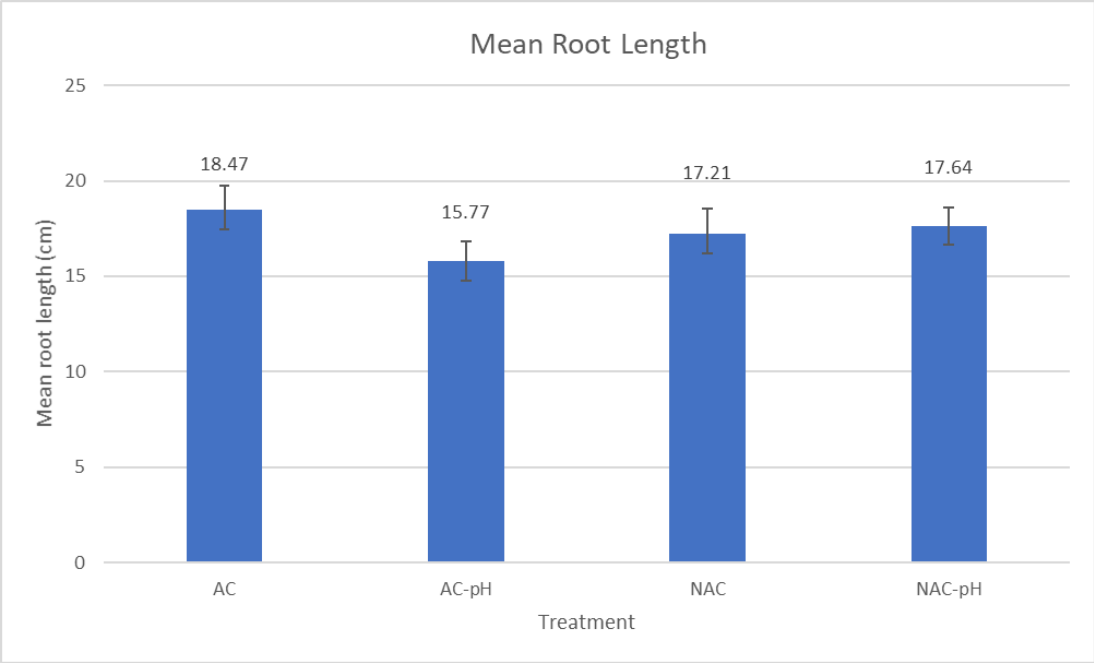

Fig. 4-4 Mean root length (cm) of the roots from each of four experimental treatments.

Treatments were AC = Autoclaved soil, AC-pH = Autoclaved soil + pH, NAC = non-autoclaved soil, NAC-pH = non-autoclaved soil + pH. Error bars are 1 standard error above and below the mean root length. Numbers above bars are the mean root length values per treatment.

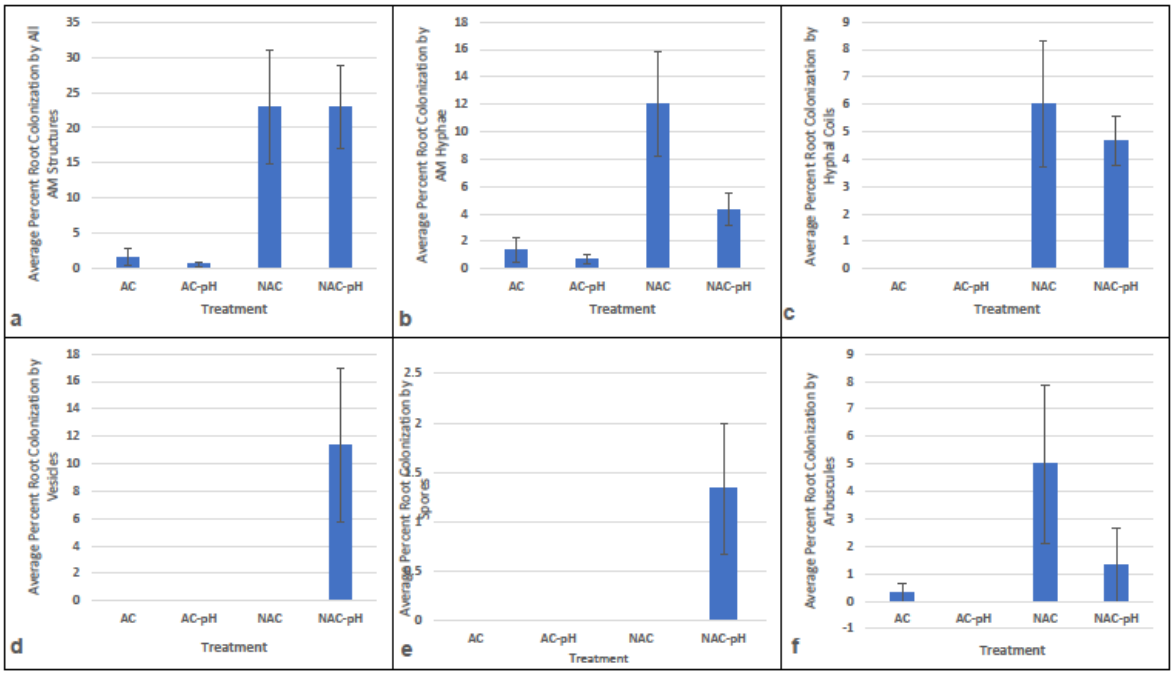

Fig. 4-5 Average percent root colonization by various AM fungal structures.

Colonization by all AM structures (a), AM hyphae (b), hyphal coils (c), vesicles (d), AM fungal spores (e), and arbuscules (f) are shown. Error bars represent 1 standard error.

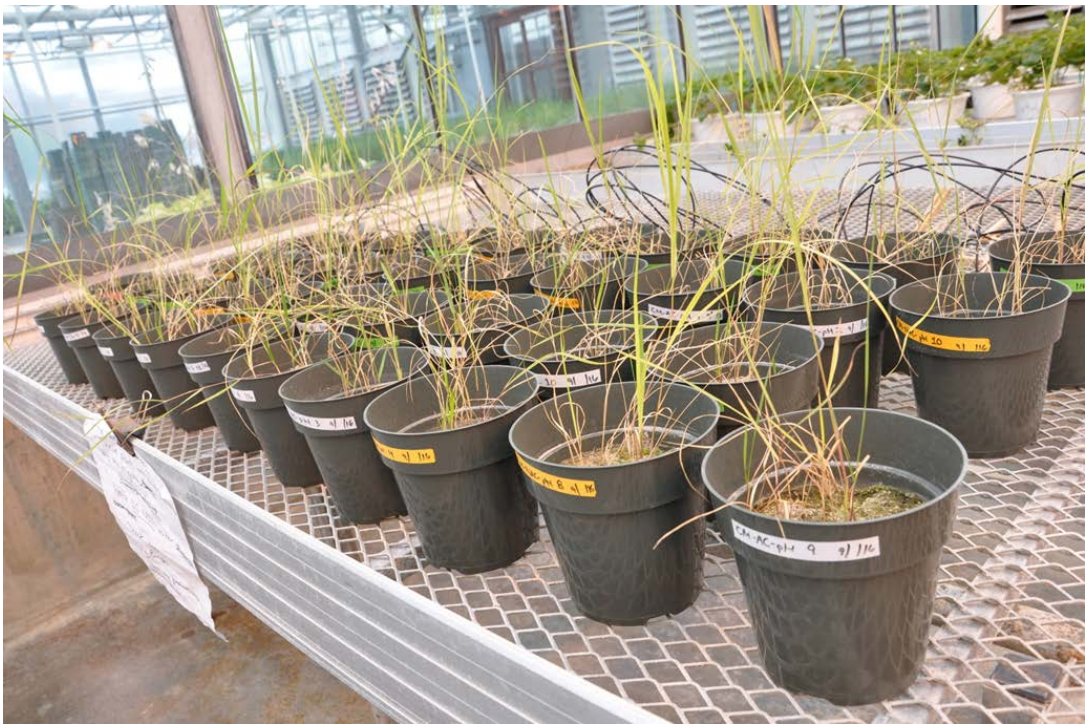

Fig. 4-6 Photograph of all switchgrass pots before harvest.

Chapter 5

Summary

This dissertation aimed to uncover the diversity of AM fungi associated with switchgrass populations in the Pine Barrens forests of New Jersey. It also aimed to compare the AM fungal communities from the Pine Barrens of New Jersey to those of Long Island, New York, as well as, contrast these communities with those from managed agricultural fields. Many studies attempt to elucidate the reasons why AM fungi exist in certain places and not in others. Despite the fact that there is a plethora of studies and data indicating which factors contribute most towards AM fungal diversity, the main drivers of AM fungal diversity remain nebulous. This dissertation aimed to help close this knowledge gap by surveying AM fungal communities from different sites and determining whether soil properties or land use influence AM fungal diversity in those sites.

Understanding what influences AM fungal diversity is important for many reasons. AM fungi associate with most terrestrial plants and provide them with otherwise inaccessible nutrients. This mycorrhizal relationship is particularly essential for plants living in oligotrophic conditions. Humans and other animals rely on plants for food and for oxygen to breathe: two absolutely essential parts of life. Therefore, in order to ensure ample healthy plants on our planet, we must ensure ample AM fungal and other microbial communities to help sustain plant populations.

The last decade has seen many advances in the ways we discover and define AM fungal species. Instead of just morphological descriptions, NGS technology has

revolutionized the sequence-based approach to AM fungal species delineation. This dissertation aimed to test commonly used bioinformatic workflow decisions in AM fungal studies to see whether making various workflow decisions, such as, using alternate reference databases for DNA queries, would alter the resulting AM fungal diversity. Hopefully some of the findings herein will aid future AM fungal diversity studies that use a NGS approach to species identification.

Lastly, in order to best understand why AM fungi exist in certain places and not in others, it is essential to learn about how they function in nature. Many questions regarding AM fungal function and diversity linger, such as: Which AM fungi are most important in crop production? Which are most tolerant of soil pH fluctuations? How many different species of AM fungi are needed to aid in plant production and pathogen protection? How do certain AM fungi improve plant health while others do not? Do AM fungi interact or compete with pathogens in plant roots and soil? Do endemic species have a purpose outside their native habitat? Are there implications for conservation efforts? Hopefully this dissertation helps others gain footing to address some of these vital questions.

Literature Cited

Abbott LK, Robson AD (1991) Factors influencing the occurrence of vesicular arbuscular mycorrhizas. Agric Ecosyst Environ 35:121-150 doi:10.1016/0167-8809(91)90048-3

Adler PR, Del Grosso SJ, Parton WJ (2007) Life-cycle assessment of net greenhouse-gas flux for bioenergy cropping systems. Ecol Appl 17:675-691

Agriculture USDo Conservation Plant Characteristics for *Panicum virgatum*.

Aguilera P, Cumming J, Oehl F, Cornejo P, Borie F (2015) Diversity of Arbuscular Mycorrhizal Fungi in Acidic Soils and Their Contribution to Aluminum Phytotoxicity Alleviation. In: Panda S., Baluška F. (eds) Aluminum Stress Adaptation in Plants. Signaling and Communication in Plants, vol 24. Springer, Cham. 24:203-228 doi:10.1007/978-3-319-19968-9_11

Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K, Mahon A (2018) Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods Ecol Evol 9:134-147 doi:10.1111/2041-210x.12849

Amijee F, Tinker PB, Stribley DP (1989) The development of endomycorrhizal root systems: detailed study of effects of soil-phosphorus on colonization. New Phytol 111:435-446 doi:10.1111/j.1469-8137.1989.tb00706.x

An G, Miyakawa S, Kawahara A, Osaki M, Ezawa T (2008) Community structure of arbuscular mycorrhizal fungi associated with pioneer grass species *Miscanthus sinensis* in acid sulfate soils: habitat segregation along pH gradients. Soil Sci Plant Nutr 54:517-528

Appoloni S, Lekberg Y, Tercek MT, Zabinski CA, Redecker D (2008) Molecular community analysis of arbuscular mycorrhizal fungi in roots of geothermal soils in Yellowstone National Park (USA). Microb Ecol 56:649-659 doi:10.1007/s00248-008-9384-9

Arnold AE, Lutzoni F (2007) Diversity and host range of foliar fungal endophytes: Are tropical leaves biodiversity hotspots? Ecology 88:541-549 doi:10.1890/05-1459

Ashworth AJ, Taylor AM, Reed DL, Allen FL, Keyser PD, Tyler DD (2015) Environmental impact assessment of regional switchgrass feedstock production comparing nitrogen input scenarios and legume-intercropping systems. J of Clean Prod 87:227-234 doi:10.1016/j.jclepro.2014.10.002

Ashworth AJ et al. (2016) Switchgrass composition and yield response to alternative soil amendments under intensified heat and drought conditions. Agr Ecosyst Environ 233:415-424 doi:10.1016/j.agee.2016.09.041

Auge RM (2001) Water relations, drought and vesicular-arbuscular mycorrhizal symbiosis. Mycorrhiza 11:3-42 doi:10.1007/s005720100097

Bago B, Pfeffer PE, Shachar-Hill Y (2000) Carbon metabolism and transport in arbuscular mycorrhizas. Plant Physiol 124:949-958

Bailly J, Fraissinet-Tachet L, Verner MC, Debaud JC, Lemaire M, Wesolowski-Louvel M, Marmeisse R (2007) Soil eukaryotic functional diversity, a metatranscriptomic approach. ISME J 1:632-642 doi:10.1038/ismej.2007.68

Ban YH, Jiang YH, Li M, Zhang XL, Zhang SY, Wu Y, Xu ZY (2017) Homogenous stands of a wetland grass living in heavy metal polluted wetlands harbor diverse consortia of arbuscular mycorrhizal fungi. Chemosphere 181:699-709

Baon JB, Smith SE, Alston AM, Wheeler RD (1992) Phosphorus efficiency of 3 cereals as related to indigenous mycorrhizal infection. Aust J Agr Res 43:479-491 doi:10.1071/ar9920479

Barinaga M (1997) Plant science - Making plants aluminum tolerant. Science 276:1497-1497

Barnes N (2010) Publish your computer code: it is good enough. Nature 467:753-753 doi:10.1038/467753a

Behjati S, Tarpey PS (2013) What is next generation sequencing? Arch Dis Child Educ Pract Ed 98:236-238 doi:10.1136/archdischild-2013-304340

Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, Kauserud H (2010) ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. Bmc Microbiology 10 doi:10.1186/1471-2180-10-189

Bennett S (2004) Solexa Ltd Pharmacogenomics 5:433-438 doi:10.1517/14622416.5.4.433

Bever JD, Morton JB, Antonovics J, Schultz PA (1996) Host-dependent sporulation and species diversity of arbuscular mycorrhizal fungi in a mown grassland. J Ecol 84:71-82 doi:10.2307/2261701

Biel ER (1958) The climate of New Jersey. In: Flink SJ (ed) The economy of New Jersey. Rutgers University Press, New Brunswick, NJ, pp 53-98

Blackwell M (2011) The fungi: 1, 2, 3 ... 5.1 million species? Am J Bot 98:426-438 doi:10.3732/ajb.1000298

Blaszkowski J et al. (2013) *Septoglomus fuscum* and *S. furcatum*, two new species of arbuscular mycorrhizal fungi (Glomeromycota). Mycologia 105:670-680 doi:10.3852/12-127

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

Borie F, Rubio R (1999) Effects of arbuscular mycorrhizae and liming on growth and mineral acquisition of aluminum-tolerant and aluminum-sensitive barley cultivars. J Plant Nutr 22:121-137

Bouffaud M-L et al. (2016) Indicator species and co-occurrence in communities of arbuscular mycorrhizal fungi at the European scale. Soil Biol Biochem 103:464-470 doi:10.1016/j.soilbio.2016.09.022

Brejda JJ, Moore K, Anderson B (2000) Fertilization of native warm-season grasses. Native warm-season grasses: research trends and issues. Proceedings of the Native Warm-Season Grass Conference and Expo, Des Moines, IA, USA, 12-13 September 1996.

Brock PM, Doring H, Bidartondo MI (2009) How to know unknown fungi: the role of a herbarium. New Phytol 181:719-724 doi:10.1111/j.1469-8137.2008.02703.x

Brown JH, Maurer BA (1989) Macroecology: the division of food and space among species on continents. Science 243:1145-1150

Bruce A, Smith SE, Tester M (1994) The development of mycorrhival infection in cucumber: effects of P-supply on root growth, formation of entry points and growth of infection units. New Phytol 127:507-514 doi:10.1111/j.1469-8137.1994.tb03968.x

Bruns TD, Corradi N, Redecker D, Taylor JW, Öpik M (2017) Glomeromycotina: what is a species and why should we care? New Phytol

Bushnell B (2016) BBMap short read aligner University of California, Berkeley, California URL http://sourceforge.net/projects/bbmap

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421 doi:10.1186/1471-2105-10-421

Cao J, Feng Y, Lin X, Wang J, Xie X (2016) Iron oxide magnetic nanoparticles deteriorate the mutual interaction between arbuscular mycorrhizal fungi and plant. J Soil Sediment 17:841-851

Caporaso JG et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335-336 doi:10.1038/nmeth.f.303

Castillo CG, Borie F, Godoy R, Rubio R, Sieverding E (2006) Diversity of mycorrhizal plant species and arbuscular mycorrhizal fungi in evergreen forest, deciduous forest and grassland ecosystems of Southern Chile. Journal of Applied Botany and Food Quality-Angewandte Botanik 80:40-47

Central Pine Barrens Overview. (2018) Central Pine Barrens Joint Planning & Policy Commission. https://pb.state.ny.us/central-pine-barrens/overview/. Accessed February 6, 2018 2018

Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, Ellison AM (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecol Monogr 84:45-67

Chaudhary VB, Cuenca G, Johnson NC (2018) Tropical-temperate comparison of landscape-scale arbuscular mycorrhizal fungal species distributions. Divers Distrib 24:116-128 doi:10.1111/ddi.12664

Ciccolini V, Ercoli L, Davison J, Vasar M, Opik M, Pellegrino E (2016) Land-use intensity and host plant simultaneously shape the composition of arbuscular mycorrhizal fungal communities in a Mediterranean drained peaMtland. FEMS Microbiol Ecol 92

Clapp JP, Fitter AH, Young JPW (1999) Ribosomal small subunit sequence variation within spores of an arbuscular mycorrhizal fungus, *Scutellospora* sp. Mol Ecol 8:915-921 doi:10.1046/j.1365-294x.1999.00642.x

Clark RB (1997) Arbuscular mycorrhizal adaptation, spore germination, root colonization, and host plant growth and mineral acquisition at low pH. Plant Soil 192:15-22

Clark RB (2002) Differences among mycorrhizal fungi for mineral uptake per root length of switchgrass grown in acidic soil. J Plant Nutr 25:1753-1772

Clark RB, Baligar VC, Zobel RW (2005) Response of mycorrhizal switchgrass to phosphorus fractions in acidic soil. Commun Soil Sci Plan 36:1337-1359

Clark RB, Zobel RW, Zeto SK (1999) Effects of mycorrhizal fungus isolates on mineral acquisition by *Panicum virgatum* in acidic soil. Mycorrhiza 9:167-176

CLC Genomics Workbench. (2017). https://www.qiagenbioinformatics.com/.

Commission P (2015) Pinelands Maps. State of New Jersey. http://www.nj.gov/pinelands/home/maps/. Accessed September 27, 2017 2017

Conrad R (1996) Soil microorganisms as controllers of atmospheric trace gases (H-2, CO, CH4, OCS, N2O, and NO). Microbiol Rev 60:609-640

Cui X, Hu J, Wang J, Yang J, Lin X (2016) Reclamation negatively influences arbuscular mycorrhizal fungal community structure and diversity in coastal saline-alkaline

land in Eastern China as revealed by Illumina sequencing. Appl Soil Ecol 98:140-149

Cumming JR, Ning J (2003) Arbuscular mycorrhizal fungi enhance aluminium resistance of broomsedge (*Andropogon virginicus* L.). J Exp Bot 54:1447-1459

Cuomo GJ, Anderson BE, Young LJ (1998) Harvest frequency and burning effects on vigor of native grasses. J Range Manage 51:32-36 doi:10.2307/4003560

Dai M, Bainard LD, Hamel C, Gan Y, Lynch D (2013) Impact of land use on arbuscular mycorrhizal fungal communities in rural Canada. Appl Environ Microbiol 79:6719-6729 doi:10.1128/AEM.01333-13

Daniell TJ, Husband R, Fitter AH, Young JPW (2001) Molecular diversity of arbuscular mycorrhizal fungi colonising arable crops. FEMS Microbiol Ecol 36:203-209

Davison J, Opik M, Zobel M, Vasar M, Metsis M, Moora M (2012) Communities of arbuscular mycorrhizal fungi detected in forest soil are spatially heterogeneous but do not vary throughout the growing season. PLoS One 7:e41938 doi:10.1371/journal.pone.0041938

de Carcer DA, Denman SE, McSweeney C, Morrison M (2011) Evaluation of subsampling-based normalization strategies for tagged high-throughput sequencing data sets from gut microbiomes. Appl Environ Microbiol 77:8795-8798

Debiane D, Calonne M, Fontaine J, Laruelle F, Grandmougin-Ferjani A, Lounes-Hadj Sahraoui A (2011) Lipid content disturbance in the arbuscular mycorrhizal, *Glomus irregulare* grown in monoxenic conditions under PAHs pollution. Fungal Biol 115:782-792 doi:10.1016/j.funbio.2011.06.003

DeLong EF et al. (2006) Community genomics among stratified microbial assemblages in the ocean's interior. Science 311:496-503 doi:10.1126/science.1120250

Diagne N et al. (2011) Uvitex2B: a rapid and efficient stain for detection of arbuscular mycorrhizal fungi within plant roots. Mycorrhiza 21:315-321 doi:10.1007/s00572-010-0357-8

Diez B, Pedros-Alio C, Massana R (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. Appl Environ Microbiol 67:2932-2941 doi:10.1128/AEM.67.7.2932-2941.2001

Dighton J (2016) 2nd ed. Fungi in ecosystem processes, vol. 31CRC Press, Boca Raton, USA

Dighton J, Gordon T, Mejia R, Sobel M (2013) Mycorrhizal status of Knieskern's beaked sedge (*Rhynchospora knieskernii*) in the New Jersey Pine Barrens. Bartonia:24-27

Dimitrov MR, Veraart AJ, de Hollander M, Smidt H, van Veen JA, Kuramae EE (2017) Successive DNA extractions improve characterization of soil microbial communities. PeerJ 5:e2915

Douds DD, Jr. (2002) Increased spore production by *Glomus intraradices* in the split-plate monoxenic culture system by repeated harvest, gel replacement, and resupply of glucose to the mycorrhiza. Mycorrhiza 12:163-167 doi:10.1007/s00572-002-0174-9

Douglas LA, Trela JJ (1979) Morphology of Pine Barrens soils. In: Forman RTT (ed) Pine Barrens Ecosystem and Landscape. Academic Presss, New York, NY,

Dumbrell AJ et al. (2011) Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. New Phytol 190:794-804 doi:10.1111/j.1469-8137.2010.03636.x

Edgar R (2016) UCHIME2: improved chimera prediction for amplicon sequencing bioRxiv doi:10.1101/074252

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797

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

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics (Oxford, England) 27:2194-2200 doi:10.1093/bioinformatics/btr381

Egan CP, Callaway RM, Hart MM, Pither J, Klironomos J (2017) Phylogenetic structure of arbuscular mycorrhizal fungal communities along an elevation gradient. Mycorrhiza 27:273-282

Egerton-Warburton LM, Johnson NC, Allen EB (2007) Mycorrhizal community dynamics following nitrogen fertilization: a cross-site test in five grasslands. Ecol Monogr 77:527-544 doi:10.1890/06-1772.1

Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P (2010) Experimental factors affecting PCR-based estimates of microbial species richness and evenness. ISME J 4:642-647 doi:10.1038/ismej.2009.153

Eom AH, Wilson GWT, Hartnett DC (2001) Effects of ungulate grazers on arbuscular mycorrhizal symbiosis and fungal community structure in tallgrass prairie. Mycologia 93:233-242 doi:10.2307/3761643

Fajardo L, Loveral M, Arrindell P, Aguilar VH, Hasmy Z, Cuenca G (2015) Morphotype-based characterization of arbuscular mycorrhizal fungal communities in a restored tropical dry forest, Margarita Island, Venezuela. Rev Biol Trop 63:859-870

Field KJ et al. (2015) First evidence of mutualism between ancient plant lineages (Haplomitriopsida liverworts) and Mucoromycotina fungi and its response to simulated Palaeozoic changes in atmospheric CO2. New Phytol 205:743-756 doi:10.1111/nph.13024

Forman RTT (1998) Pine Barrens: ecosystems and landscape. Rutgers University Press, New Brunswick, NJ

French KE, Tkacz A, Turnbull LA (2017) Conversion of grassland to arable decreases microbial diversity and alters community composition. Appl Soil Ecol 110:43-52 doi:10.1016/j.apsoil.2016.10.015

Gamalero E, Trotta A, Massa N, Copetta A, Martinotti MG, Berta G (2004) Impact of two fluorescent pseudomonads and an arbuscular mycorrhizal fungus on tomato plant growth, root architecture and P acquisition. Mycorrhiza 14:185-192 doi:10.1007/s00572-003-0256-3

Gamble JF (1963) A study of strontium, barium, and calcium relationships in soils and vegetation. AEC Final Report, New Brunswick, NJ

Gamper HA, Walker C, Schüßler A (2009) *Diversispora celata* sp. nov: molecular ecology and phylotaxonomy of an inconspicuous arbuscular mycorrhizal fungus. New Phytol 182:495-506 doi:10.1111/j.1469-8137.2008.02750.x

Gange AC, Bower E, Stagg PG, Aplin DM, Gillam AE, Bracken M (1999) A comparison of visualization techniques for recording arbuscular mycorrhizal colonization. New Phytol 142:123-132

Garcia de Leon D et al. (2018) Anthropogenic disturbance equalizes diversity levels in arbuscular mycorrhizal fungal communities. Glob Chang Biol 24:2649-2659 doi:10.1111/gcb.14131

Gehrig H, Schüßler A, Kluge M (1996) *Geosiphon pyriforme*, a fungus forming endocytobiosis with *Nostoc* (Cyanobacteria), is an ancestral member of the Glomales: evidence by SSU rRNA analysis. J Mol Evol 43:71-81 doi:10.1007/bf02352301

Geller M (2002) Why the Pine Barrens are barren. In: Geller M (ed) A key to the woody plants of the New Jersey Pine Barrens. Rutgers University Press, New Brunswick, NJ,

Gerdemann JW (1955) Relation of a large soil-borne spore to phycomycetous mycorrhizal infections. Mycologia 47:619-632

Gerdemann JW (1968) Vesicular-arbuscular mycorrhiza and plant growth. Annu Rev Phytopathol 6:397-418

Gerz M, Bueno CG, Zobel M, Moora M (2016) Plant community mycorrhization in temperate forests and grasslands: relations with edaphic properties and plant diversity. J Veg Sci 27:89-99 doi:10.1111/jvs.12338©

Ghimire SR, Charlton ND, Craven KD (2009) The mycorrhizal fungus, *Sebacina vermifera*, enhances seed germination and biomass production in switchgrass (*Panicum virgatum* L). BioEnergy Res 2:51-58 doi:10.1007/s12155-009-9033-2

Gianinazzi S, Gollotte A, Binet MN, van Tuinen D, Redecker D, Wipf D (2010) Agroecology: the key role of arbuscular mycorrhizas in ecosystem services. Mycorrhiza 20:519-530 doi:10.1007/s00572-010-0333-3

Gilmore AE (1968) Phycomycetous mycorrhizal organisms collected by open-pot culture methods. Hilgardia 39:87-105

Giovannetti M, Mosse B (1980) Evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots. New Phytol 84:489-500 doi:10.1111/j.1469-8137.1980.tb04556.x

Gomez-Alvarez V, Teal TK, Schmidt TM (2009) Systematic artifacts in metagenomes from complex microbial communities. ISME J 3:1314-1317 doi:10.1038/ismej.2009.72

Gosling P, Mead A, Proctor M, Hammond JP, Bending GD (2013) Contrasting arbuscular mycorrhizal communities colonizing different host plants show a similar response to a soil phosphorus concentration gradient. New Phytol 198:546-556

Grace C, Stribley DP (1991) A safer procedure for routine staining of vesicular-arbuscular mycorrhizal fungi. Mycol Res 95:1160-1162

Graw D, Moawad M, Rehm S (1979) Untersuchungen zur wirts und wirkungs spezifitat der VA mykorrhiza. Zeitschrift fur Aker und Pflanzenbau 148:85-98

Grube M et al. (2017) The next generation fungal diversity researcher. Fungal Biol-UK Reviews 31:124-130 doi:10.1016/j.fbr.2017.02.001

Hall KE, George JR, Riedl RR (1982) Herbage dry-matter yields of switchgrass, big bluestem, and indiangrass with N-fertilization. Agron J 74:47-51

Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 5:R245-249

Harrell Jr FE (2017) Package 'Hmisc'

Harrier LA, Watson CA (2004) The potential role of arbuscular mycorrhizal (AM) fungi in the bioprotection of plants against soil-borne pathogens in organic and/or other sustainable farming systems. Pest Manag Sci 60:149-157 doi:10.1002/ps.820

Harris JK, Sahl JW, Castoe TA, Wagner BD, Pollock DD, Spear JR (2010) Comparison of normalization methods for construction of large, multiplex amplicon pools for next-generation sequencing. Appl Environ Microbiol 76:3863-3868

Hart MM et al. (2015) Navigating the labyrinth: a guide to sequence-based, community ecology of arbuscular mycorrhizal fungi. New Phytol 207:235-247

Hart MM, Reader RJ (2002) Taxonomic basis for variation in the colonization strategy of arbuscular mycorrhizal fungi. New Phytol 153:335-344

Hartnett DC, Samenus RJ, Fischer LE, Hetrick BAD (1994) Plant demographic responses to mycorrhizal symbiosis in tallgrass prairie. Oecologia 99:21-26 doi:10.1007/bf00317079

Hartnett DC, Wilson GWT (1999) Mycorrhizae influence plant community structure and diversity in tallgrass prairie. Ecology 80:1187-1195 doi:10.1890/0012-9658(1999)080[1187:mipcsa]2.0.co;2

Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. Mycol Res 105:1422-1432 doi:Doi 10.1017/S0953756201004725

Hawksworth DL, Rossman AY (1997) Where are all the undescribed fungi? Phytopathology 87:888-891 doi:Doi 10.1094/Phyto.1997.87.9.888

Hayman DS (1974) Plant growth responses to vesicular-arbuscular mycorrhiza. VI. Effect of light and temperature New Phytol 73:71-80

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001) Molecular evidence for the early colonization of land by fungi and plants. Science 293:1129-1133 doi:10.1126/science.1061457

Helgason T, Daniell TJ, Husband R, Fitter AH, Young JPW (1998) Ploughing up the wood-wide web? Nature 394:431-431

Herrmann L et al. (2016) Diversity of root-associated arbuscular mycorrhizal fungal communities in a rubber tree plantation chronosequence in Northeast Thailand. Mycorrhiza 26:863-877 doi:10.1007/s00572-016-0720-5

Hetrick BA (1989) Acquisition of phosphorus by VA mycorrhizal fungi and the growth responses of their host plants. In: Boddy L, Marchant R, Reid DJ (eds) Nitrogen, phosphorus and sulphur utilization by fungi. Cambridge University Press, New York, pp 205-226

Hiiesalu I et al. (2014) Species richness of arbuscular mycorrhizal fungi: associations with grassland plant richness and biomass. New Phytol 203:233-244 doi:10.1111/nph.12765

Hijri I, Sykorova Z, Oehl F, Ineichen K, Mader P, Wiemken A, Redecker D (2006) Communities of arbuscular mycorrhizal fungi in arable soils are not necessarily low in diversity. Mol Ecol 15:2277-2289

Hijri M, Sanders IR (2005) Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. Nature 433:160 doi:10.1038/nature03069

https://www.nature.com/articles/nature03069#supplementary-information

Hitchcock AS (1935) Manual of the grasses of the United States. Washington, D. C.

Hoeksema JD et al. (2010) A meta-analysis of context-dependency in plant response to inoculation with mycorrhizal fungi. Ecol Lett 13:394-407 doi:10.1111/j.1461-0248.2009.01430.x

House GL et al. (2016) Phylogenetically structured differences in rRNA gene sequence variation among species of arbuscular mycorrhizal fungi and their implications for sequence clustering. Appl Environ Microbiol 82:4921-4930 doi:10.1128/aem.00816-16

Huson DH et al. (2016) MEGAN Community Edition - interactive exploration and analysis of large-scale microbiome sequencing data, PLoS Comput Biol 12:e1004957 doi:10.1371/journal.pcbi.1004957

Iqbal MT (2012) Acid tolerance mechanisms in soil grown plants. Malaysian Journal of Soil Science 16:21

James TY et al. (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443:818 doi:10.1038/nature05110

https://www.nature.com/articles/nature05110#supplementary-information

Jansa J, Mozafar A, Anken T, Ruh R, Sanders IR, Frossard E (2002) Diversity and structure of AMF communities as affected by tillage in a temperate soil. Mycorrhiza 12:225-234 doi:10.1007/s005572-002-0163-z

Jansa J, Mozafar A, Frossard E (2005) Phosphorus acquisition strategies within arbuscular mycorrhizal fungal community of a single field site. Plant Soil 276:163-176

Jansa J, Mozafar A, Kuhn G, Anken T, Ruh R, Sanders IR, Frossard E (2003) Soil tillage affects the community structure of mycorrhizal fungi in maize roots. Ecol Appl 13:1164-1176

Jansa J, Smith FA, Smith SE (2008) Are there benefits of simultaneous root colonization by different arbuscular mycorrhizal fungi? New Phytol 177:779-789 doi:10.1111/j.1469-8137.2007.02294.x

Ji B, Gehring CA, Wilson GW, Miller RM, Flores-Renteria L, Johnson NC (2013) Patterns of diversity and adaptation in Glomeromycota from three prairie grasslands. Mol Ecol 22:2573-2587

Johansen RB, Johnston P, Mieczkowski P, Perry GLW, Robeson MS, Burns BR, Vilgalys R (2016) A native and an invasive dune grass share similar, patchily distributed, root-associated fungal communities. Fungal Ecol 23:141-155

Johnson NC (1993) Can fertilization of soil select less mutualistic mycorrhizae? Ecol Appl 3:749-757 doi:10.2307/1942106

Johnson NC (2010) Resource stoichiometry elucidates the structure and function of arbuscular mycorrhizas across scales. New Phytol 185:631-647

Johnson NC, Angelard C, Sanders IR, Kiers ET (2013) Predicting community and ecosystem outcomes of mycorrhizal responses to global change. Ecol Lett 16 Suppl 1:140-153

Johnson NC, Wilson GW, Bowker MA, Wilson JA, Miller RM (2010) Resource limitation is a driver of local adaptation in mycorrhizal symbioses. Proc Natl Acad Sci U S A 107:2093-2098

Johnson NC, Zak DR, Tilman D, Pfleger FL (1991) Dynamics of vesicular-arbuscular mycorrhizae during old field succession. Oecologia 86:349-358 doi:10.1007/bf00317600

Jonsson LM, Dighton J, Lussenhop J, Koide RT (2006) The effect of mixing ground leaf litters to soil on the development of pitch pine ectomycorrhizal and soil arthropod communities in natural soil microcosm systems. Soil Biol Biochem 38:134-144

Jumpponen A, Trowbridge J, Mandyam K, Johnson L (2005) Nitrogen enrichment causes minimal changes in arbuscular mycorrhizal colonization but shifts community composition-evidence from rDNA data. Biol Fert Soils 41:217-224

Jung GA, Shaffer JA, Stout WL (1988) Switchgrass and big bluestem responses to amendments on strongly acid soil. Agron J 80:669-676

Karandashov VE, Kuzovkina IN, George E, Marschner H (1999) Monoxenic culture of arbuscular mycorrhizal fungi and plant hairy roots. Russion Journal of Plant Physiol 46:87-92

Karl TR, Melillo JM, Peterson TC (2009) Global climate change impacts in the US. Cambridge University Press, New York, NY

Kawahara A, An GH, Miyakawa S, Sonoda J, Ezawa T (2016) Nestedness in arbuscular mycorrhizal fungal communities along soil pH gradients in early primary succession: acid-tolerant fungi are pH generalists. PLoS One 11:e0165035 doi:10.1371/journal.pone.0165035

Keith SA et al. (2012) What is macroecology? Biol Lett 8:904-906 doi:10.1098/rsbl.2012.0672

Kelly CN, Morton JB, Cumming JR (2005) Variation in aluminum resistance among arbuscular mycorrhizal fungi. Mycorrhiza 15:193-201

Kivlin SN, Hawkes CV, Treseder KK (2011) Global diversity and distribution of arbuscular mycorrhizal fungi. Soil Biol Biochem 43:2294-2303

Klabi R, Bell TH, Hamel C, Iwaasa A, Schellenberg M, Raies A, St-Arnaud M (2015) Plant assemblage composition and soil P concentration differentially affect communities of AM and total fungi in a semi-arid grassland. FEMS Microbiol Ecol 91:1-13

Kleczewski N, Bauer J, D. Bever J, Clay K, L. Reynolds H (2012) A survey of endophytic fungi of switchgrass (*Panicum virgatum*) in the Midwest, and their putative roles in plant growth vol 5. doi:10.1016/j.funeco.2011.12.006

Kluber LA et al. (2012) Mycorrhizal response to experimental pH and P manipulation in acidic hardwood forests. PLoS One 7:e48946 doi:10.1371/journal.pone.0048946

Klugh KR, Cumming JR (2007) Variations in organic acid exudation and aluminum resistance among arbuscular mycorrhizal species colonizing *Liriodendron tulipifera*. Tree Physiol 27:1103-1112

Kochian LV (1995) Cellular mechanisms of aluminum toxicity and resistance in plants. Annu Rev Plant Phys 46:237-260

Kochian LV, Pineros MA, Hoekenga OA (2005) The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. Plant Soil 274:175-195

Kohout P et al. (2014) Comparison of commonly used primer sets for evaluating arbuscular mycorrhizal fungal communities: Is there a universal solution? Soil Biol Biochem 68:482-493 doi:10.1016/j.soilbio.2013.08.027

Koljalg U et al. (2013) Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol 22:5271-5277 doi:10.1111/mec.12481

Kollmeier M, Felle HH, Horst WJ (2000) Genotypical differences in aluminum resistance of maize are expressed in the distal part of the transition zone. Is reduced basipetal auxin flow involved in inhibition of root elongation by aluminum? Plant Physiol 122:945-956

Koltai H, Kapulnik Y (2010) Arbuscular mycorrhizas: physiology and function. 2 edn. Springer, Berlin

Koorem K et al. (2017) Arbuscular mycorrhizal fungal communities in forest plant roots are simultaneously shaped by host characteristics and canopy-mediated light availability. Plant Soil 410:259-271 doi:10.1007/s11104-016-3004-0

Kuczynski J, Stombaugh J, Walters WA, Gonzalez A, Caporaso JG, Knight R (2011) Using QIIME to analyze 16S rRNA gene sequences from microbial communities. Curr Protoc Bioinformatics Chapter 10: Unit 10 17 doi:10.1002/0471250953.bi1007s36

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ Microbiol 12:118-123

Larsen BB, Miller EC, Rhodes MK, Wiens JJ (2017) Inordinate fondness multiplied and redistributed: the number of species on Earth and the new pie of life. Q Rev Biol 92:229-265

Lee J, Lee S, Young JPW (2008) Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. FEMS Microbiol Ecol 65:339-349 doi:10.1111/j.1574-6941.2008.00531.x

Lekberg Y et al. (2018) More bang for the buck? Can arbuscular mycorrhizal fungal communities be characterized adequately alongside other fungi using general fungal primers? New Phytol doi:10.1111/nph.15035

LeVeque RJ (2012) Top ten reasons to not share your code (and why you should anyway). https://sinews.siam.org/Details-Page/top-ten-reasons-to-not-share-your-code-and-why-you-should-anyway

Liang M, Liu X, Etienne RS, Huang F, Wang Y, Yu S (2015) Arbuscular mycorrhizal fungi counteract the Janzen-Connell effect of soil pathogens. Ecology 96:562-574

Lin K et al. (2014) Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. PLoS Genet 10:e1004078 doi:10.1371/journal.pgen.1004078

Liu JY, Maldonado-Mendoza I, Lopez-Meyer M, Cheung F, Town CD, Harrison MJ (2007) Arbuscular mycorrhizal symbiosis is accompanied by local and systemic alterations in gene expression and an increase in disease resistance in the shoots. Plant J 50:529-544 doi:10.1111/j.1365-313X.2007.03069.x

Liu M, Zheng R, Bai S, Bai YE, Wang J (2017) Slope aspect influences arbuscular mycorrhizal fungus communities in arid ecosystems of the Daqingshan Mountains, Inner Mongolia, North China. Mycorrhiza 27:189-200

Lumini E, Orgiazzi A, Borriello R, Bonfante P, Bianciotto V (2010) Disclosing arbuscular mycorrhizal fungal biodiversity in soil through a land-use gradient using a pyrosequencing approach. Environ Microbiol 12:2165-2179

Luo C, Tsementzi D, Kyrpides N, Read T, Konstantinidis KT (2012) Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. PLoS One 7:e30087

Luo J, Walsh E, Blystone D, Zhang N (2015) Five new *Pseudophialophora* species from grass roots in the oligotrophic pine barrens ecosystem. Fungal Biol-UK 119:1205-1215

Luo J, Walsh E, Miller S, Blystone D, Dighton J, Zhang N (2017) Root endophytic fungal communities associated with pitch pine, switchgrass, and rosette grass in the pine barrens ecosystem. Fungal Biol-UK 121:478-487

Luo J, Walsh E, Naik A, Zhuang W, Zhang K, Cai L, Zhang N (2014a) Temperate pine barrens and tropical rain forests are both rich in undescribed fungi. PLoS One 9:e103753

Luo J, Walsh E, Zhang N (2014b) Four new species in Magnaporthaceae from grass roots in New Jersey Pine Barrens. Mycologia 106:580-588

Ma JF, Ryan PR, Delhaize E (2001) Aluminium tolerance in plants and the complexing role of organic acids. Trends Plant Sci 6:273-278

Magrou MJ (1936) Culture et inoculation du chapignon symbiotique d'Arum maculatum. Comptes Rendues hebd Seanc Acad Sci 203:887-888

Manoharachary C, Kunwar IK, Tilak K, Adholeya A (2010) Arbuscular mycorrhizal fungi - taxonomy, diversity, conservation and multiplication. P Natl A Sci India B 80:1-13

Mao Y, Li X, Smyth EM, Yannarell AC, Mackie RI (2014) Enrichment of specific bacterial and eukaryotic microbes in the rhizosphere of switchgrass (*Panicum virgatum* L.) through root exudates. Environ Microbiol Rep 6:293-306 doi:10.1111/1758-2229.12152

Margulies M et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376-380 doi:10.1038/nature03959

Marleau J, Dalpe Y, St-Arnaud M, Hijri M (2011) Spore development and nuclear inheritance in arbuscular mycorrhizal fungi. BMC Evol Biol 11:51 doi:10.1186/1471-2148-11-51

Marschner H (1991) Mechanisms of adaptation of plants to acid soils. Plant Soil 134:1-20 doi:10.1007/bf00010712

Marschner H, Dell B (1994) Nutrient-uptake in mycorrhizal symbiosis. Plant Soil 159:89-102

Martin FM, Uroz S, Barker DG (2017) Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. Science 356:819-+ doi:10.1126/science.aad4501

Martinez-Garcia LB, Richardson SJ, Tylianakis JM, Peltzer DA, Dickie IA (2015) Host identity is a dominant driver of mycorrhizal fungal community composition during ecosystem development. New Phytol 205:1565-1576 doi:10.1111/nph.13226

McCormick J, Buell MF (1957) Natural revegetation of a plowed field in the New Jersey Pine Barrens. Bot Gaz 118:261-264 doi:10.1086/335952

McGonigle TP, Miller MH, Evans DG, Fairchild GL, Swan JA (1990) A new method which gives an objective measure of colonization of roots by vesicular-arbuscular mycorrhizal fungi. New Phytol 115:495-495

McKenna JR, Wolf DD (1990) No-till switchgrass establishment as affected by limestone, phosphorus, and carbofuran. J Prod Agric 3:475-479

McLaughlin SB New switchgrass biofuels research program for the southeast. In: 1992 Annual Automotive Technology Development Contractor's Coordinating Meeting, Dearborn, MI, November 2-5, 1992 1993. pp 111-115

McLaughlin SB, Bransby DI, Parrish D Perennial grass production for biofuels: soil conservation considerations. In: Bioenergy '94, Proceedings of the 6th International Bioenergy Conference, Reno, NV, October 2-6, 1994 1994. pp 359-367

McNeill J, Turland NJ (2012) International Code of Nomenclature for algae, fungi, and plants (Melbourne Code) In: McNeill J et al. (eds) International Code of Nomenclature for Algae, Fungi, and Plants, vol 154. Regnum Vegetabile. p 240

Melo CD et al. (2017) Communities of arbuscular mycorrhizal fungi under *Picconia azorica* in native forests of Azores. Symbiosis 74:43-54 doi:10.1007/s13199-017-0487-2

Mendoza J, Borie F (1998) Effect of *Glomus etunicatum* inoculation on aluminum, phosphorus, calcium, and magnesium uptake of two barley genotypes with different aluminum tolerance. Commun Soil Sci Plan 29:681-695

Mesirov JP (2010) Computer science. Accessible reproducible research. Science 327:415-416 doi:10.1126/science.1179653

Mohan JE et al. (2014) Mycorrhizal fungi mediation of terrestrial ecosystem responses to global change: mini-review. Fungal Ecol 10:3-19

Moora M et al. (2014) Anthropogenic land use shapes the composition and phylogenetic structure of soil arbuscular mycorrhizal fungal communities. FEMS Microbiol Ecol 90:609-621

Moora M, Opik M, Davison J, Jairus T, Vasar M, Zobel M, Eckstein RL (2016) AM fungal communities inhabiting the roots of submerged aquatic plant *Lobelia dortmanna* are diverse and include a high proportion of novel taxa. Mycorrhiza 26:735-745 doi:10.1007/s00572-016-0709-0

Mora C, Tittensor DP, Adl S, Simpson AG, Worm B (2011) How many species are there on Earth and in the ocean? PLoS Biol 9:e1001127 doi:10.1371/journal.pbio.1001127

Morgan BST, Egerton-Warburton LM (2017) Barcoded NS31/AML2 primers for sequencing of arbuscular mycorrhizal communities in environmental samples. Appl Plant Sci 5

Morton JB, Bentivenga SP, Bever JD (1995) Discovery, measurement, and interpretation of diversity in arbuscular endomycorrhizal fungi (Glomales, Zygomycetes). Can J Bot 73:S25-S32 doi:10.1139/b95-221

Muthukumar T, Priyadharsini P, Uma E, Jaison S, Pandey RR (2014) Role of Arbuscular Mycorrhizal Fungi in Alleviation of Acidity Stress on Plant Growth. In: Use of plant growth-promoting rhizobacteria to alleviate salinity stress in plants, vol 1. pp 43-71. doi:10.1007/978-1-4614-9466-9_3

Nelson KE et al. (2010) A catalog of reference genomes from the human microbiome. Science 328:994-999 doi:10.1126/science.1183605

Newsham KK (2011) A meta-analysis of plant responses to dark septate root endophytes. New Phytol 190:783-793 doi:10.1111/j.1469-8137.2010.03611.x

Nicolson TH, Schenck NC (1979) Endogonaceous mycorrhizal endophytes in Florida. Mycologia 71:178-198

O'Callaghan EB (1853) Documents relative to the colonial history of the state of New York. Vol. 1. Weed, Parsons, & Co., Albany, NY

Oehl F, Laczko E, Oberholzer HR, Jansa J, Egli S (2017) Diversity and biogeography of arbuscular mycorrhizal fungi in agricultural soils. Biol Fert Soils 53:777-797 doi:10.1007/s00374-017-1217-x

Oehl F, Sanchez-Castro I, Palenzuela J, da Silva GA, Sieverding E (2014) *Glomus compressum*, a new arbuscular mycorrhizal fungus from different agro-ecosystems in Central Europe. Nova Hedwigia 99:429-439 doi:10.1127/0029-5035/2014/0200

Oehl F, Schneider D, Sieverding E, Burga CA (2011) Succession of arbuscular mycorrhizal communities in the foreland of the retreating Morteratsch glacier in the Central Alps. Pedobiologia 54:321-331 doi:10.1016/j.pedobi.2011.07.006

Oehl F, Sieverding E, Ineichen K, Ris EA, Boller T, Wiemken A (2005) Community structure of arbuscular mycorrhizal fungi at different soil depths in extensively and intensively managed agroecosystems. New Phytol 165:273-283 doi:10.1111/j.1469-8137.2004.01235.x

Oehl F, Sieverding E, Mader P, Dubois D, Ineichen K, Boller T, Wiemken A (2004) Impact of long-term conventional and organic farming on the diversity of arbuscular mycorrhizal fungi. Oecologia 138:574-583 doi:10.1007/s00442-003-1458-2

Ohsowski BM, Zaitsoff PD, Opik M, Hart MM (2014) Where the wild things are: looking for uncultured Glomeromycota. New Phytol 204:171-179 doi:10.1111/nph.12894

Oksanen J et al. (2017) vegan: Community Ecology Package. R package version 24-4

Öpik M, Davison J (2016) Uniting species- and community-oriented approaches to understand arbuscular mycorrhizal fungal diversity. Fungal Ecol 24:106-113 doi:10.1016/j.funeco.2016.07.005

Opik M, Davison J, Moora M, Partel M, Zobel M (2016) Response to Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". Science 351:826 doi:10.1126/science.aad5495

Öpik M, Davison J, Moora M, Zobel M (2014) DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. Botany 92:135-147 doi:10.1139/cjb-2013-0110

Opik M, Metsis M, Daniell TJ, Zobel M, Moora M (2009) Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. New Phytol 184:424-437 doi:10.1111/j.1469-8137.2009.02920.x

Opik M, Moora M, Liira J, Koljalg U, Zobel M, Sen R (2003) Divergent arbuscular mycorrhizal fungal communities colonize roots of *Pulsatilla* spp. in boreal Scots

pine forest and grassland soils. New Phytol 160:581-593 doi:10.1046/j.1469-8137x.2003.00917.x

Opik M, Moora M, Liira J, Zobel M (2006) Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. J Ecol 94:778-790 doi:10.1111/j.1365-2745.2006.01136.x

Opik M et al. (2010) The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). New Phytol 188:223-241 doi:10.1111/j.1469-8137.2010.03334.x

Opik M et al. (2013) Global sampling of plant roots expands the described molecular diversity of arbuscular mycorrhizal fungi. Mycorrhiza 23:411-430 doi:10.1007/s00572-013-0482-2

Orchard S et al. (2017a) Fine endophytes (*Glomus tenue*) are related to Mucoromycotina, not Glomeromycota. New Phytol 213:481-486 doi:10.1111/nph.14268

Orchard S, Standish RJ, Dickie IA, Renton M, Walker C, Moot D, Ryan MH (2017b) Fine root endophytes under scrutiny: a review of the literature on arbuscule-producing fungi recently suggested to belong to the Mucoromycotina. Mycorrhiza doi:10.1007/s00572-017-0782-z

Oulhen N, Schulz BJ, Carrier TJ (2016) English translation of Heinrich Anton de Bary's 1878 speech, 'Die Erscheinung der Symbiose' ('De la symbiose'). Symbiosis 69:131-139 doi:10.1007/s13199-016-0409-8

Ouzounidou G, Skiada V, Papadopoulou KK, Stamatis N, Kavvadias V, Eleftheriadis E, Gaitis F (2015) Effects of soil pH and arbuscular mycorrhiza (AM) inoculation on growth and chemical composition of chia (*Salvia hispanica* L.) leaves. Braz J Bot 38:487-495 doi:10.1007/s40415-015-0166-6

Paine LK et al. (1996) Some ecological and socio-economic considerations for biomass energy crop production. Biomass Bioenerg 10:231-242 doi:10.1016/0961-9534(95)00072-0

Palenzuela J, Azcon-Aguilar C, Barea JM, da Silva GA, Oehl F (2013) *Acaulospora pustulata* and *Acaulospora tortuosa*, two new species in the Glomeromycota from Sierra Nevada National Park (southern Spain). Nova Hedwigia 97:305-319 doi:10.1127/0029-5035/2013/0129

Parr CL, Lehmann CE, Bond WJ, Hoffmann WA, Andersen AN (2014) Tropical grassy biomes: misunderstood, neglected, and under threat. Trends Ecol Evol 29:205-213 doi:10.1016/j.tree.2014.02.004

Parrish DJ, Fike JH (2005) The biology and agronomy of switchgrass for biofuels. CRC Cr Rev Plant Sci 24:423-459

Partel M et al. (2017) Historical biome distribution and recent human disturbance shape the diversity of arbuscular mycorrhizal fungi. New Phytol 216:227-238 doi:10.1111/nph.14695

Paul EA, Clark FE (1996) Soil microbiology and biochemistry. 2 edn., San Diego, CA

Pendleton RL, Smith BN (1983) Vesicular-arbuscular mycorrhizae of weedy and colonizer plant species at disturbed sites in Utah. Oecologia 59:296-301 doi:10.1007/BF00378852

Pirozynski KA, Malloch DW (1975) The origin of land plants: a matter of mycotrophism. Biosystems 6:153-164

Pitet M, Camprubi A, Calvet C, Estaun V (2009) A modified staining technique for arbuscular mycorrhiza compatible with molecular probes. Mycorrhiza 19:125-131 doi:10.1007/s00572-008-0206-1

Porras-Alfaro A, Herrera J, Natvig DO, Sinsabaugh RL (2007) Effect of long-term nitrogen fertilization on mycorrhizal fungi associated with a dominant grass in a semiarid grassland. Plant Soil 296:65-75 doi:10.1007/s11104-007-9290-9

Powell JR, Parrent JL, Hart MM, Klironomos JN, Rillig MC, Maherali H (2009) Phylogenetic trait conservatism and the evolution of functional trade-offs in arbuscular mycorrhizal fungi. P Roy Soc B-Biol Sci 276:4237-4245 doi:10.1098/rspb.2009.1015

Powell JR, Welsh A, Hallin S (2015) Microbial functional diversity enhances predictive models linking environmental parameters to ecosystem properties. Ecology 96:1985-1993

Pozo MJ, Lopez-Raez JA, Azcon-Aguilar C, Garcia-Garrido JM (2015) Phytohormones as integrators of environmental signals in the regulation of mycorrhizal symbioses. New Phytol 205:1431-1436 doi:10.1111/nph.13252

Qin J et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464:59-65 doi:10.1038/nature08821

Quince C et al. (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. Nat Methods 6:639-641 doi:10.1038/nmeth.1361

Rath M, Grolig F, Haueisen J, Imhof S (2014) Combining microtomy and confocal laser scanning microscopy for structural analyses of plant-fungus associations. Mycorrhiza 24:293-300 doi:10.1007/s00572-013-0530-y

Read DJ (1991) Mycorrhizas in ecosystems. Experientia 47:376-391

Redecker D, Kodner R, Graham LE (2000) Glomalean fungi from the Ordovician. Science 289:1920-1921

Redecker D, Schüßler A, Stockinger H, Sturmer SL, Morton JB, Walker C (2013) An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). Mycorrhiza 23:515-531 doi:10.1007/s00572-013-0486-y

Reeder J, Knight R (2009) The 'rare biosphere': a reality check. Nat Methods 6:636-637 doi:10.1038/nmeth0909-636

Remy W, Taylor TN, Hass H, Kerp H (1994) Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proc Natl Acad Sci U S A 91:11841-11843

Research for Sustainable Bioenergy: Linking Genomic and Ecosystem Sciences (2014). U.S. Department of Energy Office of Science

Rice EL, Parenti RL (1978) Causes of decreases in productivity in undisturbed tall grass prairie. Am J Bot 65:1091-1097

Rillig MC (2004) Arbuscular mycorrhizae and terrestrial ecosystem processes. Ecol Lett 7:740-754 doi:10.1111/j.1461-0248.2004.00620.x

Rillig MC, Wright SF, Eviner VT (2002) The role of arbuscular mycorrhizal fungi and glomalin in soil aggregation: comparing effects of five plant species. Plant Soil 238:325-333

Robichaud B, Buell MF (1973) Vegetation of New Jersey. A study of landscape diversity. Rutgers Univ. Press, New Brunswick, NJ

Robson AD, Abbott LK (1989) The effect of soil acidity on microbial activity in soils. In: Robson AD (ed) Soil acidity and plant growth. Academic Press, New York, NY, pp 139-165

Rocha LA et al. (2014) Specimen collection: an essential tool. Science 344:814-815 doi:10.1126/science.344.6186.814

Rodriguez A, Clapp JP, Robinson L, Dodd JC (2005) Studies on the diversity of the distinct phylogenetic lineage encompassing *Glomus claroideum* and *Glomus etunicatum*. Mycorrhiza 15:33-46 doi:10.1007/s00572-003-0291-0

RStudio T (2016) RStudio: integrated development environment for R. Boston, MA

Sanders FE, Tinker PB (1971) Mechanism of absorption of phosphate from soil by *Endogone* mycorrhizas. Nature 233:278-279

Sanders IR, Alt M, Groppe K, Boller T, Wiemken A (1995) Identification of ribosomal DNA polymorphisms among and within spores of the Glomales: application to studies on the genetic diversity of arbuscular mycorrhizal fungal communities. New Phytol 130:419-427 doi:10.1111/j.1469-8137.1995.tb01836.x

Sanderson MA et al. (1996) Switchgrass as a sustainable bioenergy crop. Bioresour Technol 56:83-93 doi:10.1016/0960-8524(95)00176-x

Santos-Gonzalez JC, Finlay RD, Tehler A (2007) Seasonal dynamics of arbuscular mycorrhizal fungal communities in roots in a seminatural grassland. Appl Environ Microbiol 73:5613-5623 doi:10.1128/aem.00262-07

Santos JC, Finlay RD, Tehler A (2006) Molecular analysis of arbuscular mycorrhizal fungi colonising a semi-natural grassland along a fertilisation gradient. New Phytol 172:159-168 doi:10.1111/j.1469-8137.2006.01799.x

Sato K, Suyama Y, Saito M, Sugawara K (2005) A new primer for discrimination of arbuscular mycorrhizal fungi with polymerase chain reaction-denature gradient gel electrophoresis. Grassl Sci 51:179-181 doi:10.1111/j.1744-697X.2005.00023.x

Schechter SP, Bruns TD (2012) Edaphic sorting drives arbuscular mycorrhizal fungal community assembly in a serpentine/non-serpentine mosaic landscape. Ecosphere 3 doi:10.1890/es12-00059.1

Scheublin TR, Ridgway KP, Young JP, van der Heijden MG (2004) Nonlegumes, legumes, and root nodules harbor different arbuscular mycorrhizal fungal communities. Appl Environ Microbiol 70:6240-6246 doi:10.1128/AEM.70.10.6240-6246.2004

Schlaeppi K et al. (2016) High-resolution community profiling of arbuscular mycorrhizal fungi. New Phytol 212:780-791 doi:10.1111/nph.14070

Schmidt P-A, Bálint M, Greshake B, Bandow C, Römbke J, Schmitt I (2013) Illumina metabarcoding of a soil fungal community. Soil Biol Biochem 65:128-132 doi:10.1016/j.soilbio.2013.05.014

Schoch CL et al. (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci U S A 109:6241-6246 doi:10.1073/pnas.1117018109

Schroeder-Moreno MS, Greaver TL, Wang SX, Hu SJ, Rufty TW (2012) Mycorrhizal-mediated nitrogen acquisition in switchgrass under elevated temperatures and N enrichment. GCB Bioenergy 4:266-276 doi:10.1111/j.1757-1707.2011.01128.x

Schüßler A, Schwarzott D, Walker C (2001) A new fungal phylum, the Glomeromycota: phylogeny and evolution. Mycol Res 105:1413-1421 doi:Doi 10.1017/S0953756201005196

Seguel A, Cumming JR, Klugh-Stewart K, Cornejo P, Borie F (2013) The role of arbuscular mycorrhizas in decreasing aluminium phytotoxicity in acidic soils: a review. Mycorrhiza 23:167-183 doi:10.1007/s00572-013-0479-x

Self-Davis ML et al. (2003) Forage species and canopy cover effects on runoff from small plots. J Soil Water Conserv 58:349-359

Sepp SK, Jairus T, Vasar M, Zobel M, Opik M (2018) Effects of land use on arbuscular mycorrhizal fungal communities in Estonia. Mycorrhiza 28:259-268 doi:10.1007/s00572-018-0822-3

Shah V et al. (2011) Bacterial and archaea community present in the Pine Barrens Forest of Long Island, NY: unusually high percentage of ammonia oxidizing bacteria. PLoS One 6:e26263 doi:10.1371/journal.pone.0026263

Shendure J et al. (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. Science 309:1728-1732 doi:10.1126/science.1117389

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research. Mol Ecol 21:1794-1805 doi:10.1111/j.1365-294X.2012.05538.x

Sieverding E (1991) Vesicular-arbuscular mycorrhiza management in tropical agrosystems. Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ) GmbH, Eschborn, Germany

Simon L, Lalonde M, Bruns TD (1992) Specific amplification of 18S fungal ribosomal genes from vesicular-arbuscular endomycorrhizal fungi colonizing roots. Appl Environ Microbiol 58:291-295

Siqueira JO, Hubbell DH, Mahmud AW (1984) Effect of liming on spore germination, germ tube growth and root colonization by vesicular-arbuscular mycorrhizal fungi. Plant Soil 76:115-124

Smilauer P (2001) Communities of arbuscular mycorrhizal fungi in grassland: seasonal variability and effects of environment and host plants. Folia Geobotanica 36:243-263 doi:10.1007/bf02803179

Smith SE, Dickson S (1991) Quantification of active vesicular-arbuscular mycorrhizal infection using image analysis and other techniques. Aust J Plant Physiol 18:637-648

Smith SE, Gianinazzi-Pearson V (1990) Phosphate uptake and arbuscular activity in mycorrhizal *Allium cepa* L.: effects of photon irradiance and phosphate nutrition. Aust J Plant Physiol 17:177-188

Smith SE, Read DJ (2008) Mycorrhizal symbiosis. Elsevier, Academic Press, Amsterdam [etc.]

Son CL, Smith SE (1988) Mycorrhizal growth responses: interactions between photon irradiance and phosphorus nutrition. New Phytol 108:305-314 doi:10.1111/j.1469-8137.1988.tb04167.x

Southworth D, Tackaberry LE, Massicotte HB (2014) Mycorrhizal ecology on serpentine soils. Plant Ecol Divers 7:445-455

Spatafora JW et al. (2016) A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia 108:1028-1046 doi:10.3852/16-042

Stahl M (1949) Die mycorrhiza de lebermoose mit besonderer beruchsichtigung der thallosen formen. Planta 37:103-148

Stajich JE, Berbee ML, Blackwell M, Hibbett DS, James TY, Spatafora JW, Taylor JW (2009) The fungi. Curr Biol 19:R840-845 doi:10.1016/j.cub.2009.07.004

Stockinger H, Kruger M, Schüßler A (2010) DNA barcoding of arbuscular mycorrhizal fungi. New Phytol 187:461-474 doi:10.1111/j.1469-8137.2010.03262.x

Stockinger H, Walker C, Schüßler A (2009) '*Glomus intraradices* DAOM197198', a model fungus in arbuscular mycorrhiza research, is not *Glomus intraradices*. New Phytol 183:1176-1187 doi:10.1111/j.1469-8137.2009.02874.x

Stürmer SL, Oliveira LZ, Morton JB (2018) Gigasporaceae versus Glomeraceae (phylum Glomeromycota): a biogeographic tale of dominance in maritime sand dunes. Fungal Ecol 32:49-56 doi:10.1016/j.funeco.2017.11.008

Sun W, Zhang FL, He LM, Li ZY (2014) Pyrosequencing reveals diverse microbial community associated with the zoanthid *Palythoa australiae* from the South China Sea. Microb Ecol 67:942-950 doi:10.1007/s00248-014-0395-4

Suzuki K, Turgay OC, Akca MO, Harada N, Nonaka M (2014) Molecular diversity of indigenous arbuscular mycorrhizal fungi in three different agricultural regions of Turkey. Soil Sci Plant Nutr 60:367-376 doi:10.1080/00380768.2014.890916

Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+ C-content biases. Mol Biol Evol 9:678-687

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725-2729

Taylor AFS, Alexander IAN (2005) The ectomycorrhizal symbiosis: life in the real world. Mycologist 19:102-112 doi:10.1017/s0269-915x(05)00303-4

Tedersoo L et al. (2010) 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. New Phytol 188:291-301 doi:10.1111/j.1469-8137.2010.03373.x

Tedersoo L et al. (2018) High-level classification of the Fungi and a tool for evolutionary ecological analyses. Fungal Divers 90:135-159 doi:10.1007/s13225-018-0401-0

Tester M, Smith SE, Smith FA, Walker NA (1986) Effects of photon irradiance on the growth of shoots and roots, on the rate of initiation of mycorrhizal infection and on the growth of infection units in *Trifolium subterraneum* L. New Phytol 103:375-390 doi:10.1111/j.1469-8137.1986.tb00623.x

Titus JH, Leps J (2000) The response of arbuscular mycorrhizae to fertilization, mowing, and removal of dominant species in a diverse oligotrophic wet meadow. Am J Bot 87:392-401 doi:10.2307/2656635

Torres-Cruz TJ et al. (2017) *Bifiguratus adelaidae*, gen. et sp. nov., a new member of Mucoromycotina in endophytic and soil-dwelling habitats. Mycologia:1-16 doi:10.1080/00275514.2017.1364958

Treseder KK (2004) A meta-analysis of mycorrhizal responses to nitrogen, phosphorus, and atmospheric CO2 in field studies. New Phytol 164:347-355

Treseder KK (2016) Model behavior of arbuscular mycorrhizal fungi: predicting soil carbon dynamics under climate change. Botany 94:417-423 doi:10.1139/cjb-2015-0245

Treseder KK, Allen MF (2000) Research review: mycorrhizal fungi have a potential role in soil carbon storage under elevated CO2 and nitrogen deposition. New Phytol 147:189-200

Treseder KK, Cross A (2006) Global distributions of arbuscular mycorrhizal fungi. Ecosystems 9:305-316 doi:10.1007/s10021-005-0110-x

Tuininga AR, Dighton J (2004) Changes in ectomycorrhizal communities and nutrient availability following prescribed burns in two upland pine-oak forests in the New Jersey pine barrens. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 34:1755-1765 doi:10.1139/x04-037

Unterseher M, Jumpponen A, Opik M, Tedersoo L, Moora M, Dormann CF, Schnittler M (2011) Species abundance distributions and richness estimations in fungal metagenomics - lessons learned from community ecology. Mol Ecol 20:275-285 doi:10.1111/j.1365-294X.2010.04948.x

USDA N (2017) The PLANTS Database. National Plant Data Team. https://plants.usda.gov/core/profile?symbol=pavi2. Accessed October 2, 2017 2017

Valyi K, Rillig MC, Hempel S (2015) Land-use intensity and host plant identity interactively shape communities of arbuscular mycorrhizal fungi in roots of grassland plants. New Phytol 205:1577-1586 doi:10.1111/nph.13236

van der Gast CJ, Gosling P, Tiwari B, Bending GD (2011) Spatial scaling of arbuscular mycorrhizal fungal diversity is affected by farming practice. Environ Microbiol 13:241-249 doi:10.1111/j.1462-2920.2010.02326.x

van der Heijden MG, Martin FM, Selosse MA, Sanders IR (2015) Mycorrhizal ecology and evolution: the past, the present, and the future. New Phytol 205:1406-1423 doi:10.1111/nph.13288

van der Heijden MGA et al. (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. Nature 396:69-72

van der Heijden MGA et al. (2006) The mycorrhizal contribution to plant productivity, plant nutrition and soil structure in experimental grassland. New Phytol 172:739-752 doi:10.1111/j.1469-8137.2006.01862.x

van Diepen LTA, Lilleskov EA, Pregitzer KS (2011) Simulated nitrogen deposition affects community structure of arbuscular mycorrhizal fungi in northern hardwood forests. Mol Ecol 20:799-811 doi:10.1111/j.1365-294X.2010.04969.x

Van Geel M, Busschaert P, Honnay O, Lievens B (2014) Evaluation of six primer pairs targeting the nuclear rRNA operon for characterization of arbuscular mycorrhizal fungal (AMF) communities using 454 pyrosequencing. J Microbiol Methods 106:93-100 doi:10.1016/j.mimet.2014.08.006

van Geel M, Ceustermans A, van Hemelrijck W, Lievens B, Honnay O (2015) Decrease in diversity and changes in community composition of arbuscular mycorrhizal fungi in roots of apple trees with increasing orchard management intensity across a regional scale. Mol Ecol 24:941-952 doi:10.1111/mec.13079

Vandenkoornhuyse P, Ridgway KP, Watson IJ, Fitter AH, Young JPW (2003) Co-existing grass species have distinctive arbuscular mycorrhizal communities. Mol Ecol 12:3085-3095 doi:10.1046/j.1365-294X.2003.01967.x

Varela-Cervero S, Vasar M, Davison J, Barea JM, Opik M, Azcon-Aguilar C (2015) The composition of arbuscular mycorrhizal fungal communities differs among the

roots, spores and extraradical mycelia associated with five Mediterranean plant species. Environ Microbiol 17:2882-2895 doi:10.1111/1462-2920.12810

Varga S, Finozzi C, Vestberg M, Kytoviita MM (2015) Arctic arbuscular mycorrhizal spore community and viability after storage in cold conditions. Mycorrhiza 25:335-343 doi:10.1007/s00572-014-0613-4

Vasar M et al. (2017) Increased sequencing depth does not increase captured diversity of arbuscular mycorrhizal fungi. Mycorrhiza doi:10.1007/s00572-017-0791-y

Vogel KP (2004) Switchgrass. In: Moser LE, Burson BL, Sollenberger LE (eds) Warm-season (C4) grasses. ASA, CSSA, and SSSA, Madison, WI, pp 561-588

von Uexküll HR, Mutert E (1995) Global extent, development and economic impact of acid soils. Plant Soil 171:1-15 doi:10.1007/bf00009558

Wacker PO A preliminary view of the possible associations between cultural background and agriculture in New Jersey during the latter part of the eighteenth century. In: Proc. N.Y.-N.J. Div., Assoc. Am. Geographers, 1971. pp 41-57

Walker C, Gollotte A, Redecker D (2018) A new genus, *Planticonsortium* (Mucoromycotina), and new combination (*P. tenue*), for the fine root endophyte, *Glomus tenue* (basionym *Rhizophagus tenuis*). Mycorrhiza 28:213-219 doi:10.1007/s00572-017-0815-7

Walker C, Trappe JM (1993) Names and epithets in the Glomales and Endogonales. Mycol Res 97:339-344 doi:10.1016/s0953-7562(09)81132-5

Walsh E, Luo J, Naik A, Preteroti T, Zhang N (2015) *Barrenia*, a new genus associated with roots of switchgrass and pine in the oligotrophic pine barrens. Fungal Biol-UK 119:1216-1225

Walsh E, Luo J, Zhang N (2014) *Acidomelania panicicola* gen. et sp. nov. from switchgrass roots in acidic New Jersey pine barrens. Mycologia 106:856-864

Wang B, Qiu YL (2006) Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza 16:299-363 doi:10.1007/s00572-005-0033-6

Wang B, Yeun LH, Xue JY, Liu Y, Ane JM, Qiu YL (2010) Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. New Phytol 186:514-525 doi:10.1111/j.1469-8137.2009.03137.x

Wang C, White PJ, Li C (2017) Colonization and community structure of arbuscular mycorrhizal fungi in maize roots at different depths in the soil profile respond differently to phosphorus inputs on a long-term experimental site. Mycorrhiza 27:369-381 doi:10.1007/s00572-016-0757-5

Wang GM, Stribley DP, Tinker PB, Walker C (1993) Effects of pH on arbuscular mycorrhiza. I. Field pbservations on the long- term liming experiments at Rothamsted and Woburn. New Phytol 124:465-472

Wilson GWT, Hartnett DC (1998) Interspecific variation in plant responses to mycorrhizal colonization in tallgrass prairie. Am J Bot 85:1732-1738 doi:10.2307/2446507

Wilson GWT, Rice CW, Rillig MC, Springer A, Hartnett DC (2009) Soil aggregation and carbon sequestration are tightly correlated with the abundance of arbuscular mycorrhizal fungi: results from long-term field experiments. Ecol Lett 12:452-461 doi:10.1111/j.1461-0248.2009.01303.x

Wright SF and Upadhyaya A (1996) Extraction of an abundant and unusual protein from soil and comparison with hyphal protein of arbuscular mycorrhizal fungi. Soil Sci 161 (9):575-586 10.1097/00010694-199609000-00003

Wyss T, Masclaux FG, Rosikiewicz P, Pagni M, Sanders IR (2016) Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus *Rhizophagus irregularis*. ISME J 10:2514-2526 doi:10.1038/ismej.2016.29

Yang H et al. (2017) Taxonomic resolution is a determinant of biodiversity effects in arbuscular mycorrhizal fungal communities. J Ecol 105:219-228 doi:10.1111/1365-2745.12655

Young JL, Davis EA, Rose SL (1985) Endomycorrhizal fungi in breeder wheats and triticale cultivars field-grown on fertile soil. Agron J 77:219-224 doi:10.2134/agronj1985.00021962007700020011x

Zhao H, Li XZ, Zhang ZM, Zhao Y, Yang JT, Zhu YW (2017) Species diversity and drivers of arbuscular mycorrhizal fungal communities in a semi-arid mountain in China. Peerj 5 doi:10.7717/peerj.4155

Appendix

Table A2-1 Information on sampling sites.

Site name, year sampled, type of landscape (Managed, Pine Barrens, Natural Prairie), and associated soil properties are included. Nutrients measured in parts per million (ppm) represent extracted nutrients (Mehlich-3). Sites include AF14 - Adelphia Research Field 2014, AF17 - Adelphia Research Field 2017, CM - Colliers Mills, EC - EARTH Center, IO - Iowa, FAA - Federal Aviation Administration, LIPB14 - Long Island Pine Barrens 2014, LIPB16 - Long Island Pine Barrens 2016, RP - Rocky Point, SO - Somerset Research Field WSF - Wharton State Forest.

Site		EC	AF14	AF17	CM	IO	SO	WSF	LIPB14	LIPB16	RP	FAA
Sampling Year		2014	2014	2017	2014	2014	2017	2014	2014	2016	2017	2017
Site Type		Managed	Managed	Managed	Pine Barrens	Natural Prairie	Managed	Pine Barrens	Pine Barrens	Pine Barrens	Pine Barrens	Pine Barrens
Soil Properties	pH	6.6	6.35	6.7	5	7.8	5.5	5.2	4.9	4.9	4.8	4.9
	SOM[a] (%)	1.6	8.5	0.5	0.2	2.3	1.7	2.6	0.4	0.1	1.8	0.4
	CEC[b]	7.8	3.2	4.5	0.6	18.8	7.6	7.3	0.5	0.5	6.8	0.5
	P (ppm)	310	280.5	203	11	30	29	41	8	6	14	8
	K (ppm)	178	133.5	147	9	137	80	43	13	12	25	13
	Ca (ppm)	976	462.5	559	112	4707	638	484	89	77	130	89
	Mg (ppm)	300	158.5	177	26	476	195	82	23	20	38	23
	NO3 (ppm)	13	8.5	2	2	3	7	11	3	2	2	3
	NH4 (ppm)	5	5	2	1	8	3	8	8	2	3	8
	N (total %)	0.36	0.37	0.10	0.16	0.27	0.17	0.11	0.30	0.14	0.08	0.30
	Al (ppm)	975	804	3368	82	364	18820	323	918	562	4832	918
	Al Exch[c] (ppm)	48.7	17.4	0.38	4.1	2.3	2.4	15.9	45.7	14.4	10.4	45.7
	Fe (ppm)	159	216	193	48	122	121	442	93	134.2	194	93

Zn (ppm)	5.1	3.2	1.9	0.8	3.2	1.8	15.5	2.8	0.5	1.2	2.8
P (mg/Kg)	1846	2590	1263	39	703	271	374	11	33	142	11

[a] Soil organic matter content

[b] Cation exchange capacity

[c] Exchangeable aluminum

Table A2-2 Average percent AM fungal colonization (out of 100%) of switchgrass roots by different AM fungal structures, observed via root staining and microscopic observations.

Sites include AF14 - Adelphia Research Field 2014, AF17 - Adelphia Research Field 2017, CM - Colliers Mills, EC - EARTH Center, IO - Iowa, FAA - Federal Aviation Administration, LIPB14 - Long Island Pine Barrens 2014, LIPB16 - Long Island Pine Barrens 2016, RP - Rocky Point, SO - Somerset Research Field WSF - Wharton State Forest.

Site	Am[a]	Arbuscules	Coils	Vesicles	AM hyphae	Spores
AF14	33.61	6.03	9.84	1.86	15.35	0.53
AF17	6.50	0.83	3.50	0.17	2.00	0.00
CM	37.03	11.44	13.56	0.80	18.84	0.60
EC	20.17	5.86	2.63	0.10	11.58	0.00
FAA	3.00	0.30	0.40	0.20	2.20	0.00
IO	7.29	2.00	2.31	0.61	2.20	0.17
LIPB14	12.24	2.30	2.10	0.10	7.64	0.00
LIPB16	1.70	0.00	0.20	0.10	1.40	0.00
RP	3.40	1.00	1.10	0.00	1.30	0.00
SO	4.33	0.50	1.50	0.17	2.00	0.17
WSF	52.21	15.29	4.55	1.44	30.62	0.30

[a] All AM fungal structures found in the roots

Table A2-3 Operational taxonomic units (OTUs) of *Acaulospora* spp. and *Ambispora* spp. detected in this study.

OTU	Sites	Genus	GenBank Best Hit	Percent Identity	MaarjAM Best Hit	Percent Identity	Reads
CM1_1135	CM	Acaulospora	Acaulospora sp. (EU332727.1)	87	Acaulospora sp. (VTX00028)	94	87
CM1_15492	CM	Acaulospora	Acaulospora sp. (EU332732.1)	88	Acaulospora sp. (VTX00028)	98	15
CM1_182	CM	Ambispora	Archaeospora leptoticha (AB047306.1)	95	Ambispora leptoticha (VTX00242)	95	1
CM2_101561	CM	Acaulospora	Acaulospora sp. (AJ418884.1)	99	Acaulospora sp. (KC211936)	99	6
CM2_104704	CM, LIPB16	Acaulospora	Acaulospora lacunosa (HE610427.1)	98	Acaulospora sp. (VTX00024)	99	87
CM2_14643	CM	Ambispora	Ambispora leptoticha (AB047306.1)	97	Ambispora leptoticha (VTX00242)	97	4
CM2_152015	CM	Acaulospora	Acaulospora sp. (EU332732.1)	97	Acaulospora sp. (VTX00029)	97	7
CM2_367	FAA, CM	Acaulospora	Acaulospora sp. (EU332727.1)	93	Acaulospora sp. (VTX00028)	99	36
CM2_45291	CM	Acaulospora	Acaulospora sp. (EU332732.1)	97	Acaulospora sp. (VTX00029)	97	1
CM2_50062	CM	Acaulospora	Acaulospora lacunosa (HE610427.1)	95	Acaulospora lacunosa (VTX00024)	95	1
CM2_5058	CM, RP	Ambispora	Ambispora leptoticha (AB047306.1)	99	Ambispora leptoticha (VTX00242)	100	57
CM3_3478	CM	Acaulospora	Acaulospora sp. (EU332732.1)	97	Acaulospora sp. (VTX00028)	98	78
CM3_49126	CM	Ambispora	Archaeospora leptoticha (AB047309.1)	96	Ambispora leptoticha (VTX00242)	96	3
CM3_72559	CM	Acaulospora	Acaulospora sp. (EU332727.1)	98	Acaulospora sp. (VTX00028)	98	16
CM3_81523	CM	Acaulospora	Acaulospora sp. (EU332727.1)	87	Acaulospora sp. (VTX00029)	100	6
FAA1_133850	FAA	Acaulospora	Acaulospora sp. (AJ418884.1)	98	Acaulospora sp. (KC211936)	98	33
FAA1_959	FAA	Acaulospora	Acaulospora sp. (EU332727.1)	92	Acaulospora sp. (VTX00029)	92	132
FAA3_1044	FAA	Acaulospora	Acaulospora sp. (AJ418884.1)	98	Acaulospora sp. (KC211936)	98	1
FAA3_1501	FAA	Acaulospora	Acaulospora sp. (EU332732.1)	94	Acaulospora sp. (VTX00028)	95	12
FAA3_1979	FAA	Acaulospora	Acaulospora sp. (AJ418884.1)	97	Acaulospora sp. (KC211936)	98	28
FAA3_21109	FAA	Acaulospora	Acaulospora sp. (EU332732.1)	98	Acaulospora sp. (VTX00029)	98	21
FAA3_25	FAA	Acaulospora	Acaulospora sp. (EU332732.1)	95	Acaulospora sp. (VTX00029)	95	871
FAA3_32	FAA	Acaulospora	Acaulospora sp. (EU332727.1)	97	Acaulospora sp. (VTX00029)	97	812
FAA3_324	FAA, CM	Acaulospora	Acaulospora sp. (EU332727.1)	87	Acaulospora sp. (VTX00028)	98	708
FAA3_48	FAA	Acaulospora	Acaulospora sp. (EU332732.1)	97	Acaulospora sp. (VTX00029)	97	171

FAA3_500	FAA	Acaulospora	Acaulospora sp. (EU332727.1)	96	Acaulospora sp. (VTX00028)	97	132
FAA3_78949	FAA, CM	Acaulospora	Acaulospora sp. (AJ418884.1)	97	Acaulospora sp. (KC211936)	98	1
FAA3_94	IO, CM, FAA, WSF, RP	Acaulospora	Acaulospora sp. (EU332727.1)	100	Acaulospora sp. (VTX00029)	100	16,908
FAA3_9624	FAA, RP, CM, IO	Acaulospora	Acaulospora sp. (AJ418884.1)	99	Acaulospora sp. (KC211936)	100	477
IO2_130323	IO	Acaulospora	Acaulospora sp. (EU332727.1)	97	Acaulospora sp. (VTX00028)	98	2
LIPB162_25574	CM, LIPB16	Acaulospora	Acaulospora lacunosa (HE610427.1)	95	Acaulospora sp. (VTX00024)	97	19
LIPB162_32384	CM, LIPB16	Acaulospora	Acaulospora sp. (AY129616.1)	95	Acaulospora sp. (VTX00024)	97	17
LIPB163_292	CM, LIPB16	Acaulospora	Acaulospora lacunosa (HE610427.1)	97	Acaulospora sp. (VTX00024)	100	909
LIPB163_744	CM, LIPB16	Acaulospora	Acaulospora sp. (AY129616.1)	98	Acaulospora sp. (VTX00024)	100	669
RP2_1392	CM, FAA, RP	Acaulospora	Acaulospora sp. (EU332727.1)	94	Acaulospora sp. (VTX00028)	99	205
WSF1_265	WSF	Acaulospora	Glomeromycota sp. (JF414178.1)	99	Acaulospora sp. (VTX00328)	100	353
WSF2_9521	WSF	Acaulospora	Glomeromycota sp. (JF414178.1)	95	Acaulospora sp. (VTX00328)	96	10
WSF3_204765	WSF	Acaulospora	Glomeromycota sp. (JF414178.1)	94	Acaulospora sp. (VTX00328)	93	1
WSF3_230	WSF	Acaulospora	Acaulospora brasiliensis (FN825899.1)	100	Acaulospora sp. (KC211979.1)	100	277
WSF3_3283	FAA, WSF	Acaulospora	Acaulospora sp. (AJ418885.1)	97	Acaulospora sp. (VTX00028)	97	34
WSF3_3667	WSF	Acaulospora	Glomeromycota sp. (JF414178.1)	95	Acaulospora sp. (VTX00328)	95	14
WSF3_818	WSF, FAA	Acaulospora	Glomeromycota sp. (JF414178.1)	97	Acaulospora sp. (VTX00328)	97	528

Table A2-4 Soil properties correlated with arbuscular mycorrhizal (AM) fungal diversity.

Included are the number of operational taxonomic units (OTUs), both rarefied and observed values, diversity indices, and average percent colonization by microscopically observed and quantified AM fungal structures. Numbers represent Pearson's r values. Boldface values represent correlations > 0.30 and with significance of $P < 0.05$.

	OTUs (rarefied)	OTUs (observed)	Shannon	Simpson	Fisher	Reads (rarefied)	Reads (observed)	AM[b]	Arbuscules	Coils	Vesicles	AM Hyphae	Spores
pH	**0.394***	**0.351***	**0.614*****	**0.488****	**0.511****	-0.039	0.075	-0.16	-0.2	-0.072	0.109	-0.224	-0.062
NO_3	-0.227	-0.24	0.028	0.118	-0.099	-0.145	-0.33	**0.556****	**0.513****	0.091	**0.417***	**0.583*****	0.175
Al Exch[a]	**-0.625*****	**-0.600*****	**-0.587*****	**-0.568****	**-0.727*****	**-0.468****	**-0.386***	0.179	0.15	-0.091	-0.095	0.235	-0.227
Al Total	0.18	0.118	0.26	0.332	0.3	0.244	0.036	**-0.422***	**-0.418***	-0.339	-0.332	**-0.408***	-0.183
Extractable Ca	**0.412***	0.28	**0.635*****	**0.451****	**0.533****	-0.074	-0.167	-0.182	-0.147	-0.168	0.048	-0.221	-0.036
CEC	0.327	0.222	**0.698*****	**0.614*****	**0.493****	0.07	-0.058	-0.127	-0.072	-0.279	0.027	-0.148	-0.109
Fe	-0.081	0.01	0.339	**0.476****	0.091	0.155	0.182	**0.500****	**0.476****	-0.106	**0.487****	**0.549****	0.027
Organic Matter	0.116	0.015	**0.514****	**0.569****	0.206	-0.052	-0.217	**0.366***	0.182	0.337	**0.764*****	0.283	**0.487****
Rarefied Species Richness	1	**0.946*****	**0.525****	**0.447****	**0.809*****	**0.507****	0.308	0.078	0.07	0.33	0.273	0.003	**0.407***
Observed Species Richness	**0.946*****	1	**0.415***	**0.375***	**0.736*****	**0.542****	**0.546****	0.087	0.086	0.303	0.209	0.022	0.307
Shannon	**0.526****	**0.415***	1	**0.962*****	**0.726*****	0.277	0.08	-0.046	-0.088	-0.087	**0.345***	-0.097	0.141
Simpson	**0.447****	**0.375***	**0.962*****	1	**0.649*****	**0.357***	0.19	0.012	-0.04	-0.085	**0.370***	-0.035	0.134
Fisher	**0.809*****	**0.736*****	**0.726*****	**0.649*****	1	**0.374***	0.179	0.125	0.111	0.262	**0.373***	0.056	**0.424***

[a] Exchangeable aluminum

[b] Any AM fungal structures

* P < 0.05

** P < 0.01

*** P < 0.001

Table A3-1 Differences in arbuscular mycorrhizal (AM) fungal reads detected through Illumina MiSeq sequencing when using different workflow decisions.

Clustering thresholds (95% vs. 97%), reference databases (NCBI vs. AM database), and rare sequence removal (singletons, doubletons, neither). Samples include those from Colliers Mills (CM), Long Island Pine Barrens (LIPB), and Wharton State Forest (WSF). Raw (observed) reads and OTUs, as well as, rarefied reads and OTUs (rarefied to the minimum number of reads per sample) are included.

Clustering Threshold	Reference Database	Rare sequence removal	OTUs and Reads										Total
Sites				CM	CM	CM	LIPB	LIPB	LIPB	WSF	WSF	WSF	-
Replicates				1	2	3	1	2	3	1	2	3	9
95%	NCBI	N[a]	Raw Reads	49,538	72,413	62,697	8,601	7,023	24,028	74,517	18,044	116,978	433,839
			Rarified Reads	7,023	7,023	7,023	7,023	7,023	7,023	7,023	7,023	7,023	63,207
			Raw OTUs	88	102	97	47	39	26	53	25	88	565
			Rarified OTUs	61	57	66	47	39	17	27	18	42	374
		NS[b]	Raw Reads	39,754	60,363	51,250	6,709	5,157	20,097	65,202	14,418	101,423	364,373
			Rarified Reads	5,157	5,157	5,157	5,157	5,157	5,157	5,157	5,157	5,157	46,413
			Raw OTUs	33	38	41	12	8	6	28	12	41	219
			Rarified OTUs	23	24	32	12	8	4	20	11	29	163
		ND[c]	Raw Reads	36,388	54,962	47,395	5,434	4,281	18,091	59,067	12,938	91,074	329,630
			Rarified Reads	4,281	4,281	4,281	4,281	4,281	4,281	4,281	4,281	4,281	38,529

			Raw OTUs	22	28	26	8	6	4	17	11	30	152
			Rarified OTUs	20	19	21	8	6	3	15	11	23	126
		N[a]	Raw Reads	82,558	124,623	104,155	147,295	159,133	182,636	188,446	88,748	267,969	1,345,563
			Rarified Reads	82,558	82,558	82,558	82,558	82,558	82,558	82,558	82,558	82,558	743,022
			Raw OTUs	272	314	309	350	243	186	240	96	431	2,441
			Rarified OTUs	272	291	296	301	204	154	201	94	332	2,145
	MaarjAM	NS[b]	Raw Reads	64,476	101,816	83,713	126,317	136,901	158,476	161,608	76,514	230,318	1,140,139
			Rarified Reads	64,476	64,476	64,476	64,476	64,476	64,476	64,476	64,476	64,476	580,284
			Raw OTUs	119	146	138	157	112	94	151	61	260	1,238
			Rarified OTUs	119	143	135	145	108	86	143	61	234	1,174
		ND[c]	Raw Reads	59,052	92,876	77,514	116,532	128,389	148,217	146,896	72,746	209,176	1,051,398
			Rarified Reads	59,052	59,052	59,052	59,052	59,052	59,052	59,052	59,052	59,052	531,468
			Raw OTUs	82	100	89	95	83	72	111	47	187	866
			Rarified OTUs	82	100	89	95	83	72	111	47	187	866
		N[a]	Raw Reads	52,018	75,851	65,788	10,168	7,669	24,197	73,887	18,177	111,078	438,833
			Rarified Reads	7,669	7,669	7,669	7,669	7,669	7,669	7,669	7,669	7,669	69,021
			Raw OTUs	208	240	263	91	50	54	134	68	253	1,361
			Rarified OTUs	109	99	127	81	50	29	50	41	96	682
	NCBI	NS[b]	Raw Reads	41,696	62,964	53,916	7,923	5,804	20,253	62,439	14,393	93,706	363,094
			Rarified Reads	5,804	5,804	5,804	5,804	5,804	5,804	5,804	5,804	5,804	52,236
			Raw OTUs	50	78	70	18	9	9	45	17	90	386
97%			Rarified OTUs	37	46	37	18	9	6	23	16	52	244
		ND[c]	Raw Reads	38,630	58,253	50,410	6,936	5,043	18,367	56,497	12,921	84,323	331,380
			Rarified Reads	5,043	5,043	5,043	5,043	5,043	5,043	5,043	5,043	5,043	45,387
			Raw OTUs	27	41	34	11	7	5	23	15	46	209
			Rarified OTUs	22	22	25	11	7	5	21	14	37	164

	N[a]	Raw Reads	84,765	127,686	107,368	138,363	156,224	180,230	186,948	88,927	260,145	1,330,656
		Rarified Reads	84,765	84,765	84,765	84,765	84,765	84,765	84,765	84,765	84,765	762,885
		Raw OTUs	587	694	660	658	431	356	394	157	775	4,712
		Rarified OTUs	587	598	615	558	346	280	306	156	556	4,002
MaarjAM	NS[b]	Raw Reads	66,632	104,863	87,089	118,281	134,696	156,807	158,278	76,500	220,975	1,124,121
		Rarified Reads	66,632	66,632	66,632	66,632	66,632	66,632	66,632	66,632	66,632	599,688
		Raw OTUs	193	256	231	254	158	128	182	71	399	1,872
		Rarified OTUs	193	244	223	237	146	117	164	71	339	1,734
	ND[c]	Raw Reads	61,338	96,657	81,004	108,835	126,397	146,648	143,833	72,736	200,897	1,038,345
		Rarified Reads	61,338	61,338	61,338	61,338	61,338	61,338	61,338	61,338	61,338	552,042
		Raw OTUs	110	148	128	129	100	85	131	55	244	1,130
		Rarified OTUs	110	145	128	128	99	84	127	54	232	1,107

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

Table A3-2 Differences in arbuscular mycorrhizal (AM) fungal OTUs detected through Illumina MiSeq sequencing when different workflow decisions were made, using raw data.

Different clustering thresholds (95% vs. 97%), reference databases (NCBI vs. AM database), and rare sequence removal (singletons, doubletons, neither) were used. Colors represent different AM fungal orders.

Clustering Threshold				95%						97%					
Reference Database				NCBI			MaarjAM			NCBI			MaarjAM		
Rare Sequence Removal				N[a]	NS[b]	ND[c]	N	NS	ND	N	NS	ND	N	NS	ND
OTUs	Order	Family	Genus												
	Archaeosporales	Ambisporaceae	Ambispora	0	0	0	14	5	4	0	0	0	32	12	6
		Archaeosporaceae	Archaeospora	5	3	3	18	12	11	4	3	3	30	16	13
		Geosiphonaceae	Geosiphon	0	0	0	0	0	0	0	0	0	2	0	0
		N/A	Undescribed Archaeosporales	5	4	3	7	3	2	9	5	4	12	6	5
	Diversisporales	Acaulosporaceae	Acaulospora	29	16	10	33	22	13	54	22	15	74	31	19
		Claroideoglomeraceae	Claroideoglomus	1	0	0	24	15	13	1	0	0	32	16	13
		N/A	Undescribed Diversisporales	1	0	0	2	1	0	1	0	0	8	2	0
		Diversisporaceae	Diversispora	0	0	0	1	0	0	0	0	0	2	0	0
		Gigasporaceae	Undescribed Gigasporaceae	1	1	1	5	2	2	10	2	2	15	4	4
		Gigasporaceae	Gigaspora	1	1	1	0	0	0	1	1	1	0	0	0
		Gigasporaceae	Scutellospora	2	2	1	3	4	3	4	2	2	5	3	3
	Glomerales	Glomeraceae	Glomus	82	28	21	199	92	54	309	79	33	638	186	83
		Glomeraceae	Undescribed Glomeraceae	32	7	2	0	0	0	109	16	7	0	0	0
		N/A	Undescribed Glomerales	1	0	0	20	10	6	1	0	0	25	13	8

	Paraglomerales	Paraglomeraceae	Undescribed Paraglomeraceae	1	0	0	0	0	0	3	0	0	0	0	0
		Paraglomeraceae	Paraglomus	2	2	1	297	160	92	6	6	2	841	326	167
	N/A	N/A	Undescribed Glomeromycetes	14	10	6	352	201	150	40	13	8	601	298	188
	N/A	N/A	Undescribed Glomeromycota	24	17	7	0	0	0	102	36	10	0	0	0
Total OTUs	-	-	-	201	91	56	975	527	350	654	185	87	2317	913	509

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

Table A3-3 Differences in arbuscular mycorrhizal (AM) fungal reads detected through Illumina MiSeq sequencing when different workflow decisions were made, using raw data.

Different clustering thresholds (95% vs. 97%), reference databases (NCBI vs. AM database), and rare sequence removal (singletons, doubletons, neither) were used. Colors represent different AM fungal orders.

Clustering Threshold				95%						97%					
Reference Database				NCBI			MaarjAM			NCBI			MaarjAM		
Rare Sequence Removal				N[a]	NS[b]	ND[c]	N	NS	ND	N	NS	ND	N	NS	ND
Reads	Order	Family	Genus												
	Archaeosporales	Ambisporaceae	Ambispora	0	0	0	1,199	249	256	0	0	0	608	319	244
		Archaeosporaceae	Archaeospora	729	545	461	3,862	2,856	2,380	724	543	461	5,264	3,884	3,263
		Geosiphonaceae	Geosiphon	0	0	0	0	0	0	0	0	0	24	0	0
		N/A	Undescribed Archaeosporales	373	229	197	17,904	15,642	14,752	362	219	193	18,301	15,913	14,950
	Diversisporales	Acaulosporaceae	Acaulospora	12,251	8,580	7,319	14,245	10,280	8,598	10,203	7,104	6,066	13,835	9,710	8,121
		Claroideoglomeraceae	Claroideoglomus	1	0	0	4,985	3,974	3,369	1	0	0	4,942	3,951	3,363
		N/A	Undescribed Diversisporales	2	0	0	44	2	0	2	0	0	303	10	0
		Diversisporaceae	Diversispora	0	0	0	1	0	0	0	0	0	2	0	0
		Gigasporaceae	Undescribed Gigasporaceae	2,054	1,700	1,508	2,077	1,710	1,516	2,160	1,742	1,524	2,193	1,757	1,535
		Gigasporaceae	Gigaspora	18	10	8	0	0	0	16	10	8	0	0	0

	Gigasporaceae	Scutellospora	171	87	51	1,088	832	686	57	31	27	972	771	659
	Glomeraceae	Glomus	276,351	231,558	213,891	437,086	367,777	333,088	238,998	198,308	180,393	441,432	365,907	334,204
Glomerales	Glomeraceae	Undescribed Glomeraceae	38,504	29,001	21,403	0	0	0	91,882	74,610	68,677	0	0	0
	N/A	Undescribed Glomerales	2	0	0	2,059	1,600	1,324	1	0	0	2,396	1,748	1,370
	Paraglomeraceae	Undescribed Paraglomeraceae	71	0	0	0	0	0	58	0	0	0	0	0
Paraglomerales	Paraglomeraceae	Paraglomus	230	159	129	586,203	508,353	477,567	239	161	122	572,829	498,075	466,912
N/A	N/A	Undescribed Glomeromycetes	20,807	16,499	14,567	274,810	226,864	207,862	22,021	17,257	14,904	267,555	222,076	203,724
N/A	N/A	Undescribed Glomeromycota	82,275	76,005	70,096	0	0	0	72,109	63,109	59,005	0	0	0
Total Reads	-	-	-	433,839	364,373	329,630	1,345,563	1,140,139	1,051,398	438,833	363,094	331,380	1,330,656	1,124,121	1,038,345

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

Table A3-4 Differences in arbuscular mycorrhizal (AM) fungal OTUs detected in this study through Illumina MiSeq sequencing when different workflow decisions were made, using rarefied data.

Different clustering thresholds (95% vs. 97%), reference databases (NCBI vs. AM database), and rare sequence removal (singletons, doubletons, neither) were used. Colors represent different AM fungal orders.

Clustering Threshold				95%						97%					
Reference Database				NCBI			MaarjAM			NCBI			MaarjAM		
Rare Sequence Removal				N[a]	NS[b]	ND[c]	N	NS	ND	N	NS	ND	N	NS	ND
OTUs	Order	Family	Genus												
	Archaeosporales	Ambisporaceae	Ambispora	0	0	0	12	5	4	0	0	0	24	11	6
		Archaeosporaceae	Archaeospora	3	3	3	17	12	11	3	3	3	25	16	13
		Geosiphonaceae	Geosiphon	0	0	0	0	0	0	0	0	0	1	0	0
		N/A	Undescribed Archaeosporales	4	4	3	5	3	2	7	3	3	11	4	5
	Diversisporales	Acaulosporaceae	Acaulospora	16	12	7	30	22	13	28	13	9	68	29	19
		Claroideoglomeraceae	Claroideoglomus	0	0	0	19	15	13	0	0	0	25	16	13
		N/A	Undescribed Diversisporales	0	0	0	2	1	0	0	0	0	6	2	0

		Diversisporaceae	Diversispora	0	0	0	0	0	0	0	0	0	1	0	0
		Gigasporaceae	Undescribed Gigasporaceae	1	1	1	5	2	2	2	2	2	14	4	4
		Gigasporaceae	Gigaspora	0	1	0	0	0	0	1	1	1	0	0	0
		Gigasporaceae	Scutellospora	1	1	1	3	4	3	1	1	1	5	2	3
		Glomeraceae	Glomus	57	20	16	172	82	50	156	46	27	532	171	82
	Glomerales	Glomeraceae	Undescribed Glomeraceae	24	1	2	0	0	0	62	11	4	0	0	0
		N/A	Undescribed Glomerales	1	0	0	18	10	6	1	0	0	21	12	8
	Paraglomerales	Paraglomeraceae	Undescribed Paraglomeraceae	1	0	0	0	0	0	2	0	0	0	0	0
		Paraglomeraceae	Paraglomus	1	2	1	262	147	89	2	4	1	709	290	162
	N/A	N/A	Undescribed Glomeromycetes	18	8	6	308	190	150	19	9	8	496	274	185
	N/A	N/A	Undescribed Glomeromycota	13	8	5	0	0	0	53	22	8	0	0	0
Total OTUs	-	-	-	140	61	45	853	493	343	337	115	67	1938	831	500

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

Table A3-5 Differences in arbuscular mycorrhizal (AM) fungal reads detected through Illumina MiSeq sequencing when different workflow decisions were made, using rarefied data.

Different clustering thresholds (95% vs. 97%), reference databases (NCBI vs. AM database), and rare sequence removal (singletons, doubletons, neither) were used. Colors represent different AM fungal orders.

Clustering Threshold				95%						97%					
Reference Database				NCBI			MaarjAM			NCBI			MaarjAM		
Rare Sequence Removal				N[a]	NSb	ND[c]	N	NS	ND	N	NS	ND	N	NS	ND
Reads	Order	Family	Genus												
	Archaeosporales	Ambisporaceae	Ambispora	0	0	0	639	134	145	0	0	0	380	175	138
		Archaeosporaceae	Archaeospora	65	34	36	1,698	1,159	906	67	49	47	2,427	1,658	1,399
		Geosiphonaceae	Geosiphon	0	0	0	0	0	0	0	0	0	19	0	0
		N/A	Undescribed Archaeosporales	58	36	15	6,358	5,033	4,705	58	34	26	6,914	5,629	5,230
	Diversisporales	Acaulosporaceae	Acaulospora	1,250	759	611	9,495	6,221	5,235	1,123	708	561	9,013	5,824	4,918
		Claroideoglomeraceae	Claroideoglomus	0	0	0	1,943	1,426	1,240	0	0	0	2,028	1,502	0
		N/A	Undescribed Diversisporales	0	0	0	23	1	0	0	0	0	162	4	0
		Diversisporaceae	Diversispora	0	0	0	0	0	0	0	0	0	1	0	0
		Gigasporaceae	Undescribed Gigasporaceae	211	146	98	944	709	647	223	167	144	1,053	758	674
		Gigasporaceae	Gigaspora	0	2	0	0	0	0	1	1	1	0	0	0

	Gigasporaceae	Scutellospora	18	10	6	392	276	202	9	1	3	349	240	232
	Glomeraceae	Glomus	43,732	32,839	28,288	254,374	200,189	181,734	38,201	29,146	25,322	265,432	208,285	190,406
Glomerales	Glomeraceae	Undescribed Glomeraceae	8,615	5,688	3,661	0	0	0	19,760	14,863	12,975	0	0	0
	N/A	Undescribed Glomerales	2	0	0	814	531	437	1	0	0	968	632	489
Paraglomerales	Paraglomeraceae	Undescribed Paraglomeraceae	12	0	0	0	0	0	7	0	0	0	0	0
	Paraglomeraceae	Paraglomus	11	12	7	298,206	235,551	218,605	14	10	2	304,976	243,747	226,761
N/A	N/A	Undescribed Glomeromycetes	2,018	1,361	1,108	168,136	129,054	117,612	2,436	1,664	1,323	169,163	131,234	120,478
N/A	N/A	Undescribed Glomeromycota	7,215	5,526	4,699	0	0	0	7,121	5,593	4,983	0	0	0
Total reads			63,207	46,413	38,529	743,022	580,284	531,468	69,021	52,236	45,387	762,885	599,688	550,725

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

Table A3-6 Shannon, Simpson, and Fisher's α diversity indices resulting from all workflow decisions, using raw data.

The 12 workflow decisions include: different clustering thresholds (95% vs. 97%), reference databases (NCBI vs. AM database), and rare sequence removal (singletons, doubletons, neither). Raw read data was used.

Clustering Threshold	Reference Database	Rare sequence removal	Diversity Indices	Site		
				Colliers Mills	Long Island Pine Barrens	Wharton State Forest
95%	NCBI	N[a]	Shannon	1.33	0.78	1.98
			Simpson	0.52	0.36	0.78
			Fisher's α	12.85	7.46	10.73
		NS[b]	Shannon	0.99	0.62	1.89
			Simpson	0.43	0.30	0.76
			Fisher's α	4.51	1.39	5.27
		ND[c]	Shannon	0.88	0.47	1.83
			Simpson	0.39	0.23	0.75
			Fisher's α	2.67	0.97	3.44
	MaarjAM	N	Shannon	2.58	2.04	3.12
			Simpson	0.81	0.69	0.90
			Fisher's α	47.44	46.18	55.11
		NS	Shannon	2.23	1.85	2.99
			Simpson	0.77	0.65	0.89
			Fisher's α	20.04	19.99	32.82
		ND	Shannon	2.12	1.75	2.90
			Simpson	0.75	0.63	0.89
			Fisher's α	11.74	12.67	22.75
	NCBI	N	Shannon	1.42	1.06	2.33
			Simpson	0.62	0.54	0.83
			Fisher's α	43.17	16.73	35.09
		NS	Shannon	1.19	0.84	2.16
			Simpson	0.59	0.50	0.81
			Fisher's α	9.57	2.41	11.22
		ND	Shannon	1.11	0.80	2.07
			Simpson	0.57	0.49	0.80

			Fisher's α	4.10	1.29	5.36
97%						
		N	Shannon	2.68	2.21	3.34
			Simpson	0.84	0.71	0.91
			Fisher's α	127.79	102.61	111.32
	MaarjAM	NS	Shannon	2.38	1.99	3.17
			Simpson	0.82	0.67	0.90
			Fisher's α	38.21	33.35	51.25
		ND	Shannon	2.26	1.87	3.06
			Simpson	0.82	0.65	0.90
			Fisher's α	19.08	17.63	30.56

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

Table A3-7 Shannon, Simpson, and Fisher's α diversity indices resulting from all workflow decisions, using raw data.

The 12 workflow decisions include: different clustering thresholds (95% vs. 97%), reference databases (NCBI vs. AM database), and rare sequence removal (singletons, doubletons, neither). Rarified read data was used.

Clustering Threshold	Reference Database	Rare sequence removal	Diversity Indices	Site		
				Colliers Mills	Long Island Pine Barrens	Wharton State Forest
95%	NCBI	N[a]	Shannon	1.33	1.00	1.94
			Simpson	0.52	0.47	0.79
			Fisher's α	10.52	7.71	7.28
		NS[b]	Shannon	0.98	0.84	1.88
			Simpson	0.43	0.42	0.78
			Fisher's α	4.41	1.52	4.69
		ND[c]	Shannon	0.86	0.68	1.83
			Simpson	0.38	0.35	0.77
			Fisher's α	2.58	1.06	3.82
	MaarjAM	N	Shannon	2.58	2.02	2.93
			Simpson	0.81	0.68	0.88
			Fisher's α	46.63	43.35	47.58
		NS	Shannon	2.23	1.83	2.81
			Simpson	0.77	0.64	0.87
			Fisher's α	20.42	20.55	32.70
		ND	Shannon	2.12	1.72	2.73

			Simpson	0.75	0.62	0.87
			Fisher's α	11.98	13.73	24.41
		N	Shannon	1.43	1.18	2.24
			Simpson	0.62	0.59	0.83
			Fisher's α	26.09	15.47	18.04
	NCBI	NS	Shannon	1.17	0.96	2.09
			Simpson	0.59	0.55	0.81
			Fisher's α	7.93	2.48	8.08
		ND	Shannon	1.10	0.91	2.01
			Simpson	0.57	0.55	0.80
			Fisher's α	3.46	1.40	5.85
97%						
		N	Shannon	2.68	2.20	3.12
			Simpson	0.84	0.70	0.89
			Fisher's α	120.89	89.83	89.11
	MaarjAM	NS	Shannon	2.38	1.96	2.97
			Simpson	0.82	0.67	0.88
			Fisher's α	38.32	34.12	47.72
		ND	Shannon	2.25	1.85	2.88
			Simpson	0.82	0.64	0.88
			Fisher's α	19.56	19.07	32.26

[a] no singleton sequences removed

[b] singleton sequences removed

[c] singleton and doubleton sequences removed

Table A3-8 Twenty select arbuscular mycorrhizal (AM) fungal operational taxonomic units (OTUs) detected after using 95% clustering with singleton and doubleton removal.

Columns show how the MaarjAM and NCBI database queries found different matches for some OTUs. Included differences are: percent query coverage, identity matches, whether or not the BLAST results listed were all the same or different, the top BLAST hit, the abundance of each of the OTUs (number of reads), the reference sequence length, and how each database identified the OTUs.

Reference Database		MaarjAM						NCBI					
OTU ID	Raw reads	Identification	Top hit	Query coverage (%)	Identity match (%)	Different BLAST results?	Ref. seq. size (bp)	Identification	Top hit	Query coverage (%)	Identity match (%)	Different BLAST results?	Ref. seq. size (bp)
CM2_24191	51	Scutellospora projecturata	Scutellospora projecturata	100	100	same	506	Scutellospora	Scuttelospora	100	99	same	506
WSF3_42	45,355	Paraglomus occultum	Paraglomus occultum	94	84	diff	485	Dikarya	Basidiomycota	100	98	diff	258
LIPB143_162	991	Paraglomus laccatum	Paraglomus laccatum	29	88	same	520	Amoebozoa	Amoebozoa	100	90	diff	749
CM2_98	7,572	Paraglomus	Paraglomus	100	88	diff	520	Dikarya	Odonticium	100	94	diff	1,771
CM2_373	499	Paraglomus	Paraglomus	68	90	diff	798	Dikarya	Bourdotia	100	93	diff	1,486
LIPB143_2777	1,803	Paraglomus	Glomus	53	90	same	522	Umbelopsis	Umbelopsis	100	100	same	1,115
CM1_15586	7	Paraglomus	Paraglomus	76	91	diff	798	Dikarya	Talbotiomyces	98	91	diff	1,134
LIPB142_37	4,093	Paraglomus	Paraglomus	68	93	diff	798	Dikarya	Bourdotia	100	93	diff	1,486
LIPB143_21	5,056	Glomeromycetes	Archaeospora	73	88	diff	518	Dikarya	Powellomycetaceae	99	97	diff	1,664
CM1_31250	105	Glomeromycetes	Archaeospora	80	85	diff	795	Dikarya	Camptobasidium	94	97	diff	1,581
WSF2_3	82,565	Glomeromycetes	Paraglomus	51	92	diff	798	Dikarya	Rhodotorula	100	100	diff	1,613
WSF1_1229	14,389	Glomeromycetes	Paraglomus	52	93	diff	798	Dikarya	Bensingtonia	100	99	diff	2,623
CM2_172	3,915	Glomeromycetes	Archaeospora	69	88	diff	795	Dikarya	Phialophora	100	98	diff	507
CM2_3	131,186	Glomus	Glomus	100	100	same	522	Glomus	Rhizophagus	100	97	diff	838
CM3_7284	42	Glomus	Glomus	54	93	same	489	Chytridiomycetes	Catenomyces	99	93	diff	1,767
WSF1_101	1,592	Archaeospora	Paraglomus	57	92	diff	520	Chytridiales	Neokarlingia	78	93	diff	1,663
WSF3_7231	327	Archaeospora	Archaeospora	72	89	same	502	Chytriomyces	Chytriomyces	100	90	diff	1,733
CM2_5058	120	Ambispora leptoticha	Ambispora leptoticha	100	100	same	1,785	Archaeosporales	Archaeospora	100	99	diff	1,794
CM3_58	4,989	Acaulospora	Acaulospora	100	100	same	505	Acaulospora	Acaulospora	100	100	same	742
CM2_5200	102	Acaulospora	Acaulospora	100	100	same	435	Acaulospora	Acaulospora	100	99	same	714

Fig. A2-1 Detailed information regarding sampling sites and sampling protocol

Panicum virgatum roots and root-abiding soil was collected in the summer months of 2014, 2016, and 2017 from: the Pine Barrens forests of New Jersey, Wharton State Forest (°39.7557667, °-74.6947333) and Colliers Mills (°40.0680667, °-74.4449333), Pine Barrens forests of Long Island, NY, Rocky Point Pine Barrens Preserves (°40.907136, °-72.916676) and David A. Sarnoff Pine Barrens Preserve (°40.8975333, °-72.6586500), forest floors of the Federal Aviation Administration of New Jersey (°39.4443089, °-74.5632541), and 3 switchgrass research plots in New Jersey: Switchgrass research field plot in Somerset (°40.4748833, °-74.5310333), Rutgers Plant Science Research and Extension Farm in Adelphia (°40.2279333, °-74.2517000), and E.A.R.T.H Center at the Rutgers Cooperative Extension of Middlesex County in North Brunswick (°39.8787667, °-74.3767167) were also sampled in order to compare AM colonization and community structure from both natural and managed switchgrass populations. The NJ Pine Barrens forests have canopies dominated by pitch pine trees (*Pinus rigida*) and several oak species (*Quercus* sp.) and understories of shrubby oaks, heath plants, and an herbaceous ground layer of grasses and sedges (McCormick and Buell 1957; Forman 1998). The NJ Pine Barrens' are the largest of its kind worldwide, stretching ~1.4 million acres (Forman 1998). The Long Island Pine Barrens (LIPB) forests are similar in plant and soil makeup but we noticed more lichen, moss, and wintergreen present on the forest floors during sampling. The LIPB forests make up ~105,000 acres of the island (Central Pine Barrens Overview 2018).

Samples were also collected from Doolittle Prairie State Preserve in Ames, Iowa. This site is a 40-acre preserved prairie wetland with 223 native plant species, no trails, and is prescribed burned every 2-3 years. Sites were considered 'managed' if the soil was amended with fertilizers in any way and/or mowed annually. Sites were considered 'natural' if not managed, as we define it. Root samples were collected from similarly sized plants. However, the plants at the Adelphia Research Field were much larger than elsewhere, most likely because they were specially cared for when the plot was originally established and because it has full sun exposure. Other grass species were grown adjacent to switchgrass in these research field plots.

Fig. A2-2 Protocols used to confirm proper plant host via sequencing of plant leaf DNA.

In order to confirm our identification of *P. virgatum*, the leaf sheath from one *P. virgatum* plant from each sampled site was used for DNA analysis. Total genomic DNA was extracted from 0.125g of leaf tissue using a PowerSoil DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, CA) according to manufacturer's instructions. PCR was carried out with primer pairs MATKF-MATKR and rbcLAF-rbcLAR, using the protocols of Hollingsworth et al. 2009. PCR products were purified with ExoSAP-IT (Affymetrix, California) as per manufacturer's instructions and sequenced by GenScript (Piscataway, New Jersey). The sequence was manually checked and trimmed for quality using Finch TV (PerkinElmer Inc., Seattle, Washington) and compared against a known database in the database for the National Center for Biotechnology Information using the BLASTn function (http://blast.ncbi.nlm.nih.gov).

Fig. A2-3. Protocols for staining switchgrass roots for the microscopic observation and quantification of AM fungi.

Ten to fifteen root segments (1-2 cm each) from each site were stained with 0.05% aniline blue following a modified version of the procedures of Grace and Stribley (1991). Roots were first cleared in 3% KOH for 45 minutes in an 80°C water bath, rinsed with 3-4 changes of water, and acidified in 1% HCl for 15 minutes at room temperature. Roots were then immersed in 0.05% aniline blue stain at 80°C for 25 minutes and destained with 85% lactic acid or 25% glycerol for 15 minutes and stored in 25% glycerol until microscopic observation.

Fig. A3-1 Annotated bioinformatic workflow.

Annotated bioinformatic workflow

All steps were performed using Perl and bash scripts on the Linux platform, with the following additional programs: USEARCH v. 8.0, USEARCH v. 9.0 (Edgar 2010), BLAST+ v. 2.2.28 (Camacho et al. 2009), BBMap v.36 (Bushnell 2016),

All reads were obtained through Illumina MiSeq sequencing (v3 cycle 600)

Databases used were those of NCBI nucleotide database and MaarjAM AM fungal database

MaarjAM 18S virtual taxon sequence database (MaarjAM_18s_sequences_long.txt) downloaded from http://maarjam.botany.ut.ee and converted to .udb database file (18s_AM_db.udb) using 'makeblastdb' command using Blast+. This database is used for chimeric checking.

NCBI full nucleotide sequence database accessed through ftp://ftp.ncbi.nlm.nih.gov/blast/db/: ncbi_nt_database.nal

MaarjAM type virtual taxon sequence database, downloaded from http://maarjam.botany.ut.ee: MaarjAMdb.txt. This database is used for queries of template sequences for identity matches in known database.

##
############
FASTA header names were edited from their original name (overly lengthy numbers and letters) to the proper sample name (i.e. CM1) with the BBMap package within the BBTools bioinformatic program suite (Bushnell 2016)
##
############
./rename.sh program within BBMap package that renames reads to a <prefix of your choosing>_<ordered number>
seqs.fa is original fasta file of sequences
seqsC.fa is fasta file of all sequences with edited headers
CM1 represents the new header name, bearing sample name information

./rename.sh in=seqs.fa out=seqsC.fa prefix=CM1

Note: this step was done separately for all 12 Illumina DNA samples.

##
############
##Create ISUs (Individual Sequence Units), dereplicate all sequences using USEARCH (Edgar 2010)
##
############
-fastx_uniques finds the unique set of sequences within input file
seqsC.fa is fasta file of all sequences, from previous step

-fastaout option specifies that fasta output file list sequences in descending order of abundance
unique.fa is fasta output file
-sizeout option specifies that the abundance of each unique read should be printed onto the header line as 'size='

./usearch9.0 -fastx_uniques seqsC.fa -fastaout unique.fa -sizeout

##
Concatenate all dereplicated fasta sequence files into one file
##
cat concatenates all fasta input files
*.fa represents all the fasta input files in the directory
unique_con.fa is the concatenated output fasta file

cat *.fa > unique_con.fa

##
Sort reads by size, remove singletons/doubletons/keep all rare sequences
##
-sortbysize sorts sequences in input file into descending order of abundance (size)
unique_con.fa is the input file (output from previous step)
-fastaout creates fasta output file
sortedN.fa is the sorted fasta output file, with all rare sequences included
-minsize option specifies a minimum abundance (size) for reads to continue to next step in workflow
sortedNS.fa is the sorted fasta output file with singleton reads removed
sortedND.fa is the sorted fasta output file with singleton and doubleton reads removed

./usearch9.0 -sortbysize unique_con.fa -fastaout sortedN.fa

./usearch9.0 -sortbysize unique_con.fa -fastaout sortedNS.fa -minsize 2

./usearch9.0 -sortbysize unique_con.fa -fastaout sortedND.fa -minsize 3

##
Cluster reads at 95%, 97%, and 99% similarity, make OTU abundance table
##
-cluster_otus clusters OTUs (operational taxonomic units)

sorted.fa is 1 of 3 fasta input files (output from previous step)
-otus specifies a fasta output file with header names taken directly from input file headers and abundance annotations removed
clustered95.fa output file with OTU's clustered at 95%
-uparseout option creates a tabbed text output file detailing how the input sequences were clustered
Notu95.up is a tabbed text output file
otu_radius_pct option specifies a clustering radius (percent different) other than the default of 3% (97% similar) (i.e. otu_radius_pct 5 for 95% similarity)

./usearch8.0 -cluster_otus sorted.fa -otus clustered95.fa -uparseout Notu95.up otu_radius_pct 5.0

./usearch8.0 -cluster_otus sorted.fa -otus clustered97.fa -uparseout Notu97.up

Note: Repeat these steps for sortedNS.fa and sorted sortedND.fa

##
Check for chimeric sequences against the 18S AM fungal sequences in the MaarjAM database (Öpik et al. 2010) using the uchim2_ref command (Edgar 2016)
##
-uchime2_ref detects chimeric sequences
clustered97.fa is the input file (output file from previous step)
-db option specifies a database file
18s_AM_db.udb is the MaarjAM database file
-notmatched option specifies sequences that did not match up to the sequences of the input database
nonchim97.fa is the output file
-mode specific option reports the non-chimeric sequence predictions with a lower false negative rate
-strand plus option specifies searching for database hits on the forward DNA strand only

./usearch9.0 -uchime2_ref clustered97.fa -db 18s_AM_db.udb -notmatched nonchim97.fa -mode specific -strand plus

Note: Repeat these steps for clustered97NS.fa, clustered97ND.fa, clustered95.fa, clustered95NS.fa, clustered95ND.fa, clustered99.fa, clustered99NS.fa, and clustered99ND.fa

##
############
BLAST sequences against known database (NCBI and MaarjAM) using BLAST+ (Camacho et al. 2009)
##
############
blastn specifies using BLAST+ program
-query designates the queried file
nonchim97.fa is the input file (output file from previous step)
-task megablast option optimizes the query parameters for intraspecies comparison, with large word sizes
-db option specifies a database file
ncbi_nt_database is the NCBI database file
-out creates output file
ncbi97blast.blastn is the output file
-outfmt 0 option specifies pairwise alignment
-num_threads 8 option specifies using 8 CPUs (threads) in BLAST search
-num_alignments 10 option specifies output to show 10 database sequence alignments
-num_descriptions 10 option specifies showing single line descriptions for 10 database sequences
MaarjAMdb is the MaarjAM type sequence database file

blastn -query nonchim97.fa -task megablast -db ncbi_nt_database -out ncbi97blast.blastn -outfmt 0 -num_threads 8 -num_alignments 10 -num_descriptions 10

blastn -query nonchim97.fa -task megablast -db MaarjAMdb -out ncbi97blast.blastn -outfmt 0 -num_threads 8 -num_alignments 10 -num_descriptions 10

Note: after completing all BLASTs, BLAST files and corresponding non-chimera files are imported into MEGAN Community Edition, v. 6.7 (Huson et al. 2016) and the following Perl scripts are run.

Fig. A3-2 Perl scripts.

```
##############################################################################
############
## Script: Cluster_sum_multiple.pl
## Function: Links OTU header name with associated abundance (number of reads)
## Usage: perl Cluster_sum_multiple.pl >cluster_sum_multiple_output.txt
##############################################################################
############

use strict;
use warnings;

my $outfile = "Notu95.up";
my (%cluster, %data);

open FILEB, $outfile;
while (<FILEB>) {
   chomp;
   my @cols1 = split; #print "$cols[1]\t";
   my ($read1) = $cols1[0] =~ /(\S+);\S+;$/; #print "$read1\t";
   my ($size1) = $cols1[0] =~ /(\d+);$/; #print "$size1\t";

   my @array = (0, 0, 0, 0, 0, 0, 0, 0, 0);

   if ($cols1[1] =~ /otu/) {
      $cluster{$read1} = [@array];
      if ($cols1[0] =~ /CM1/) {
         $cluster{$read1}[0] = $size1;
      } elsif ($cols1[0] =~ /CM2/) {
         $cluster{$read1}[1] = $size1;
      } elsif ($cols1[0] =~ /CM3/) {
         $cluster{$read1}[2] = $size1;
      } elsif ($cols1[0] =~ /LIPB141/) {
         $cluster{$read1}[3] = $size1;
      } elsif ($cols1[0] =~ /LIPB142/) {
         $cluster{$read1}[4] = $size1;
      } elsif ($cols1[0] =~ /LIPB143/) {
         $cluster{$read1}[5] = $size1;
      } elsif ($cols1[0] =~ /WSF1/) {
         $cluster{$read1}[6] = $size1;
      } elsif ($cols1[0] =~ /WSF2/) {
         $cluster{$read1}[7] = $size1;
      } elsif ($cols1[0] =~ /WSF3/) {
         $cluster{$read1}[8] = $size1;  #print "$reads\t";
      }
   }
```

```
}
close FILEB;

#foreach (keys %cluster) {
#       print "$_\t$cluster{$_}[0]\t$cluster{$_}[1]\t$cluster{$_}[2]\n";
#}

open FILEB, $outfile;
while (<FILEB>) {
   chomp;
   my @cols2 = split; #print "$cols[1]\t";
   next if $#cols2 != 4;
#    my ($read2) = $cols2[0] =~ /(\S+);\S+;$/; #print "$read2\t";
   my ($size2) = $cols2[0] =~ /(\d+);$/; #print "$size2\n";

   my @array = (0, 0, 0, 0, 0, 0, 0, 0, 0);

   if ($cols2[1] =~ /match/) {
      $data{$cols2[4]} = [@array];
   }
}
close FILEB;

open FILEB, $outfile;
while (<FILEB>) {
   chomp;
   my @cols2 = split; #print "$cols[1]\t";
   next if $#cols2 != 4;
#    my ($read2) = $cols2[0] =~ /(\S+);\S+;$/; #print "$read2\t";
   my ($size2) = $cols2[0] =~ /(\d+);$/; #print "$size2\n";

   if ($cols2[1] =~ /match/) {
      if (exists $data{$cols2[4]}) {
         if ($cols2[0] =~ /CM1/) {
            $data{$cols2[4]}[0] += $size2;
         } elsif ($cols2[0] =~ /CM2/) {
            $data{$cols2[4]}[1] += $size2;
         } elsif ($cols2[0] =~ /CM3/) {
            $data{$cols2[4]}[2] += $size2;
         } elsif ($cols2[0] =~ /LIPB141/) {
            $data{$cols2[4]}[3] += $size2;
         } elsif ($cols2[0] =~ /LIPB142/) {
            $data{$cols2[4]}[4] += $size2;
         } elsif ($cols2[0] =~ /LIPB143/) {
            $data{$cols2[4]}[5] += $size2;
         } elsif ($cols2[0] =~ /WSF1/) {
```

```
            $data{$cols2[4]}[6] += $size2;
          } elsif ($cols2[0] =~ /WSF2/) {
            $data{$cols2[4]}[7] += $size2;
          } elsif ($cols2[0] =~ /WSF3/) {
            $data{$cols2[4]}[8] += $size2;
          }
        } else {
          if ($cols2[0] =~ /CM1/) {
            $data{$cols2[4]}[0] = $size2;
          } elsif ($cols2[0] =~ /CM2/) {
            $data{$cols2[4]}[1] += $size2;
          } elsif ($cols2[0] =~ /CM3/) {
            $data{$cols2[4]}[2] += $size2;
          } elsif ($cols2[0] =~ /LIPB141/) {
            $data{$cols2[4]}[3] += $size2;
          } elsif ($cols2[0] =~ /LIPB142/) {
            $data{$cols2[4]}[4] += $size2;
          } elsif ($cols2[0] =~ /LIPB143/) {
            $data{$cols2[4]}[5] += $size2;
          } elsif ($cols2[0] =~ /WSF1/) {
            $data{$cols2[4]}[6] += $size2;
          } elsif ($cols2[0] =~ /WSF2/) {
            $data{$cols2[4]}[7] += $size2;
          } elsif ($cols2[0] =~ /WSF3/) {
            $data{$cols2[4]}[8] += $size2;
          }
        }
    }
}

#foreach (keys %data) {
#       print "$_\t$data{$_}[0]\t$data{$_}[1]\t$data{$_}[2]\n";
#}

foreach my $key2 (keys %data) {
    if (exists $cluster{$key2}) {
        $cluster{$key2}[0] += $data{$key2}[0];
        $cluster{$key2}[1] += $data{$key2}[1];
        $cluster{$key2}[2] += $data{$key2}[2];
        $cluster{$key2}[3] += $data{$key2}[3];
        $cluster{$key2}[4] += $data{$key2}[4];
        $cluster{$key2}[5] += $data{$key2}[5];
        $cluster{$key2}[6] += $data{$key2}[6];
        $cluster{$key2}[7] += $data{$key2}[7];
        $cluster{$key2}[8] += $data{$key2}[8];
    } else {
```

```
        $cluster{$key2}[0] += $data{$key2}[0];
        $cluster{$key2}[1] += $data{$key2}[1];
        $cluster{$key2}[2] += $data{$key2}[2];
        $cluster{$key2}[3] += $data{$key2}[3];
        $cluster{$key2}[4] += $data{$key2}[4];
        $cluster{$key2}[5] += $data{$key2}[5];
        $cluster{$key2}[6] += $data{$key2}[6];
        $cluster{$key2}[7] += $data{$key2}[7];
        $cluster{$key2}[8] += $data{$key2}[8];
    }
}

foreach (keys %cluster) {
    print
"$_\t$cluster{$_}[0]\t$cluster{$_}[1]\t$cluster{$_}[2]\t$cluster{$_}[3]\t$cluster{$_}[4]\
t$cluster{$_}[5]\t$cluster{$_}[6]\t$cluster{$_}[7]\t$cluster{$_}[8]\n";
}

close FILEB;


### Note: Input file for this script is the *.up file created at the clustering step.
### Note: Repeat this script for all 6 .up files (Notu95.up, NSotu95.up, NDotu95.up,
Notu97.up, NSotu97.up, NDotu97.up).


##############################################################################
############
## Script: Link_megan_cluster.pl
## Function: Links output from cluster_sum script with classification output from
MEGAN
## Usage: perl link_megan_cluster.pl >link_megan_cluster_output.txt
##############################################################################
############

use strict;
use warnings;

my $file1 = "Cluster_sum_multiple_outputN95.txt";
my $file2 = "NcbiN95blast-ex.txt";

my %hash1;
my %hash2;

open FILEA, $file1;
while (<FILEA>) {
```

```
   chomp;
   my ($cola1, $cola2) = split(/\t+/, $_, 2); #print "$cola1\t";
   $hash1{$cola1} = $cola2; #print "$hash1{$cola1}\n";
}

open FILEB, $file2;
while (<FILEB>) {
   chomp;
   my ($colb1, $colb2) = split(/\t+/, $_, 2);
   $hash2{$colb1} = $colb2; #print "$hash2{$colb1}\n";
}

foreach my $key2 (keys %hash2) {
   #print "$key2\n";
   if (exists $hash1{$key2}) {
      $hash1{$key2} .= "\t$hash2{$key2}";
      print "$key2\t$hash1{$key2}\n";
      #print "$_\n";
   }
}

foreach (keys %hash1) {
```

Note: $file1 input for this script is the output from the cluster_sum_multiple script.
Note: $file2 input for this script is the classification out from MEGAN. File must be two columns, first with OTU_ID and second with classification information.
Note: Repeat this script for all 12 classification files (Nncbi95blast-ex.txt, NSncbi95blast-ex.txt, NDncbi95blast-ex.txt, Nncbi97blast-ex.txt, NSncbi97blast-ex.txt, NDncbi97blast-ex.txt, Nmaarj95blast-ex.txt, NSmaarj95blast-ex.txt, NDmaarj95blast-ex.txt, Nmaarj97blast-ex.txt, NSmaarj97blast-ex.txt, NDmaarj97blast-ex.txt).

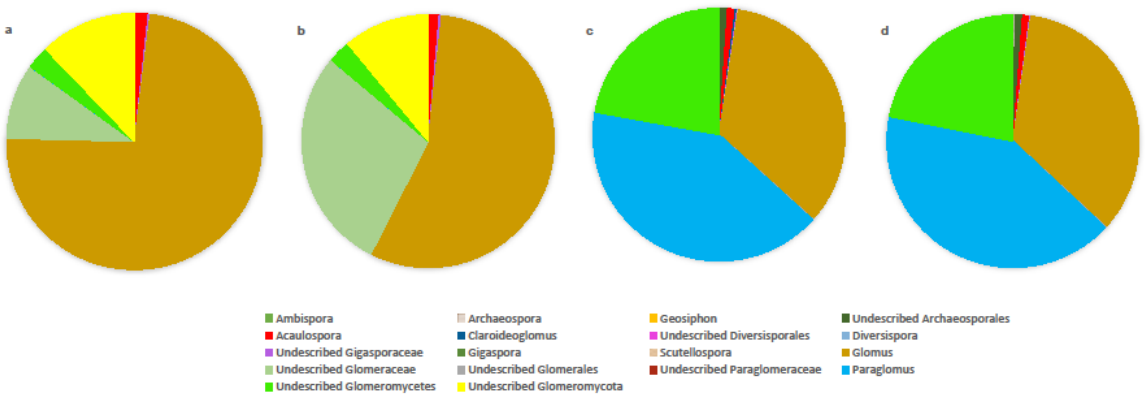

Fig. A3-3 Arbuscular mycorrhizal (AM) fungal diversity detected when using different workflow decisions.

95% clustering threshold, NCBI reference database, singleton and doubleton removal (a), 97% clustering threshold, NCBI reference database, singleton and doubleton removal (b), 95% clustering threshold, MaarjAM reference database, singleton and doubleton removal (c), and 97% clustering threshold, MaarjAM reference database, singleton and doubleton removal (d). Read data used are rarefied to the minimum number of reads per sample.

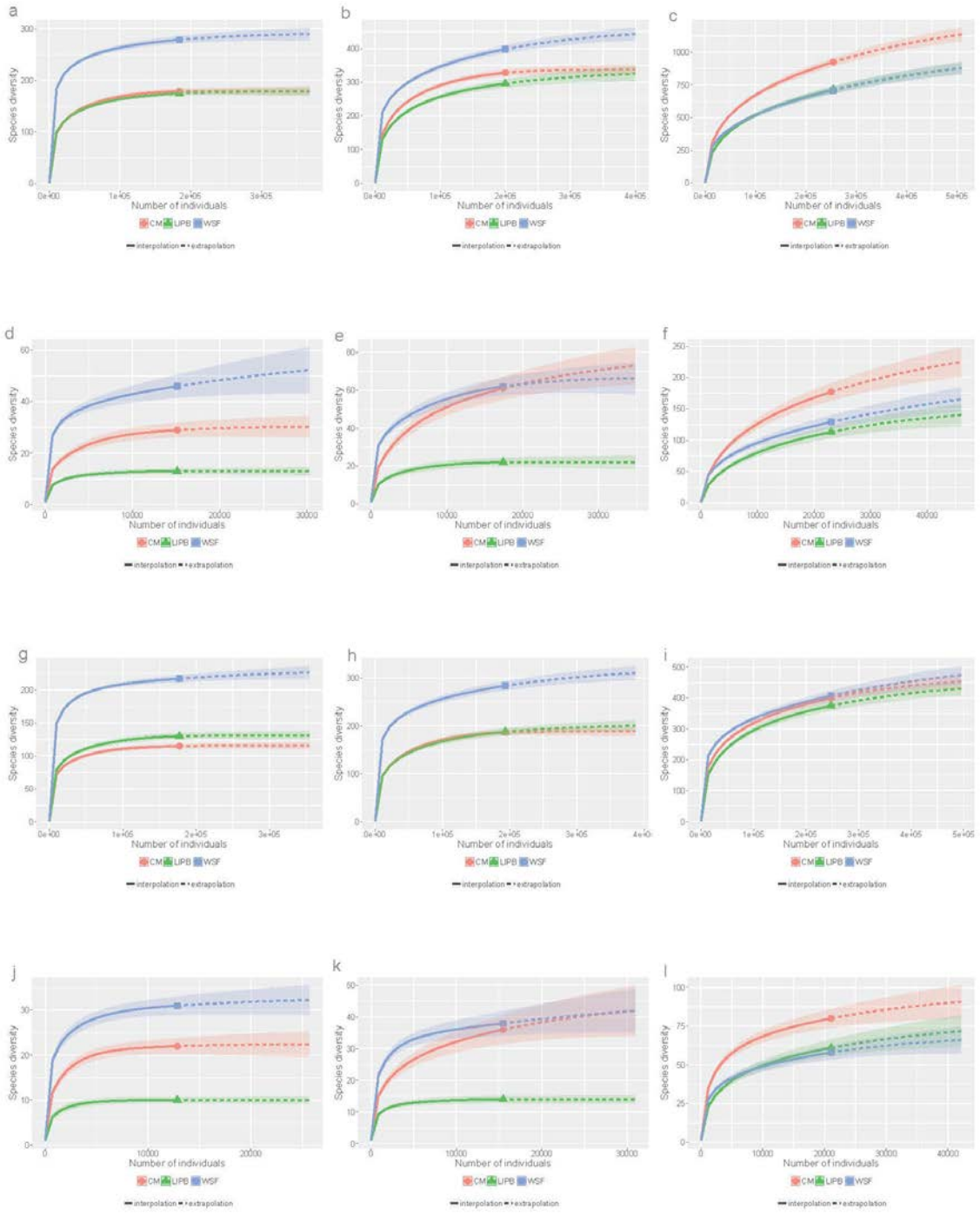
a
b
c
d
e
f
g
h
i
j
k
l
Species diversity
Number of individuals
CM
LIPB
WSF
interpolation
extrapolation

Fig. A3-4 Rarefaction curves of operational taxonomic units (OTUs) from all 12 workflow decisions, using rarefied read data to the minimum number of reads per sample.

Curves represent samples from Colliers Mills (CM) (red), Long Island Pine Barrens (LIPB) (green) and Wharton State Forest (WSF) (blue). Workflow decisions were: 97% clustering, singletons and doubletons removed, MaarjAM database (a), 97% clustering, singletons removed, MaarjAM database (b), 97% clustering, no rare sequences removed, MaarjAM database (c), 97% clustering, singletons and doubletons removed, NCBI database (d), 97% clustering, singletons removed, NCBI database (e), 97% clustering, no rare sequences removed, NCBI database (f), were 95% clustering, singletons and doubletons removed, MaarjAM database (g), 95% clustering, singletons removed, MaarjAM database (h), 95% clustering, no rare sequences removed, MaarjAM database (i), were 95% clustering, singletons and doubletons removed, NCBI database (j), 95% clustering, singletons removed, NCBI database (k), 95% clustering, no rare sequences removed, NCBI database (l).