INVESTIGATION OF THE ROLE OF MIRNAS AND SYMBIOTIC FUNGI IN BIOFUEL SORGHUM PRODUCTION

by

RACHEL MARION RACKERS

(Under the Direction of Katrien M. Devos)

ABSTRACT

Growing biofuel sorghum on marginal lands could be an efficient way to produce sustainable liquid fuels while avoiding competition with food crops. Microbial symbioses with both mycorrhizal and non-mycorrhizal fungi could be promising tools for optimizing marginal land growth under nutrient deficient conditions. Here I identified miRNAs, small non-coding RNAs involved in gene regulation, that could be useful in regulating genes important for optimum biofuel sorghum production, such as those involved in AMF symbiosis. MiRNA eQTL were also identified to give insights into the regulation of miRNA expression. Additionally, I explored how two endophytes*, A. alternata* and *Neopestalotiopsis* behaved across different sorghum varieties under drought stress. While microbial contamination confounds the results, this still represents the first effort to observe the role of these endophytes across multiple sorghum genotypes, providing a basis to inform future experimentation.

INDEX WORDS: Biofuels, Sorghum, MicroRNA, MiRNA, Symbiosis, Arbuscular mycorrhizal fungi, AMF, Endophyte, *Alternaria alternata*, *Neopestalotiopsis*, Expression genome wide association study (eGWAS), Expression quantitative trait loci (eQTL)

INVESTIGATION OF THE ROLE OF MIRNAS AND SYMBIOTIC FUNGI IN BIOFUEL SORGHUM PRODUCTION

by

RACHEL MARION RACKERS

B.S., North Carolina State University, 2021

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2025

INVESTIGATION OF THE ROLE OF MIRNAS AND SYMBIOTIC FUNGI IN BIOFUEL SORGHUM PRODUCTION

by

RACHEL MARION RACKERS

Major Professor:	Katrien M. Devos
Committee:	Jeffrey L. Bennetzen Li Yang

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
University of Georgia
May 2025

DEDICATION

I dedicate this work to my late grandmother, Judy Crawford, the pioneering female scientist of my family. I hope to make her proud by following in her footsteps and dedicating my energy to scientific discovery and advancement.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Dr. Katrien M. Devos. Her guidance and patience have helped me grow significantly as a scientist. I would also like to thank my committee members, Jeffrey Bennetzen and Li Yang for their willingness to serve on my committee.

I would also like to thank everyone I've overlapped with in the Devos lab over the years, all of whom contributed to a wonderful working environment. I want to specifically thank Ching-Ting Huang, Thomas H. Pendergast, Gurjot Singh Sidhu, and Fiona Wacera Wahinya. Additionally, I would be remiss not to thank Shreena Pradhan, whose companionship during the busy graduation season was much appreciated.

Thank you to the Institute of Plant Breeding, Genetics, and Genomics and all of its students. Being a member of this department has afforded me many invaluable experiences that I greatly appreciate.

I also want to thank my parents, Kevin and Holly, as well as my siblings, Nathan, Hope, and Sarah, for their support. Even if you didn't always understand what I was researching, you were always ready to offer words of encouragement.

And lastly, I'd like to thank my two cats, Emrys and Percy, for their companionship throughout my degree—especially during all the late nights.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVEIW

Introduction

The current global economy relies heavily on the use of fossil fuels. But while fossil fuels are one of the most convenient sources of energy, increasing scarcity, rising prices, and high CO_2 emissions may push governments and concerned parties to find more renewable alternatives (International Energy Agency, 2022; Lincoln, 2005). Continued reliance on fossil fuels is not sustainable for most countries, especially those that don't have local sources of fossil fuels and rely on imports (Ediger et al., 2007).

In response to the need for more sustainable fuel sources, biofuels have arisen as a promising alternative to fossil fuels. Biofuels are a renewable energy source derived from plants, algae or animal waste processed to produce biodiesel or bioethanol. Since the transportation sector relies heavily on liquid fossil fuels, biofuels are an ideal replacement as they are also energy dense liquid fuels (Khanna et al., 2021).

Biofuels could be especially useful in the context of aviation. Unlike other transportation sectors that can more easily benefit from other renewable energy sources like electricity, aviation can't easily transition to electric energy. This is due to the energy requirements of flight that make utilizing electricity impractical as current batteries aren't able to store enough energy to power a commercial plane. As other sectors make advancements in reducing their carbon footprint and the demand for aviation increases, the carbon dioxide emissions resulting from aviation could increase from the current

~2.5% of global emissions to ~11% over the next two decades if changes aren't made (Adu-Gyamfi & Good, 2022). Given this, biofuels could serve as a useful energy dense liquid fuel alternative to fossil fuels (Yao et al., 2017).

In order to invest in innovations that allow for increased production of sustainable aviation fuels on a commercial scale, the U.S. Department of Energy (DOE), the U.S. Department of Transportation (DOT), and the U.S. Department of Agriculture (USDA) created the Sustainable Aviation Fuel Grand Challenge. The main goal of this challenge is to produce 3 billion gallons of sustainable aviation fuel by 2030, as well as produce 35 billion gallons by 2050, which should be enough to cover 100% of the domestic aviation fuel demand. Ultimately, this project seeks to make the U.S. aviation sector have net zero emissions by 2050 (U.S. Department of Energy, 2024).

However, while increasing renewable energy is important, the increasing global population and rising food demand puts biofuel crops in competition with food crops. Using food crops, such as corn, for biofuel takes away from food production both directly by reducing the amount of corn available for food and feed, and indirectly by reducing the amount of arable land available for growing food crops. This increases food prices and exacerbates global food shortages (Service, 2007). As a result, there has been a push to limit biofuel crop cultivation to marginal lands in order to maintain food crop production levels.

Marginal lands are defined as non-profitable for the production of staple crops such as maize, soybean and wheat, and may be degraded, drought-prone or have low soil fertility (Khan et al., 2021). Cai et al. (2011) estimate that there could be up to 1107 million hectares (mha) of available marginal lands, and that this land has the potential to

produce sufficient feedstocks to supply 10-52% of the world's current liquid fuel consumption.

First-generation biofuel crops, like maize, are often unsuitable for growing on marginal lands because they require significant water and fertilizer inputs (Khan et al., 2021). In fact, ethanol production from maize was estimated to only provide about 25% more energy than is used in its production (Somerville, 2007).

In an effort to avoid competition with food crops, focus has been shifted towards second generation biofuel crops that are dedicated bioenergy crops grown for their non-edible plant biomass (Khan et al., 2021). However, while growing lignocellulosic feedstocks for biofuel does help avoid using edible yields for bioenergy, there is still the issue of producing economically viable yields without taking over land needed for food production.

The current challenge then becomes identifying crops able to sustain profitable biomass production on marginal lands with minimal inputs.

Study System: Sorghum

Sorghum (*Sorghum bicolor* (L.) Moench) is a C4 diploid (2n=20) grass in the Panicoideae subfamily. It is a hardy annual crop well suited to semi-arid, tropical, and subtropical areas due to its heat and drought tolerance. It is also very versatile and can be grown for a variety of uses such as grain, forage, syrup, and bioenergy (Maqbool et al., 2001; Rooney, 2014). It is largely self-pollinating. This, in addition to its relatively diverse germplasm and long history of cultivation, make breeding of improved inbred

and hybrid sorghum cultivars easier (Olson et al., 2012; Rooney, 2014; Rooney et al., 2007).

In terms of using sorghum for biofuel, there are several ways to obtain energy from the plant. Grain sorghum varieties are grown to generate starch and can be utilized for ethanol production similarly to corn. Sweet sorghum is characterized by juicy stalks with high sugar content and has historically been used for the production of sweeteners, but can also be used to produce ethanol similarly to sugarcane. However, while the stalk juice contains mostly sucrose and glucose and can be easily fermented, the sugars can be unstable, meaning ethanol production must be initiated quickly to avoid energy loss (Rooney, 2014). The main sorghum biofuel focus, however, is on energy or biomass sorghum, which is grown for producing large quantities of biomass which can be used for producing cellulosic biofuel (Rooney, 2014; Rooney et al., 2007). These biomass varieties are bred to be photoperiod sensitive, where a threshold day shortness is required to induce reproductive growth. When grown in regions with longer days, reproductive growth is delayed, and plants continue vegetative growth for longer, leading to more biomass (Mullet et al., 2014; Rooney, 2014; Rooney et al., 2007). Sorghum plants are also less susceptible to heat and drought stress during vegetative growth compared to reproductive growth. Plants respond to water shortages by halting vegetative growth and then resuming it as soon as there is enough moisture (Mullet et al., 2014; Olson et al., 2012; Rooney, 2014). This makes biomass sorghum a good crop to cultivate on marginal lands with harsher heat or drought stress.

Sorghum has many genotypes that are well suited for bioenergy production on marginal lands due to their ability to produce high biomass yields while also requiring

fewer nutrient and water inputs Stamenković et al. (2020). Maw et al. (2017) conducted a five-year study comparing ethanol yield from sweet sorghum, high biomass sorghum, and maize grown on marginal lands and subjected to drought and heat stress. They found that the estimated yields from sweet sorghum and high biomass sorghum were higher than the ethanol yield from maize. Additionally, they found that biomass sorghum was more tolerant of environmental stress than both maize and sweet sorghum.

Using sorghum as a biofuel crop can have many advantages despite not being a perennial grass such as switchgrass and sugarcane, which are commonly grown as dedicated biofuel crops. Similar to the other biofuel grasses, sorghum does C4 photosynthesis. This allows it to better accumulate large amounts of biomass while also being more drought tolerant (Mullet et al., 2014; Olson et al., 2012; Rooney et al., 2007). Sorghum has a smaller, diploid genome (~800 Mbp) compared to its complex polyploid grass relatives, switchgrass and sugarcane. The sorghum genome has also been fully sequenced which makes studying and breeding for genetic improvements significantly easier (Mullet et al., 2014; Paterson et al., 2009). And while sorghum is grown as an annual instead of a perennial, an annual growth habit allows for faster testing of varieties across multiple locations and treatments without having to wait multiple seasons for stands to establish and reach full biomass production potential (Mullet et al., 2014).

Overall, biomass sorghum is a promising second generation biofuel crop well suited for growing on marginal lands. Adoption of sorghum will help prevent land resources being taken away from food crop production, while maintaining or increasing biofuel production.

Introduction to Endophytes

Despite sorghum being a promising biofuel candidate for marginal lands, there is potential for improvements that could make the crop more productive without increasing inputs. To this end, genetic analyses have been conducted on traits such as biomass, drought and salinity tolerance, and disease resistance as a means to target breeding improvements (Allwright & Taylor, 2016). However, root architecture is less researched despite its importance in mediating stressful conditions such as water deficient or lower fertility soils (Del Bianco & Kepinski, 2018; Lynch, 2022).. A contributor to root nutrient uptake efficiency is the symbiosis between plant roots and the soil microbiome. More specifically, endophytes represent a class of microorganisms that could play a large role in optimal nutrient utilization.

The word endophyte means ‘in the plant’. The relationship between a host plant and its endophyte can range from mutualistic, to exploitive, to parasitic, to facultatively saprotrophic (Schulz & Boyle, 2005). Here, the term endophyte is used to specifically describe fungi that colonize a host plant.

Endophytic fungi are often categorized into three major groups: Mycorrhizal fungi, Clavicipitaceous endophytes (C-endophytes), and Non-Clavicipitaceous endophytes (NC-endophytes).

Mycorrhizal fungi symbiosis represents one of the plant-microbe relationships that have a significant effect on essential nutrient absorption from the soil (Smith & Read, 2010). Mycorrhizal fungi are present in many diverse environments and form symbiotic relationships with a variety of plant species. Their major role is to absorb nutrients, such as phosphorus and nitrogen, from the soil and then transfer them to the

host plant (Mitra et al., 2020; Smith & Read, 2010). Arbuscular Mycorrhizal Fungi (AMF), a subclass of mycorrhizal fungi, will be further discussed.

C-endophytes interact with grasses and are generally present in shoots and are transmitted vertically (passed from one host plant to the next through seeds). They can play roles in conferring drought tolerance and increasing plant biomass, as well as discouraging herbivory through the production of toxic chemicals. Endophytic advantages vary by host species, host genotype, and environmental conditions (Rodriguez et al., 2009). C-endophytes were not used in my research and are not further discussed.

NC-endophytes are much more diverse and can interact with a large range of hosts, including grasses. This group of endophytes has been isolated from every major lineage of land plants across many diverse ecosystems and are found in almost every plant organ (Arnold & Lutzoni, 2007). They are often horizontally transmitted, and the endophyte presence in the host plant usually increases in species diversity and density as the plant ages (Arnold & Herre, 2003). Two genera of NC-endophytes will be further discussed, *Alternaria* and *Neopestalotiopsis*.

Arbuscular Mycorrhizal Fungi (AMF)

AMF, the most widespread class of mycorrhizal fungi, are estimated to form mutualistic interactions with around 80% of vascular plants (Mitra et al., 2020; Smith & Read, 2010). AMF are considered obligate biotrophs and require a host to obtain carbon resources and complete their life cycle (Hodge & Fitter, 2010; Jansa et al., 2013). AMF can perceive plant root signals at distances of up to 13 mm. When in close enough

proximity, AMF will begin extending hyphae toward the roots to establish a connection with the host (Smith et al., 1986).

Upon contact with a plant root, the fungal hyphae adhere to the root surface. AMF can then penetrate into the root cortex and form arbuscules. Arbuscules are highly branched structures that form intracellular networks inside the cortical cells of roots and facilitate the transfer of nutrients between AMF and their hosts (Luginbuehl & Oldroyd, 2017; Rosewarne et al., 1997; Smith et al., 1986). However, while arbuscules can penetrate the cell wall, they do not penetrate the cellular membrane. Instead, the cell membrane invaginates into the cell, creating an apoplastic compartment for the hyphae. This space is called the periarbuscular space, and it serves as the site of nutrient exchange between the two symbionts (Krajinski et al., 2014; Wang et al., 2014). Maintaining some separation between symbionts allows each organism to more closely regulate the materials they release and absorb. As such, transfer of nutrients is achieved by a symbiont effluxing nutrients into the periarbuscular space for the other symbiont to absorb (Smith & Read, 2010).

After colonization, mycorrhizae will begin extending hyphae out into the soil to retrieve nutrients beyond the nutrient depletion zone near the plant roots. AMF absorb soil nutrients through their extraradical mycelium (hyphal network existing outside the root). The nutrients can then be transported to the intraradical mycelium (hyphal network existing inside the root) to be transferred to the host plant (Smith & Read, 2010). AMF can help provide many nutrients from the soil to the host plant, including phosphorus, nitrogen, potassium, calcium, copper, iron, magnesium, zinc, and sulfur (Chen et al., 2017). However, of these nutrients, phosphorus uptake represents the most significant

contribution of the fungus to its host, especially in low phosphorus soils (Birhane et al., 2012; Mitra et al., 2020; Smith & Read, 2010). Nitrogen transfer from the fungi to the host can also be beneficial. However, AMF themselves require a significant amount of nitrogen resources, and therefore transfer a smaller percentage of absorbed nitrogen to the host. This means that the uptake of nitrogen for their own use could put AMF in competition with their host plants when nitrogen levels are more limited (Chen et al., 2017; Hodge & Fitter, 2010; Paterson et al., 2009).

The absorption of nutrients like phosphorus and nitrogen by AMF is often facilitated by high affinity phosphate and ammonium transporters (Govindarajulu et al., 2005; Maldonado-Mendoza et al., 2001). AMF can also modulate the expression of its transporters in response to nutrient availability. For example, Maldonado-Mendoza et al. (2001) determined that the phosphate transporter gene *GiPT* had increased expression under high phosphorus conditions to accommodate increased phosphorus intake. The absorbed phosphorus is then transported in the form of phosphate.

The uptake of nitrogen from the soil is a little more complex than phosphorus absorption. Govindarajulu et al. (2005) utilized isotope labeling to determine the process of nitrogen translocation from the soil to the host. This revealed that the process likely starts with the assimilation of inorganic nitrogen by the fungus. The nitrogen is synthesized into arginine in the extraradical mycelium prior to being transported into intraradical mycelium. AMF likely transport nitrogen in the form of arginine due to it being non-toxic and more nitrogen concentrated. Additionally, arginine may also be able to bind to polyphosphates in order to leverage phosphorus transport to aid in nitrogen translocation. Once in the intraradical mycelium, arginine is broken down into ammonia

which is the nitrogen form that is transferred to the host plant. Breaking down arginine into ammonia prior to transfer to the host ensures that the fungi do not lose the valuable carbon present in arginine (Govindarajulu et al., 2005).

Once nutrients have been acquired by the fungus, they need to be transferred to the host plant. As mentioned previously, nutrients are transferred between symbionts through the periarbuscular space, which separates the host cytoplasm from the fungal arbuscules. This transfer is likely facilitated by a proton gradient (Krajinski et al., 2014; Wang et al., 2014). This is supported by the fact that both Krajinski et al. (2014) and Wang et al. (2014) determined that H^+ ATPases in the host plasma membrane are upregulated as a result of fungal arbuscule presence. This allows for the establishment of a proton gradient, with the periarbuscular space having a higher pH to encourage absorption of effluxed compounds into one of the symbionts. Additionally, transporters present on each of the symbionts help mediate the transfer of specific nutrients. For example, the absorption of phosphate by the host plant is facilitated by proton-coupled phosphate transporters (Banasiak et al., 2021; Javot et al., 2007).

In return for soil nutrients, the host plant provides carbon resources to the AMF (Hodge & Fitter, 2010; Jiang et al., 2017; Luginbuehl & Oldroyd, 2017; Smith & Read, 2010). The carbon transferred from the host to the fungi is typically in the form of sugars, but fatty acids represent another form of carbon that can be transferred between symbionts (Jiang et al., 2017). AMF can impose significant carbon demands on the host plant, with up to 20% of a plant's photosynthates being channeled into the fungus (Bago et al., 2000). This increased carbon demand can impact carbon accumulation in various plant tissues. For example, Wright et al. (1998) found that while carbon levels present in

a plant's leaves were not significantly changed as a result of mycorrhizae colonization, the roots of mycorrhizae plants had increased levels of carbon compared to non-mycorrhizae plants. Additionally, both Wright et al. (1998) and Chen et al. (2017) noted increased levels of photosynthesis as well as larger biomass as a result of mycorrhizae colonization. This increase in photosynthesis is likely a result of increased carbon demands by the fungi and is facilitated by the better nutrient assimilation provided by AMF (Chen et al., 2017; Wright et al., 1998).

In addition to aiding in nutrient assimilation, AMF can also help mediate biotic and abiotic stresses. For example, Birhane et al. (2012) demonstrated that AMF symbiosis is able to benefit plants under irregular watering regimes, likely through increased water storage and higher assimilation induced by AMF presence. In this way, mycorrhizae could increase water use efficiency to mediate short-term drought periods. However, Li et al. (2019) conducted a field study (in a temperate semi-arid climate) on grasses (*Leymus chinensis* and *Hemarthria altissima*) which revealed that AMF helped mediate long-term drought (~2 years) and improve plant growth under light (30% rainfall exclusion) and moderate (50% rainfall exclusion) drought conditions but didn't have significant benefits under extreme drought (70% rainfall exclusion). AMF can also help with increasing temperature stress resistance. Zhu et al. (2010) demonstrated that AMF could produce higher dry root weight in maize, even in a variety of temperature treatments. This growth increase despite temperature stress could be attributed to AMF helping the host better conserve water.

However, while most plant-AMF interactions are mutually beneficial, there are instances where the cost of supporting the fungi exceeds the benefits they provide to the

plant. This promotes the idea that the AMF-plant relationship exists on a mutualism-parasitism continuum, where varying conditions can dictate the benefit or detriment the AMF have on the plant. Determining the potential negative impact of AMF on the host plant requires assessment of whether the carbon allocated to the fungus results in a reduction in plant fitness, and whether adequate nutrient absorption could have been achieved without AMF assistance. For example, the presence of abundant soil nutrients, such as phosphorus, results in a decrease in a plant's need for AMF to provide such nutrients. If AMF colonization does not decrease as a result of increased nutrient availability, it's possible for the cost of the plant's carbon allocations to the fungi to exceed the benefit of nutrient transfer from AMF (Johnson & Graham, 2013; Johnson et al., 1997).

It's important to note that mycorrhizal interactions with host plants are not simply between one fungus and the host, but rather many mycorrhizal fungi (Smith & Read, 2010). Chen et al. (2017) explored the effects of mycorrhizal diversity on the degree of benefits received by cucumber seedlings. They discovered that colonization by a larger number of AMF from different genera had a greater effect on nutrient absorption than inoculations consisting of multiple species from a single genus. They also found that while having multiple genera had the greatest impact, inoculating with multiple species within the same genus was still more productive than just using only a single species of AMF. Ultimately, they saw a trend of increased AMF diversity resulting in increased photosynthesis and biomass production. This illustrates the fact that AMF diversity could be an important component to consider when trying to optimize AMF benefits.

Additionally, host genotype can also affect where on the mutualism-parasitism continuum a plant-AMF symbiosis falls. Cobb et al. (2016) illustrated the genotypic effect in sorghum by comparing three open-pollinated African and Latin American tropical landrace genotypes against three commercial hybrid genotypes. Historically, the landraces were grown in low input conditions whereas the hybrid cultivars were often grown under high fertilizer inputs. They found that under non-fertilized conditions the commercial hybrids did not exhibit a difference in biomass production when inoculated with mycorrhizae vs. not. Under fertilized conditions, the AMF inoculated plants actually performed worse than their non-inoculated counterparts. In contrast, the sorghum landraces had significantly increased biomass yields when inoculated with AMF vs. not. They also noted a 149% increase in the percent AMF root colonization for the open-pollinated landraces compared to the hybrids under non-fertilized conditions. This suggests that intensive breeding of hybrids in fertilizer rich systems may have diminished the plant's ability to effectively recruit beneficial AMF. It also highlights the potential for some genotypes to be more AMF responsive and thus better able to employ AMF symbiosis to grow more productively on low fertility soils without requiring fertilization.

Watts-Williams et al. (2019) determined that there are many sorghum genes that are differentially expressed in response to AMF symbiosis. While many of these AMF linked genes were conserved across different sorghum genotypes, many also had variable expression by genotype. They also determined that AMF species diversity played a role in how much a given genotype benefited, with some species and species combinations being better in certain genotypes than others. Ultimately, this highlights the importance of both sorghum genotype and AMF species composition when establishing an optimally

yielding system. Based on the Watts-Williams et al. (2019) results, mycorrhizal responsiveness is likely a heritable trait that could be mapped to genetic loci to ultimately be utilized for breeding sorghum lines with favorable AMF interactions.

Because AMF appear to be promising in terms of mediating poor soil conditions under low or no fertilizer inputs (Ortas, 2012; Sabia et al., 2015), they should be considered in the context of biomass sorghum production. However, most previous research on the interactions between AMF and their host plants has been in controlled greenhouse studies. Therefore, research is needed on which sorghum genotypes can best utilize AMF under authentic marginal land conditions so that the best-suited genotypes can be grown for biofuel production as well as used for future breeding efforts of improved biomass sorghum.

Alternaria alternata

The genus *Alternaria* (family Pleosporaceae) is made up of many diverse species that range from saprotrophs to endophytes to pathogens. Members of the genus *Alternaria* are classified as phaeodictyosporic hyphomycetes, meaning they are asexual spored fungi that produce conidia (asexual spores) (Lawrence et al., 2016).

Due to the inconsistent distribution of important traits across the genus, categorization based on variation in host specificity and nutritional modes is often inconsistent, making species boundaries difficult to discern (DeMers, 2022). Despite significant variation within the *Alternaria* genus, *Alternaria alternata*, a small-spored species, is widely accepted as the type species for the genus (Lawrence et al., 2016; Simmons, 1967). *Alternaria alternata* is considered a cosmopolitan species, meaning it's

widely distributed across many different regions and environments. Because *A. alternata* isolates display significant diversity in lifestyle (pathogenic, endophytic, saprophytic), it is often debated whether the species should be subdivided into multiple species. However, no conclusive morphological differences between isolates have been observed (DeMers, 2022)

In some systems, *A. alternata* strains identified in the roots of healthy plants have been categorized as Dark Septate Endophytes (DSEs) (Spagnoletti & Chiocchio, 2020). DSEs are characterized by their darkly pigmented, septate (having cross walls) hyphae (Ruotsalainen et al., 2022; Singara Charya, 2015). DSEs are often considered the evolutionary intermediate between mycorrhizae and free-living saprotrophs, sharing characteristics with both ends of the spectrum. However, while DSEs can form intraradical hyphal networks similarly to mycorrhizae, there is no direct proof of nutrient transfer between host and endophyte. Instead, benefits the plant may receive from DSEs are more likely indirect. For example, since DSEs can exhibit saprotrophic behavior, they may support their host by breaking down organic matter in the soil, releasing nutrients that are more readily accessible to the plant. As they can obtain carbon from the soil without requiring a host, DSEs require relatively low investment from the host plant, potentially just obtaining carbon from metabolic waste exuded from the host plant (Malicka et al., 2022; Ruotsalainen et al., 2022).

Most *Alternaria* species predominantly reproduce asexually via conidia. However, while most species only exhibit an asexual haploid phase, some species can exhibit both haploid and diploid phases and thus may be capable of sexual reproduction. Genetic evidence suggests that *Alternaria* species were likely sexual at some point in

evolutionary history due to the presence of one of the two MAT1 idiomorphs, components of the heterothallic mating system (Lawrence et al., 2016). Heterothallic species carry either the MAT1-1 or MAT1-2 idiomorph, and can reproduce sexually with another strain containing the opposite idiomorph (Lu et al., 2011). *A. alternata* as well as *A. penicillata* are among the few species that exhibit a functional heterothallic mating system and can outcross. A few other species, such as *A. papavericola*, possess both MAT1 idiomorphs (homothallic) and are capable of selfing. The rest of the genus reproduces strictly asexually (Inderbitzin et al., 2006; Lawrence et al., 2016).

As a phytopathogen, *A. alternata* can infect a wide range of hosts, sometimes causing severe yield losses. *A. alternata* can cause fruit rot and significantly impact the quality of fruits and vegetables (Ventura-Aguilar et al., 2021; M. Wang et al., 2019). *A. alternata* has also been linked to pre- and post-harvest diseases in many cereal crops, especially wheat and barley (Tralamazza et al., 2018). While infection of foliar tissues does not typically affect the vitality and yield of the plant, infections can result in reduced grain quality due to kernel discoloration as well as potential loss in nutritional value (Amatulli et al., 2013; Tralamazza et al., 2018) While there is much variety in the level of pathogenicity between different *A. alternata* strains, because they are difficult to tell apart morphologically, they are often categorized as pathotypes of *A. alternata* rather than being named new species (DeMers, 2022).

Some *A. alternata* isolates can also exist as endophytes. Because endophytic *Alternaria spp.* have been linked to enhanced plant growth, they have been labeled as having potential as bio-fertilizers. Zhou et al. (2018) identified an *Alternaria* spp. that increased root biomass as well as the accumulation of bioactive compounds like phenolic

acid in *Salvia miltiorrhiza* (Chinese sage). Ferreira et al. (2023) demonstrated that some endophytic strains of *A. alternata* were able to solubilize phosphate as well as produce plant hormones, like indole-3-acetic acid (an auxin), ultimately helping promote plant growth. *Alternaria alternata* strains may also reduce pathogen risks to the host by altering the rhizosphere microbiome. Luo et al. (2025) identified an *A. alternata* strain that when inoculated into soil planted with tomato seedlings produced metabolites that reduced the presence of harmful microorganisms in the rhizosphere while increasing the abundance of beneficial ones. These studies demonstrate that *A. alternata* isolates may have potential as bio-fertilizers.

Some pathogenic *A. alternata* isolates are able to produce Host Specific Toxins (HSTs). After cell wall penetration via appressoria, HSTs help condition the host cells for easier fungal entry. HSTs, which are typically secondary metabolites, can aid fungal invasion by impacting key processes like metabolism and membrane permeability, as well as suppressing plant defense responses (Tsuge et al., 2013). While appressoria are often associated with fungal pathogens, they can also be formed by endophytes, saprotrophs, and even some AMF (Ryder et al., 2022). In non-pathogenic isolates, appressoria can still aid in cell wall penetration, however subsequent fungal growth remains apoplastic and doesn't breach the cell membrane (Narusaka et al., 2005; Thomma, 2003; Tsuge et al., 2013). Successful cellular penetration often involves HSTs that condition the host to allow invasion (Tsuge et al., 2013). However, production of HSTs is not a requirement for a strain to be pathogenic. Some *A. alternata* strains can still infect a host plant by utilizing non-host-specific toxins or hydrolytic enzymes that can increase virulence and allow for infection (Thomma, 2003).

These examples demonstrate that while *A. alternata* can be pathogenic in some contexts, it also has the potential to benefit the host plant. Compared to mycorrhizae, knowledge regarding the potential agricultural uses for DSEs lags behind. More research is needed to determine if *A. alternata* could be useful as a bio-fertilizer for biofuel sorghum grown on marginal lands.

Neopestalotiopsis

Similar to *Alternaria*, *Neopestalotiopsis* is a diverse genus with species identified as pathogens, endophytes, and saprotrophs (Maharachchikumbura et al., 2011; Maharachchikumbura et al., 2014). With some exceptions, *Neopestalotiopsis* also has septate hyphae but, unlike *Alternaria*, the hyphae are generally hyaline (translucent or clear) (Hermawan et al., 2021; Shi et al., 2024). *Neopestalotiopsis* was previously part of the genus *Pestalotiopsis.* However, at the discretion of Maharachchikumbura et al. (2014) the broad and diverse *Pestalotiopsis* genus was subdivided into three genera: *Pestalotiopsis, Neopestalotiopsis,* and *Pseudopestalotiopsis*. Each of the newly classified genera have their own distinct conidial morphologies and can be differentiated genetically. While all three genera have 5-celled conidia, the median conidial cells of *Neopestalotiopsis* are versicolorous (having variable pigmentation). In contrast, *Pestalotiopsis* and *Pseudopestalotiopsis* are characterized by concolorous (uniformly pigmented) median cells, with *Pseudopestalotiopsis* having a darker pigmentation than *Pestalotiopsis*.

While the three new genera are distinguishable, they still share many features and are often broadly referred to as *Pestalotiopsis sensu lato* or *Pestalotiopsis*-like fungi (Hsu

et al., 2024). Because prior to 2014, the species now considered *Neopestalotiopsis spp.* were included in the genus *Pestalotiopsis*, the term *Pestalotiopsis*-like fungi will be utilized to describe information that predates the formal naming of the *Neopestalotiopsis* genus. Additionally, as there is much overlap in terms of behavior between *Neopestalotiopsis, Pestalotiopsis,* and *Pseudopestalotiopsis,* the term *Pestalotiopsis*-like fungi will also be used to describe features conserved between the genera.

Pestalotiopsis-like fungi are largely asexual, producing conidia that allow for dispersal and colonization of new environments (Maharachchikumbura et al., 2011; Maharachchikumbura et al., 2014). However, some *Pestalotiopsis*-like fungi can exhibit both teleomorph (sexual) and anamorph (asexual) phases. This led fungi observed in their teleomorph phase to often be named as separate species than when observed in their anamorph phase. The issue of dual classification of the same species was realized through DNA sequencing which allowed the teleomorph and anamorph phases to assigned to the same fungal species.

The ITS region in the ribosomal RNA operon is a highly variable sequence, making it an ideal candidate for fungal identification (Op De Beeck et al., 2014). However, Hu et al. (2007) asserted that the ITS sequence alone was insufficient for adequately categorizing *Pestalotiopsis*-like isolates. They argued that β-tubulin phylogenies were more informative, but a combined analysis of ITS and β-tubulin would be most ideal for determining isolate identities. In contrast, Liu et al. (2010) determined that the ITS region in isolation was informative in differentiating between *Pestalotiopsis*-like fungi with different pigmentations in the median conidial cells. As the color of

median cells is often used to distinguish species, the ITS region could therefore be a useful metric for classifying isolates.

Pestalotiopsis-like fungi exhibit low host specificity (Darapanit et al., 2021; Hu et al., 2007; Shi et al., 2024; Wei et al., 2007). Further, *Pestalotiopsis*-like fungi are also not tissue specific. This unfortunately undermines many of the early species classifications as, historically, species differentiation and naming were largely based on host association (Wei et al., 2007).

While morphology is a better classification parameter than host association or tissue type, there is still much morphological overlap between species that can make categorization complicated. Although most morphological characteristics remain relatively consistent after multiple subcultures in media, some features differ in an artificial culture compared to growth in a host (Bologna et al., 2018; Hu et al., 2007). Additionally, while ITS and/or β-tubulin sequencing can help increase the resolution for categorizing isolates at the genus level, it is not always reliable at distinguishing between species. As such, many previously categorized species may actually represent only one species (Maharachchikumbura et al., 2011; Wei et al., 2007).

Pestalotiopsis-like fungi are present in many environments, being commonly found in temperate and tropical regions (Maharachchikumbura et al., 2014) and growing best at 20-30 °C under alkaline conditions (Baggio et al., 2021; Hsu et al., 2024). *Pestalotiopsis*-like fungi are often found on living plants and can exhibit pathogenic or endophytic behaviors. However, they can also exist as saprotrophs on dead plant materials (Hsu et al., 2024; Maharachchikumbura et al., 2011; Maharachchikumbura et al., 2014). Additionally, some species are believed to have both endophytic and

pathogenic stages within their life cycle, as illustrated by Lee et al. (1995) who isolated *Pestalotiopsis microspora* from both diseased and healthy *Terreya taxifolia* trees. This demonstrates how the same fungus in a given host can exhibit variable nutritional modes, likely as a result of environmental conditions.

Most *Pestalotiopsis*-like species are weak, opportunistic pathogens that invade the host through natural openings or wounds (Baggio et al., 2021; Darapanit et al., 2021; Lee et al., 1995; Maharachchikumbura et al., 2011; Shi et al., 2024). They can cause a variety of diseases, including leaf spots, needle blight, fruit rots, severe chlorosis, and various post-harvest diseases (Baggio et al., 2021; Darapanit et al., 2021; Maharachchikumbura et al., 2014; Shi et al., 2024).

Fail and Langenheim (1990) explored the mechanisms through which *Pestalotiopsis subcuticularis* colonizes *Hymenaea courbaril* (a tropical tree) as an endophyte and as a pathogen. Upon inoculation on leaves, spores germinate and penetrate the epidermis via conidial germ tubes. Hyphae preferentially enter the host through wounds (but not through stomata). In the absence of wounds, fungi penetrate the epidermis. Interestingly, this penetration is accomplished without developing appressoria, although it is unclear if direct penetration results from some form of mechanical pressure or through fungal enzymes. Hyphae then grow intercellularly between the mesophyll cells, causing no damage to the host. Upon wounding, the fungal hyphae quickly invade the disrupted cells. The fungi were not observed invading any healthy, living cells (Fail & Langenheim, 1990).

Lee et al. (1995) analyzed the fungal metabolites produced by pathogenic *Pestalotiopsis microspora* in diseased *Torreya taxifolia* to determine which chemicals

may be responsible for causing chlorosis in the needles. They identified pestaloside, pestalopyrone, and hydroxypestalopyrone to be likely involved in disease symptoms. Interestingly, while pestalopyrone and hydroxypestalopyrone were only identified in diseased plants, pestaloside was also found in the symptomless plants. Further analysis revealed that pestaloside has antifungal properties that may provide *P. microspora* a competitive advantage over other endophytes. The antimicrobial compounds could also provide direct benefits to the host. For example, the secondary metabolites with strong antimicrobial activity produced by endophytic *Neopestalotiopsis* species in the leaves of *Cinnamomum loureiroi* effectively inhibited six bacterial pathogens as well as two fungal pathogens (Tanapichatsakul et al., 2019). This could ultimately help protect the host plant from disease. Because these endophytic bioactive compounds are useful for defending the host plant, they have also become of interest as less toxic alternatives to current pesticides (de Carvalho et al., 2021; Liu et al., 2016).

Pestalotiopsis-like fungi could also be useful to host plants through mobilization of nutrients to make them accessible to the plant. Munir et al. (2022) determined that an endophytic *Pestalotiopsis thailandica* isolate had phosphate-solubilizing capabilities. Similarly, Jacob et al. (2023) identified multiple *Pestalotiopsis*-like fungi that could solubilize zinc and phosphate. This could be useful in improving soil fertility as well as increasing phosphorus and zinc absorption by the plant. As a result, phosphate-solubilizing fungi, such as *Pestalotiopsis*-like fungi, could be useful as biofertilizers. This could be especially useful for mediating the poor soil fertility often found on marginal lands. However, as fungal behaviors can vary by host and environmental conditions, more research is needed to determine whether *Neopestalotiopsis* could be beneficial in

bioenergy crop production systems and, specifically, biomass sorghum production systems.

MiRNA

MicroRNAs (miRNAs) are small non-coding RNAs around 22 nucleotides long that are involved in regulating gene expression post-transcriptionally by interacting with target mRNAs. Mature miRNAs can interact with complementary mRNA sequences to regulate gene expression through mRNA degradation or translational inhibition (O'Brien et al., 2018; Sun, 2012). MiRNAs are a relatively recent discovery, having been first identified about 30 years ago. Initial research focused predominantly on miRNAs in animals, but research on plant miRNAs has increased significantly over the past decade. Plant miRNAs have now been found to be involved in many important plant functions like growth and development as well as responses to external stimuli such as light, nutrient availability, and stress (Millar, 2020; Rogers & Chen, 2013; J. Wang et al., 2019).

There have been many recent advancements in the understanding of miRNA biogenesis, but research has been mostly performed in model species like *Arabidopsis*. This has uncovered the general pathways (Budak & Akpinar, 2015). Nevertheless, research on a wider variety of plants is needed to help better decipher the functions of miRNAs.

MiRNA synthesis begins through transcription of miRNA genes (*MIRs*) by RNA Polymerase II to produce primary transcripts termed pri-miRNAs. Plant genomes generally contain a hundred to several hundreds of *MIRs*. Of these genes, many of them are present as families of miRNAs located in close proximity to each other in the

genome. They can be present in intergenic or intronic regions of the genome. The majority of miRNAs come from intergenic regions and are regulated by their own promoters (Sun, 2012; J. Wang et al., 2019). Additionally, the involvement of various transcription factors in the transcription of *MIRs* suggests that the process can be finely regulated (Budak & Akpinar, 2015; Rogers & Chen, 2013).

The pri-miRNAs form hairpin structures that are recognized by Dicer-like RNase II endonucleases (DCLs) which can cleave the hairpin. The DCL performs two cuts, the first separating the hairpin from the rest of the primary transcript to produce pre-miRNA, and the second separating the miRNA/miRNA* duplex from the hairpin loop (Budak & Akpinar, 2015; Rogers & Chen, 2013; Sun, 2012; J. Wang et al., 2019). There are multiple different DCLs that can be involved in miRNA synthesis, with different DCLs producing miRNAs of varying lengths. The most common DCL in plants is DCL1 and it produces miRNAs that are 21 nucleotides long those are probably 100 kids (Budak & Akpinar, 2015; Rogers & Chen, 2013; J. Wang et al., 2019).

The DCL cleavage of pri-miRNA produces a double stranded miRNA/miRNA* duplex with 2-nucleotide 3' overhangs. The duplex is then methylated by the small RNA methyltransferase HUA Enhancer 1 (HEN1) which helps protect the 3' terminus from uridylation and degradation (Budak & Akpinar, 2015; Rogers & Chen, 2013).

It was previously believed that the miRNA/miRNA* duplex was then transported to the cytoplasm by Hasty (HST), similarly to how its homologous protein, Exportin 5 (EXPO5), does in animals (Budak & Akpinar, 2015; Rogers & Chen, 2013). However, Park et al. (2005) found that *hst* mutants didn't have an accumulation of miRNAs in the nucleus, meaning that HST may not be the predominant mechanism for moving miRNAs

to the cytoplasm. More recent research by Bologna et al. (2018) suggests that assembly of the RNA Induced Silencing Complex (RISC) takes place in the nucleus, which is then exported to the cytoplasm by EXPO1.

Regardless of cellular location, the miRNA/miRNA* duplex must separate into single stranded mature miRNAs and interact with Argonaute (AGO) proteins to form the RISC complex that can target mRNAs. The miRNA/miRNA* duplex separates into a guide strand and a passenger strand. The guide strand binds to the AGO protein to form RISC and is used to target mRNAs, whereas the passenger strand is degraded. Determination of which strand will be the guide strand is based on the 5' stability, with the strand with lower 5' thermodynamic stability being preferentially chosen as the guide strand (Budak & Akpinar, 2015; Rogers & Chen, 2013; Sun, 2012; J. Wang et al., 2019).

The RISC is then guided to target mRNAs based on complementarity with the miRNA. In contrast to animals where imperfect pairing of the miRNA with its target mRNA is tolerated, plants are much more stringent and generally have perfect or near perfect pairing. However, while almost perfect pairing is required in the 5' region with no more than 1 mismatch tolerated, more mismatches (up to 4) are permitted in the 3' region. The target mRNA sequence is then either cleaved and degraded or translationally inhibited (Budak & Akpinar, 2015; Rogers & Chen, 2013; Sun, 2012; J. Wang et al., 2019). While mRNA cleavage and subsequent degradation is generally the predominant mode of miRNA regulation, translation inhibition can also be important in regulating some processes. For example, floral development in *Arabidopsis* can be regulated through translation inhibition by the miR172 family to trigger transition from meristem tissue growth to floral organ development (Budak & Akpinar, 2015; Chen, 2004).

Because miRNAs serve as post-transcriptional regulators that can be finely controlled, they are useful in mediating plant growth and development as well as responses to changing environmental conditions, including biotic and abiotic stresses. In terms of development, miRNAs have been found to be involved in processes such as leaf morphogenesis, vegetative phase change, and flowering time. MiRNAs are likely able to regulate these processes by targeting conserved transcription factors or other regulatory genes involved in development (Millar, 2020; Rogers & Chen, 2013).

Recent research on the role of miRNAs during stress responses showed that multiple *Arabidopsis* miRNAs had high-level fold changes when plants were exposed to heat, drought, and salt stress Pegler et al. (2019b). These miRNAs could potentially be leveraged in generating more stress tolerant plants. A study by Njaci et al. (2018) found that extreme water deficit in *Tripogon loliiformis*, a resurrection plant naturally tolerant of low water conditions, resulted in changes in miRNA abundances. Additionally, they found that many of the conserved miRNAs had different abundances in the roots vs. shoots, likely due to the changes in metabolism the plant experiences as a result of the stress.

Physiological responses to various nutrient availabilities can also be regulated by miRNAs. For instance, copper is an important plant nutrient regulated by many different miRNAs. For example, Shahbaz and Pilon (2019) found that miR397, miR398, and miR408 were inhibited under copper limiting conditions, resulting in higher levels of their target mRNAs. Regulation of phosphorus has also been shown to be regulated by miRNAs. (Pegler et al., 2019a) found that miR399 regulates *PHOSPHATE2 (PHO2),* an important gene involved in maintaining phosphate homeostasis. (Johnson et al., 2022)

found that AMF colonization resulted in changes in miRNA abundances for miRNAs targeting genes involved in copper ion binding and GRAS family transcription factors.

Overall, miRNAs are important mediators of plant physiology that could be useful in better understanding plant functions and responses.

Genome Wide Association Study (GWAS)

Domestication has had a significant impact on the genetic diversity in many crops. As global food and energy demands increase, better understanding of the genetic basis for trait variation could be invaluable in improving crop productivity. As genetic resources continue to increase, harnessing genetic tools for efficient targeted breeding is becoming increasingly prevalent. High throughput sequencing and the generation of reference genomes for many major crops, including sorghum, have enabled the characterization of genome wide variation by comparing the reference against many diverse varieties (Huang XueHui & Han Bin, 2014; Paterson et al., 2009).

Genome variations between different varieties can be identified by genotyping each variety individually. The genotype represents the hereditary information, usually represented by a specific allele pattern at various molecular markers, for a given individual. The observable characteristics influenced by the genotype (and the environment) represent the individual's phenotype. Comparing diverse lines yields variants between genomes, such as single-nucleotide polymorphisms (SNPs) and insertions/deletions. As SNPs are the predominant sequence variants identified and are widely distributed across the genome, they are often used as molecular markers (Collard et al., 2005; Gupta et al., 2019; Huang XueHui & Han Bin, 2014)

Many agronomically important traits, such as yield, are complex traits controlled by many different genes. The genetic regions associated with these quantitative traits are called Quantitative Trait Loci (QTL) (Collard et al., 2005; Gupta et al., 2019; Tanksley, 1993). QTL for traits of interest can be identified through genetic mapping in biparental populations or association mapping in panels of diverse accessions. Trait mapping relies on the assumption that a given phenotype shared by individuals can be linked to conserved genomic regions that are in close proximity to the causal mutation as a result of Linkage Disequilibrium (LD). In other words, when a beneficial mutation arises in an ancestor, the mutation and its adjacent sequences are passed down to offspring together due to LD. That means that, in related offspring, the variants around the mutation, while not conferring the actual trait, are co-inherited with the actual causal mutation, and thus are also linked to the phenotype (Collard et al., 2005; Tanksley, 1993; Yan et al., 2023).

The extent of LD can be influenced by a number of factors, such as mating patterns, genetic drift, domestication, natural selection, and population bottlenecks. For instance, outcrossing species tend to have faster LD decay, meaning LD blocks are smaller as a result of higher recombination. In contrast, selfing species tend to have larger LD blocks since recombination is less effective in introducing variation, as self-fertilization rapidly increases homozygosity. Additionally, while physical proximity between two variants increases intrachromosomal LD, interchromosomal LD can also occur. This can be a result of population structure, selection, or historical bottlenecks. For example, in populations with high selection pressure or genetic bottlenecks, some unlinked alleles may consistently co-occur in offspring (Ersoz et al., 2007; Flint-Garcia et al., 2003).

QTL mapping in a biparental population is limited to capturing allelic diversity that segregates between the parents of the population. The mapping resolution will depend on how much recombination occurs during the creation of the biparental mapping population (Korte & Farlow, 2013). The number of recombination events can be increased by increasing the population size, allowing for more meiosis events, or generating recombinant inbred lines (RILs) through multiple generations of selfing. Over subsequent generations of selfing, repeated recombination helps break up LD blocks, increasing the resolution (Ersoz et al., 2007; Flint-Garcia et al., 2003).

To overcome some of the limitations of QTL mapping in biparental populations (limited recombination, limited allelic variation), a Genome Wide Association Study (GWAS) can be conducted (Flint-Garcia et al., 2003; Korte & Farlow, 2013) GWAS allows for genome-wide identification of genetic variants associated with a trait of interest across diverse natural populations (Gupta et al., 2019; Korte & Farlow, 2013; Yan et al., 2018). While similar to QTL mapping, GWAS generally provides higher resolution since natural populations have undergone more recombination events over the course of many generations (Chen et al., 2020). Increased recombination reduces LD block sizes, meaning fewer markers remain tightly linked to a causal gene. As a result, trait-associated regions identified by GWAS are typically much smaller than those detected by traditional QTL mapping, making it easier to pinpoint the causal gene(s) by reducing the number of candidate genes within the QTL region. However, while natural populations tend to have smaller LD blocks from many historical recombination events, some factors can reduce the resolution. As mentioned previously, crops that predominantly self-pollinate could have larger LD blocks, ultimately reducing resolution and increasing the QTL size.

Over the past two decades, thousands of GWAS have been conducted in a variety of crops, such as rice (Chen et al., 2014; Huang et al., 2012), maize (Zhang et al., 2022) wheat (Gurung et al., 2014), and sorghum (Habyarimana et al., 2020; Kimani et al., 2020; Maina et al., 2022; Prom et al., 2019), and for a wide range of morphological traits including yield, grain quality, biomass, flowering time, disease resistance, and drought tolerance (Gurung et al., 2014; Habyarimana et al., 2020; Huang et al., 2012; Kimani et al., 2020; Maina et al., 2022; Prom et al., 2019; Zhang et al., 2020). In addition to morphological phenotypes, gene expression represents an important molecular phenotype that drives trait variation. This has led to interest in identifying expression QTL (eQTL), or genomic regions involved in the regulation of gene expression (Farhangi et al., 2024; Nica & Dermitzakis, 2013; Zhang et al., 2020). Conducting expression GWAS (eGWAS) to identify eQTL can be valuable in elucidating potential gene regulatory networks. While most eQTL are in *cis* and located near the genes they regulate, of particular interest are *trans* eQTL, which can regulate multiple genes. For example, Zhang et al. (2020) identified an eQTL hotspot that was linked to regulation of many genes involved in anthocyanin accumulation in sweet potatoes. From the eQTL hotspot, they identified IbMYB1-2, a transcription factor, as the causal gene. This suggests that IbMYB1-2 is a master regulator of anthocyanin accumulation.

More recently, a miRNA eGWAS was conducted in maize (Chen et al., 2020). This represents the first eGWAS in plants that uses miRNA expression levels as the phenotype. They were able to identify four significant *trans* miRNA eQTL, indicating that eGWAS is a viable method for identifying genomic regions linked to variable miRNA expression levels.

Overall, GWAS represents a useful tool for identifying genomic regions involved in regulating a trait of interest. Identified QTL can then be mined for causal genes that can ultimately be targeted in breeding programs to generate more productive varieties.

References

Adu-Gyamfi, B. A., & Good, C. (2022). Electric aviation: A review of concepts and enabling technologies. *Transportation Engineering*, *9*, 100134.

Allwright, M. R., & Taylor, G. (2016). Molecular breeding for improved second generation bioenergy crops. *Trends in Plant Science*, *21*(1), 43-54.

Amatulli, M. T., Fanelli, F., Moretti, A., Mule, G., & Logrieco, A. F. (2013). *Alternaria* species and mycotoxins associated to black point of cereals. *Mycotoxins*, *63*(1), 39-46.

Arnold, A. E., & Herre, E. A. (2003). Canopy cover and leaf age affect colonization by tropical fungal endophytes: ecological pattern and process in *Theobroma cacao* (Malvaceae). *Mycologia*, *95*(3), 388-398.

Arnold, A. E., & Lutzoni, F. (2007). Diversity and host range of foliar fungal endophytes: are tropical leaves biodiversity hotspots? *Ecology*, *88*(3), 541-549.

Baggio, J. S., Forcelini, B. B., Wang, N.-Y., Ruschel, R. G., Mertely, J. C., & Peres, N. A. (2021). Outbreak of leaf spot and fruit rot in Florida strawberry caused by *Neopestalotiopsis* spp. *Plant Disease*, *105*(2), 305-315.

Bago, B., Pfeffer, P. E., & Shachar-Hill, Y. (2000). Carbon metabolism and transport in arbuscular mycorrhizas. *Plant Physiology*, *124*(3), 949-958.

Banasiak, J., Jamruszka, T., Murray, J. D., & Jasiński, M. (2021). A roadmap of plant membrane transporters in arbuscular mycorrhizal and legume–rhizobium symbioses. *Plant Physiology*, *187*(4), 2071-2091.

Birhane, E., Sterck, F. J., Fetene, M., Bongers, F., & Kuyper, T. W. (2012). Arbuscular mycorrhizal fungi enhance photosynthesis, water use efficiency, and growth of frankincense seedlings under pulsed water availability conditions. *Oecologia*, *169*, 895-904.

Bologna, N. G., Iselin, R., Abriata, L. A., Sarazin, A., Pumplin, N., Jay, F., Grentzinger, T., Dal Peraro, M., & Voinnet, O. (2018). Nucleo-cytosolic shuttling of ARGONAUTE1 prompts a revised model of the plant microRNA pathway. *Molecular Cell*, *69*(4), 709-719. e705.

Budak, H., & Akpinar, B. A. (2015). Plant miRNAs: biogenesis, organization and origins. *Functional & integrative genomics*, *15*, 523-531.

Cai, X., Zhang, X., & Wang, D. (2011). Land availability for biofuel production. *Environmental science & technology*, *45*(1), 334-339.

Chen, S.-Y., Su, M.-H., Kremling, K. A., Lepak, N. K., Romay, M. C., Sun, Q., Bradbury, P. J., Buckler, E. S., & Ku, H.-M. (2020). Identification of miRNA-eQTLs in maize mature leaf by GWAS. *BMC genomics*, *21*(1), 1-13.

Chen, S., Zhao, H., Zou, C., Li, Y., Chen, Y., Wang, Z., Jiang, Y., Liu, A., Zhao, P., & Wang, M. (2017). Combined inoculation with multiple arbuscular mycorrhizal fungi improves growth, nutrient uptake and photosynthesis in cucumber seedlings. *Frontiers in Microbiology*, *8*, 2516.

Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., Li, Y., Liu, X., Zhang, H., & Dong, H. (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nature genetics*, *46*(7), 714-721.

Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science*, *303*(5666), 2022-2025.

Cobb, A. B., Wilson, G. W., Goad, C. L., Bean, S. R., Kaufman, R. C., Herald, T. J., & Wilson, J. D. (2016). The role of arbuscular mycorrhizal fungi in grain production and nutrition of sorghum genotypes: Enhancing sustainability through plant-microbial partnership. *Agriculture, Ecosystems & Environment*, *233*, 432-440.

Collard, B. C., Jahufer, M., Brouwer, J., & Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, *142*, 169-196.

Darapanit, A., Boonyuen, N., Leesutthiphonchai, W., Nuankaew, S., & Piasai, O. (2021). Identification, pathogenicity and effects of plant extracts on *Neopestalotiopsis* and *Pseudopestalotiopsis* causing fruit diseases. *Scientific Reports*, *11*(1), 22606.

de Carvalho, C. R., Maia, M. Q., Sobral, M., Pereira, G. M. D., da Silva, K., Vital, M. J. S., Zilli, J. É., Rosa, C. A., & Rosa, L. H. (2021). Diversity and antimicrobial activity of culturable endophytic fungi associated with the neotropical ethnomedicinal plants *Copaifera langsdorffii* and *Copaifera pubiflora*. *South African Journal of Botany*, *142*, 305-315.

Del Bianco, M., & Kepinski, S. (2018). Building a future with root architecture. In (Vol. 69, pp. 5319-5323): Oxford University Press UK.

DeMers, M. (2022). *Alternaria alternata* as endophyte and pathogen. *Microbiology*, *168*(3), 001153.

Ediger, V. Ş., Hoşgör, E., Sürmeli, A. N., & Tatlıdil, H. (2007). Fossil fuel sustainability index: An application of resource management. *Energy policy*, *35*(5), 2969-2977.

Ersoz, E. S., Yu, J., & Buckler, E. S. (2007). Applications of linkage disequilibrium and association mapping in crop plants. In *Genomics-Assisted Crop Improvement: Vol. 1: Genomics Approaches and Platforms* (pp. 97-119). Springer.

Fail, G., & Langenheim, J. (1990). Infection processes of *Pestalotia subcuticularis* on leaves of *Hymenaea courbaril*. *Phytopathology*, *80*(11), 1259-1265.

Farhangi, S., Gòdia, M., Derks, M. F., Harlizius, B., Dibbits, B., González-Prendes, R., Crooijmans, R. P., Madsen, O., & Groenen, M. A. (2024). Expression genome-wide association study identifies key regulatory variants enriched with metabolic and immune functions in four porcine tissues. *BMC genomics*, *25*(1), 684.

Ferreira, A. P., dos Santos Oliveira, J. A., Polonio, J. C., Pamphile, J. A., & Azevedo, J. L. (2023). Recombinants of *Alternaria alternata* endophytes enhance inorganic phosphate solubilization and plant growth hormone production. *Biocatalysis and Agricultural Biotechnology*, *51*, 102784.

Flint-Garcia, S. A., Thornsberry, J. M., & Buckler IV, E. S. (2003). Structure of linkage disequilibrium in plants. *Annual review of plant biology*, *54*(1), 357-374.

Govindarajulu, M., Pfeffer, P. E., Jin, H., Abubaker, J., Douds, D. D., Allen, J. W., Bücking, H., Lammers, P. J., & Shachar-Hill, Y. (2005). Nitrogen transfer in the arbuscular mycorrhizal symbiosis. *Nature*, *435*(7043), 819-823.

Gupta, P. K., Kulwal, P. L., & Jaiswal, V. (2019). Association mapping in plants in the post-GWAS genomics era. *Advances in genetics*, *104*, 75-154.

Gurung, S., Mamidi, S., Bonman, J. M., Xiong, M., Brown-Guedira, G., & Adhikari, T. B. (2014). Genome-wide association study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. *PloS one*, *9*(9), e108179.

Habyarimana, E., De Franceschi, P., Ercisli, S., Baloch, F. S., & Dall'Agata, M. (2020). Genome-wide association study for biomass related traits in a panel of *Sorghum bicolor* and *S. bicolor*× *S. halepense* populations. *Frontiers in Plant Science*, *11*, 551305.

Hermawan, R., Safitri, R. R., & Sidiq, M. R. (2021). *Neopestalotiopsis zimbabwana* isolated from *Xylaria stromata*. *Biotropika: Journal of Tropical Biology*, *9*(3), 203-209.

Hodge, A., & Fitter, A. H. (2010). Substantial nitrogen acquisition by arbuscular mycorrhizal fungi from organic material has implications for N cycling. *Proceedings of the National Academy of Sciences*, *107*(31), 13754-13759.

Hsu, S.-Y., Xu, Y.-C., Lin, Y.-C., Chuang, W.-Y., Lin, S.-R., Stadler, M., Tangthirasunun, N., Cheewangkoon, R., Al-Shwaiman, H. A., & Elgorban, A. M. (2024). Hidden diversity of *Pestalotiopsis* and *Neopestalotiopsis* (Amphisphaeriales, Sporocadaceae) species allied with the stromata of entomopathogenic fungi in Taiwan. *MycoKeys*, *101*, 275.

Hu, H., Jeewon, R., Zhou, D., Zhou, T., & Hyde, K. D. (2007). Phylogenetic diversity of endophytic Pestalotiopsis species in Pinus armandii and Ribes spp.: evidence from rDNA and β-tubulin gene phylogenies. *Fungal diversity*.

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., & Zhu, C. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature genetics*, *44*(1), 32-39.

Huang XueHui, H. X., & Han Bin, H. B. (2014). Natural variations and genome-wide association studies in crop plants.

Inderbitzin, P., Shoemaker, R. A., O'Neill, N. R., Turgeon, B. G., & Berbee, M. L. (2006). Systematics and mating systems of two fungal pathogens of opium poppy: the heterothallic *Crivellia papaveracea* with a *Brachycladium penicillatum* asexual state and a homothallic species with a *Brachycladium papaveris* asexual state. *Botany*, *84*(8), 1304-1326.

International Energy Agency. (2022). *World Energy Investment 2022*. OECD Publishing.

Jacob, J. K. S., Witzel, K., & dela Cruz, T. E. E. (2023). Comparative diversity and functional traits of fungal endophytes in response to elevated mineral content in a Mangrove ecosystem. *Journal of Fungi*, *9*(12), 1186.

Jansa, J., Bukovská, P., & Gryndler, M. (2013). Mycorrhizal hyphae as ecological niche for highly specialized hypersymbionts–or just soil free-riders? *Frontiers in Plant Science*, *4*, 134.

Javot, H., Penmetsa, R. V., Terzaghi, N., Cook, D. R., & Harrison, M. J. (2007). A *Medicago truncatula* phosphate transporter indispensable for the arbuscular

mycorrhizal symbiosis. *Proceedings of the National Academy of Sciences*, *104*(5), 1720-1725.

Jiang, Y., Wang, W., Xie, Q., Liu, N., Liu, L., Wang, D., Zhang, X., Yang, C., Chen, X., & Tang, D. (2017). Plants transfer lipids to sustain colonization by mutualistic mycorrhizal and parasitic fungi. *Science*, *356*(6343), 1172-1175.

Johnson, A. C., Pendergast IV, T. H., Chaluvadi, S., Bennetzen, J. L., & Devos, K. M. (2022). Identification of microRNAs responsive to arbuscular mycorrhizal fungi in *Panicum virgatum* (switchgrass). *BMC genomics*, *23*(1), 688.

Johnson, N. C., & Graham, J. H. (2013). The continuum concept remains a useful framework for studying mycorrhizal functioning. *Plant and soil*, *363*, 411-419.

Johnson, N. C., Graham, J. H., & Smith, F. A. (1997). Functioning of mycorrhizal associations along the mutualism–parasitism continuum. *The New Phytologist*, *135*(4), 575-585.

Khan, N., Sudhakar, K., & Mamat, R. (2021). Role of biofuels in energy transition, green economy and carbon neutrality. *Sustainability*, *13*(22), 12374.

Khanna, M., Chen, L., Basso, B., Cai, X., Field, J. L., Guan, K., Jiang, C., Lark, T. J., Richard, T. L., & Spawn-Lee, S. A. (2021). Redefining marginal land for bioenergy crop production. *Gcb Bioenergy*, *13*(10), 1590-1609.

Kimani, W., Zhang, L.-M., Wu, X.-Y., Hao, H.-Q., & Jing, H.-C. (2020). Genome-wide association study reveals that different pathways contribute to grain quality variation in sorghum (*Sorghum bicolor*). *BMC genomics*, *21*, 1-19.

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant methods*, *9*, 1-9.

Krajinski, F., Courty, P.-E., Sieh, D., Franken, P., Zhang, H., Bucher, M., Gerlach, N., Kryvoruchko, I., Zoeller, D., & Udvardi, M. (2014). The H+-ATPase HA1 of *Medicago truncatula* is essential for phosphate transport and plant growth during arbuscular mycorrhizal symbiosis. *The Plant Cell*, *26*(4), 1808-1817.

Lawrence, D. P., Rotondo, F., & Gannibal, P. B. (2016). Biodiversity and taxonomy of the pleomorphic genus Alternaria. *Mycological progress*, *15*, 1-22.

Lee, J. C., Yang, X., Schwartz, M., Strobel, G., & Clardy, J. (1995). The relationship between an endangered North American tree and an endophytic fungus. *Chemistry & Biology*, *2*(11), 721-727.

Li, J., Meng, B., Chai, H., Yang, X., Song, W., Li, S., Lu, A., Zhang, T., & Sun, W. (2019). Arbuscular mycorrhizal fungi alleviate drought stress in C3 (*Leymus chinensis*) and C4 (*Hemarthria altissima*) grasses via altering antioxidant enzyme activities and photosynthesis. *Frontiers in Plant Science*, *10*, 499.

Lincoln, S. F. (2005). Fossil fuels in the 21st century. *Ambio*, 621-627.

Liu, A.-R., Chen, S.-C., Wu, S.-Y., Xu, T., Guo, L.-D., Jeewon, R., & Wei, J.-G. (2010). Cultural studies coupled with DNA based sequence analyses and its implication on pigmentation as a phylogenetic marker in *Pestalotiopsis* taxonomy. *Molecular Phylogenetics and Evolution*, *57*(2), 528-535.

Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS genetics*, *12*(2), e1005767.

Lu, S.-W., Yun, S.-H., Lee, T., & Turgeon, B. G. (2011). Altering sexual reproductive mode by interspecific exchange of MAT loci. *Fungal Genetics and Biology*, *48*(7), 714-724.

Luginbuehl, L. H., & Oldroyd, G. E. (2017). Understanding the arbuscule at the heart of endomycorrhizal symbioses in plants. *Current Biology*, *27*(17), R952-R963.

Luo, W., Ping, X., Zhou, J., Gao, S., Huang, X., Song, S., Xu, J., & He, W. (2025). Alternaria alternata JTF001 Metabolites Recruit Beneficial Microorganisms to Reduce the Parasitism of Orobanche aegyptiaca in Tomato. *Biology*, *14*(2), 116.

Lynch, J. P. (2022). Harnessing root architecture to address global challenges. *The Plant Journal*, *109*(2), 415-431.

Maharachchikumbura, S. S., Guo, L.-D., Chukeatirote, E., Bahkali, A. H., & Hyde, K. D. (2011). *Pestalotiopsis*—morphology, phylogeny, biochemistry and diversity. *Fungal diversity*, *50*, 167-187.

Maharachchikumbura, S. S., Hyde, K. D., Groenewald, J. Z., Xu, J., & Crous, P. W. (2014). Pestalotiopsis revisited. *Studies in Mycology*, *79*(1), 121-186.

Maina, F., Harou, A., Hamidou, F., & Morris, G. P. (2022). Genome-wide association studies identify putative pleiotropic locus mediating drought tolerance in sorghum. *Plant Direct*, *6*(6), e413.

Maldonado-Mendoza, I. E., Dewbre, G. R., & Harrison, M. J. (2001). A phosphate transporter gene from the extra-radical mycelium of an arbuscular mycorrhizal fungus Glomus intraradices is regulated in response to phosphate in the environment. *Molecular Plant-Microbe Interactions*, *14*(10), 1140-1148.

Malicka, M., Magurno, F., & Piotrowska-Seget, Z. (2022). Plant association with dark septate endophytes: When the going gets tough (and stressful), the tough fungi get going. *Chemosphere*, *302*, 134830.

Maqbool, S. B., Devi, P., & Sticklen, M. B. (2001). Biotechnology: Genetic improvement of sorghum (*Sorghum bicolor (L.) Moench*). *In Vitro Cellular & Developmental Biology-Plant*, *37*, 504-515.

Maw, M. J., Houx III, J. H., & Fritschi, F. B. (2017). Maize, sweet sorghum, and high biomass sorghum ethanol yield comparison on marginal soils in Midwest USA. *Biomass and Bioenergy*, *107*, 164-171.

Millar, A. A. (2020). The function of miRNAs in plants. In (Vol. 9, pp. 198): MDPI.

Mitra, D., Uniyal, N., Panneerselvam, P., Senapati, A., & Ganeshamurthy, A. N. (2020). Role of mycorrhiza and its associated bacteria on plant growth promotion and nutrient management in sustainable agriculture. *International Journal of Life Sciences and Applied Sciences*, *1*(1), 1-1.

Mullet, J., Morishige, D., McCormick, R., Truong, S., Hilley, J., McKinley, B., Anderson, R., Olson, S. N., & Rooney, W. (2014). Energy Sorghum—a genetic model for the design of C4 grass bioenergy crops. *Journal of experimental botany*, *65*(13), 3479-3489.

Munir, E., Yurnaliza, Y., Lutfia, A., & Hartanto, A. (2022). Isolation and characterization of phosphate solubilizing activity of endophytic fungi from *Zingiberaceous* species. *OnLine J. Biol. Sci*, *22*, 149-156.

Narusaka, Y., Narusaka, M., Seki, M., Ishida, J., Shinozaki, K., Nan, Y., Park, P., Shiraishi, T., & Kobayashi, M. (2005). Cytological and molecular analyses of

non-host resistance of *Arabidopsis thaliana* to *Alternaria alternata*. *Molecular plant pathology*, *6*(6), 615-627.

Nica, A. C., & Dermitzakis, E. T. (2013). Expression quantitative trait loci: present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1620), 20120362.

Njaci, I., Williams, B., Castillo-González, C., Dickman, M. B., Zhang, X., & Mundree, S. (2018). Genome-wide investigation of the role of microRNAs in desiccation tolerance in the resurrection grass *Tripogon loliiformis*. *Plants*, *7*(3), 68.

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in endocrinology*, *9*, 402.

Olson, S. N., Ritter, K., Rooney, W., Kemanian, A., McCarl, B. A., Zhang, Y., Hall, S., Packer, D., & Mullet, J. (2012). High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops. *Biofuels, Bioproducts and Biorefining*, *6*(6), 640-655.

Op De Beeck, M., Lievens, B., Busschaert, P., Declerck, S., Vangronsveld, J., & Colpaert, J. V. (2014). Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PloS one*, *9*(6), e97629.

Ortas, I. (2012). The effect of mycorrhizal fungal inoculation on plant yield, nutrient uptake and inoculation effectiveness under long-term field conditions. *Field crops research*, *125*, 35-48.

Park, M. Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., & Poethig, R. S. (2005). Nuclear processing and export of microRNAs in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, *102*(10), 3691-3696.

Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., & Poliakov, A. (2009). The Sorghum bicolor genome and the diversification of grasses. *Nature*, *457*(7229), 551-556.

Pegler, J. L., Oultram, J. M., Grof, C. P., & Eamens, A. L. (2019a). DRB1, DRB2 and DRB4 are required for appropriate regulation of the microRNA399/PHOSPHATE2 expression module in *Arabidopsis thaliana*. *Plants*, *8*(5), 124.

Pegler, J. L., Oultram, J. M., Grof, C. P., & Eamens, A. L. (2019b). Profiling the abiotic stress responsive microRNA landscape of *Arabidopsis thaliana*. *Plants*, *8*(3), 58.

Prom, L. K., Ahn, E., Isakeit, T., & Magill, C. (2019). GWAS analysis of sorghum association panel lines identifies SNPs associated with disease response to Texas isolates of Colletotrichum sublineola. *Theoretical and Applied Genetics*, *132*, 1389-1396.

Rodriguez, R., White Jr, J., Arnold, A. E., & Redman, a. R. a. (2009). Fungal endophytes: diversity and functional roles. *New Phytologist*, *182*(2), 314-330.

Rogers, K., & Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. *The Plant Cell*, *25*(7), 2383-2399.

Rooney, W. L. (2014). Cellulosic energy cropping systems. In (pp. 110-123). John Wiley & Sons.

Rooney, W. L., Blumenthal, J., Bean, B., & Mullet, J. E. (2007). Designing sorghum as a dedicated bioenergy feedstock. *Biofuels, Bioproducts and Biorefining*, *1*(2), 147-157.

Rosewarne, G., Barker, S., & Smith, S. (1997). Production of near-synchronous fungal colonization in tomato for developmental and molecular analyses of mycorrhiza. *Mycological Research*, *101*(8), 966-970.

Ruotsalainen, A. L., Kauppinen, M., Wäli, P. R., Saikkonen, K., Helander, M., & Tuomi, J. (2022). Dark septate endophytes: mutualism from by-products? *Trends in Plant Science*, *27*(3), 247-254.

Ryder, L. S., Cruz-Mireles, N., Molinari, C., Eisermann, I., Eseola, A. B., & Talbot, N. J. (2022). The appressorium at a glance. *Journal of Cell Science*, *135*(14), jcs259857.

Sabia, E., Claps, S., Morone, G., Bruno, A., Sepe, L., & Aleandri, R. (2015). Field inoculation of arbuscular mycorrhiza on maize (*Zea mays L.*) under low inputs: preliminary study on quantitative and qualitative aspects. *Italian Journal of Agronomy*, *10*(1), 30-33.

Schulz, B., & Boyle, C. (2005). The endophytic continuum. *Mycological Research*, *109*(6), 661-686.

Service, R. F. (2007). Biofuel researchers prepare to reap a new harvest. In: American Association for the Advancement of Science.

Shahbaz, M., & Pilon, M. (2019). Conserved Cu-microRNAs in *Arabidopsis thaliana* function in copper economy under deficiency. *Plants*, *8*(6), 141.

Shi, J., Li, B., Wang, S., Zhang, W., Shang, M., Wang, Y., & Liu, B. (2024). Occurrence of Neopestalotiopsis clavispora causing Apple leaf spot in China. *Agronomy*, *14*(8), 1658.

Simmons, E. G. (1967). Typification of *Alternaria, Stemphylium*, and *Ulocladium*. *Mycologia*, *59*(1), 67-92.

Singara Charya, M. (2015). Fungi: an overview. *Plant Biology and Biotechnology: Volume I: Plant Diversity, Organization, Function and Improvement*, 197-215.

Smith, S., John, B. S., Smith, F., & Bromley, J. L. (1986). EFFECTS OF MYCORRHIZAL INFECTION ON PLANT GROWTH, NITROGEN AND PHOSPHORUS NUTRITION IN GLASSHOUSE-GROWN *ALLIUM CEPAL*. *New Phytologist*, *103*(2), 359-373.

Smith, S. E., & Read, D. J. (2010). *Mycorrhizal symbiosis*. Academic press.

Somerville, C. (2007). Biofuels. *Current Biology*, *17*(4), R115-R119.

Spagnoletti, F. N., & Chiocchio, V. M. (2020). Tolerance of dark septate endophytic fungi (DSE) to 1 agrochemicals in vitro. *Revista argentina de microbiología*, *52*(1), 43-49.

Stamenković, O. S., Siliveru, K., Veljković, V. B., Banković-Ilić, I. B., Tasić, M. B., Ciampitti, I. A., Đalović, I. G., Mitrović, P. M., Sikora, V. Š., & Prasad, P. V. (2020). Production of biofuels from sorghum. *Renewable and Sustainable Energy Reviews*, *124*, 109769.

Sun, G. (2012). MicroRNAs and their diverse functions in plants. *Plant molecular biology*, *80*, 17-36.

Tanapichatsakul, C., Khruengsai, S., Monggoot, S., & Pripdeevech, P. (2019). Production of eugenol from fungal endophytes *Neopestalotiopsis sp*. and *Diaporthe sp.* isolated from *Cinnamomum loureiroi* leaves. *PeerJ*, *7*, e6427.

Tanksley, S. D. (1993). Mapping polygenes. *Annual review of genetics*, *27*(1), 205-233.

Thomma, B. P. (2003). *Alternaria spp.*: from general saprophyte to specific parasite. *Molecular plant pathology*, *4*(4), 225-236.

Tralamazza, S. M., Piacentini, K. C., Iwase, C. H. T., & de Oliveira Rocha, L. (2018). Toxigenic Alternaria species: impact in cereals worldwide. *Current opinion in food science*, *23*, 57-63.

Tsuge, T., Harimoto, Y., Akimitsu, K., Ohtani, K., Kodama, M., Akagi, Y., Egusa, M., Yamamoto, M., & Otani, H. (2013). Host-selective toxins produced by the plant pathogenic fungus Alternaria alternata. *FEMS microbiology reviews*, *37*(1), 44-66.

U.S. Department of Energy. (2024). *Sustainable Aviation Fuel Grand Challenge*. https://www.energy.gov/eere/bioenergy/synthetic-aviation-fuel-grand-challenge?nrg_redirect=472662

Ventura-Aguilar, R. I., Bautista-Baños, S., Hernández-López, M., & Llamas-Lara, A. (2021). Detection of *Alternaria alternata* in tomato juice and fresh fruit by the production of its biomass, respiration, and volatile compounds. *International Journal of Food Microbiology*, *342*, 109092.

Wang, E., Yu, N., Bano, S. A., Liu, C., Miller, A. J., Cousins, D., Zhang, X., Ratet, P., Tadege, M., & Mysore, K. S. (2014). A H+-ATPase that energizes nutrient uptake during mycorrhizal symbioses in rice and Medicago truncatula. *The Plant Cell*, *26*(4), 1818-1830.

Wang, J., Mei, J., & Ren, G. (2019). Plant microRNAs: biogenesis, homeostasis, and degradation. *Frontiers in Plant Science*, *10*, 360.

Wang, M., Fu, H., Shen, X. X., Ruan, R., Rokas, A., & Li, H. (2019). Genomic features and evolution of the conditionally dispensable chromosome in the tangerine pathotype of *Alternaria alternata*. *Molecular plant pathology*, *20*(10), 1425-1438.

Watts-Williams, S. J., Emmett, B. D., Levesque-Tremblay, V., MacLean, A. M., Sun, X., Satterlee, J. W., Fei, Z., & Harrison, M. J. (2019). Diverse *Sorghum bicolor* accessions show marked variation in growth and transcriptional responses to arbuscular mycorrhizal fungi. *Plant, Cell & Environment*, *42*(5), 1758-1774.

Wei, J.-G., Xu, T., Guo, L.-D., Liu, A.-R., Zhang, Y., & Pan, X.-H. (2007). Endophytic *Pestalotiopsis* species associated with plants of *Podocarpaceae, Theaceae and Taxaceae* in southern China. *Fungal Divers*, *24*(1), 55-74.

Wright, D., Read, D., & Scholes, J. (1998). Mycorrhizal sink strength influences whole plant carbon balance of *Trifolium repens L*. *Plant, Cell & Environment*, *21*(9), 881-891.

Yan, S., Weng, B., Jing, L., & Bi, W. (2023). Effects of drought stress on water content and biomass distribution in summer maize (*Zea mays L.*). *Frontiers in Plant Science*, *14*, 1118131.

Yan, Y., Liu, Q., Jacobsen, S. E., & Tang, Y. (2018). The impact and prospect of natural product discovery in agriculture: New technologies to explore the diversity of secondary metabolites in plants and microorganisms for applications in agriculture. *EMBO reports*, *19*(11), e46824.

Yao, G., Staples, M. D., Malina, R., & Tyner, W. E. (2017). Stochastic techno-economic analysis of alcohol-to-jet fuel production. *Biotechnology for Biofuels*, *10*(1), 1-13.

Zhang, L., Yu, Y., Shi, T., Kou, M., Sun, J., Xu, T., Li, Q., Wu, S., Cao, Q., & Hou, W. (2020). Genome-wide analysis of expression quantitative trait loci (eQTLs) reveals the regulatory architecture of gene expression variation in the storage roots of sweet potato. *Horticulture research*, *7*.

Zhang, X., Ren, Z., Luo, B., Zhong, H., Ma, P., Zhang, H., Hu, H., Wang, Y., Zhang, H., & Liu, D. (2022). Genetic architecture of maize yield traits dissected by QTL mapping and GWAS in maize. *The Crop Journal*, *10*(2), 436-446.

Zhou, L. S., Tang, K., & Guo, S. X. (2018). The plant growth-promoting fungus (PGPF) *Alternaria sp.* A13 markedly enhances *Salvia miltiorrhiza* root growth and active ingredient accumulation under greenhouse and field conditions. *International journal of molecular sciences*, *19*(1), 270.

Zhu, X.-C., Song, F.-B., & Xu, H.-W. (2010). Arbuscular mycorrhizae improves low temperature stress in maize via alterations in host water status and photosynthesis. *Plant and soil*, *331*, 129-137.

CHAPTER 2

IDENTIFICATION OF MIRNAS EXPRESSED IN SORGHUM ROOTS ACROSS GENOTYPES AND ENVIRONMENTS

Introduction

Increases in greenhouse gas emissions have many negative impacts on the environment, leading to rising interest in developing more sustainable fuel alternatives to replace fossil fuels. One promising alternative is biofuels generated from dedicated feedstocks. As aviation cannot reliably use electricity on a commercial scale due to battery limitations, energy-dense biofuels represent a more practical sustainable alternative to traditional, non-renewable jet fuel (Adu-Gyamfi & Good, 2022; Khan et al., 2021; Yao et al., 2017).

However, while securing a renewable fuel source for the aviation industry is important, the rising global population and associated increasing food demands put biofuel crop production in competition with food crops. To avoid diverting prime croplands away from food production, there has been interest in cultivating biofuel crops on marginal lands. Marginal lands are categorized as areas non-profitable for food crop production and often have drought-prone conditions and poor soil fertility (Khanna et al., 2021).

Sorghum has been proposed as a promising biofuel crop candidate because it is adapted to hot and dry climates (Maqbool et al., 2001; Rooney, 2014). Of particular interest are the biomass sorghum varieties which are more tolerant of marginal land conditions and can be used to produce cellulosic biofuel (Mullet et al., 2014; Olson et al., 2012; Rooney, 2014; Rooney et al., 2007).

In order to avoid the need for excessive fertilizer inputs, utilization of soil microbial symbioses could provide a more sustainable solution for growing on marginal lands with poor soil fertility. Of the soil microorganisms, Arbuscular Mycorrhizal Fungi (AMF) are one of the most significant symbioses that could be invaluable for mediating poor soil conditions under low inputs (Ortas, 2012; Sabia et al., 2015). They can aid the host plant in absorption of essential nutrients such as phosphorus and nitrogen (Mitra et al., 2020; Smith & Read, 2010). However, interactions between AMF and the host plant can be complex given the differential benefits provided by different AMF species and communities (Chen et al., 2017). Additionally, the host plant is also able to mediate its interactions with AMF, meaning some genotypes are able to more efficiently recruit favorable AMF species through variable expression of AMF linked genes (Cobb et al., 2016; Watts-Williams et al., 2019). As such, exploring the genetic mechanisms linked to AMF symbiosis could be useful for identifying the ideal plant varieties to grow efficiently on marginal lands.

MicroRNAs (miRNAs) are a class of small non-coding RNAs that post-transcriptionally regulate target mRNAs to, typically, downregulate protein production. MiRNA regulation involves a mature miRNA binding to a target mRNA at a complementary binding site, and then causing mRNA cleavage and degradation or

translational inhibition (O'Brien et al., 2018; Sun, 2012). MiRNAs are important regulators of many different plant processes such as growth, development, nutrient acquisition, symbiosis, and biotic and abiotic stresses. MiRNAs often exhibit control over these processes by targeting pivotal transcription factors or regulatory genes (Johnson et al., 2022; Millar, 2020; Pegler et al., 2019a; Rogers & Chen, 2013). As miRNAs are intertwined with many different plant functions, they represent important components of trait regulatory networks. A better understanding of miRNAs can ultimately be useful in more finely controlling a trait via post-transcriptional regulation.

My research is a part of a larger project focusing on identifying sorghum genotypes that perform optimally in a given environment (Georgia or Arizona) by efficiently recruiting and utilizing microbial symbioses. Specifically, I seek to identify miRNAs from a sorghum bioenergy panel (BAP), characterize their expression levels and identify genomic regions that regulate miRNA expression in order to gain insights into their contribution in the control of important traits, such as AMF symbiosis. The current study represents preliminary miRNA results for a subset of the Georgia and Arizona field samples. A total of 98 MiRNAs were identified. I further identified 37 miRNAs that were differential expressed in a set of 12 genotypes when grown in Georgia compared to Arizona. Additionally, a miRNA expression Genome Wide Association Study (eGWAS) was conducted to gain insights into the genetic basis for variable miRNA expression patterns. Future work cross-analyzing these miRNA results with other aspects of the larger project such as AMF colonization levels and community composition information, mRNA transcript levels and biomass yield data will help generate a more holistic understanding of the genetics of biofuel sorghum production on marginal lands.

Materials and Methods

Planting and Harvesting

A total of 337 sorghum accessions make up the bioenergy association panel (BAP) that was used in this study (Brenton et al., 2016). These genotypes are predominantly biomass sorghum varieties as well as some sweet sorghum varieties. The BAP panel was planted in Watkinsville, Georgia in 2022, in and Maricopa, Arizona in 2023. For the Georgia field study, a field at the University of Georgia's Wellbrook farm that had lain fallow for at least six years was initially sprayed with glyphosate, planted with a winter wheat cover crop, then again sprayed with glyphosate on May 18th, 2022, at a rate of 2.34 L/ha to control weeds. The wheat on the field was then bush hogged. Based on soil analyses, conducted before the wheat was planted, the starting nitrate concentration was 0.75 ppm and the starting phosphorus concentration was 10.46 ppm. No nitrogen or phosphorus were added. Granular potash (0-0-60 K_2O) was applied at a rate of 89.7 kg/ha to each of the replicates.

Prior to planting, the sorghum seeds were surface sterilized with a 2% sodium hypochlorite solution and submerged and rinsed with tap water twice. Seeds were then treated with Concep III at a rate of 0.4 active ingredient/kg seed to protect seedlings from herbicide treatment in case such a treatment needed to be performed (however, herbicide treatment was never performed). Seeds were germinated in pots in the greenhouse, and 2-week-old seedlings were transferred into the field June 6th –13th, 2022. Seedlings were planted in three replicates (blocks) (See Figure 2.1 for field layout). The plants were

irrigated for the first 2 weeks following planting to ensure seedling establishment, but no irrigation was applied for the remainder of the season.

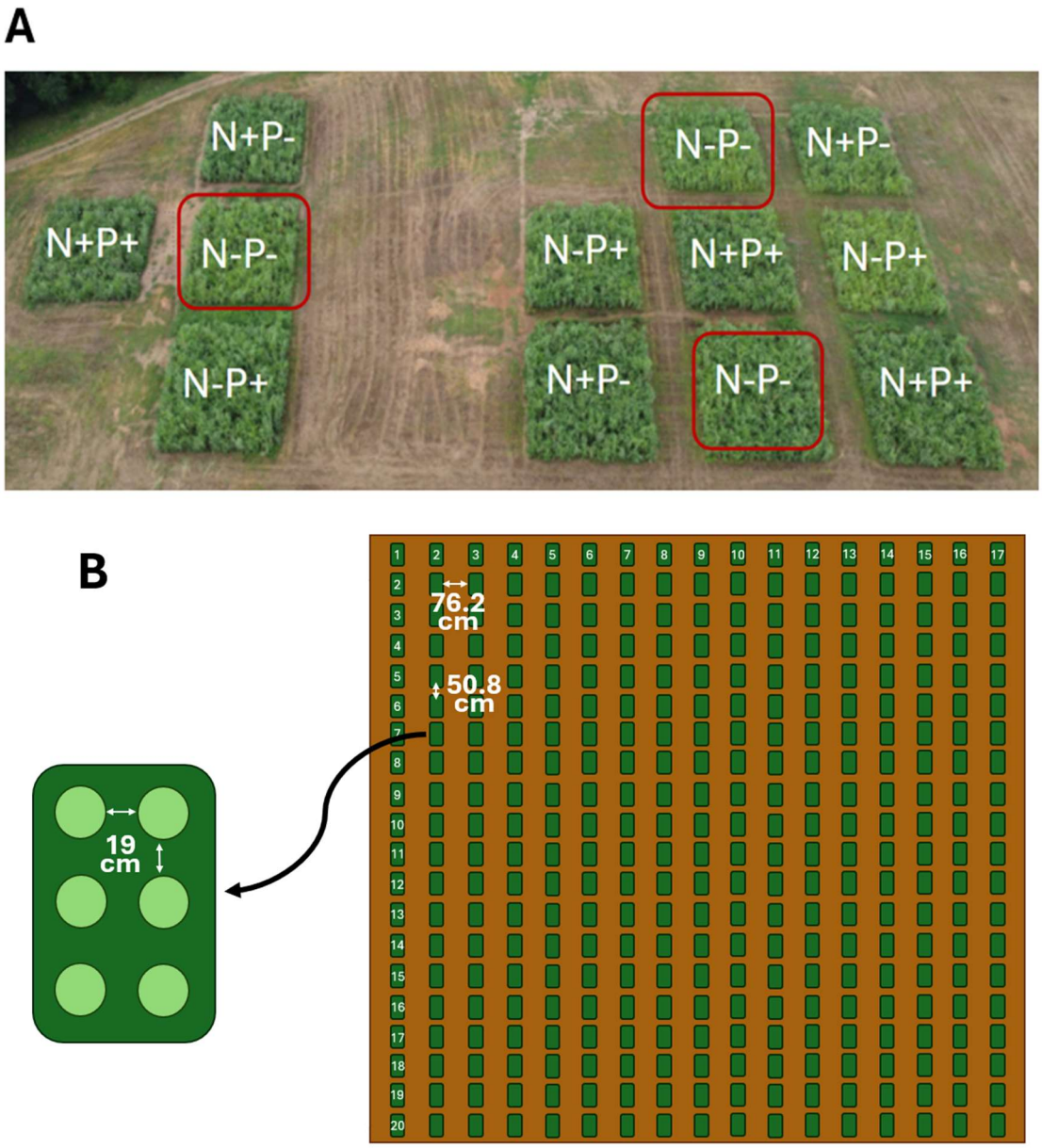

Figure 2.1 Field layout for Georgia (2022). (A) Aerial drone photo of the sorghum field (taken by Peng Qi). N and P refer to nitrogen and phosphorus respectively. The (+) indicates that nitrogen or phosphorus was applied while (-) indicates no additional nutrients were added. Blocks bordered by the red boxes indicate the low N/low P treatment blocks that were used for expression analysis. (B) Schematic showing the organization of plots in a block (17 columns, 20 rows). Each plot has six plants of the same genotype.

The BAP panel was grown in 2023 at the Maricopa Agricultural Center, Arizona, with a few experimental design modifications necessitated by the field shape, irrigation

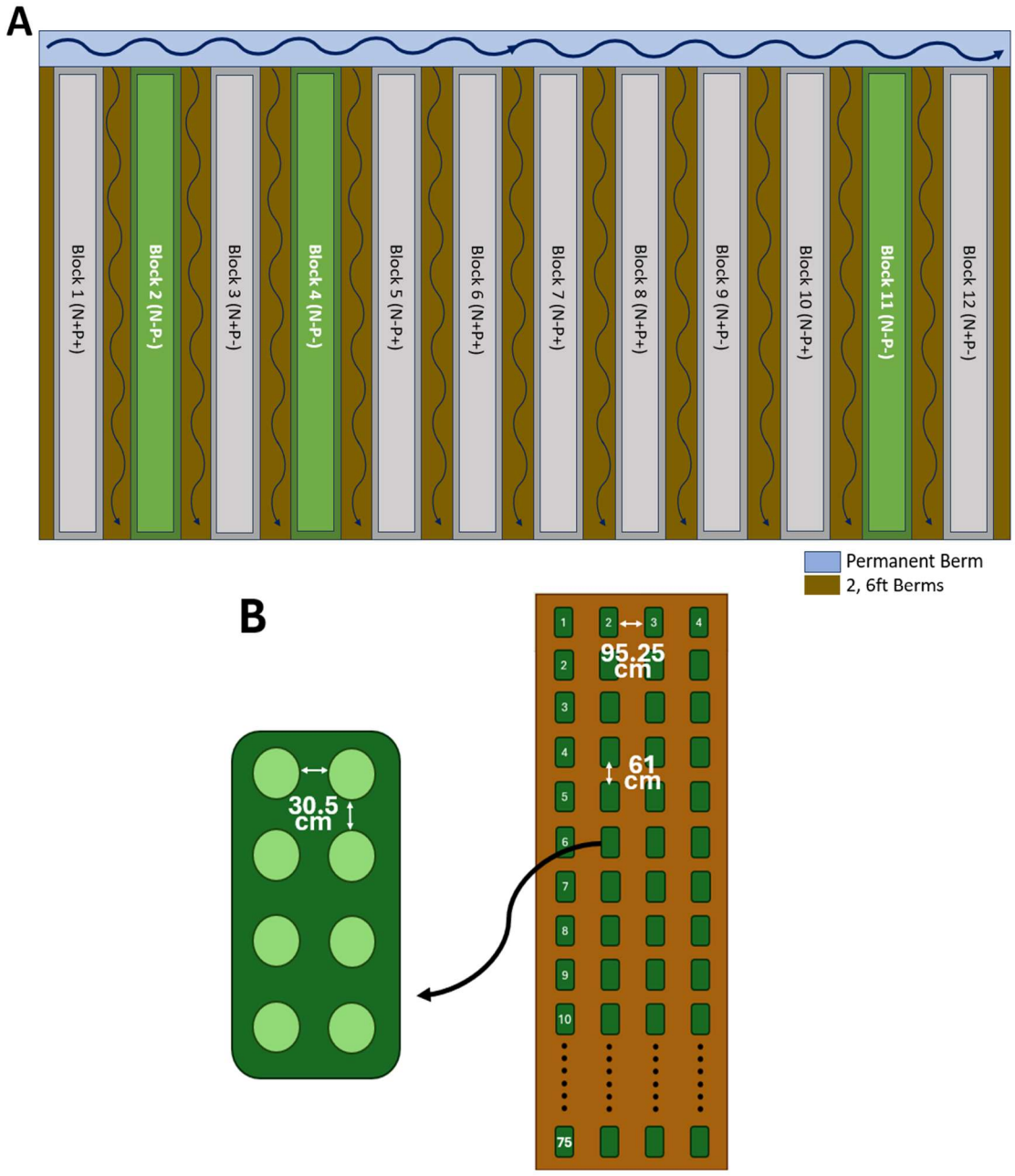

Figure 2.2. Field Layout for Arizona (2023). (A) Diagram depicting the organization of the treatment blocks. Green blocks represent the low N/low P treatments used for RNA analysis. Areas between blocks are temporary berms for irrigation (brown). The blue rectangle represents the irrigation ditch. Permanent berms are located on the edges of the field. (B) Schematic showing the organization of plots in a block (4 columns, 75 rows). Each plot has eight plants of the same genotype.

method and environmental conditions. Because of field size limitations, only 305 of the BAP genotypes were grown per replicate with a total of three replicates (See figure 2.2 for field layout). Soil testing showed the presence of 98 ppm of phosphorus with the total % soil nitrogen being ~12. No N, P or K were added to the field. Before planting, sorghum seeds were surface sterilized with a 2% sodium hypochlorite solution, rinsed twice with tap water, and dried prior to planting. Seeds were planted directly in the field from April 13th—18th, 2023. For each genotype, eight seeds were planted per replicate in a plot (Figure 2.2B). Due to the drier conditions, the blocks were periodically flooded for irrigation throughout the experiment to maintain adequate soil moisture.

For both locations, plant roots were harvested eight weeks after planting in the field. The middle plants in each plot were preferentially harvested for RNA analysis unless middle plants did not survive. Harvesting was conducted at least an hour after sunrise to ensure the plants were in their active growing phase. Plants were dug up using drain spades, being careful not to cut roots or damage them when pulling up the plants. The roots, still attached to the plants, were briefly soaked in water to loosen the soil and then gently washed with a hose to remove remaining soil. The washed roots were cut off the plant with shears, dried with paper towels, placed in a plastic zip-lock bag, and submerged in a dry ice-ethanol bath to freeze the samples. In Georgia, the samples were transferred to a -80°C freezer at the end of each day and stored at -80°C until RNA extraction. In Arizona, the samples were also stored in a -80°C freezer at the end of each day and transported on dry ice to UGA at the end of the harvest (harvesting the RNA blocks took 3 days), where they were returned to a -80°C freezer until RNA extraction.

About 3 months after the root harvest, aboveground biomass was harvested for both locations, with the second center plant being preferentially selected. Phenotypic data on plant growth stage at the time of biomass harvest was also collected for both locations. In Georgia, belowground biomass as well as plant height and stalk diameter measurements were collected. Due to time and labor constraints, these measurements were not collected in Arizona. The aboveground biomass was cut off and dried prior to measuring their dry weights. In Georgia, the plants were dried in 50°C ovens at J. Phil Campbell Sr. Research and Education Center and Iron Horse Farm for at least 7 days. In Arizona, the samples were dried at Maricopa Agricultural Center in a temperature-unregulated greenhouse that regularly gets above 37°C due to the Arizona climate. Samples were left to dry for 60 days prior to weighing.

RNA Extraction

Frozen sorghum roots were ground in liquid nitrogen with a mortar and pestle. RNA extractions were done on three aliquots of ~100 mg of root powder per sample to allow more input tissue to be processed. 1 mL of trizol was added to each aliquot and vortexed. Samples were incubated at room temperature for 1 hour. Following incubation, 200 µL of chloroform was added, and tubes were mixed by inverting several times. After incubating at room temperature for 2-3 minutes, samples were centrifuged at 12,000 RPM for 15 minutes at 4 °C. The aqueous layer was removed and put into a clean 5 mL tube, combining the aqueous layers from the three aliquots into one tube. An equal volume of room temperature 100% ethanol was added and the solution was added to a Zymo-Spin^{TM} IC Column placed in a Collection Tube, and centrifuged for 30 seconds.

In order to process the entire sample, the spin column was filled and centrifuged multiple times until all the sample had passed through the column. Flow-through was discarded. Subsequent steps closely followed the Zymo RNA Clean & Concentrator kit protocol, with the exception of eluting in 20 µL of water to increase the RNA concentration instead of the suggested 100 µL of water. RNA concentrations were measured with a nanodrop, and all samples were run on 0.8% agarose gels to verify RNA quality.

MiRNA Library Preparation

The miRNA Library Preparation protocol used was adapted from Persson et al. (2017) by Alex Johnson and Sabrena Rutledge. Protocol was further streamlined by Gurjot Sidhu. A visual overview of the protocol is shown in Figure 2.3. 500 ng of total RNA was used per sample. For the forward index, i501 (Persson et al., 2017) was used for all the samples. For the reverse index, 96 unique i7 indices were used (Glenn et al., 2019). Final library concentrations were measured with a Qubit 1X dsDNA High Sensitivity Assay Kit (Invitrogen). A total of 25 ng/µL from each sample with a concentration greater than 4 ng/µL was pooled for a total of 19-96 samples per pool. 200 µL of pooled volume was concentrated to 40 µL using a 4X bead cleanup with Sera-Mag SpeedBeads Streptavidin-Blocked Magnetic Particles (Cytiva). Concentrated pools were measured with a Qubit 1X dsDNA High Sensitivity Assay Kit (Invitrogen) to determine concentrations. 30 µL of the concentrated pools were then size selected to a target size of 156 bp (Figure 2.4) on a 3% Agarose, Dye-Free, 100 bp-250 bp, Marker P gel cassette in a PippinPrep (Sage Science) based on the manufacturers' recommendations. After size selection, concentrations were again measured with a Qubit 1X dsDNA High Sensitivity

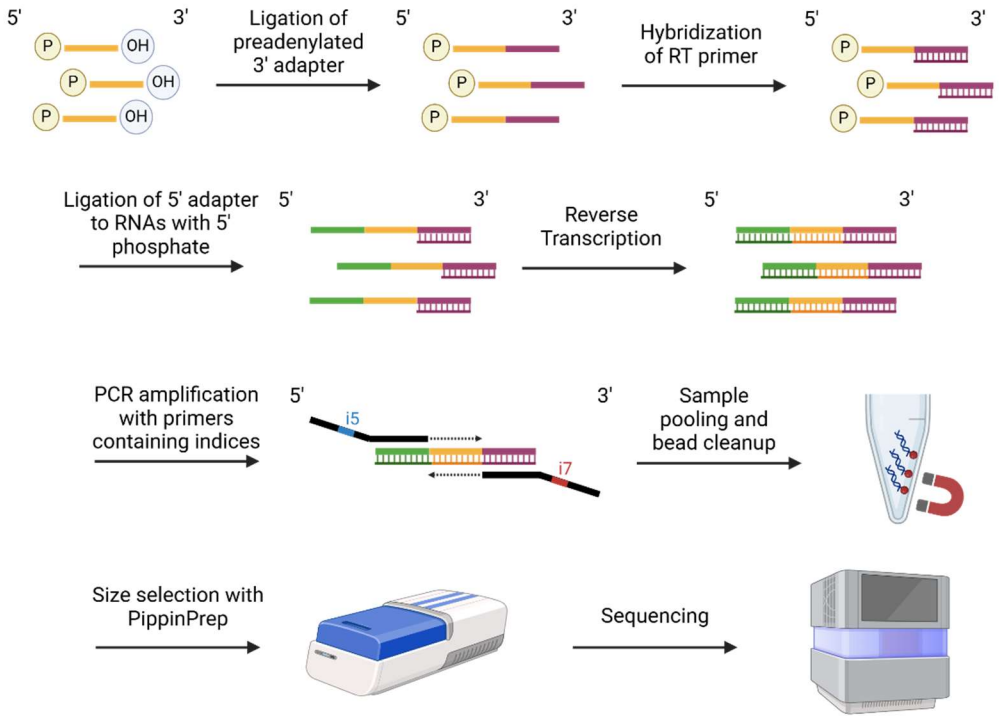

Figure 2.3. Flow diagram of the small RNA library preparation protocol. Mature miRNAs have a 5' phosphate (P) and a 3' hydroxyl (OH). The first step in library preparation is ligation of the 3' adapter. The 3' adapter has its 5' phosphate replaced with an adenyl group (App). T4 RNA ligase 2 can then catalyze the reaction between the miRNA's 3'-OH and the adapter's adenyl group, displacing the adenyl group and forming a phosphodiester bond. The reverse transcription (RT) primer hybridizes to the 3' adapter. Next, the 5' adapter's 3'-OH is attached to the 5'-P of the miRNA via T4 RNA Ligase 1. Reverse transcription then produces cDNA. PCR is performed using primers that incorporate i5 and i7 barcodes/indices. After PCR amplification, samples are pooled and then a bead cleanup is performed to concentrate the pool. Target library size for miRNAs (156 bp) is sized selected using a PippinPrep (see Figure 2.4 for breakdown of final library size/composition). Size selected pools can then be sequenced. (Diagram made in BioRender)

Assay Kit (Invitrogen). Pools were analyzed on a Fragment Analyzer Automated CE System (Agilent) at the Georgia Genomics and Bioinformatics Core (GGBC) at UGA. Once fragment analysis confirmed the presence of fragments of ~156 bp in size, pools were sent for Illumina sequencing. A total of 35 µL of each pool at 5 nM was sent to Admera Health (South Plainfield, NJ) for sequencing on a NovaSeq X Plus or to GGBC for sequencing on a NextSeq 2000 to obtain ~10 M reads/ sample.

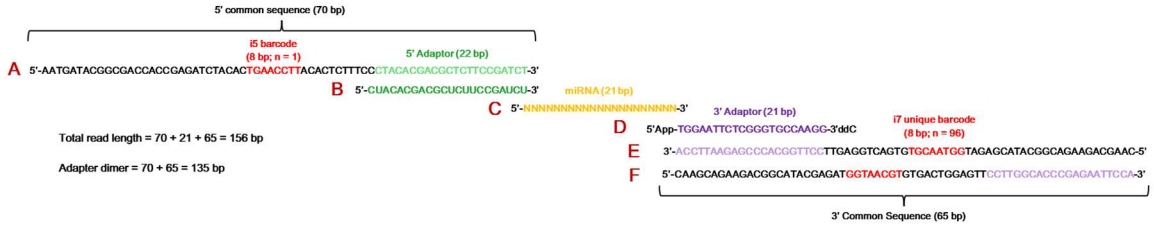

Figure 2.4. miRNA library structure. A) 5' Illumina-compatible common sequence containing the i5 barcode which is added during the PCR amplification step by the 5' PCR primer. B) 5' adapter sequence. Contains binding site for the 5' PCR primer. C) Actual small RNA being sequenced. D) 3' adapter sequence. Contains the binding site for the reverse transcription primer and 3' PCR primer. E) 3' Illumina-compatible common sequence containing the unique i7 barcode. Added during the PCR amplification step by the 3' PCR primer. F) 3' common sequence written 5'-3' to represent the coding strand in the final sequencing read.

MiRNA Identification

Preprocessing

After sequencing, samples with fewer than 5 million reads were removed from the dataset. The remaining raw small RNA reads were preprocessed to remove the adapter sequence using Cutadapt v4.1 (Martin, 2011). Size filtering was applied to retain reads between 17 and 30 nucleotides in length. Reads with a Phred quality score below 30 were discarded.

To enable alignment analysis, a Bowtie index was generated with bowtie-build v1.3.1 (Langmead et al., 2009) for the *Sorghum bicolor* BTx623 reference genome v5.0 downloaded from Phytozome (https://phytozome-next.jgi.doe.gov/info/Sbicolor_v5_1). The small RNA reads were then mapped to the sorghum reference genome using Bowtie

v1.3.1 (Langmead et al., 2009), allowing for zero mismatches. This generated .sam files as well as mapped .fastq files containing the sequences that successfully mapped to the sorghum reference genome. The .sam files were processed using SAMTools v1.21 (Li et al., 2009) to generate sorted .bam files as well as their corresponding .bai index files for downstream use.

After mapping the small RNA reads against the sorghum reference genome to filter out unmapped sequences, reads were compared against the Rfam database version 14.10 (Griffiths-Jones et al., 2003), a collection of non-coding RNA families, to filter out non-miRNAs. For this analysis, all the rRNA, tRNA, snRNA, and snoRNA sequences from Rfam were downloaded. Bowtie v1.3.1 (Langmead et al., 2009) was used to index the downloaded Rfam sequences and then map the small RNA reads against the created index. Perfectly mapped reads were discarded. The remaining unmapped reads were used as input for miRDeep-P2 miRNA identification. miRNAs were identified used two identification software: ShortStack (Axtell, 2013) and miRDeep-P2 (Kuang et al., 2019). No filtering against Rfam was done for reads processed through ShortStack.

ShortStack

All reads (in fastq format) with a perfect match to the sorghum reference genome were used for ShortStack. No filtering was carried out for known small RNAs from Rfam. ShortStack was run using default settings except that the --dn_mirna parameter was added in order to allow the software to annotate *de novo* miRNA clusters in the dataset.

ShortStack's algorithm identifies small RNA clusters by first identifying 'islands', or regions with significant alignment coverage. The significant alignment coverage is specified by the --mincov parameter, with the default being 1 read per million. Because samples with fewer than 5 million total reads were filtered out (see *Preprocessing* section), a minimum depth of 5 reads at a given small RNA cluster is required for it to be considered a potential miRNA cluster. Once islands have been identified, they're padded by adding 200 bp of upstream and downstream sequence to determine if fragments overlap. Overlapping regions are then categorized as clusters and represent potential miRNA precursor regions (see Figure 2.5).

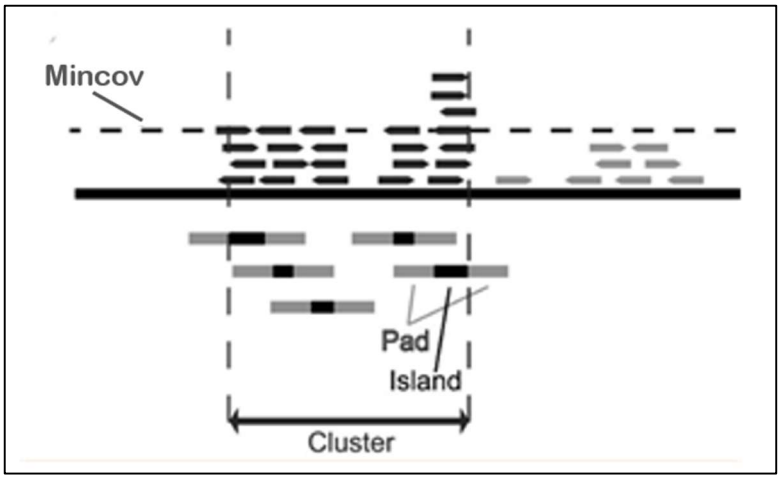

Figure 2.5. Visual showing ShortStack's cluster identification components (Axtell, 2013)

Small RNA reads in each cluster are categorized based on their size. MiRNAs falling within the size range 21-24 bp are kept for downstream miRNA identification steps. The algorithm then assesses the likelihood of hairpin structure formation. Loci that are confirmed to be able to form hairpin structures consistent with known miRNAs are annotated as miRNA loci. ShortStack then outputs the sequences that are most likely mature miRNAs.

miRDeep-P2

The input data for miRDeep-P2 (Kuang et al., 2019) consisted of the reads that successfully mapped to the sorghum reference genome, with reads that mapped to known rRNA, tRNA, snRNA, or snoRNA from Rfam removed (see *Preprocessing* section). Because miRDeep-P2 requires a single .fastq input file, reads across all the samples were concatenated into a single .fastq file. Additionally, fastx_collapser from FASTX-Toolkit v0.0.14 (Hannon, 2010) was used to collapse identical sequences to reduce redundancy and lighten the computational load for running miRDeep-P2.

miRDeep-P2 identifies potential miRNAs by first mapping the reads against the reference genome. It then extracts a 250 base pair window around the read from the reference to predict RNA secondary structures. Any reads that are included in the extracted window for sequences determined to form secondary structures likely to be miRNA precursors are then processed by the miRDeep core algorithm (Friedländer et al., 2008).

The miRDeep core algorithm assigns a score to each potential miRNA precursor based on characteristics that qualify it to give rise to a mature miRNA. One primary characteristic considered when scoring is the number of reads corresponding to the mature miRNA, miRNA*, or loop sequences within the hairpin structure, with a higher number of aligned reads increasing the likelihood of the sequence being a valid precursor (see Figure 2.6).

All miRDeep-P2 default settings were used for processing the small RNA reads in this experiment. After running, miRDeep-P2 provides likely mature miRNA sequences.

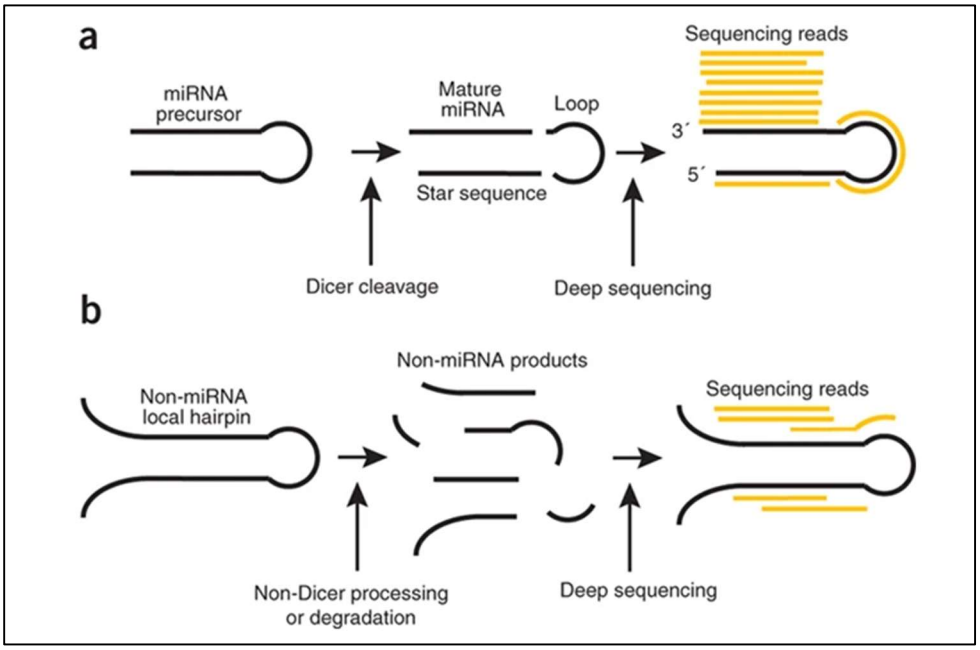

Figure 2.6. Diagram depicting miRDeep core algorithm scoring (a) When miRNA precursors are processed by Dicer (or DCL as in plants), three sequences are generated: the mature miRNA sequence, the miRNA, and the loop. When mapping reads to the reference genome, reads should accumulate in a pattern reminiscent of the three Dicer products, with the mature miRNA being the most abundant. (b) In addition to miRNA hairpins, other non-miRNA hairpins may be represented in the sequence data. However, reads mapping to these hairpins lack the configuration of Dicer-cleaved sequences and can thus be discarded. (taken from Friedländer et al. (2008))*

Postprocessing

After miRNAs were identified with miRDeep-P2 and ShortStack, the mature miRNA sequences were compared to the miRBase miRNA database (https://www.mirbase.org/). To accomplish this, miRNA sequences were downloaded from miRBase release 22.1 (Kozomara et al., 2019) and used to create a BLAST database using BLAST+ v2.14.1 (Camacho et al., 2009). A blastn analysis of the identified mature miRNAs to the miRBase BLAST database generated a list of known miRNAs. Identified

miRNAs with more than 3 base pair mismatches with miRNAs from miRBase were categorized as potentially novel miRNAs.

Transcript quantification for mature miRNAs was performed to estimate miRNA expression levels. To accomplish this, output information retrieved from ShortStack or miRDeep-P2 was used to generate .bed files with the genomic coordinates for the identified mature miRNAs (ShortStack and miRDeep-P2 outputs were processed separately). BEDTools v2.31.0 (Quinlan & Hall, 2010) was utilized to then calculate the read counts for each miRNA across all of the samples using the multicov feature and the previously generated sorted.bam files (see *Preprocessing* section) as inputs. The '-f 1' parameter was used to specify that only reads that overlapped the entire miRNA sequence were counted. This ultimately generated an output text file with the transcript abundances for each sample listed alongside their associated miRNAs.

Normalization

ShortStack and miRDeep-P2 raw transcript abundances were combined, and duplicate miRNAs removed using the dplyr package in R (Wickham, 2015). MiRNAs with an average read count across all samples of 10 or less were discarded. The remaining miRNA transcript counts were normalized using DESeq2 (Love et al., 2014). The full Georgia dataset was normalized together. A subset of Arizona samples that had been sequenced (24 samples, 12 genotypes with 2 replicates each) were normalized alongside their corresponding Georgia samples, but was kept separate from the full Georgia dataset.

MiRNA source (ShortStack, miRDeep-P2, or both) as well as proportion of known to novel miRNAs were assessed before and after filtering. Venn diagram visuals showing the miRNA sources were generated using the VennDiagram package in R (Chen & Boutros, 2011). Principal Component Analysis (PCA) was also performed in R to assess the structure of the normalized miRNA count data and examine potential block effects. PCA were done separately for the GA samples, and for the Arizona/Georgia subset. For the GA samples, Blocks 2, 8, and 10 represent the non-fertilized (N-P-) blocks used for expression analysis. The first two PCs were extracted for visualization, and points are colored by block. For the Arizona/Georgia subset, four blocks are represented in the PCA (Blocks 2 and 8 from Georgia, and Blocks 2 and 4 from Arizona).

Genotypic Effects

The effect of genotype on miRNA expression was assessed using a linear mixed-effects model (LMM). The model was run in SAS version 9.4 (SAS, 2023) using the PROC MIXED procedure. Genotype was treated as a fixed effect and block was included as a random effect to account for block variation. Type III Tests of Fixed Effects was used to determine differences between genotypes, and least squares means (LSMeans) was calculated using Tukey's post hoc adjustment for multiple comparisons.

Differential Expression Analysis

Differential Expression Analysis using DESeq2 (Love et al., 2014) was conducted on the miRNAs identified in the Arizona/Georgia subset to compare the impact of location on miRNA expression (design = ~ Location). Log2 fold changes were used to

estimate the level of differential miRNA expression between the two locations as well as determine the regulation type (up (+) or down (-) regulation). P-values generated from the Wald test were corrected for multiple hypothesis testing using the Benjamini-Hochberg correction to control for false discovery rate. A significance threshold of $p < 0.05$ and Log2-fold change > |1| was used to determine miRNAs that were significantly differentially expressed by location.

Target mRNA Identification

The potential target mRNA sequences were predicted bioinformatically using psRNATarget (Dai et al., 2018) and TargetFinder (Bo & Wang, 2005). The sorghum genome was selected as the reference. The sorghum genome version available in the psRNATarget database, a web-based tool, came from Phytozome version 10 (Sbicolor_255_v2.1.cds.fa.gz). Dai et al. (2018) determined that when assessing the psRNATarget outputs, an Expectation (E) value of ≤ 3 represents a relaxed significance threshold, whereas E ≤ 2 would represent a more stringent cutoff that helps avoid false positives. The current study used a E ≤ 2.5 cutoff to avoid missing potential target sequences while also trying to minimize false positives.

For the second program, TargetFinder, sorghum cDNA sequences were downloaded from Phytozome (CDS File: Sbicolor_730_v5.1.cds_primaryTranscriptOnly.fa.gz) and used as the reference. The recommended TargetFinder default settings were used, with a cutoff prediction score of 4 being used for higher prediction confidence (scores represent the level of complementarity; 0: perfect match). The identities of select target genes were determined

by using the gene peptide sequence for doing a BLASTp search against the UniProtKB/Swiss-Prot NCBI database.

Expression Genome Wide Association Study

An expression Genome Wide Association Study (eGWAS) was run using the miRNA expression levels (determined by normalized miRNA read counts) from the GA trial as the phenotypes. Genotypic data came from Hu et al. (2019)'s 459,304-SNP dataset for 10,323 sorghum genotypes generated with version 3.1 of the sorghum reference genome. SNP dataset was filtered (Minor Allele Frequency ≥ 0.05; Missing Data ≤ 0.5) to produce subset of 92,455 SNPs for the BAP genotypes. GWAS was run on the normalized samples from Block 2 (92 genotypes) and Block 8 (88 genotypes) separately since they each represented a different subset of genotypes (only 27 overlapping genotypes between Block 2 and 8). As such, each GWAS run differed somewhat in terms of the SNP density, with the Block 2 eGWAS run having 92,136 SNPs and the Block 8 having 91,905 SNPs.

Two GWAS models were used: Fixed and Random Model Circulating Probability Unification (FarmCPU) (Liu et al., 2016) and Bayesian information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) (Huang et al., 2019). For both BLINK and FarmCPU, all the default parameters were used. FarmCPU and BLINK were run in R (version 4.4.1) using the GAPIT (Genome Association and Prediction Integrated Tool) package (Wang & Zhang, 2021). The Bonferroni multiple test threshold ($\alpha = 0.01$) was used to assess significance. Additionally, to help control population structure, the

first four principal components (PCs) were included in the model as covariates as they accounted for a majority of the genetic variance in the dataset.

Results

Small RNA Sequencing

A total of 388 samples from the Georgia field experiment were sequenced. Additionally, a smaller subset of 24 samples were sequenced from the Arizona field experiment. After removing 161 GA samples with a total read number < 5 million, the average read number for the remaining 227 GA samples was 18,300,847 (range 5,099,530—86,318,340). The 24 AZ samples had an average read number of 30,865,354 (range 20,241,497—41,497,813).

MiRNA Identification: Georgia Samples

A total of 98 unique miRNAs were identified across both ShortStack and miRDeep-P2. ShortStack identified a total of 52 miRNAs and miRDeep-P2 identified 87. 11 miRNAs were uniquely identified by ShortStack, 46 uniquely by miRDeep-P2, and 41 were identified by both programs (see Figure 2.7). Of the 98 identified miRNAs, 78 had matches to know miRNA sequences from miRBase. The remaining 20 miRNAs were categorized as novel miRNAs.

After removing lowly expressed miRNAs (with average read count across all samples being 10 or less), 54 miRNAs, including eight novel miRNAs, remained for downstream analysis (see Figure 2.8 for filtered miRNAs identified by ShortStack and miRDeep-P2). The 54 miRNAs and their sequences alongside the program that identified them are listed in Table 2.1 (Full data in Supplemental S2.1.1).

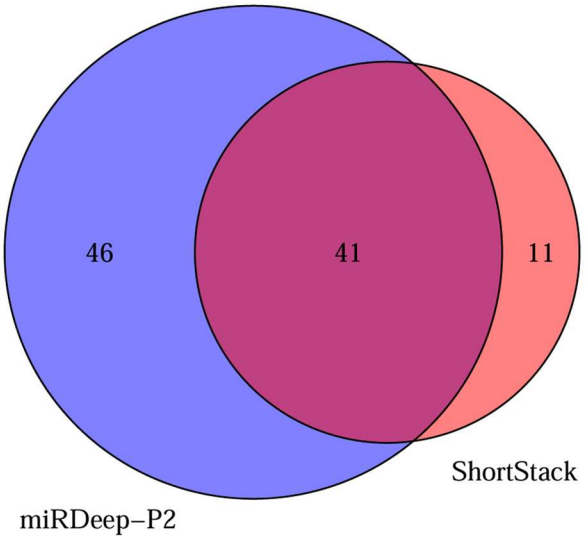

Figure 2.7 Venn Diagram showing the number of miRNAs identified uniquely by miRDeep-P2 or ShortStack as well as overlapping miRNAs before filtering. Made in R using the VennDiagram and grid packages (Chen & Boutros, 2011; Murrell, 2005)

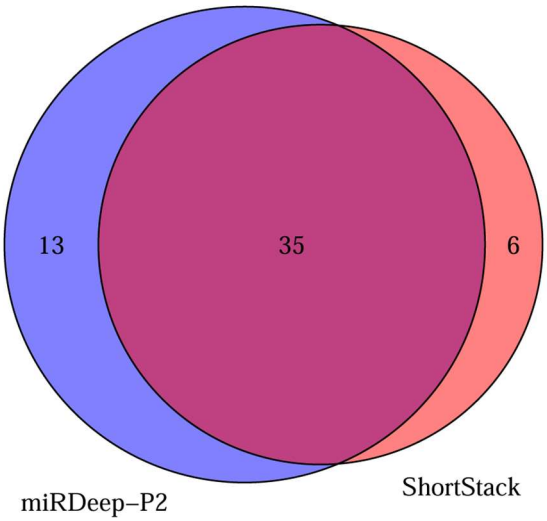

Figure 2.8. Venn Diagram showing the number of miRNAs identified uniquely by miRDeep-P2 or ShortStack as well as overlapping miRNAs after filtering out lowly expressed miRNAs. Made in R using the VennDiagram and grid packages (Chen & Boutros, 2011; Murrell, 2005)

Table 2.1 Identified Georgia miRNAs (after filtering)

miRNA Name	Sequence	Program
miR156	TGACAGAAGAGAGTGAGCAC	Both
miR164a-5p	TGGAGAAGCAGGGCACGTGCT	Both
miR171a-3p-1	TGAGCCGAACCAATATCACTC	Both
miR393-5p	CTCCAAAGGGATCGCATTGAT	Both
miR160b-3p	GCGTGCAAGGAGCCAAGCATG	Both
miR166f	TCTCGGACCAGGCTTCATTCC	Both
miR319b-3p-1	CTTGGACTGAAGGGTGCTCCCT	ShortStack
miR319b-3p-2	TTGGACTGAAGGGTGCTCCCT	Both
miR529-5p	AGAAGAGAGAGAGTACAGCCT	Both
miR164a	TGGAGAAGCAGGGCACGTGCA	Both
miR171a-3p-2	TTGAGCCGCGCCAATATCTCT	ShortStack
miR167a-3p	AGATCGTCTGGCAGTTTCATC	ShortStack
miR172b-5p	GCAGCATCATCAAGATTCACA	Both
miR167g	TGAAGCTGCCAGCATGATCTGA	Both
miR399b	TGCCAAAGGAGAGTTGCCCTG	Both
miR166-3p	TCGGACCAGGCTTCATTCCCC	Both
miR171a	TTGAGCCGCGTCAATATCTCC	Both
miR397a	TTGAGTGCAGCGTTGATGAGC	Both
Novel-1	TCGATGTCACACTCGTTCGCA	Both
Novel-2	GCCTTCAGGAGGGGGGGATCC	Both
Novel-3	TTGGCCAAAGTTAGAGAAGTT	Both
Novel-4	TGAGCCGAGCCAATATCACTT	Both
miR167h-3p	AGGTCATGCTGTAGTTTCATC	Both
miR396c-5p	TTCCACAGCTTTCTTGAACTT	Both
miR827	TTAGATGACCATCAGCAAACA	Both
miR1030j	AGGTGCAGGGGCAGATGCAGT	Both
miR167a-1	TGAAGCTGCCAGCATGATCTA	ShortStack
miR167c	TGAAGCTGCCAGCATGATCTG	Both
miR171d	TGATTGAGCCGTGCCAATATC	Both
miR396d	TCCACAGGCTTTCTTGAACTG	Both
miR5385	CCACCAACACCACCGCTTCTC	Both
miR5564a-1	TGGGGAAGCAATTCGTCGAAC	Both
miR5564b	AGCAATTCGTCGAACAGCTTG	Both
miR159	CTTGGATTGAAGGGAGCTCC	ShortStack
miR159a	TTTGGATTGAAGGGAGCTCTG	ShortStack

miR528	TGGAAGGGGCATGCAGAGGAG	Both
miR390b	AAGCTCAGGAGGGATAGCGCC	Both
miR1432-5p	CTCAGGAGAGATGACACCGA	Both
miR156k-5p	TGACAGAAGAGAGCGAGCAC	Both
miR160a-3p	GCGTGCAAGGGGCCAAGCATG	Both
miR168b-5p	TCGCTTGGTGCAGATCGGGAC	Both
Novel-5	TTGGACCACAGAAGCTACAGC	miRDeep-P2
Novel-6	CGAGCGGCTGTCATATCGACC	miRDeep-P2
Novel-7	AGGATGGATGTATCTCATGGCC	miRDeep-P2
Novel-8	TTGCATGGGAGATGAGTTTGC	miRDeep-P2
miR5568d-5p	AGTTATGTATCTAGAAAAGCC	miRDeep-P2
miR5564a-2	TGGGAAGCAATTCGTCGAACA	miRDeep-P2
miR169d-3p	TGGGCGGTCACCTTGGCTAGC	miRDeep-P2
miR395c	TGAAGTGTTTGGGGGAACTC	miRDeep-P2
miR394a	TTGGCATTCTGTCCACCTCC	miRDeep-P2
miR167a-2	TGAAGCTGCCAGCATGATCT	miRDeep-P2
miR393	TCCAAAGGGATCGCATTGATC	miRDeep-P2
miR172d	AGAATCTTGATGATGCTGCAT	miRDeep-P2
miR171g	TGATTGAGCCGCGCCAATATC	miRDeep-P2

The identified miRNAs represent 26 miRNA families. MiR166 is the most highly expressed miRNA family across the dataset, and miR5385 the most lowly expressed (after filtering). Figure 2.9 shows the top 25 most highly expressed miRNA families in

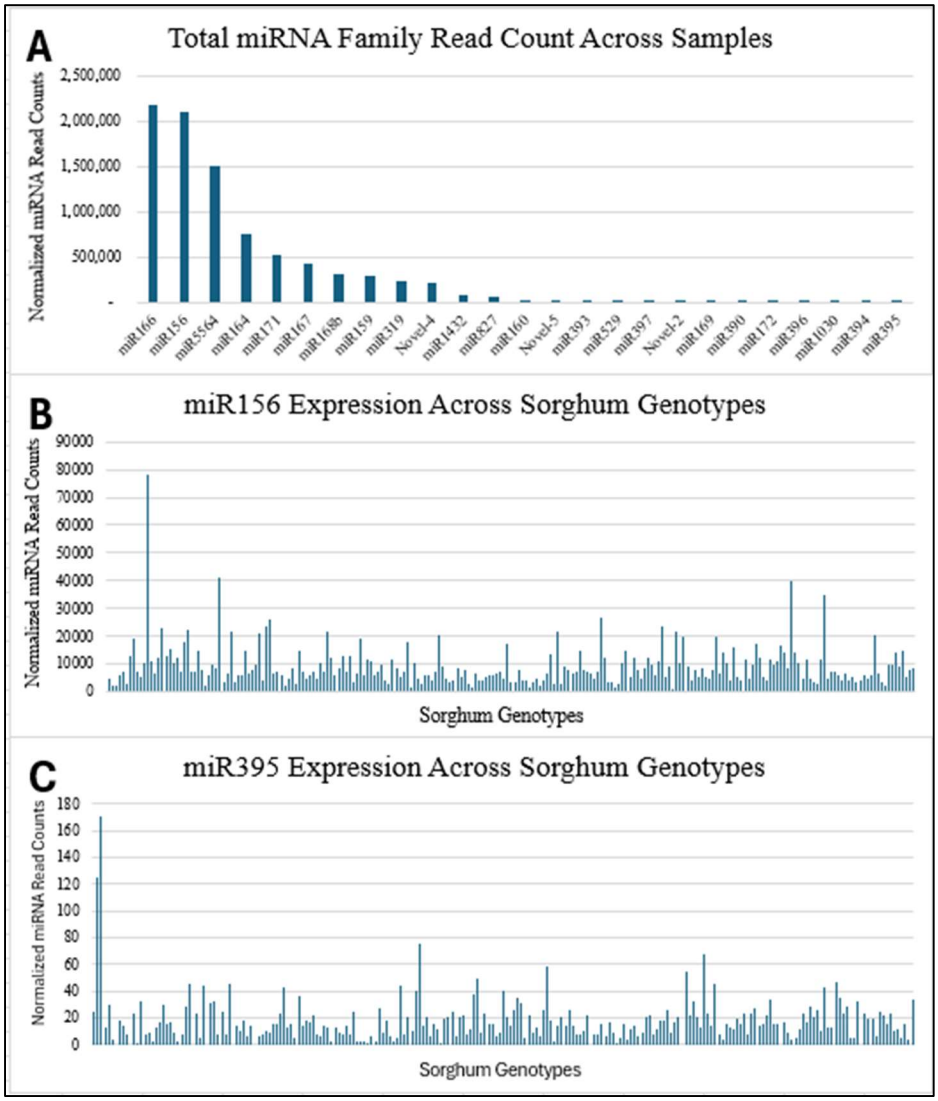

Figure 2.9. (A) Graph showing the top 25 most highly expressed miRNA families based on total miRNA read counts across the dataset for a given miRNA. (B) Graph of miR156 (highly expressed miRNA). (C) Graph of miR395 (lowly expressed miRNA).

the dataset as well as an example of genotypic variation in a highly (miR156) and more lowly (miR395) expressed miRNA. Figure 2.10 shows the PCA plot for the normalized

miRNA count data with colors representing the blocks (B2, B8, and B10). Note that the genotypes present in each of the blocks only partially overlap. There were also 14 miRNAs that had significant ($P > 0.05$) genotypic effects. In other words, the expression of these miRNAs varied significantly by genotype. Significant miRNAs are listed in Table 2.2. However, it should be noted that not all genotypes had replicates, so some significance may be attributable to plant variation by environment/block.

Table 2.2. miRNAs with Significant Genotypic Effects (GA)

miRNA Name	P-value
Novel-7	0.0002
miR396c-5p	0.0006
miR827	0.0013
Novel-5	0.0014
miR5564a-1	0.0042
miR156k-5p	0.0067
miR167a-2	0.007
Novel-1	0.0082
Novel-8	0.0143
miR156	0.0189
miR390b	0.0192
miR167g	0.0292
Novel-4	0.042
miR164a	0.0484

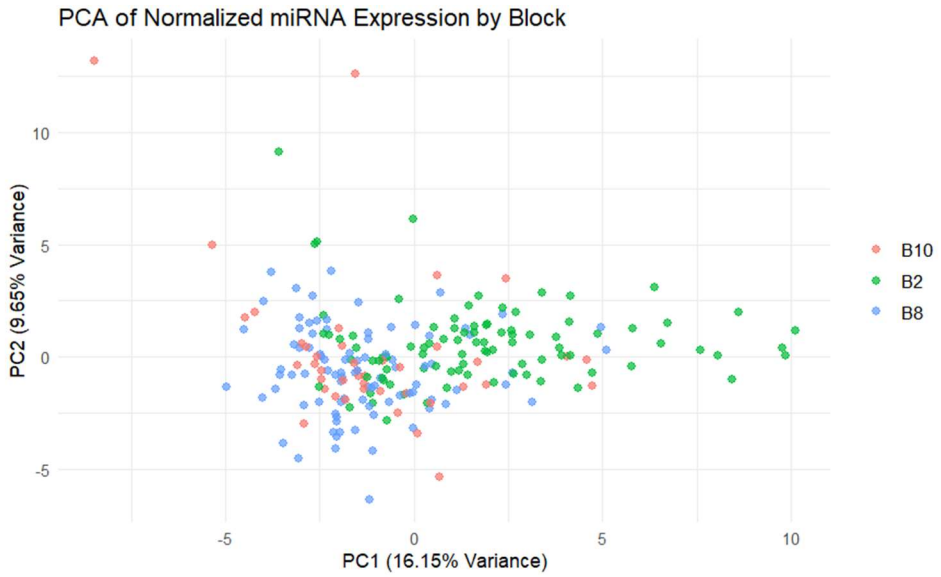

Figure 2.10. Principal Component Analysis plot showing the first two PCs for the normalized miRNA count data. Created in R using ggplot2, dplyr, tidyr, and RcolorBrewer packages (Neuwirth, 2014; Villanueva & Chen, 2019; Wickham, 2015; Wickham et al., 2023)

MiRNA Identification: 12 genotypes common to Arizona and Georgia trials

Prior to filtering, the number of miRNAs identified overall and by each program was the same as for the Georgia data. However, filtering of the data for the 12 genotypes common to the Arizona and Georgia trials produced a different subset of miRNAs compared to the overall set of Georgia samples. The breakdown of the number of miRNAs identified by ShortStack (51 miRNAs), miRDeep-P2 (61 miRNAs), or both (41 miRNAs) is visualized in Figure 2.11.

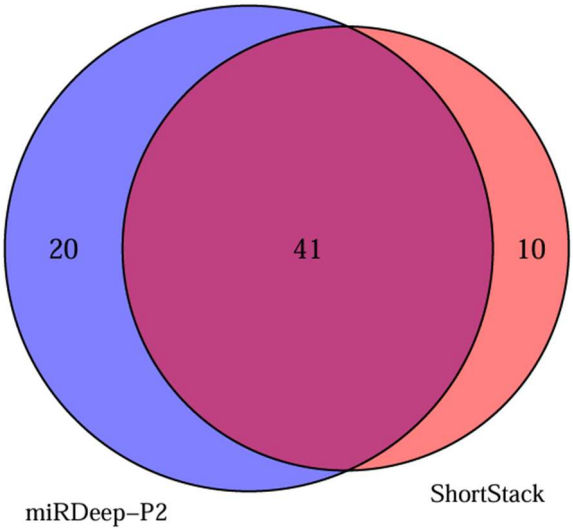

Figure 2.11. Venn Diagram showing the number of miRNAs identified uniquely by miRDeep-P2 or ShortStack as well as overlapping miRNAs after filtering has been done. Made in R using the VennDiagram and grid packages (Chen & Boutros, 2011; Murrell, 2005)

Arizona vs. Georgia Differential Expression Analysis

A PCA analysis of the miRNA expression levels across 12 genotypes in AZ and GA (2 replicates per location) showed clustering by location (Figure 2.12) with location explaining ~33% of the variation, indicating location has a strong effect on miRNA expression levels. Differential expression analysis revealed that transcript levels of 23 miRNAs were significantly higher in Arizona compared to Georgia, and that 14 miRNAs had transcript levels that were significantly lower (Figure 2.13). The top-5 most significantly up- and down-regulated miRNAs in the two environments are listed in Table 2.3 (Full list in Supplemental S2.1.4).

Table 2.3 Top Up- and Down-regulated miRNAs from GA/AZ Differential Expression Analysis

miRNA Name	Log2 Fold Change	P-Value	Regulation
miR399g	3.121246123	2.69E-10	Upregulated in AZ
miR394a	2.776889576	2.02E-35	Upregulated in AZ
Novel-11	2.65869642	5.40E-08	Upregulated in AZ
miR399b-2	2.631725895	5.92E-05	Upregulated in AZ
miR396e	2.500737228	4.76E-05	Upregulated in AZ
Novel-4	-6.235467552	1.36E-52	Downregulated in AZ
Novel-6	-3.501837503	4.58E-13	Downregulated in AZ
Novel-2	-2.920723255	7.33E-09	Downregulated in AZ
miR156	-2.803331926	4.50E-35	Downregulated in AZ
miR166-3p	-2.73746974	2.57E-11	Downregulated in AZ

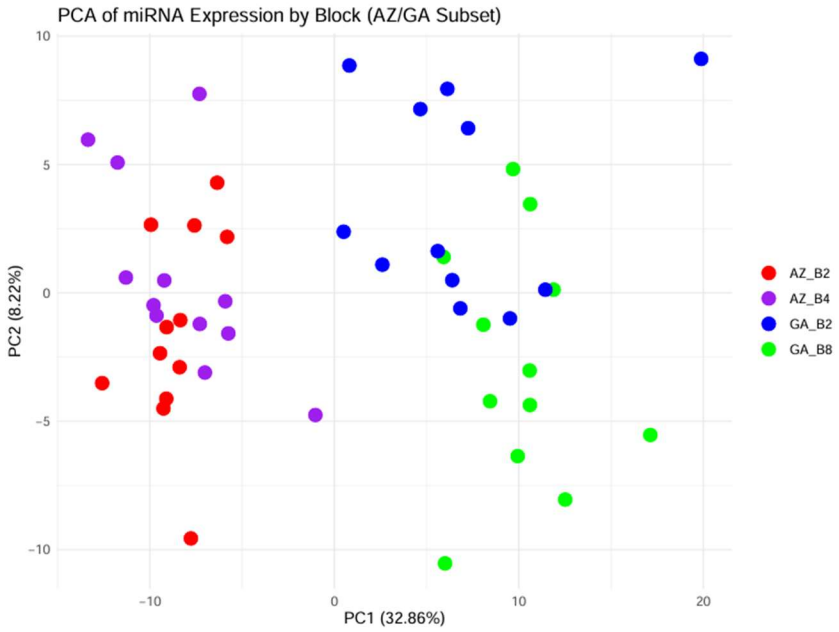

*Figure 2.12. Principal Component Analysis plot of miRNA expression levels in Georgia and Arizona subsets (12 genotypes * 2 replications * 2 environments). Created in R using ggplot2, dplyr, and tidyr, (Villanueva & Chen, 2019; Wickham, 2015; Wickham et al., 2023)*

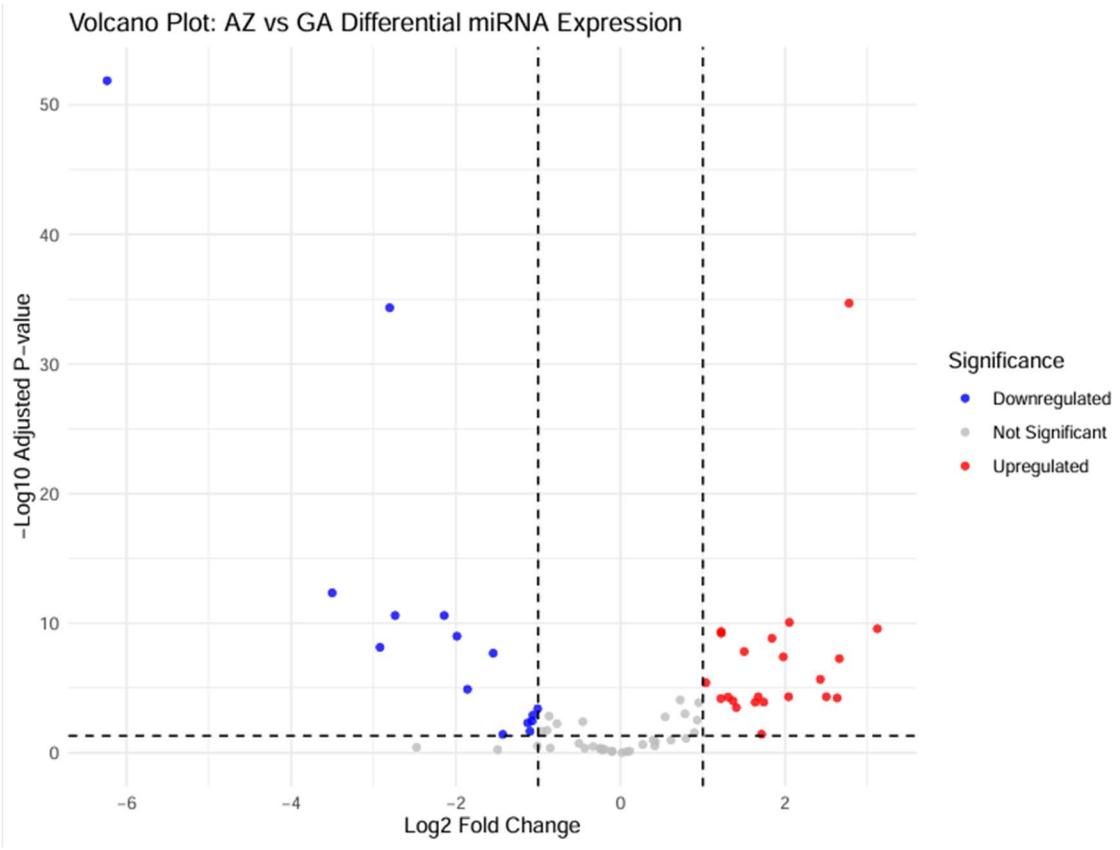

Figure 2.13. Volcano plot showing the distribution of miRNAs that are downregulated or upregulated in Arizona compared to Georgia. Image generated in R using the ggplot2 and dplyr packages (Villanueva & Chen, 2019; Wickham, 2015)

Target Gene Identification

A total of 1,378 unique potential target genes were identified across psRNATarget and TargetFinder for a set of 71 miRNAs (combined, non-redundant list of miRNAs from Georgia and Arizona). psRNATarget identified 946 target genes, TargetFinder 596, and 164 genes were identified by both softwares (Figure 2.14). The target genes identified for each mRNA are listed in Supplementary S2.1.6. The identities of the target genes of a

few of the miRNAs identified as significant based on previous analysis (Differential Expression, Genotypic Effect) are shown in Table 2.4.

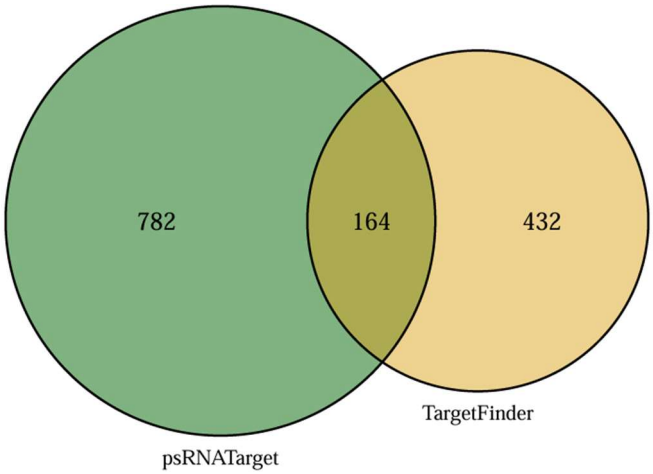

Figure 2.14. Venn Diagram showing the target mRNAs identified uniquely by psRNATarget or TargetFinder as well commonly identified mRNAs. Made in R using the VennDiagram and grid packages (Chen & Boutros, 2011; Murrell, 2005)

Table 2.4. Identified Target Genes

MiRNA Name	Target Gene ID	Target Gene	Query Cover	Percent Identity	Species
miR399	Sobic.007G164400	Inorganic phosphate transporter 1-6	96%	85.31%	*Oryza sativa*
miR156	Sobic.004G058900	Squamosa promoter-binding-like protein 11	94%	51.64%	*Oryza sativa*
miR396	Sobic.001G139800	Growth-regulating factor 9	96%	69.88%	*Oryza sativa*
Novel-4	Sobic.004G290800	Scarecrow-like protein 6	50%	43.96%	*Arabidopsis thaliana*

Expression GWAS

eGWAS was run using the Georgia samples. The number of genotypes analyzed per replicate (block) varied and the genotypes only partially overlapped. Block 2 had 92 samples, Block 8 had 88 samples, and Block 10 had 42 samples. Because Block 10 had so few samples, it was not used for eGWAS. Twenty-seven of the analyzed genotypes were common between Block 2 and Block 8. The eGWAS was run separately for Block 2 and 8, resulting in 70 significant eQTL for Block 2 and 80 for Block 8 (Supplementary S2.1.3). While there were five miRNAs that had eQTL identified from both Block 2 and Block 8, only one miRNA (miR528) had eQTL from different blocks that mapped within 1 Mb of each other (348,700 bp apart). A total of 34 eQTL (across B2 and B8) were commonly identified between both BLINK and FarmCPU. All of the identified eQTL were in *trans*, with the shortest distance between a miRNA (miR5385) and its eQTL being 3,729,503 bp.

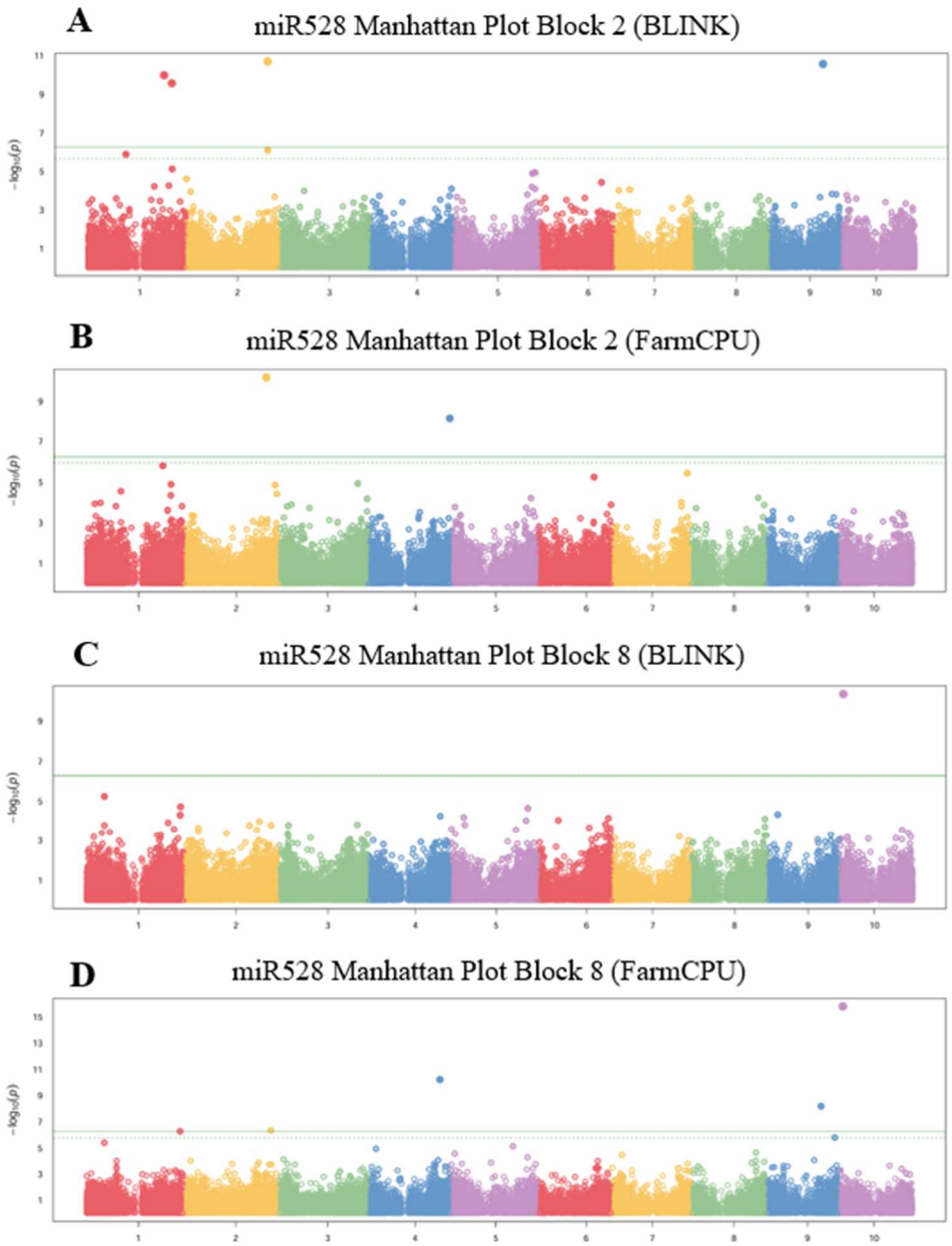

Figure 2.15. Shows the Manhattan Plots for miR528 (A) Block 2 BLINK, (B) Block 2 FarmCPU, (C) Block 8 BLINK, and (D) Block 8 FarmCPU. Solid line represents Bonferroni correction (α = 0.01) and the dotted line (α = 0.05).

Discussion

MiRNA Identification

In the current study, the two programs, ShortStack (Axtell, 2013) and miRDeep-P2 (Kuang et al., 2019) were used in conjunction, allowing for increased miRNA discovery. Each program was able to identify unique miRNAs not found by the other. This is consistent with a previous comparison between ShortStack and miRDeep-P2 which had also determined that each program identifies some unique miRNAs not found by the other (Hammond et al., 2023). This supports the use of diverse miRNA identification software for future miRNA studies. It should be noted that the input files for MiRDeep-P2 and ShortStack were different. Prior to running miRDeep-P2, reads were mapped against known rRNA, tRNA, snRNA, or snoRNAs from Rfam. This approach, recommended by the developer, sought to eliminate previously identified small RNA types so that remaining sequences would consist of a higher proportion of real miRNA sequences. Additionally, filtering out reads mapping to Rfam helped reduce the number of reads being processed by miRDeep-P2. As miRDeep-P2 is computationally intensive, reducing the input reads helps with faster processing.

On the other hand, ShortStack is less resource intensive. As such, reducing the number of input reads isn't as necessary for optimal processing. In the current experiment, reads were not mapped to Rfam sequences prior to running ShortStack. It is possible that Rfam filtering could remove potential miRNAs that overlap with other non-coding RNAs. For example, Qu et al. (2015) identified two dicistronic genes that encoded for both small nucleolar RNA (snoRNA) and miRNA precursors. In this case, a single precursor sequence could be processed differently to either produce mature snoRNA or

miRNAs. While this overlap between miRNAs and other non-coding RNAs is not well characterized yet and may only occur infrequently, it could mean that filtering out Rfam sequences may be eliminating sequence data containing real miRNAs. Testing is underway to determine whether filtering inputs for rRNA, tRNA, snRNA and snoRNA affects miRNA discovery with ShortStack.

Arizona vs. Georgia Differential Expression Analysis

There were many miRNAs that were differentially expressed when genotypes were grown in Georgia compared to Arizona. MiR399 represents the most differentially expressed miRNA family and is upregulated in Arizona compared to Georgia genotypes. MiR399 family members target genes involved in phosphate homeostasis. MiR399 suppresses the ubiquitin-conjugating E2 enzyme Phosphate 2 (PHO2), which limits phosphate uptake to avoid inorganic phosphate (Pi) toxicity in shoot tissue. In response to phosphate starvation, miR399 is upregulated and suppresses PHO2, which allows for Pi levels to increase (Gelaw & Sanan-Mishra, 2021; Ledford et al., 2024). Given the fact that the Georgia field was more phosphorus limiting (10.46 ppm) compared to Arizona (98 ppm), upregulation of miR399 in AZ plants is unexpected. Wortmann et al. (2013) notes that soil phosphorus levels around 10 ppm and lower require phosphorus fertilizer inputs for optimum grain sorghum production. This suggests that Georgia's soils are phosphorus deficient. As such, and contrary to our observations, we would expect high levels of miR399 in GA leading to suppression of PHO2 and increased Pi uptake. However, the situation may be more complex. It has been shown that the gene encoding the non-coding protein INDUCED BY PHOSPHATE STARVATION1 (IPS1) in

Arabidopsis can bind to but is not degraded by miR399 due incomplete complementarity (Franco-Zorrilla et al., 2007). *IPS1* can thus sequester miR399. It will be interesting to assess expression of both *PHO2* and *IPS* (or similar non-coding proteins) in AZ- and GA-grown plants, work that is ongoing as part of a larger collaboration. Sorghum-AMF interactions may also provide a potential partial explanation. For example, if the low soil Pi in Georgia promoted more extensive plant - AMF interactions, the AMF symbiosis could help with mediating the low phosphorus levels. If AMF are extremely efficient at absorbing phosphate and contributing them to the sorghum host, then low levels of miR399 and thus increased phosphate uptake limitation by PHO2 would be logical. In contrast, if the Arizona plants had weaker AMF interactions and thus less efficient phosphate absorption, upregulation of miR399 in order to encourage phosphate uptake may have been necessary. However, this can't be validated until AMF data from the BAP is available to see if there are AMF abundance or composition differences between locations. Again, data on AMF abundance and species composition are being generated as part of a larger collaboration.

MiR156 was identified as one of the miRNAs significantly downregulated in Arizona. MiR156 is a conserved miRNA involved in many aspects of growth and development, usually by targeting SQUAMOSA promoter-binding protein-like (SPL) transcription factors (Wu et al., 2009). Interestingly, miR156 has been linked to AMF symbiosis, with Pradhan et al. (2023b) showing that miR156 overexpression resulted in a 50% increase in AMF colonization. Lower expression levels of miR156 in Arizona compared to Georgia plants fits with the hypothesis of lower AMF colonization because of a higher Pi level in Arizona compared to Georgia. MiR156 has also been linked to

drought-stress response, with overexpression further enhancing drought tolerance in plants. While, initially, it might be assumed that the Arizona plants would be more drought stressed due to the hot, dry climate, plants were regularly watered. The Georgia field was watered during seedling establishment, but then rainfed conditions were used for the remainder of the experiment. Increased miR156 expression could potentially be caused by the more intermittent water access in Georgia.

MiR166 has been identified as a negative drought stress regulator, meaning decreased expression is correlated with better drought tolerance (Pradhan et al., 2023a). However, miR166 is downregulated in Arizona, which contradicts the hypothesis that Georgia plants may have been more drought-stressed. So, while the literature provides some insights into miRNA functions, further experimental validation is needed to more directly determine miRNA functions in this specific context.

Genotypic Effects

In addition to some of the miRNA that were differential expressed between AZ and GA plants being linked to AMF symbiosis, some of the miRNAs that showed significant genotypic effects on miRNA expression in Georgia were also linked to AMF. For example, miR396 regulates GROWTH-REGULATING FACTOR (GRF) transcription factors, which are involved in growth and stress responses. More specifically, miR396 has been identified as a negative regulator of AMF symbiosis (Ledford et al., 2024). This could mean that variation in miR396 levels between genotypes could be correlated with variable AMF abundances. But again, AMF data for

the same sorghum roots used for generating the miRNA expression data is needed to validate this theory.

Target Gene Prediction

Potential target genes were bioinformatically predicted and gene identities were assessed for miRNAs of interest. Target gene identities were useful in corroborating the functions of key miRNAs inferred by previous literature by linking them to sorghum specifically. For example, miR399 discussed previously for its involvement in phosphate homeostasis (Gelaw & Sanan-Mishra, 2021; Ledford et al., 2024), had a target gene identified as an inorganic phosphate transporter. MiR156 has been commonly linked to regulation of SQUAMOSA promoter-binding protein-like (SPL) transcription factors (Wu et al., 2009). Here, it had target genes also identified as SPL proteins. And miR396, having been identified as a regulator of GROWTH-REGULATING FACTOR (GRF) transcription factors, also had a GRF target in the current experiment.

Additionally, target gene prediction can be useful for elucidating novel miRNA functions. For example, Novel-4 was determined to target a Scarecrow-like (SCL) protein. SCL protein expression has been linked to mycorrhizal colonization, with increased SCL expression being observed in cells with mycorrhizal arbuscules in tomatoes (Ho-Plágaro et al., 2019). This suggests that Novel-4 may play a role in downregulating AMF symbiosis in sorghum. Novel-4 was one of the miRNAs identified to be significantly downregulated in Arizona. Decreased Novel-4 may derepress SCL, allowing for increased mycorrhizal arbuscule formation in Arizona compared to Georgia. Novel-4 was also identified as showing significant genotypic variation in terms of

miRNA expression. This could mean that modulation of Novel-4 could be one mechanism through which a genotype regulates its interactions with AMF. However, these patterns need to be validated with AMF data. Additionally, while all of these proposed target genes may be legitimate, they will need to be validated against mRNA expression data to ensure miRNAs are indeed impacting their expression.

Expression GWAS

The present study identified many different eQTL across the two different experimental blocks. MiR528 is the only miRNA for which an eQTL was identified in the same region in both replicates analyzed. MiR528 has been linked to enhanced nitrogen absorption through the targeting of OsNRT2 transporter genes (Zhao et al., 2022). Interestingly, miR528 expression increased in response to nitrogen with the increase being more pronounced in response to NO_3^--N (Nitrate-Nitrogen) compared to NH_4^+-N (ammonium) (Zhao et al., 2022). This suggests that, while the presence of nitrogen in general results in increased miR528 expression, the form of nitrogen can also affect the degree to which miR528 varies (Zhao et al., 2022). Further, the fact that eQTL were identified for miR528 indicates that the miR528 response to nitrogen varies by cultivar.

All of the eQTL were *trans* regulating eQTL, meaning they were not in the same genomic region as the miRNA transcript. The identification of *trans* miRNA eQTL is consistent with Chen et al. (2020), who identified four trans miRNA eQTL in maize. *Trans*-eQTL indicate that a gene's expression is regulated through intermediary genes encoding for regulators such as transcription factors. As many of these intermediary genes are more responsive to the environment, it suggests that miRNA expression is more

linked to variable environmental conditions (Chen et al., 2020). This allows the miRNAs to be able to more finely regulate their target genes depending on what the situation dictates.

Environmental effects may, in part, explain the lack of consistency in eQTL identification across Blocks 2 and 8. For example, miR528 is one of the miRNAs for which multiple eQTL were identified across Block 2 and 8, but only one eQTL region was potentially conserved between the blocks (eQTL from each block were 348,700 bp apart). While efforts were made to make field conditions as consistent as possible, PCA results for miRNA expression in Georgia (Figure 2.10) suggest that miRNA expression in Block 2-plants varied some from Block 8 plants. This could mean that slightly variable conditions (nutrient levels, water drainage, etc.) were enough to influence miRNA expression. This variation between blocks could help explain why there was little overlap in the eQTL identified for each block (2 or 8).

Further, each block had a set of sorghum genotypes that only partially overlapped. Associated with the differences in the genetic composition of the subsets analyzed in Blocks 2 and 8 is the fact that different genotypes progress through their developmental stages at different rates. This could be another source of variation. MiRNAs have been found to have different functions depending on a plant's growth stage (Dong et al., 2022). So, while all the genotypes were harvested at a uniform timepoint, they were not all at a uniform growth stage. The growth stage at harvest was recorded and might need to be taken into account when considering the significance of genotypic effects on miRNA expression.

It is also important to consider that the number of samples run in the eGWAS was less than 100, meaning that the statistical power is relatively low. Reduced power means there is an increased likelihood of both false positives and false negatives (Ko et al., 2024).

Conclusion

This study represents a preliminary identification and analysis of miRNAs from a sorghum BAP. Here, we demonstrated the usefulness of employing two miRNA identification software; ShortStack and miRDeep-P2. We identified miRNAs that were differentially expressed in the same genotypes grown in Georgia compared to Arizona, indicating environmental effects on miRNA expression patterns. Additionally, significant genotypic effects were identified in some miRNAs, suggesting that different genotypes can vary in how they modulate miRNA expression. Lastly, we conducted a miRNA expression GWAS to identify miRNA eQTL. MiRNA eQTL provide a promising avenue to elucidating entire miRNA pathways, from potential environmental triggers, eQTL regulatory gene expression, miRNA synthesis, target gene regulation, and trait effect. As this study is part of a larger project that is generating other data, there will be future opportunities for cross-analyzing miRNA data with AMF abundances, AMF species compositions, mRNA expression, and mRNA eGWAS.

References

Adu-Gyamfi, B. A., & Good, C. (2022). Electric aviation: A review of concepts and enabling technologies. *Transportation Engineering*, *9*, 100134.

Axtell, M. J. (2013). ShortStack: comprehensive annotation and quantification of small RNA genes. *Rna*, *19*(6), 740-751.

Bo, X., & Wang, S. (2005). TargetFinder: a software for antisense oligonucleotide target site selection based on MAST and secondary structures of target mRNA. *Bioinformatics*, *21*(8), 1401-1402.

Brenton, Z. W., Cooper, E. A., Myers, M. T., Boyles, R. E., Shakoor, N., Zielinski, K. J., Rauh, B. L., Bridges, W. C., Morris, G. P., & Kresovich, S. (2016). A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. *Genetics*, *204*(1), 21-33.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC bioinformatics*, *10*, 1-9.

Chen, H., & Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC bioinformatics*, *12*, 1-7.

Chen, S.-Y., Su, M.-H., Kremling, K. A., Lepak, N. K., Romay, M. C., Sun, Q., Bradbury, P. J., Buckler, E. S., & Ku, H.-M. (2020). Identification of miRNA-eQTLs in maize mature leaf by GWAS. *BMC genomics*, *21*(1), 1-13.

Chen, S., Zhao, H., Zou, C., Li, Y., Chen, Y., Wang, Z., Jiang, Y., Liu, A., Zhao, P., & Wang, M. (2017). Combined inoculation with multiple arbuscular mycorrhizal

fungi improves growth, nutrient uptake and photosynthesis in cucumber seedlings. *Frontiers in Microbiology*, *8*, 2516.

Cobb, A. B., Wilson, G. W., Goad, C. L., Bean, S. R., Kaufman, R. C., Herald, T. J., & Wilson, J. D. (2016). The role of arbuscular mycorrhizal fungi in grain production and nutrition of sorghum genotypes: Enhancing sustainability through plant-microbial partnership. *Agriculture, Ecosystems & Environment*, *233*, 432-440.

Dai, X., Zhuang, Z., & Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic acids research*, *46*(W1), W49-W54.

Dong, Q., Hu, B., & Zhang, C. (2022). microRNAs and their roles in plant development. *Frontiers in Plant Science*, *13*, 824240.

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., Leyva, A., Weigel, D., García, J. A., & Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature genetics*, *39*(8), 1033-1037.

Friedländer, M. R., Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., & Rajewsky, N. (2008). Discovering microRNAs from deep sequencing data using miRDeep. *Nature biotechnology*, *26*(4), 407-415.

Gelaw, T. A., & Sanan-Mishra, N. (2021). Non-coding RNAs in response to drought stress. *International journal of molecular sciences*, *22*(22), 12519.

Glenn, T. C., Nilsen, R. A., Kieran, T. J., Sanders, J. G., Bayona-Vásquez, N. J., Finger, J. W., Pierson, T. W., Bentley, K. E., Hoffberg, S. L., & Louha, S. (2019). Adapterama I: universal stubs and primers for 384 unique dual-indexed or

147,456 combinatorially-indexed Illumina libraries (iTru & iNext). *PeerJ*, *7*, e7755.

Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., & Eddy, S. R. (2003). Rfam: an RNA family database. *Nucleic acids research*, *31*(1), 439-441.

Hammond, R. K., Gupta, P., Patel, P., & Meyers, B. C. (2023). miRador: a fast and precise tool for the prediction of plant miRNAs. *Plant Physiology*, *191*(2), 894-903.

Hannon, G. n. (2010). Fastx-toolkit. In.

Ho-Plágaro, T., Molinero-Rosales, N., Fariña Flores, D., Villena Díaz, M., & García-Garrido, J. M. (2019). Identification and expression analysis of GRAS transcription factor genes involved in the control of arbuscular mycorrhizal development in tomato. *Frontiers in Plant Science*, *10*, 268.

Hu, Z., Olatoye, M. O., Marla, S., & Morris, G. P. (2019). An integrated genotyping-by-sequencing polymorphism map for over 10,000 sorghum genotypes. *The Plant Genome*, *12*(1), 180044.

Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience*, *8*(2), giy154.

Johnson, A. C., Pendergast IV, T. H., Chaluvadi, S., Bennetzen, J. L., & Devos, K. M. (2022). Identification of microRNAs responsive to arbuscular mycorrhizal fungi in *Panicum virgatum* (switchgrass). *BMC genomics*, *23*(1), 688.

Khan, N., Sudhakar, K., & Mamat, R. (2021). Role of biofuels in energy transition, green economy and carbon neutrality. *Sustainability*, *13*(22), 12374.

Khanna, M., Chen, L., Basso, B., Cai, X., Field, J. L., Guan, K., Jiang, C., Lark, T. J., Richard, T. L., & Spawn-Lee, S. A. (2021). Redefining marginal land for bioenergy crop production. *Gcb Bioenergy*, *13*(10), 1590-1609.

Ko, B. S., Lee, S. B., & Kim, T.-K. (2024). A brief guide to analyzing expression quantitative trait loci. *Molecules and Cells*, *47*(11), 100139.

Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic acids research*, *47*(D1), D155-D162.

Kuang, Z., Wang, Y., Li, L., & Yang, X. (2019). miRDeep-P2: accurate and fast analysis of the microRNA transcriptome in plants. *Bioinformatics*, *35*(14), 2521-2522.

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*, *10*, 1-10.

Ledford, W. C., Silvestri, A., Fiorilli, V., Roth, R., Rubio-Somoza, I., & Lanfranco, L. (2024). A journey into the world of small RNAs in the arbuscular mycorrhizal symbiosis. *New Phytologist*, *242*(4), 1534-1544.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Subgroup, G. P. D. P. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079.

Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS genetics*, *12*(2), e1005767.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*, 1-21.

Maqbool, S. B., Devi, P., & Sticklen, M. B. (2001). Biotechnology: Genetic improvement of sorghum (*Sorghum bicolor (L.) Moench*). *In Vitro Cellular & Developmental Biology-Plant*, *37*, 504-515.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17*(1), 10-12.

Millar, A. A. (2020). The function of miRNAs in plants. In (Vol. 9, pp. 198): MDPI.

Mitra, D., Uniyal, N., Panneerselvam, P., Senapati, A., & Ganeshamurthy, A. N. (2020). Role of mycorrhiza and its associated bacteria on plant growth promotion and nutrient management in sustainable agriculture. *International Journal of Life Sciences and Applied Sciences*, *1*(1), 1-1.

Mullet, J., Morishige, D., McCormick, R., Truong, S., Hilley, J., McKinley, B., Anderson, R., Olson, S. N., & Rooney, W. (2014). Energy Sorghum—a genetic model for the design of C4 grass bioenergy crops. *Journal of experimental botany*, *65*(13), 3479-3489.

Murrell, P. (2005). *R graphics*. Chapman and Hall/CRC.

Neuwirth, E. (2014). RColorBrewer: colorbrewer palettes.

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in endocrinology*, *9*, 402.

Olson, S. N., Ritter, K., Rooney, W., Kemanian, A., McCarl, B. A., Zhang, Y., Hall, S., Packer, D., & Mullet, J. (2012). High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops. *Biofuels, Bioproducts and Biorefining*, *6*(6), 640-655.

Ortas, I. (2012). The effect of mycorrhizal fungal inoculation on plant yield, nutrient uptake and inoculation effectiveness under long-term field conditions. *Field crops research*, *125*, 35-48.

Pegler, J. L., Oultram, J. M., Grof, C. P., & Eamens, A. L. (2019). DRB1, DRB2 and DRB4 are required for appropriate regulation of the microRNA399/PHOSPHATE2 expression module in Arabidopsis thaliana. *Plants*, *8*(5), 124.

Persson, H., S∅ kilde, R., Pirona, A. C., & Rovira, C. (2017). Preparation of highly multiplexed small RNA sequencing libraries. *BioTechniques*, *63*(2), 57-64.

Pradhan, M., Baldwin, I. T., & Pandey, S. P. (2023a). Argonaute7 (AGO7) optimizes arbuscular mycorrhizal fungal associations and enhances competitive growth in Nicotiana attenuata. *New Phytologist*.

Pradhan, M., Baldwin, I. T., & Pandey, S. P. (2023b). Argonaute7 (AGO7) optimizes arbuscular mycorrhizal fungal associations and enhances competitive growth in *Nicotiana attenuata*. *New Phytologist*, *240*(1), 382-398.

Qu, G., Kruszka, K., Plewka, P., Yang, S.-Y., Chiou, T.-J., Jarmolowski, A., Szweykowska-Kulinska, Z., Echeverria, M., & Karlowski, W. M. (2015). Promoter-based identification of novel non-coding RNAs reveals the presence of dicistronic snoRNA-miRNA genes in *Arabidopsis thaliana*. *BMC genomics*, *16*, 1-15.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841-842.

Rogers, K., & Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. *The Plant Cell*, *25*(7), 2383-2399.

Rooney, W. L. (2014). Cellulosic energy cropping systems. In (pp. 110-123). John Wiley & Sons.

Rooney, W. L., Blumenthal, J., Bean, B., & Mullet, J. E. (2007). Designing sorghum as a dedicated bioenergy feedstock. *Biofuels, Bioproducts and Biorefining*, *1*(2), 147-157.

Sabia, E., Claps, S., Morone, G., Bruno, A., Sepe, L., & Aleandri, R. (2015). Field inoculation of arbuscular mycorrhiza on maize (*Zea mays L.*) under low inputs: preliminary study on quantitative and qualitative aspects. *Italian Journal of Agronomy*, *10*(1), 30-33.

SAS. (2023). *SAS® 9.4M8 Global Statements: Reference*. SAS Institute Inc.

Smith, S. E., & Read, D. J. (2010). *Mycorrhizal symbiosis*. Academic press.

Sun, G. (2012). MicroRNAs and their diverse functions in plants. *Plant molecular biology*, *80*, 17-36.

Villanueva, R. A. M., & Chen, Z. J. (2019). ggplot2: elegant graphics for data analysis. In: Taylor & Francis.

Wang, J., & Zhang, Z. (2021). GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genomics, proteomics & bioinformatics*, *19*(4), 629-640.

Watts-Williams, S. J., Emmett, B. D., Levesque-Tremblay, V., MacLean, A. M., Sun, X., Satterlee, J. W., Fei, Z., & Harrison, M. J. (2019). Diverse *Sorghum bicolor*

accessions show marked variation in growth and transcriptional responses to arbuscular mycorrhizal fungi. *Plant, Cell & Environment*, *42*(5), 1758-1774.

Wickham, H. (2015). dplyr: A grammar of data manipulation. *R package version 04.*, *3*, p156.

Wickham, H., Vaughan, D., & Girlich, M. (2023). Tidyr: Tidy messy data [R package tidyr version 1.3. 0]. *Comprehensive R Archive Network (CRAN)*.

Wortmann, C. S., Ferguson, R. B., Hergert, G. W., Shapiro, C. A., & Shaver, T. M. (2013). Nutrient management suggestions for grain sorghum. *Bulletin of Nebraska University*, 206-215.

Wu, G., Park, M. Y., Conway, S. R., Wang, J.-W., Weigel, D., & Poethig, R. S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *cell*, *138*(4), 750-759.

Yao, G., Staples, M. D., Malina, R., & Tyner, W. E. (2017). Stochastic techno-economic analysis of alcohol-to-jet fuel production. *Biotechnology for Biofuels*, *10*(1), 1-13.

Zhao, Y., Liu, J., Li, Y., Wen, H., Teotia, S., Zhang, X., Zhang, J., Sun, H., Cheng, Y., & Zhao, S. (2022). osa-miR528 promotes seedling growth by enhancing nitrate uptake under nitrogen deficiency in rice. *Environmental and Experimental Botany*, *202*, 105040.

CHAPTER 3

EVALUATION OF THE IMPACT OF ALTERNARIA ALTERNATA AND NEOPESTALOTIOPSIS SPP. ACROSS DIFFERENT SORGHUM GENOTYPES UNDER DROUGHT STRESS

Introduction

Endophytic fungi are a diverse group of microorganisms that grow inside many different plant species. These endophytes are often grouped into three major categories: mycorrhizal fungi, Clavicipitaceous (C-endophytes), and Non-Clavicipitaceous endophytes (NC-endophytes). Mycorrhizal fungi are a well-studied group of fungi that can benefit their host plants through nutrient absorption (Smith & Read, 2010). C-endophytes represent a smaller group of fungi that interact with grasses and can benefit their hosts by conferring stress tolerance and discouraging herbivory through toxin production (Rodriguez et al., 2009). And lastly, NC-endophytes represent a diverse group of endophytes that can interact with a large range of hosts and have been isolated from many ecosystems (Arnold & Lutzoni, 2007).

There has been recent interest in utilizing beneficial endophytes as biofertilizers to help increase plant performance while requiring fewer inputs. Mycorrhizal fungi have been of particular interest in this context due to their assistance with nutrient absorption (Ortas, 2012; Sabia et al., 2015). However, the non-mycorrhizal fungi are of growing

interest for their potential in providing a fitness advantage for their host (Ferreira et al., 2023; Jacob et al., 2023; Munir et al., 2022; Zhou et al., 2018). Of the NC-endophytes, two will be further explored here for their potential advantages: *Alternaria alternata* and *Neopestalotiopsis.*

A. alternata is septate fungi that's widely distributed across many different environments and can interact with many diverse hosts. However, *A. alternata* is not strictly a beneficial endophyte and has also been isolated as a saprotroph and a pathogen (DeMers, 2022). As a phytopathogen, *A. alternata* can infect many different plants and has been linked to pre- and post-harvest diseases (Tralamazza et al., 2018). But despite its negative effects on its host as a pathogen, *A. alternata* isolates can also exist as useful endophytes. For example, Ferreira et al. (2023) determined that an endophytic *A. alternata* strain was able to solubilize phosphate, which could be useful in mobilizing nutrients for the host plant. Another potential benefit was determined by Luo et al. (2025) which identified an *A. alternata* strain that was able to produce metabolites that reduced the presence of harmful microorganisms in the rhizosphere for its tomato host. These advantages of *A. alternata* have increased interest in potentially utilizing them as bio-fertilizers. However, more work needs to be done to determine the efficacy of inoculating with *A. alternata*.

Neopestalotiopsis represents a diverse genus of septate fungi that has species with variable lifestyles ranging from free-living saptrotrophs, to endophytes, to pathogens (Maharachchikumbura et al., 2011; Maharachchikumbura et al., 2014). *Neopestalotiopsis spp.* generally exhibit low host specificity and have also been noted colonizing many different tissue types (Darapanit et al., 2021; Hu et al., 2007; Shi et al., 2024; Wei et al.,

2007). As pathogens, *Neopestalotiopsis spp*. are often characterized as weak, opportunistic pathogens that predominantly invade the host through natural openings or wounds (Baggio et al., 2021; Darapanit et al., 2021; Lee et al., 1995; Maharachchikumbura et al., 2011; Shi et al., 2024). Additionally, the boundary between endophyte and pathogen is not always clear-cut, with some species being observed to exhibit both endophytic and pathogenic stages within their life cycle (Lee et al., 1995)(Lee et al., 1995). In terms of the benefits endophytic *Neopestalotiopsis* can provide their host plant, they have been observed aiding in nutrient mobilization of phosphate and zinc (Jacob et al., 2023; Munir et al., 2022). As a result, similarly to *A. alternata*, there has also been interest in utilizing *Neopestalotiopsis spp*. as biofertilizers.

Sorghum has been proposed as an ideal candidate for cellulosic biofuel production on marginal lands (Rooney, 2014; Rooney et al., 2007). However, while biomass sorghum varieties suited for biomass production are generally well adapted to the often hot and dry conditions characteristic of many marginal lands (Maw et al., 2017; Stamenković et al., 2020), optimizing production while minimizing water and fertilizer inputs will be invaluable for making biofuel sorghum production economically profitable on marginal lands. As many marginal lands also have low fertility soils (Khanna et al., 2021), identifying endophytes able to confer sorghum fitness advantages, especially through better access to usable nutrients, could be useful in determining an optimal system for sorghum biofuel production on marginal lands.

Endophytic *A. alternata* has been identified from the roots of grain sorghum (Janet, 1983). However, the impact on sorghum health and biomass accumulation as a result of inoculation with *A. alternata* is not well characterized. *Neopestalotiopsis* has not

been noted colonizing sorghum root tissues before. That means its ability to colonize sorghum roots and confer either fitness advantages or disadvantages has not been explored.

Here, *A. alternata* and *Neopestalotiopsis spp.* were inoculated into soil with sorghum seeds to encourage endophytic interactions. Additionally, the effects of endophyte inoculation was measured across different sorghum genotypes. As different genotypes have garnered variable benefits from mycorrhizal fungi (Cobb et al., 2016), *A. alternata* and *Neopestalotiopsis* may also exhibit different interaction patterns depending on the sorghum genotype. To also assess if the endophytes could be useful in mediating stress, drought was imposed on the sorghum plants. Plant heights, stalk diameters, and aboveground biomass measurements were taken to gauge plant performance under different endophyte/drought treatments.

Materials and Methods

Sorghum Genotype Selection:

Five sorghum genotypes were selected for initial testing in Greenhouse Experiment 1. These genotypes were selected from the larger Bioenergy Association Panel (BAP) used in the sorghum field experiments described in Chapter 2. To achieve the greatest diversity among the five genotypes, they were each selected from different sorghum subpopulations represented in the BAP as defined in Brenton et al. (2016). Additional filtering was done on potential genotypes to only include ones that had seeds in stock and had good germination rates (at least 70% germination) based on small-scale

germination tests. All genotypes selected had been grown in both Georgia (2022) and

Table 3.1. Genotype Information for Greenhouse Experiments 1 and 2

Genotype[1]	Photoperiod Sensitivity	Type	Origin	Race	Germination Rate (%)
BIP068_PI 196598	**Photoperiod Insensitive**	**Grain**	**NA**	**NA**	**99**
BIP121_PI 329541	Photoperiod Sensitive	Cellulosic	Ethiopia	NA	77
BIP223_PI 562985	**Photoperiod Sensitive**	**Cellulosic**	**Nigeria**	**Guinea**	**97**
BIP289_PI 570090	**Photoperiod Sensitive**	**Cellulosic**	**NA**	**Guinea-caudatum**	**85**
BIP359_PI 152971	Photoperiod Insensitive	Sweet	Sudan	NA	98

[1] Genotypes selected for Greenhouse Experiments 1 (all genotypes) and 2 (red colored genotypes). Photoperiod sensitivity, type, origin, and race information are from Brenton et al. 2016. Germination rates determined by Trudi Thomas.

Arizona (2023) field experiments. Based on the results from Greenhouse Experiment 1, three sorghum genotypes out of the five were selected for further testing in Greenhouse Experiment 2 (see Table 3.1).

Alternaria alternata Strain:

The *Alternaria alternata* strain used as inoculum in this experiment was provided by Beatrice Bock (Johnson Lab at Northern Arizona University). The *A. alternata* strain was initially isolated by Dr. Ron Deckert from sorghum seeds, and was identified through sequencing of the ITS1 region by Beatrice Bock. The fungus was cultured on Potato Dextrose Agar (PDA) plates, and each replating is recorded as a subsequent generation.

The generation sent by Beatrice Bock was Generation IV. The fungus was replated multiple times in order to increase stock for using as inoculum and for archiving. The generation used for inoculation in Greenhouse Experiment 1 was VI. The generation used for inoculation in Greenhouse Experiment 2 was Generation VII.

Neopestalotiopsis sp. Strain:

The *Neopestalotiopsis sp.* strain used in Greenhouse Experiment 2 was isolated from roots from an AP13 switchgrass plant grown as part of a GWAS panel at the Iron Horse Farm in Watkinsville, GA (C23_1 G20). Roots were surface sterilized (see *Surface Sterilization* section below) and cut into sections. Cut roots were placed on PDA plates, allowing fungal endophytes to grow out from the roots for about a week. Fungal growth was then sampled and replated (see *Replating Fungi* section below). As the endophytes emerging from the roots were not all uniform, samples were taken from multiple unique fungal growths. Plates were then further sampled in subsequent generations in an effort to isolate a morphologically homogeneous strain. To identify strain purity and identity, DNA was extracted and the ITS region was sequenced (See *DNA Extraction* and *ITS region Sequencing* sections). Using the ITS sequences as queries in megablast searches against the NCBI's Core Nucleotide Database (core_nt), the strains were determined to fall under the *Neopestalotiopsis* genus. However, none of the strains mapped uniformly to one species, so no species designation was assigned to the strain. The *Neopestalotiopsis sp.* strain with the fastest growth was selected to be used as inoculum in Greenhouse Experiment 2 (Generation VI) (Supplemental S3.3).

Fungal DNA Extraction

In order to identify the identity of the cultured AP13 Switchgrass endophytes (mentioned in *Neopetalotiosis Strain* section above), DNA was extracted for ITS amplification and sequencing. The Promega Wizard DNA Extraction kit was used for extracting DNA (Promega). The protocol from the kit was used with modifications made by the Chung Lab (UGA) to adapt it for fungal DNA extraction. From a fungal plate, a ~1 cm x 2 cm rectangle of growth from the growing edge of the fungi was carefully scraped off the surface of the agar with forceps, trying to take up hyphae from the agar without taking too much agar with it. The fungal sample was placed in a clean 1.5 mL tube, and 100 µL of nuclei lysis solution was added. The sample was then ground up by twisting a clean plastic pestle (sterilized with 70% ethanol) against the inner surface of the 1.5 mL tube for 1 minute. An additional 500 µL of nuclei lysis solution was added, and the sample was incubated in a water bath at 65° C for 15 minutes. 3 µL of RNAse solution was added, and the sample was mixed by inverting several times. The sample was then incubated in a 37° C water bath for 15 minutes, and then allowed to stand at room temperature for 5 minutes. Next, 200 µL of protein precipitation solution was added, and the tube was vortexed for 5 seconds. The tube was then incubated on ice for 5 minutes followed by centrifugation at 10,000 rpm for 4 minutes. 600 µL of the supernatant was transferred into a clean 1.5 mL tube, taking care not to disturb the pelleted debris at the bottom of the tube. To the new tube with 600 µL of sample, 600 µL of isopropanol was added and mixed by inverting the tube. The sample was then centrifuged at 10,000 rpm for 1 minute to pellet the DNA. After centrifugation, the supernatant was discarded and 600 µL of 70% ethanol was added. The tube was closed and flicked several times to wash

the pellet, then centrifuged again at 10,000 rpm for 1 minute. The supernatant was again discarded, using a pipette to remove as much ethanol as possible without disturbing the pellet. The sample was then air dried for ~30 minutes to evaporate out the remaining ethanol. 10 µL of the DNA rehydration solution was added, and the DNA pellet was dissolved by pipetting up and down several times. Lastly, the sample was incubated in a 65° C water bath for ~1 hour, and then stored at -20° C until further use.

ITS Sequencing

To determine fungal identity, the ITS (Internal Transcribed Spacer) region of fungal DNA was sequenced. The ITS1F and ITS2 primers were used to amplify the ITS1 region (Kumar & Shukla, 2005; Op De Beeck et al., 2014). The PCR conditions used were the same as in Kumar and Shukla (2005). Amplicons were verified on a 1% TAE agarose gel, then cleaned using Qiagen's PCR clean and concentrator kit (Qiagen). Cleaned PCR products were sent for Sanger sequencing (with ITS1F and ITS2 primers) through Genewiz (Azenta Life Sciences). Resulting sequences were used as queries in megablast searches against NCBI's Core Nucleotide Database (core_nt) to determine fungal identities.

PDA Plates

Fungi for this experiment were all cultured on Potato Dextrose Agar (PDA) plates. Plates were made per the manufacturer's instructions by mixing 39 g of PDA powder (MilliporeSigma) with 1000 mL of deionized water and then autoclaving at 121°C for 1 hour. After allowing it to cool slightly after autoclaving, the medium was

poured into plates. 25 mL of medium was measured into each plate to ensure equal nutrients were present across all plates. After allowing the plates to cool, they were sealed with parafilm and stored at room temperature for short term storage or 4°C for longer term storage.

Replating Fungi

The fungi used in this experiment (*Alternaria alternata* and *Neopetalotiopsis*) generally needed to be replated every 1-2 weeks to maintain the cultures. Plates were usually kept at room temperature for growth but, if needed, growth could be slowed by placing in cold room (4°C). To replate fungi, a cork borer was used to punch a hole (~1 cm x 1 cm) in the growing edge of the fungus. The plug could then be transferred to a new plate using forceps. Both the cork borer and forceps were sterilized in between each fungal sample by cleaning with 70% ethanol and heat sterilizing with a bead sterilizer. Replating was always conducted in a laminar flow hood to ensure sterile conditions.

Archiving Fungi

In order to preserve fungal strains or generations, fungi could be archived for future use. To archive a sample, a plug was taken from the growing edge of the fungus (similarly to the *Replating Fungi* section) and submerged in a 2 mL tube filled with 1500 µL of sterile deionized water. Archived fungi can then be stored at room temperature. To reactivate fungi, simply remove the plug from the water and place on a new PDA plate. Archiving fungi allows for preservation of specific fungal generations which can be

useful when trying to repeat experiments with the same fungi while avoiding generational changes that may arise from keeping the fungi in culture for months.

Seed/Root Surface Sterilization

Prior to planting, seeds were surface sterilized to minimize contamination with microorganisms other than the intended inoculum. To this end, the seeds were soaked in 70% ethanol for 3 minutes, then in 2.5% bleach for 5 minutes, followed by three rinses in sterile DI water (surface sterilization method from Xiomy Pinchi-Davila—Chung Lab, UGA). This sterilization method was also used for surface sterilizing roots prior to plating them for extracting root endophytes.

Soil Preparation

Prior to inoculating sorghum seeds with fungi, the soil mixture had to be prepared. To assess the potential benefits that root endophytes (like *Alternaria alternata* and *Neopetalotiopsis)* may provide to their host plant, a low-nutrient soil mix consisting of 50% field soil from the Wellbrook Farm in Watkinsville, GA and 50% sand was used. Both the soil and sand were autoclaved for 1-2 hours at 121°C (until autoclave tape buried in the soil/sand changed color). The sand and soil were mixed using a soil mixer. Samples from the starting soil/sand mixture were sent to the University of Georgia Agricultural & Environmental Services Laboratories for soil nutrient testing. To ensure the pots (9-inch diameter) were sterile, they were soaked in 5% bleach for 10 minutes and then rinsed in reverse osmosis (RO) water prior to filling with soil to ~4/5ths of the way to the top. Thin mesh sheets, briefly soaked in 5% bleach and rinsed in RO, were placed

in the bottom of the pots to prevent the soil/sand mixture from falling out the drainage holes in the bottom of the pots.

Fungal Inoculum Preparation

Fungal inoculum was prepared the day prior to planting. Two plugs were taken from the growing edge of a fungal culture with a cork borer (~1 cm x 1 cm size plug), one from the very edge of the growth and one slightly further in to maximize the amount of inoculum that could be harvested from a single plate (see Figure 3.1). Plugs were submerged in 900 µL sterile DI water in 2 mL tubes, and ground up using plastic pestles

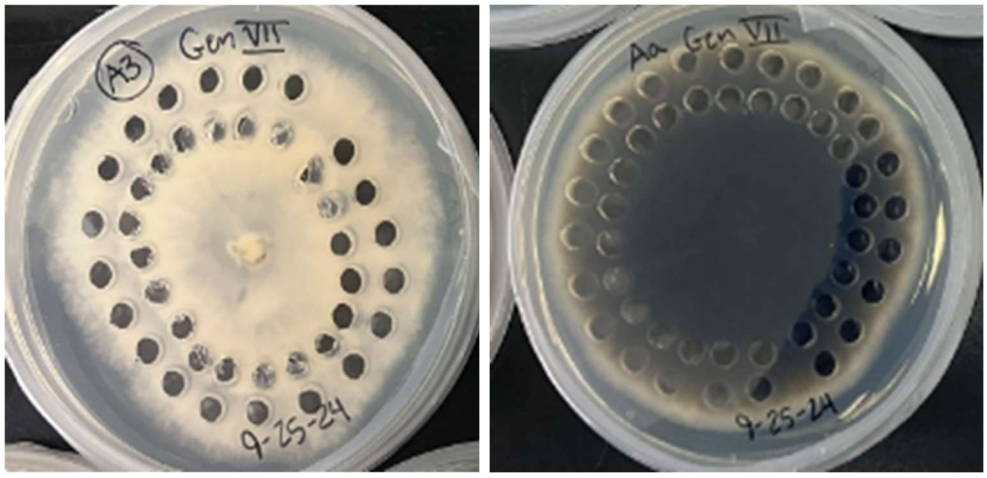

Figure 3.1. Images of fungal plates used for generating inoculum in Greenhouse Experiment 2. The left image shows the Neopestalotiopsis sp. fungi and the right image shows Alternaria alternata. Two plugs were used for generating each inoculum, with one plug taken from the exterior edge of the fungus and one plug taken slightly more interior. Images taken on 10-2-24.

that had been sterilized with 70% ethanol. Because agar pieces were included in the fungal plugs, blank inoculum consisting of ground-up plugs from sterile PDA plates was generated to inoculate control plants to keep conditions across treatments consistent. Tubes with fungal inoculum were left at room temperature overnight until planting and inoculation the next day.

Planting and Inoculation

For each pot (prepared as described in the *Soil Preparation* section), the soil/sand mixture was saturated with Reverse Osmosis (RO) water. For Greenhouse Experiment 1, the seeds were planted prior to saturating the soil/sand, and pots were watered until water started draining out the bottom of the pot. As watering post-planting in Greenhouse Experiment 1 resulted in shifting the seeds around in the pot, watering was done prior to planting seeds for Greenhouse Experiment 2. Additionally, in Greenhouse Experiment 2, we watered all the pots with the same volume (600 mL) of RO water. Two sterile sorghum seeds (prepared as described in the *Seed/Root Surface Sterilization* section) were planted a couple of inches apart from each other in the pots. Seeds were buried ~½ inch into the soil. Planting two seeds increased the likelihood of at least one seed successfully germinating. If both seeds germinated, the extra plant was removed three weeks after planting for Greenhouse Experiment 1 and 1 ½ weeks after planting for Greenhouse Experiment 2. The roots of the plants removed from Greenhouse Experiment 1 were stored at -80 ^{0}C for later DNA extraction. Once seeds were planted, the tubes containing the water and ground fungal mixture (prepared according to the *Fungal Inoculation Preparation* section) were dumped into the center of the pots. Additional sterile sand/soil mixture was then sprinkled in a thin layer across the top of the pots to cover the fungi.

Watering Schedule

Plants were watered with 400 mL of reverse osmosis (RO) water every couple of days until all plants had germinated. Then, plants were watered less frequently. For

Greenhouse Experiment 1, which was conducted during the summer (starting in June), plants were given 400 mL of water every 3-4 days. For Greenhouse Experiment 2, which was conducted in the fall (starting in October), plants were given 400 mL every 6-7 days.

Greenhouse Experiment 1: Design

Greenhouse Experiment 1 consisted of five sorghum genotypes, BIP068, BIP121, BIP223, BIP289, BIP359 (Table 3.1) and two inoculation treatments, *Alternaria alternata* and control. Each inoculation treatment was replicated six times. Plants were organized on a bench in the greenhouse in five columns (five genotypes) of 12 plants. The first three plants (rows) were used as control (non-inoculated, non-drought) plants, the next three plants as non-inoculated drought plants, the next three represented the *A. alternata* inoculated non-drought plants, and the last three the *A. alternata* inoculated drought plants. About four weeks into the experiment, the 'drought plants' were subjected to a drought treatment (Figure 3.2). Water was withheld completely from the replicates under

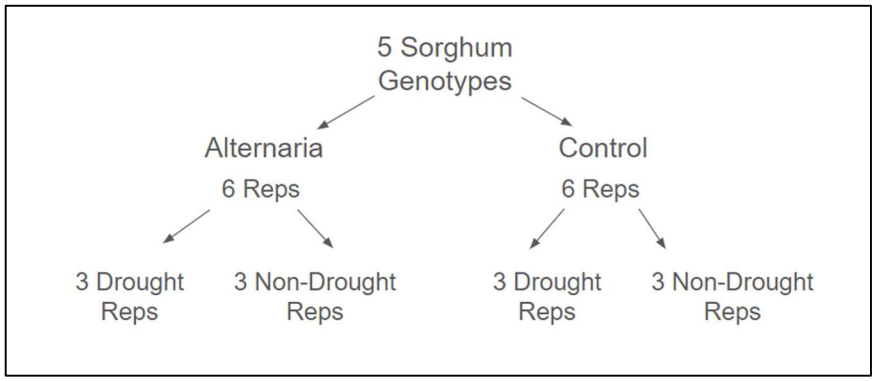

Figure 3.2. Experimental design for Greenhouse Experiment 1.

drought stress, and the level of wilting was scored on a scale of 1-5 with 1 representing no wilting and 5 being very wilted but likely to recover upon rewatering. Once plants reached a score of 5, we resumed watering. Scores continued to be taken after watering

resumed to assess speed of recovery. This design allowed us to determine how long the plants could survive without water as well as how many days it took the plants to recover after the drought. After the initial severe drought cycle, plants were watered normally for ~1.5 weeks and then a moderate drought treatment was imposed, with the drought plants being watered once a week compared to the control plants twice a week. This continued for 3 weeks. A total of three moderate drought cycles were imposed.

Plant heights and stalk diameters were measured once a week throughout the experiment (measurements taken by Mary Beth Lowe). Plant heights were measured from the soil level to the top of the sorghum stalk (not including leaves). As no flowering was observed, this approach remained consistent throughout the experiment. To account for oblong stalks, stalk diameter was determined by measuring the stalk diameter from two angles at the base of the stalk with calipers and then averaging the measurements. Due to signs of nutrients deficits impacting plant health, plants were fertilized with 30 mL of 1x Hoagland solution at 8 weeks to prevent plant mortality.

Greenhouse Experiment 1 continued for ~11 weeks prior to harvest (1 severe drought cycle, and 3 moderate drought cycles). Aboveground biomass was harvested and dried, then weighed to determine biomass yields. Roots were collected, washed, dried and put on ice until the end of the harvest day when samples were placed at -80 °C until further analysis.

Greenhouse Experiment 2: Design

Greenhouse Experiment 2 consisted of three genotypes, BIP068, BIP223, and BIP289 (Table 3.1), and three inoculation treatments, *Alternaria alternata, Neopestalotiopsis sp.*, and control. Each inoculation treatment was replicated 14 times (Figure 3.3). About five weeks into the experiment, drought was imposed on half the plants, subdividing the 14 reps per inoculation treatment into subgroups of seven drought reps and seven well-watered reps (Figure 3.3). Plants were randomized across the greenhouse bench by assigning each sample a number and then using a random number generator (https://www.random.org/sequences/) to assign locations to each number. The randomized layout is shown in Supplemental S3.2. The drought treatment for Greenhouse Experiment 2 differed from Experiment 1 in that all drought-treated plants were watered

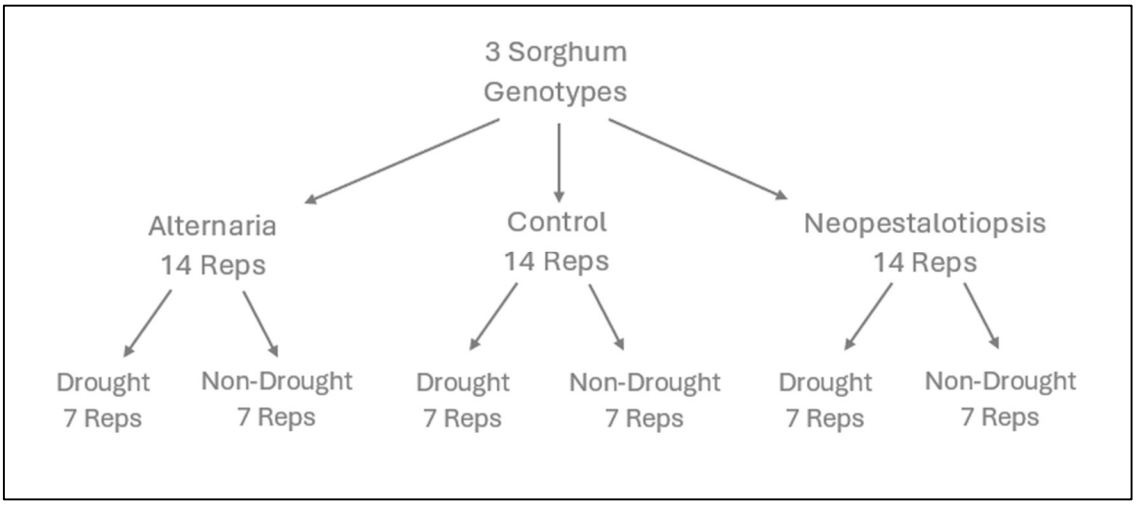

Figure 3.3. Experimental design for Greenhouse Experiment 2.

at a set interval. The drought cycle length was determined by wilting progress. Once ~3-5 plants (across all genotypes) showed distinct signs of wilting (scores of 3-4 based on the 1-5 scale described previously), all droughted plants were watered. The first drought

cycle was 13 days but was shortened to 10 days for the next 2 drought cycle intervals (total of three drought cycles). Plant heights and stalk diameter measurements were taken once a week throughout the experiment similarly to Greenhouse Experiment 1. Greenhouse Experiment 2 continued for ~10 weeks, after which aboveground biomass was harvested, dried and weighed to determine biomass yields. Similar to Greenhouse Experiment 1, roots were collected, washed, dried and stored at -80 °C until further analysis. Additionally, subsamples of the sorghum roots were placed in 5 mL tubes filled with 50% ethanol, and stored in the cold room at 4°C for later microscopic analysis.

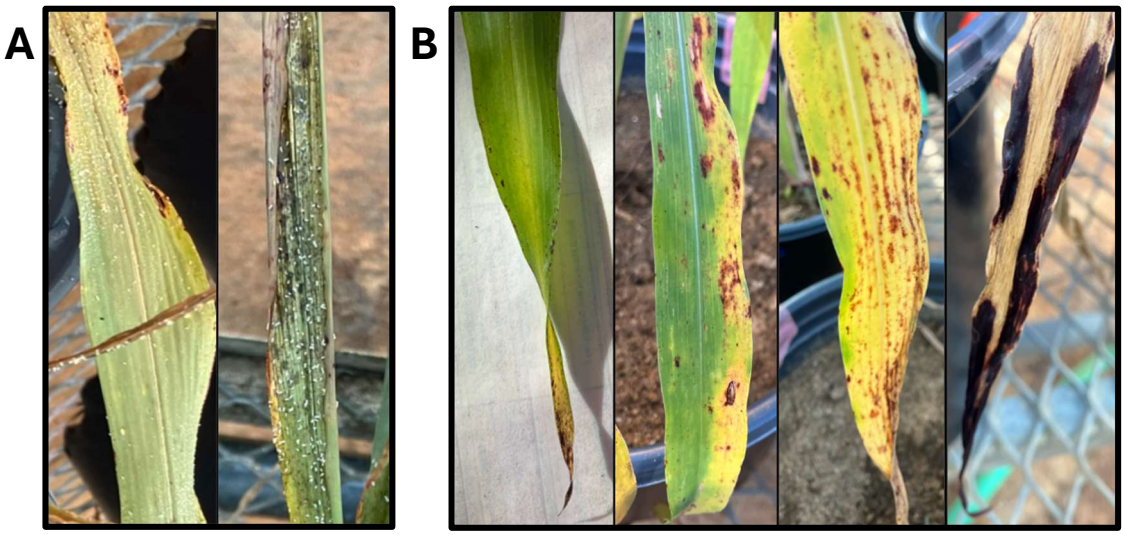

Figure 3.4. Leaf damage as a result of pests and disease. (A) Leaves after aphid infestation in Greenhouse Experiment 1. (B) Leaves infected with anthracnose in Greenhouse Experiment 2. Images are organized in increasing disease severity. The leftmost image represents a lowly infected leaf. The rightmost image shows severe infection.

Disease Scoring

Unfortunately, both Greenhouse Experiments 1 and 2 had issues with pests and disease. Greenhouse Experiment 1 was impacted significantly by aphids. Plants were sprayed to control aphids, but plants suffered damage prior to spraying (see Figure 3.4A).

Greenhouse Experiment 2 also had to be sprayed for aphids, but, fortunately, only five plants were mildly affected by the aphids and completely recovered after pesticide treatment. More notably, Greenhouse Experiment 2 suffered from a fungal disease, likely anthracnose based on visual disease symptoms (Figure 3.4B). Fungicide application has historically been effective against anthracnose. However, because this experiment is a fungal inoculation experiment, no fungicides could be applied to control disease as it could negatively impact the *A. alternata* and *Neopestalotiopsis sp.* endophytes. As a result, the experiment was ended earlier than originally planned (only 3 drought cycles were completed, instead of 5+ as originally planned). Mild anthracnose symptoms were also noted at the very end of Greenhouse Experiment 1, but the disease didn't significantly affect the plants before the end of the experiment (allowing for the completion of 1 severe drought cycle, and 3 moderate drought cycles). To assess the impact of pests and diseases on plant growth, symptoms were scored on a scale of 1-5, with 1 representing no disease and 5 representing high disease levels (see Figure 3.4B). Scores were taken for aphids in Greenhouse Experiment 1 prior to spraying pesticides. Scores were taken for anthracnose in Greenhouse Experiment 2 one week prior to harvest as well as the day before harvest.

Measuring Photosynthetic Gas Exchange

Photosynthetic gas exchange across different treatment groups in Greenhouse Experiment 2 was measured using the LI-COR Portable Photosynthesis System (LI-COR Biosciences, 2017). Initial data were collected on two reps per treatment group at the end of drought cycle 1 when plants were showing signs of drought stress (plants were ~7

weeks old). At the end of drought cycle 3, LI-COR measurements were taken on all samples (seven reps per treatment group). Because the large number of samples, measurements were conducted over the course of two days (the two days prior to harvest). CO_2 assimilation (A score) measurements were taken in the center of the second fully developed leaf between 10 am and 4 pm.

PCR Analysis

As mentioned in the *Planting and Inoculation* section, extra plants that germinated in Greenhouse Experiment 1 were removed three weeks after planting. To test if *A. alternata* was present in the plant roots, DNA was extracted from a subset of samples. This subset consisted of one plant from each genotype for the control group and one plant for each genotype in the *A. alternata* inoculated group. The DNA extraction protocol used was adapted from (Xin & Chen, 2012) by Philip Brailey-Crane for use in the larger DOE AMF-Sorghum Project. Further modifications were made by Ashton Brinkley and Ndenum Shitta to optimize the protocol for sorghum roots. The modified protocol differed from the original in that phenol:chloroform:isoamyl alcohol (25:24:1) was used rather than chloroform:isoamyl alcohol (24:1), three phenol:chloroform washes were performed instead of one, 100% ethanol was used for overnight precipitation at -20 °C rather than CTAB-based precipitation, centrifugation at 3,000 g for 20 minutes instead of 15 for DNA pelleting, DNA pellet was washed with 70% ethanol as a opposed to CTAB wash buffer, and DNA was resuspended with Tris-HCL instead of TE buffer. Additionally, a magnetic bead cleanup with Sera-Mag SpeedBeads (Cytiva) was used rather than the MagAttract kit (Qiagen).

PCR was conducted on the DNA samples (~100 ng of DNA) using primers ITS1F and ITS2 using the same conditions referenced in the *ITS Sequencing* section. Additionally, as a reference to gauge endophyte levels and to ensure the DNA extracted was of sufficient quality, a sorghum primer set was run as a control. The sorghum primer set was for *EIF4α*, a housekeeping gene commonly used as a reference (Sudhakar Reddy et al., 2016).

DNA from the sorghum roots harvested at the end of Greenhouse Experiment 1 were also tested for fungal presence using the ITS1F and ITS2 primers as well as the sorghum *EIF4α* primer. Amplicons were sent for Sanger Sequencing through Genewiz (Azenta Life Sciences). For Greenhouse Experiment 2, DNA was also extracted from the sorghum roots and ITS primers were tested on a subset of samples to confirm fungal presence, however, PCR products were not sent for sequencing. In addition to the ITS primers, *A. alternata* (Konstantinova et al., 2002) and *Neopestalotiopsis* (Rebello et al., 2023) primers were tested.

Root Microscopy

Root clearing, staining and visual scoring of endophyte colonization levels in the sorghum roots was conducted by Tom H. Pendergast (with assistance from Mary Lowe and Sofia Fachisthers). The clearing/staining protocol was optimized for sorghum roots by Tom H. Pendergast based on the original protocol from Koske and Gemma (1989). The roots stored in 50% ethanol were rinsed with water, and placed into biopsy cassettes (each sample was put into its own cassette) in acidic glycerol (500 mL glycerol, 450 mL H_2O, 50 mL 1% HCl) for at least 24 hours. The roots were then placed for 15 minutes in

10% KOH, heated to 90°C on a hot plate, followed by five rinses with water. If roots were not translucent, they were soaked in 5% bleach for about 5 minutes or until translucent. Next, roots were soaked in 1-2% HCl for at least 1 hour followed by soaking in an acidic glycerol/trypan blue solution (0.5 g trypan blue crystals into 1000 mL acidic glycerol ~ 0.05%) for 7 minutes at 90°C on a hot plate. To destain and store, the roots were placed in an acidic glycerol solution (without trypan blue). If the roots appeared too dark, they were rinsed before storing in glycerol. For microscopy, roots were removed from cassettes and placed on slides. Clear nail polish was spread over the roots and a cover slip was placed on top. The roots were scored and imaged using a light microscope with a digital camera at 10x (or 20x) objective and 10x optical zoom. Scoring was done using the gridline intersection method (McGonigle et al., 1990). The number of roots associated with the sample and endophyte features such as hyphae, microsclerotia, and vesicles were recorded.

Statistical Analysis

ANOVA was used to statistically compare variation between treatment groups for a given trait. Anova results were generated in SAS (Statistical Analysis System) (SAS, 2023) using scripts generated by Tom H. Pendergast. For traits measured before drought treatment, a two-way ANOVA was run with sorghum genotype and endophyte treatment, as well as the genotype by endophyte interaction as fixed factors. For traits measured after drought treatment, a three-way ANOVA was run where genotype, endophyte treatment, drought treatment, as well as all the two-way and three-way interactions, were

treated as fixed factors. After ANOVA, Tukey's post-hoc test was used to adjust p-values to control for false positives.

Additionally, correlation analysis between different traits (plant height, stalk diameter, biomass, disease, etc) was conducted in R version 4.4.1 (R Core Team, 2024) using the GGally package (Schloerke et al., 2021). For each pairwise correlation between traits, Pearson's correlation coefficient (r) and coefficients of determination (R^2) were calculated to gauge the strength of the relationship (r) and the proportion of variance explained (R^2). Scatterplots with fitted linear regression lines were also generated for each trait pair to visualize relationships. Statistical significance thresholds for Pearson's correlation coefficient were defined as p-values < 0.05 (*), 0.01 (**), and 0.001 (***).

For graphical outputs (generated in Excel), the standard error of the mean (SEM) was used to calculate the error bars. SEM is calculated by dividing the standard deviation by the square root of the sample size. This is ideal for the current experiment, as some of the genotype/endophyte/drought treatment combinations had plants die prior to the end of the experiment (with zero mortality, there are 7 reps per treatment combination). Since SEM accounts for sample size, it gives a more accurate comparison between treatment groups even when they vary in sample number.

Results

Starting Soil Nutrients

Nutrient levels of the starting (autoclaved) sand/field soil mixture were assessed for both Greenhouse Experiment 1 and 2 (see Table 3.2).

Table 3.2. Initial Soil Nutrients

		Mehlich 1 mg/kg (ppm)						%
Sample	**pH** [2]	**Ca**	**K**	**Mg**	**Mn**	**P**	**Zn**	**N**
Greenhouse Experiment 1	7.28	544	26.5	59.2	43.6	4.58	1.47	0.07
Greenhouse Experiment 2	6.92	482	34.7	57.2	53.9	9.76	1.51	0.08

Greenhouse Experiment 1

Before Drought

Irrespective of *A. alternata* inoculation, the plant height prior to drought treatment (Table 3.3A) was not significantly different between the five genotypes ($P = 0.1179$). However, there was a general trend of *A. alternata* inoculated plants having larger heights (Figure 3.5). And while the impact of *A. alternata* inoculation across all the genotypes was not statistically significant ($P = 0.6266$), the interaction between *A. alternata* inoculation and genotype was significant ($P < 0.0001$). This suggests that while some genotypes don't benefit from *A. alternata*, others likely do. The stalk diameter measurements from before the drought treatment (Table 3.3B) had no significant variation between genotypes ($P = 0.156$). There was a trend of *A. alternata* inoculated plants having thicker stalks overall ($P = 0.0127$). Additionally, the interaction between *A. alternata* and genotype was

significant (P = 0.033), suggesting the impact of *A. alternata* inoculation varied by genotype

Table 3.3. Greenhouse Experiment 1 P-values

Source	P-Value
A) Height Before Drought	
Genotype	0.1179
A. alternata	0.6266
Genotype**A. alternata*	<.0001
B) Stalk Diameter Before Drought	
Genotype	0.156
A. alternata	0.0127
Genotype**A. alternata*	0.033
C) Height After Drought	
Genotype	0.1355
A. alternata	0.111
Drought	<.0001
Genotype**A. alternata*	<.0001
Genotype*Drought	0.055
*A. alternata**Drought	0.1505
Genotype**A. alternata**Drought	0.2951
D) Stalk Diameter after Drought	
Genotype	0.4137
A. alternata	0.0057
Drought	<.0001
Genotype**A. alternata*	0.0313
Genotype*Drought	0.5732
*A. alternata**Drought	0.0026
Genotype**A. alternata**Drought	0.1102
E) Wilting Progression	
Genotype	0.0001
A. alternata	0.0047

Genotype*_A. alternata_	0.3049
F) Biomass	
Genotype	0.0011
A. alternata	0.014
Drought	<.0001
Genotype*_A. alternata_	0.002
Genotype*Drought	0.0174
_A. alternata_*Drought	0.8701
Genotype*_A. alternata_*Drought	0.1725
G) Aphid Infestation	
Genotype	0.0002
A. alternata	<.0001
Drought	<.0001
Genotype*_A. alternata_	0.0062
Genotype*Drought	0.1092
_A. alternata_*Drought	<.0001
Genotype*_A. alternata_*Drought	0.358

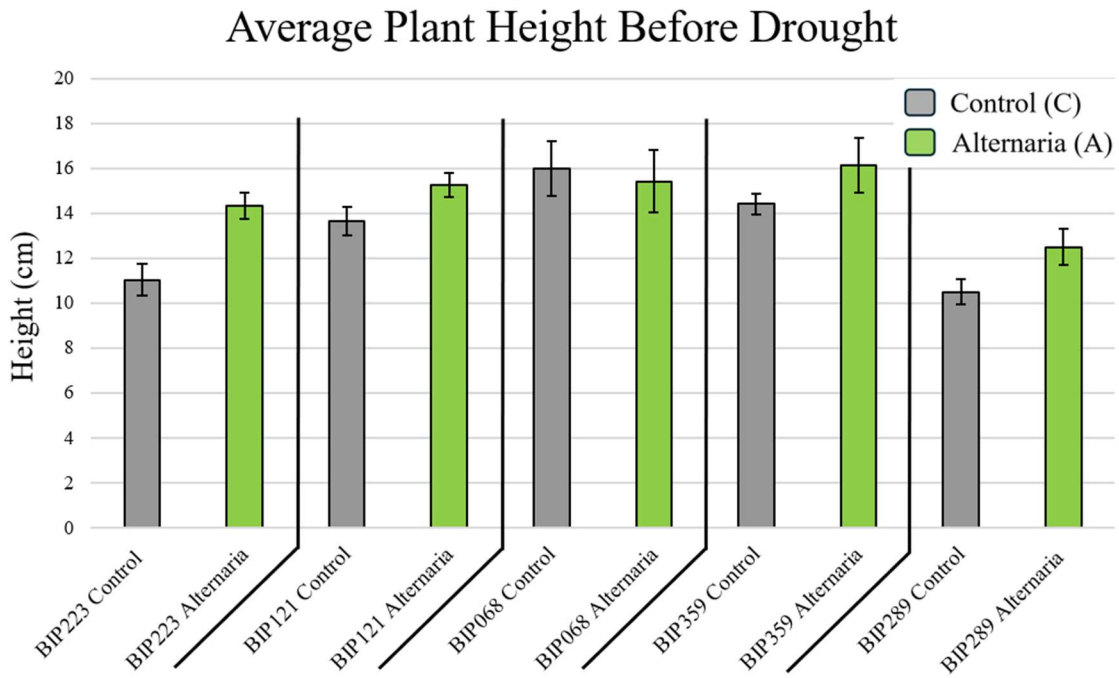

Figure 3.6. Average plant height for each genotype before drought treatment. Each bar represents the average of all 6 replicates within a given treatment/genotype combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean.

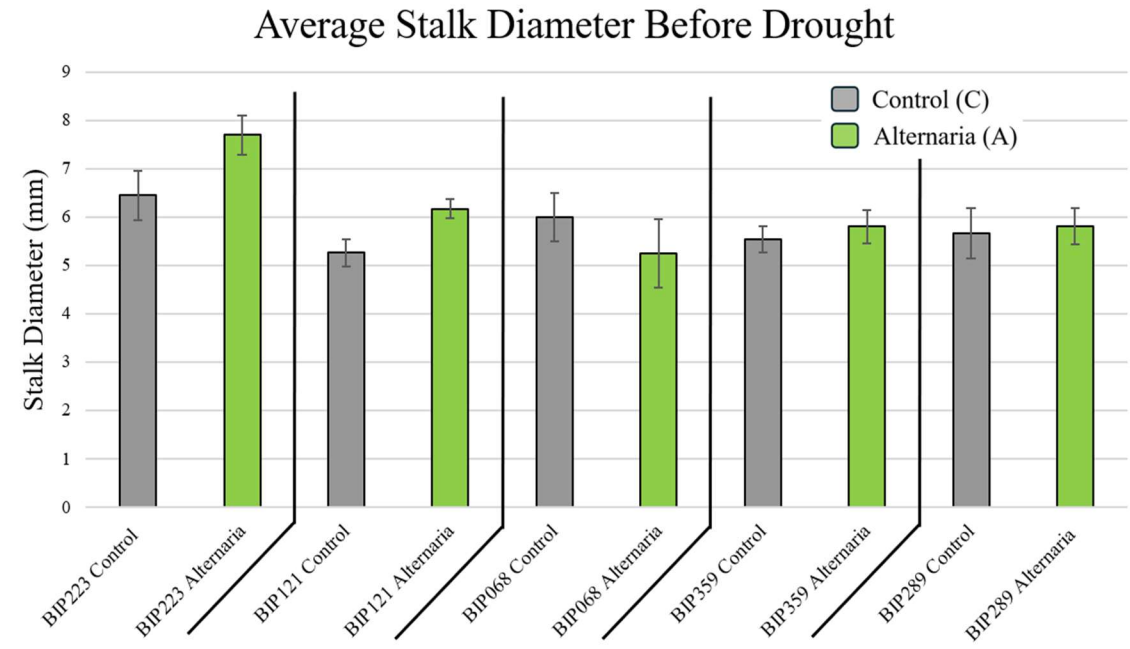

Figure 3.5. Average stalk diameter for each genotype before drought treatment. Each bar represents the average of all 6 replicates within a given treatment/genotype combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean.

After Drought

The post-drought measurements were taken one week after the last plant resumed regular watering after the severe drought treatment. This timepoint still shows variation due to the drought treatment but was before the aphid infestation.

The plant height measurements taken after the drought treatment (Table 3.3C) did not show statistically significant variation between genotypes when looking at height irrespective of *A. alternata* inoculation or drought treatment ($P = 0.1355$), meaning genotype alone does not determine heights. Drought did have a significant effect on plant height ($P < 0.0001$), with a trend of lower heights in droughted plants (Figure 3.7). The genotype by drought interaction is not quite significant ($P = 0.055$), but there is a

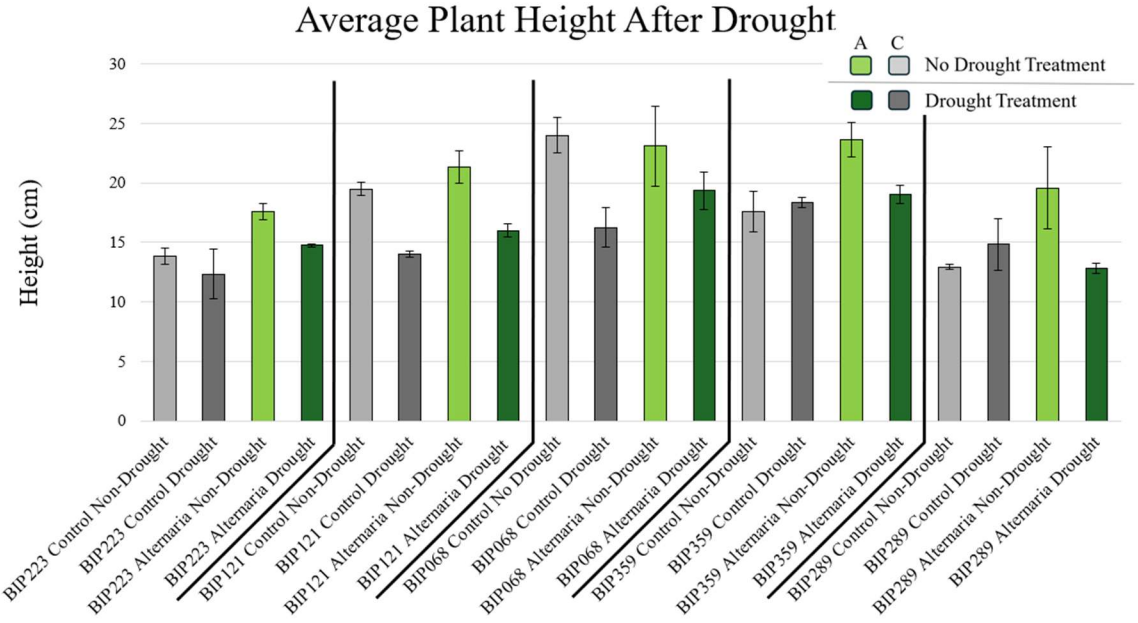

Figure 3.7. Average plant height for each genotype after drought treatment (regular watering has been resumed for all plants, though some had not fully recovered from drought and had wilting scores greater than 1 when height measurements were conducted). Each bar represents the average of 3 replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The gray bars represent the controls (C) and the green bars the Alternaria alternata (A) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the well-watered plants.

graphical trend of some genotypes being more affected by drought than others. The impact of *A. alternata* inoculation overall (not considering drought treatment or genotype) was not statistically significant, suggesting that there was no universal benefit of *A. alternata* across all genotypes. However, the genotype by *A. alternata* inoculation interaction is significant ($P < 0.0001$), meaning some genotypes were more responsive to *A. alternata* than others. But the interaction of *A. alternata* inoculation and drought was not significant ($P = 0.1505$). So, while *A. alternata* does seem to help some genotypes retain height even under drought stress, the effect is not strong enough to be statistically significant across the entire dataset. Additionally, the three-way interaction between genotype, *A. alternata*, and drought is not significant ($P = 0.2951$), suggesting that the effect of *A. alternata* more likely depends on specific genotype-endophyte interactions rather than a universal effect across the entire experiment.

Like for plant height, stalk diameter after drought (Table 3.3D) did not show statistically significant variation between genotypes ($P = 0.4137$). Stalks did show a significant decrease in diameter as a result of drought ($P < 0.0001$). The interaction between genotype and drought was not significant ($P = 0.5732$), indicating drought affected different genotypes similarly. Based on the significant effect of *A. alternata* ($P = 0.0057$) and the graphical trends (Figure 3.8), it appears that *A. alternata* may be able to help plants maintain thicker stalks. Again, the *A. alternata* effect varies by genotype ($P = 0.0313$), with some genotypes likely responding more positively to inoculation than others. Additionally, the interaction between *A. alternata* and drought was significant ($P = 0.0026$), indicating *A. alternata* significantly impacts response to drought, with a trend towards increased stalk thickness (Figure 3.8). In contrast, the three-way interaction

between genotype, *A. alternata* inoculation, and drought was not significant (P = 0.1102), meaning the impact of *A. alternata* under drought stress is not dependent on genotype.

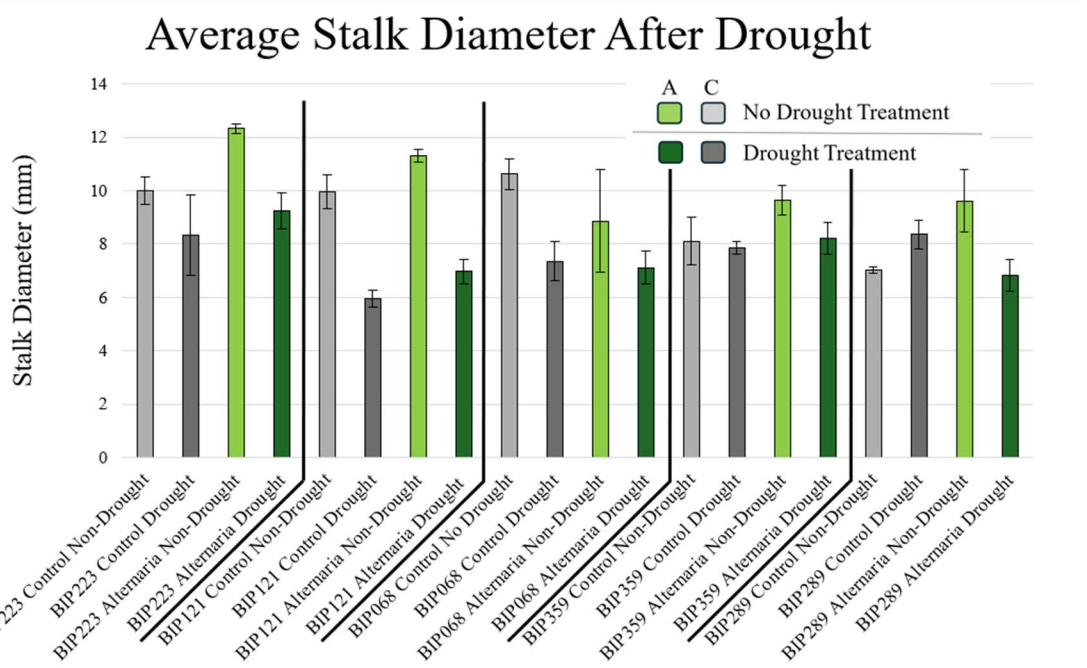

Figure 3.8. Average plant stalk diameter for each genotype after drought treatment (regular watering has been resumed for all plants, though some are still not fully recovered from drought and have scores greater than 1 for wilting at the time the measurements were made). Each bar represents the average of 3 replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The gray bars represent the controls (C) and the green bars the Alternaria alternata *(A) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.*

Measurements on the average number of days before plants reached a wilting score of five (Table 3.3E) revealed a genotypic effect (P=0.0001). In other words, genotype significantly impacted how quickly plants wilted under water limiting conditions, suggesting some genotypes may be more drought tolerant than others regardless of *A. alternata* treatment. The *A. alternata* inoculation effect was marginally significant (P = 0.047), with inoculation resulting in a trend of delayed wilting across

genotypes (Figure 3.9). However, the genotype by *A. alternata* inoculation interaction was not significant (P = 0.3049).

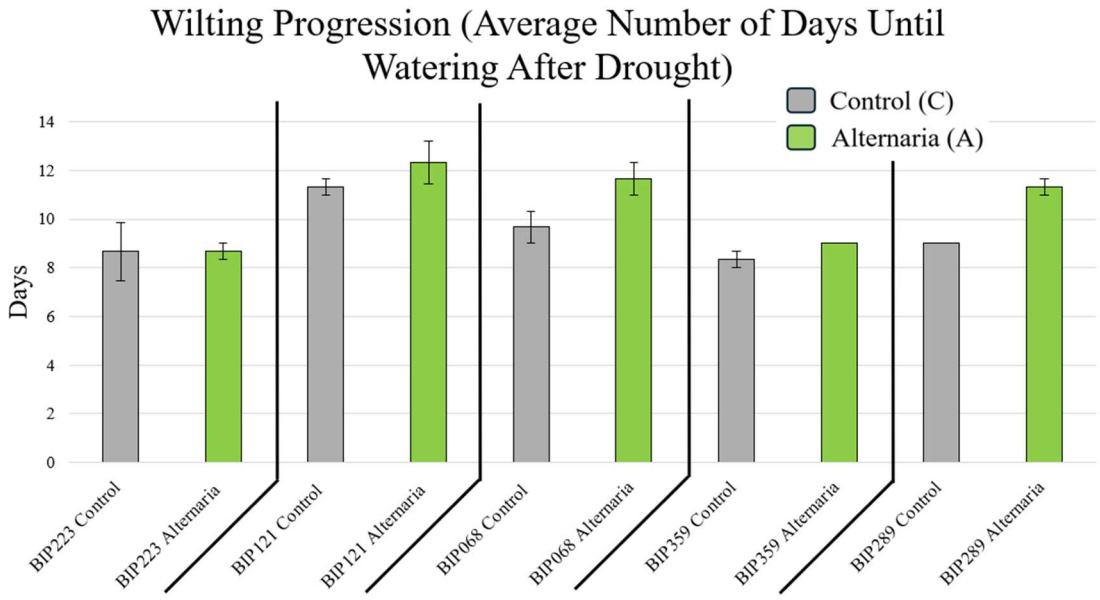

Figure 3.9. Wilting progression (average number of days water was withheld until plants reached a wilting score of 5). Bars represent the average of the 3 replicates within a given treatment/genotype combination that were subjected to drought stress. Vertical lines separate the different genotypes. Error bars represent the standard error of the mean.

Biomass

Aboveground biomass measurements taken at the end of the experiment (Table 3.3F) revealed that genotype had a significant effect on biomass accumulation (P = 0.0011), suggesting some genotypes naturally produce more biomass than others. Drought was also highly significant (P < 0.0001), with droughted plants generally having lower biomass than non-droughted plants (Figure 3.10). Some genotypes were also able to better maintain biomass yields under drought as illustrated by the significant interaction between genotype and drought (P = 0.0174). *A. alternata* also had a significant effect (P = 0.014), suggesting that *A. alternata* might impact biomass

accumulation. Additionally, the genotype by *A. alternata* inoculation interaction was also significant (P = 0.002), indicating that some genotypes may respond more positively to *A.*

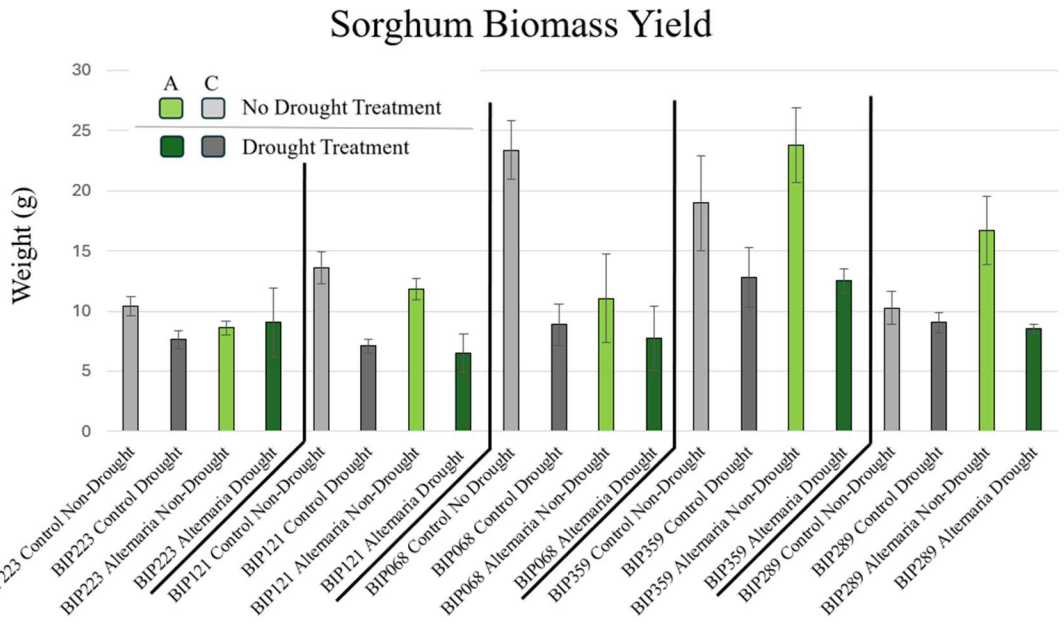

Figure 3.10. Average sorghum aboveground dry biomass for each genotype at the end of the experiment. Each bar represents the average of 3 replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The gray bars represent the controls (C) and the green bars the Alternaria alternata (A) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

alternata. However, the interaction between *A. alternata* and drought was not significant (P = 0.8701), meaning that *A. alternata* did not consistently affect biomass (positively or negatively) under drought conditions. The three-way interaction between genotype, *A. alternata* inoculation, and drought was again not statistically significant (P = 0.1725), suggesting that while *A. alternata* might provide biomass accumulation benefits for certain genotypes, there wasn't a global trend of *A. alternata* impacts being correlated with drought tolerance across genotypes.

Aphid Infestation

Despite aphid infestation (Table 3.3G) being unintentional, the disease pressure revealed some interesting trends. As expected, genotype had a significant effect on sustained aphid damage ($P = 0.0002$). Interestingly, *A. alternata* inoculation also had a significant impact on disease severity ($P < 0.0001$) with the *A. alternata* inoculated plants showing higher infestation levels than control plants (Figure 3.11). The effect of *A. alternata* on aphid severity was dependent on genotype ($P = 0.0062$), meaning some genotypes may be more negatively impacted by inoculation than others. While there wasn't significant variation in infestation levels among genotypes as a result of drought

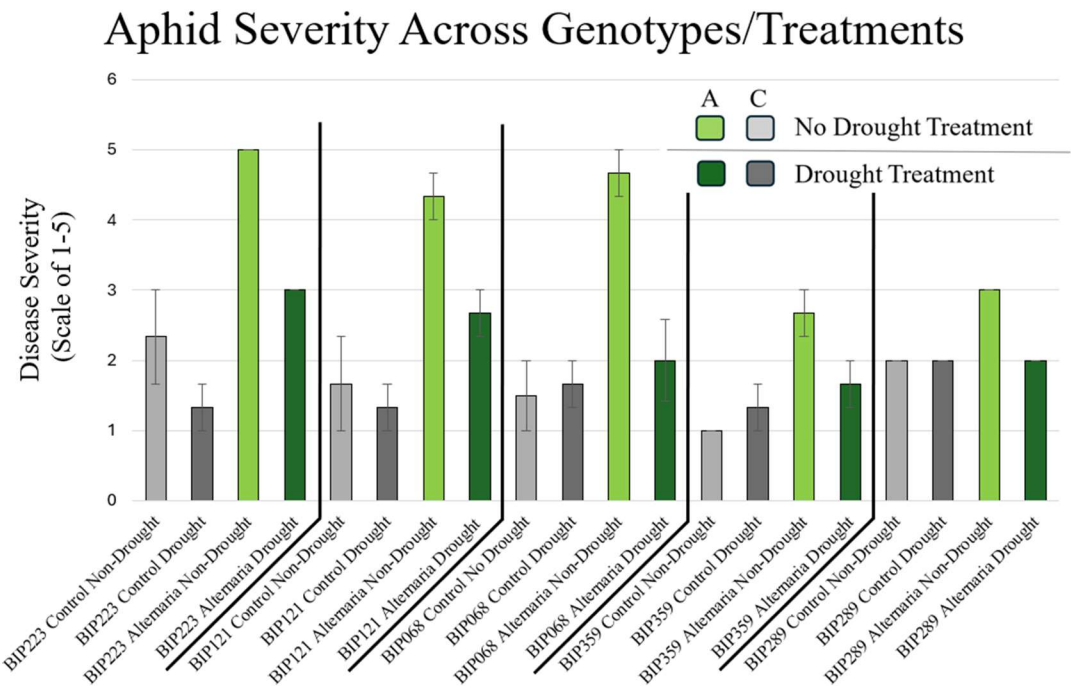

Figure 3.11. Severity of aphid pressure (scored on a scale of 1-5. 1: no infestation; 5: very severe infestation). Each bar represents the average of 3 replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The gray bars represent the controls (C) and the green bars the Alternaria alternata (A) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

(P = 0.1092), the effect of *A. alternata* on disease severity did vary as a result of drought conditions (P < 0.0001). Figure 3.11 suggests that while *A. alternata* generally increased aphid severity, disease symptoms were worse under non-drought conditions. Based on the three-way interaction (P = 0.358), the effects of *A. alternata* inoculation on aphid severity under drought did not depend strongly on genotype. It should also be noted that 10 plants died prior to the end of the experiment, likely as a result of aphid infestation. However, this mortality occurred in the last 1-2 weeks of the experiment. This is after the severe drought cycle had been completed and post-severe drought cycle height and stalk diameter measurements were taken.

Correlation Analysis

The following correlation analysis was done on the full dataset (*A. alternata*-inoculated and non-inoculated plants; drought and non-drought). The correlation matrix between height and stalk diameters measurements taken before the drought treatment (Figure 3.12) doesn't show any significant correlation (R = -0.04, P > 0.05). In other words, a taller plant does not necessarily have a thicker stalk.

Figure 3.13 shows the correlation matrix comparing traits (height, stalk diameter, biomass, aphid severity, and wilting progression/days until watering) after the drought treatment. In contrast to the pre-drought data, post-drought height and stalk diameter are positively correlated (R = 0.49, P < 0.05). In other words, taller plants more commonly had greater stalk thickness. Biomass was strongly correlated with plant height (R = 0.69, P < 0.001), with taller plants having increased biomass. Thicker stalk diameter was moderately correlated with increased biomass (R= 0.40, P > 0.05), however the

relationship was not statistically significant. Biomass ($R = -0.67$, $P < 0.05$) and stalk diameter ($R = -0.80$, $P < 0.01$) were both negatively correlated with wilting progression. In other words, plants with lower biomass and/or thinner stalks survived longer without water before reaching a wilting score of 5. However, plant height was not significantly correlated with delayed wilting ($R = -0.02$). While there was a weak trend of taller plants having more aphid damage ($R = 0.33$, Not Significant), thicker stalks were more strongly correlated with increased aphid severity.

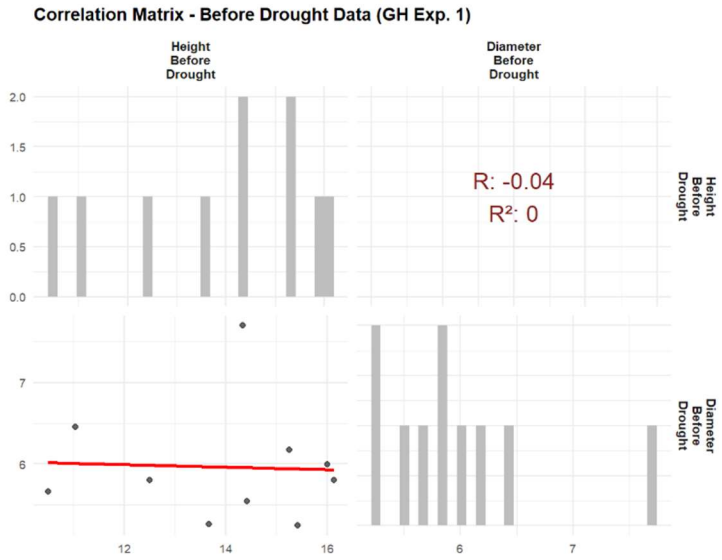

Figure 3.12. The correlation matrix comparing the traits (height and stalk diameter) before the drought treatment. For each pairwise correlation between traits, Pearson's correlation coefficient (r) and coefficients of determination (R^2) were calculated to gauge the strength of the relationship (r) and the proportion of variance explained (R^2). Statistical significance thresholds for Pearson's correlation coefficient were defined as p-values < 0.05 (*), 0.01 (**), and 0.001 (***).

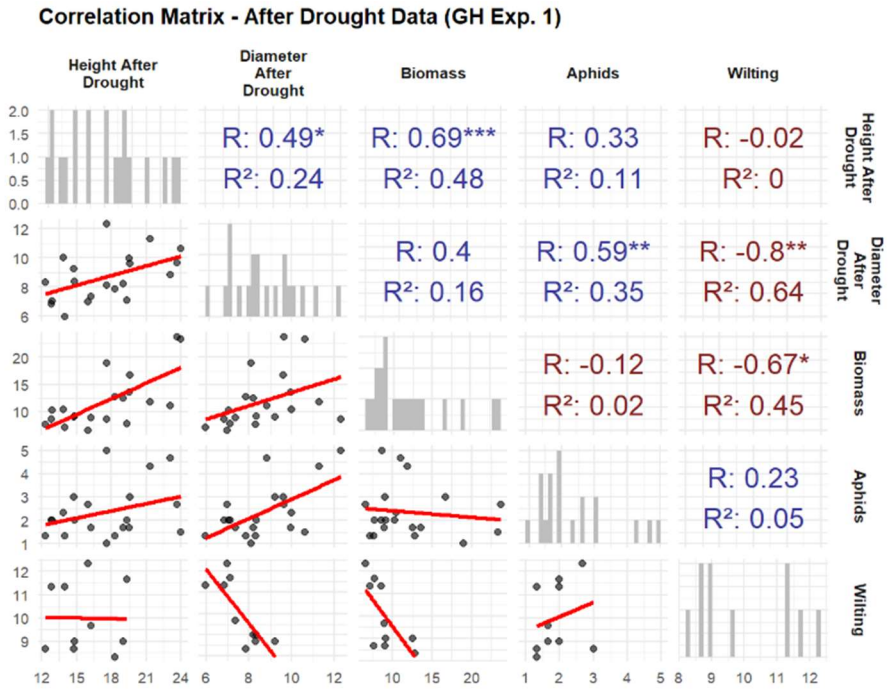

Figure 3.13. The correlation matrix comparing the traits (height, stalk diameter, biomass, aphid severity, and wilting progression/days until watering) after the drought treatment. For each pairwise correlation between traits, Pearson's correlation coefficient (r) and coefficients of determination (R^2) were calculated to gauge the strength of the relationship (r) and the proportion of variance explained (R^2). Statistical significance thresholds for Pearson's correlation coefficient were defined as p-values < 0.05 (*), 0.01 (**), and 0.001 (***).

Greenhouse Experiment 2

Before Drought

The plant height prior to drought treatment (Table 3.4A) was significantly influenced by genotype alone (P = 0.0039), with BIP223 showing a trend of shorter plants (Figure 3.14). However, endophyte (*A. alternata* or *Neopestalotiopsis*) inoculation did not have a significant impact on plant height (P = 0.9041). Additionally, the interaction between genotype and endophyte inoculation was not significant (P = 0.4065), meaning endophyte presence did not differentially affect plant heights across genotypes.

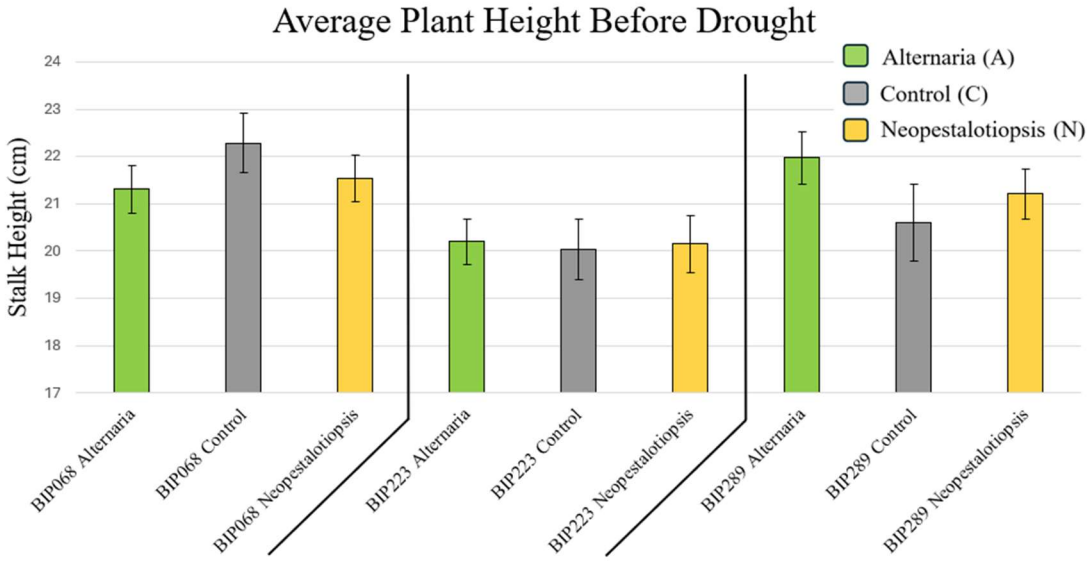

Figure 3.14. Average plant height for each genotype before drought treatment. Each bar represents the average of all 14 replicates within a given treatment/genotype combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean.

Table 3.4. Greenhouse Experiment 2 P-values

Source	P-Value
A) Height Before Drought	
Genotype	0.0039
Endophyte	0.9041
Genotype*Endophyte	0.4065
B) Stalk Diameter Before Drought	
Genotype	<.0001
Endophyte	0.7152
Genotype*Endophyte	0.4474
C) Height After Drought	
Genotype	<.0001
Endophyte	0.7706
Drought	0.0202
Genotype*Endophyte	0.3738
Genotype*Drought	0.087
Endophyte*Drought	0.088
Genoty*Endoph*Drough	0.0605
D) Stalk Diameter After Drought	
Genotype	0.0007
Endophyte	0.8145
Drought	<.0001
Genotype*Endophyte	0.1131
Genotype*Drought	0.3198
Endophyte*Drought	0.0304
Genoty*Endoph*Drough	0.3927
E) Wilting	
Genotype	0.0014
Endophyte	0.6919
Genotype*Endophyte	0.5636
F) Biomass	
Genotype	0.0138
Endophyte	0.4654
Drought	<.0001
Genotype*Endophyte	0.5634
Genotype*Drought	0.2123
Endophyte*Drought	0.0512

Genoty*Endoph*Drough	0.4478
G) Anthracnose Disease	
Genotype	<.0001
Endophyte	0.4941
Drought	0.5663
Genotype*Endophyte	0.7601
Genotype*Drought	0.6826
Endophyte*Drought	0.9392
Genoty*Endoph*Drough	0.9516
H) CO_2 Assimilation (After Drought Cycle 1)	
Genotype	0.0007
Endophyte	0.2839
Drought	0.0518
Genotype*Endophyte*Drought	0.4028
I) CO_2 Assimilation (After Drought Cycle 3)	
Genotype	0.5456
Endophyte	0.8023
Drought	0.6884
Genotype*Endophyte*Drought	0.6394

Similarly to plant height, stalk diameter prior to the drought treatment (Table 3.4B) also had a strong genotypic effect irrespective of endophyte inoculation ($P < 0.0001$). Interestingly, BIP223, which had the shortest plants (Figure 3.14), showed a trend of larger stalk diameters (Figure 3.15). Neither endophyte inoculation ($P = 0.7152$) nor the interaction between endophyte and genotype ($P = 0.4474$) significantly affected stalk diameter.

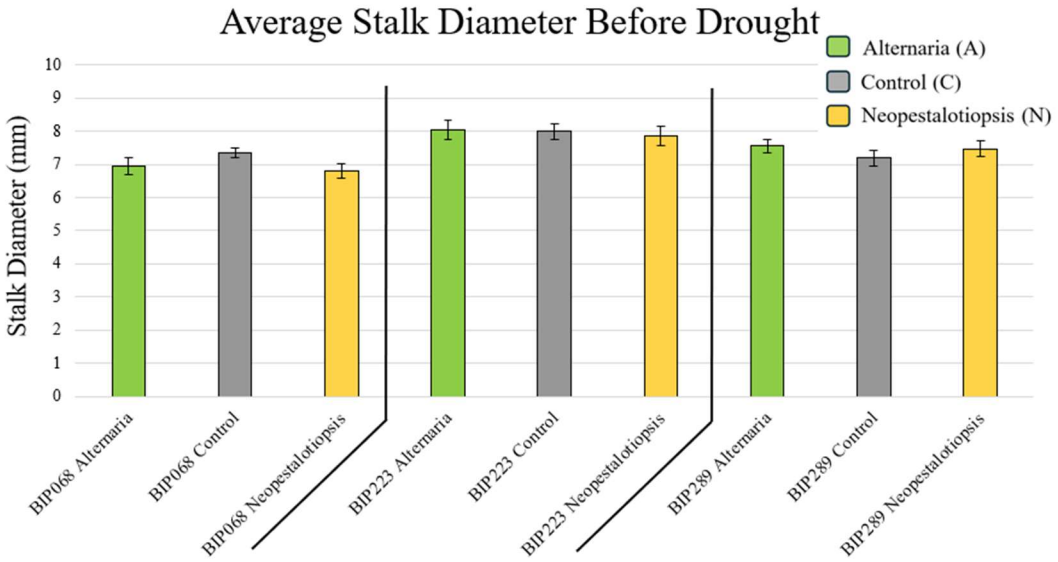

Figure 3.15. Average plant diameter for each genotype before drought treatment. Each bar represents the average of all 14 replicates within a given treatment/genotype combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean.

After Drought

The post-drought measurements were taken at the end of the experiment (after three drought cycles), a day before harvest. After the drought treatment (three drought cycles), plant height (Table 3.4C) remained strongly linked to genotype alone ($P < 0.0001$). Drought also had a significant effect on plant height ($P = 0.0202$), with droughted plants generally having decreased heights (Figure 3.16). However, the genotype by drought interaction was not statistically significant ($P= 0.087$). Endophyte inoculation did not significantly impact plant height ($P = 0.7706$), suggesting that the presence of *A. alternata* or *Neopestalotiopsis* neither enhanced or reduced plant height.

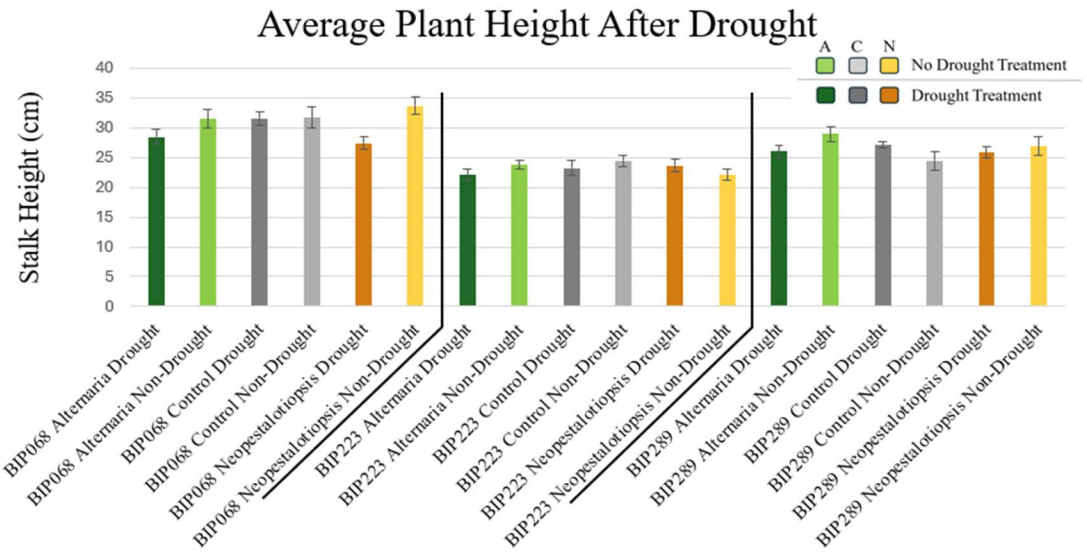

Figure 3.16. Average plant height for each genotype at the end of the experiment, showing the cumulative effect of drought on height. Each bar represents the average of seven replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The green bars the Alternaria alternata (A) inoculated plants, gray bars represent the controls (C), and the orange bars represent Neopestalotiopsis (N) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

The three-way interaction between genotype, endophyte inoculation, and drought was not significant (P = 0.0605). However, while the interaction wasn't quite significant, Figure 3.16A does show some interesting trends. For example, BIP068 shows a larger decrease in height under drought when inoculated with an endophyte compared to the other genotypes.

Similar to the post-drought height measurements, there was a significant genotypic effect (P = 0.0007) and drought effect (P < 0.0001) in post-drought stalk diameter (Table 3.4D), irrespective of endophyte inoculation (Figure 3.17). However, while drought resulted in decreased stalk diameters overall, the drought by genotype effect was not significant (P = 0.3198). This means that while genotypes varied in overall stalk diameter, all genotypes responded similarly to drought with thinner stalks. While the overall endophyte effect (P = 0.8145) and the genotype by endophyte effect (P = 0.1131) were not significant, the endophyte by drought effect was (P = 0.0304). This means that the effect of endophyte inoculation on stalk thickness was drought-dependent.

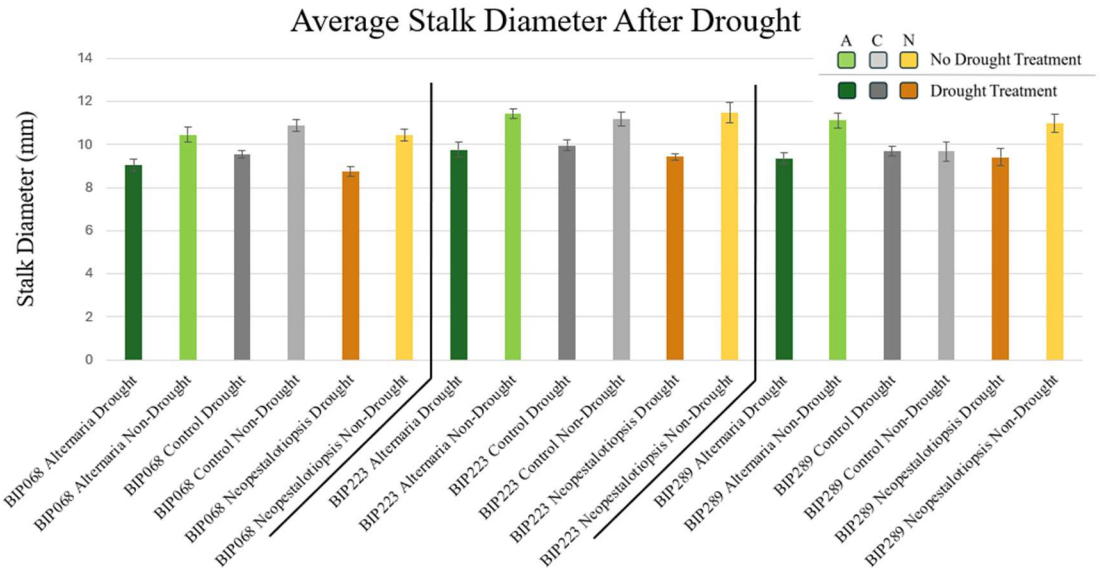

Figure 3.17. Average stalk diameter for each genotype at the end of the experiment, showing the cumulative effect of drought on stalk diameter. Each bar represents the average of seven replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The green bars represent the https://www.random.org/sequences/ (A) inoculated plants, gray bars represent the controls (C), and the orange bars represent Neopestalotiopsis (N) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

Plants showed the most severe wilting after drought cycle two, so average wilting scores for this drought cycle were used as the representative wilting scores (Table 3.4E). There was a significant genotypic effect on wilting, irrespective of endophyte inoculation or drought treatment (P = 0.0014). This trend can be visualized in Figure 3.18 which suggests that BIP068 is the most drought tolerant as it generally has the lowest wilting scores. However, there were not any significant endophytic effects (P = 0.6919) or genotype by endophyte interactions (P = 0.5636). This suggests that endophyte inoculation, overall or across genotypes, did not significantly impact wilting.

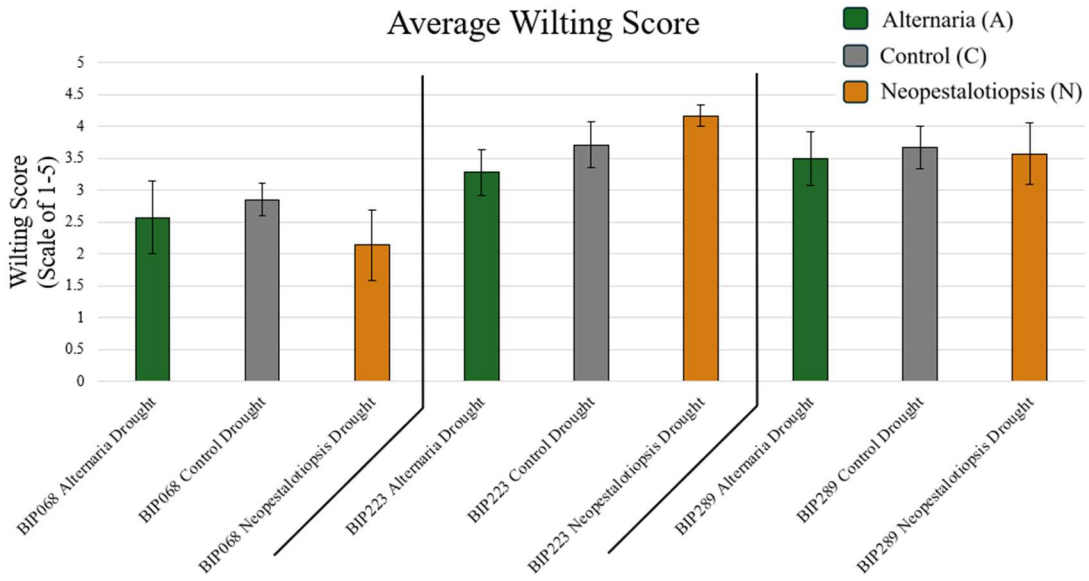

Figure 3.18. Shows the average wilting score after the second drought cycle (Scale of 1-5). Bars represent the average of the seven replicates within a given treatment/genotype combination that were put under drought stress. Error bars represent the standard error of the mean.

Biomass

For the aboveground biomass collected at the end of the experiment (Table 3.4F), there was once again a significant overall genotypic effect (P = 0.0138) on biomass accumulation. Drought also had a significant effect (P < 0.0001), with drought stress

resulting in reduced biomass yields. Endophyte inoculation did not significantly impact overall aboveground biomass (P = 0.4654), however, the endophyte by drought interaction was marginally significant (P = 0.0512). This could mean that endophyte inoculation may significantly impact biomass accumulation under drought conditions. Based on Figure 3.19, there seems to be a trend of decreased biomass when inoculated with an endophyte under drought conditions, with *Neopestalotiopsis* causing larger decreases in BIP068 and BIP223 (Figure 3.19).

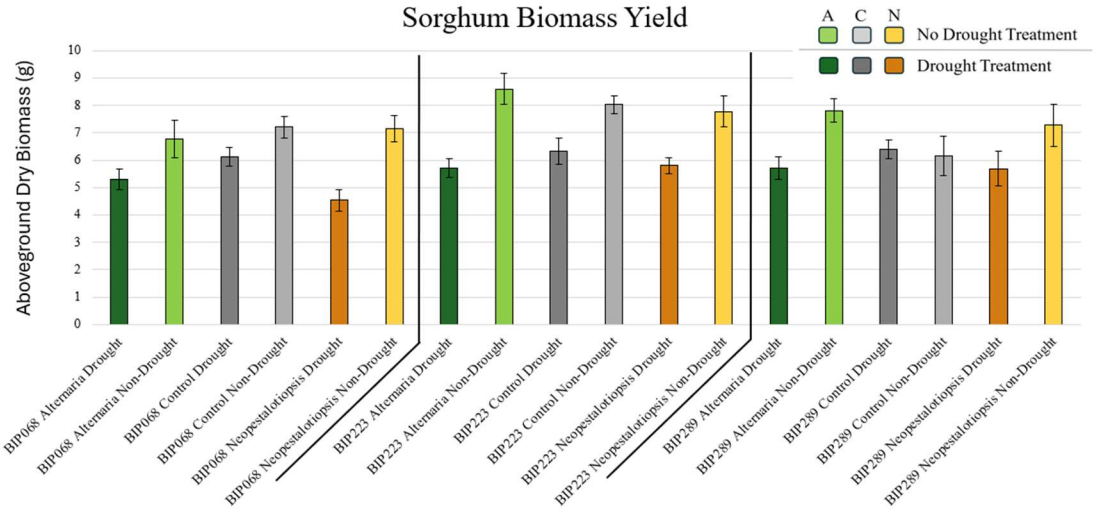

Figure 3.19. Average aboveground dry biomass for each genotype at the end of the experiment, showing the cumulative effect of drought on biomass. Each bar represents the average of seven replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The green bars represent the Alternaria alternata (A) inoculated plants, gray bars represent the controls (C), and the orange bars represent Neopestalotiopsis (N) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

Anthracnose Disease

Greenhouse Experiment 2 suffered from unintentional anthracnose infection (Table 3.4G). While the experiment was ended relatively soon after anthracnose symptoms became noticeable to ensure they didn't skew results, measurements were still

taken on anthracnose disease severity across the treatment groups. Anthracnose disease severity varied significantly by genotype ($P < 0.0001$). However, there were no significant trends as a result of endophyte inoculation or drought. Four plants were dead/missing at the end of the experiment, however, disease was not the cause. Two plants never germinated, one plant germinated but died a week later, and one plant was damaged while taking height and stalk measurements.

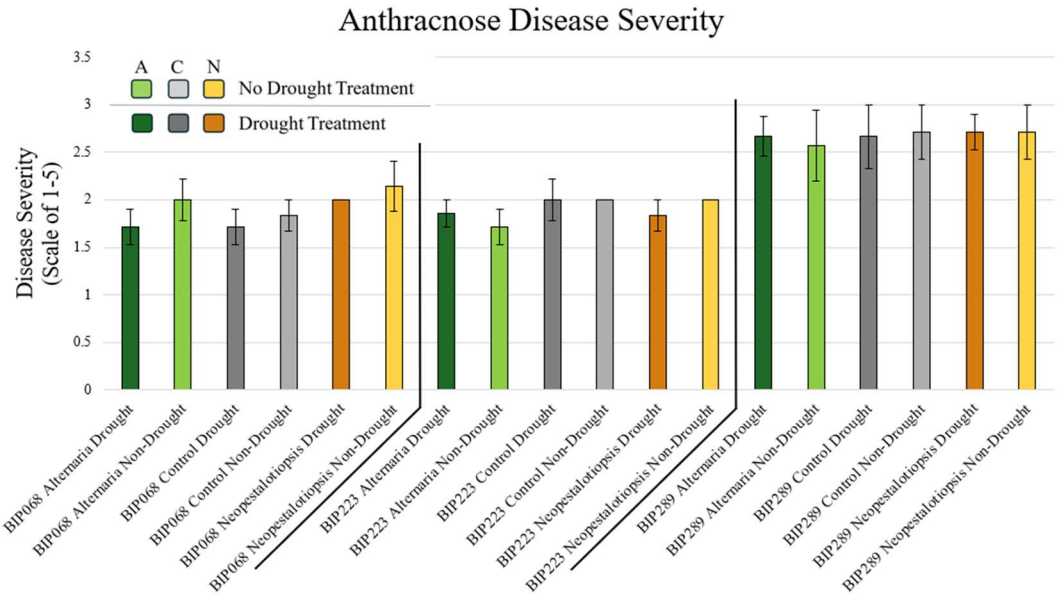

Figure 3.20. Severity of anthracnose disease pressure (scored on a scale of 1-5. 1: no disease; 5: very severe disease). Each bar represents the average of seven replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The green bars represent the (A) inoculated plants, gray bars represent the controls (C), and the orange bars represent Neopestalotiopsis (N) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

Photosynthetic Gas Exchange

LI-COR measurements were taken at two timepoints (end of drought cycle 1 and end of drought cycle 3). For the first timepoint, only two reps per genotype/endophyte/drought treatment group were measured (Table 3.4H)). This revealed that genotype had a significant impact on CO_2 assimilation (P = 0.0007). Drought treatment was marginally significant (P = 0.0518), indicating a potential effect of drought treatment on CO_2 assimilation. However, endophyte inoculation did not have a significant effect on CO_2 assimilation (P = 0.2839).

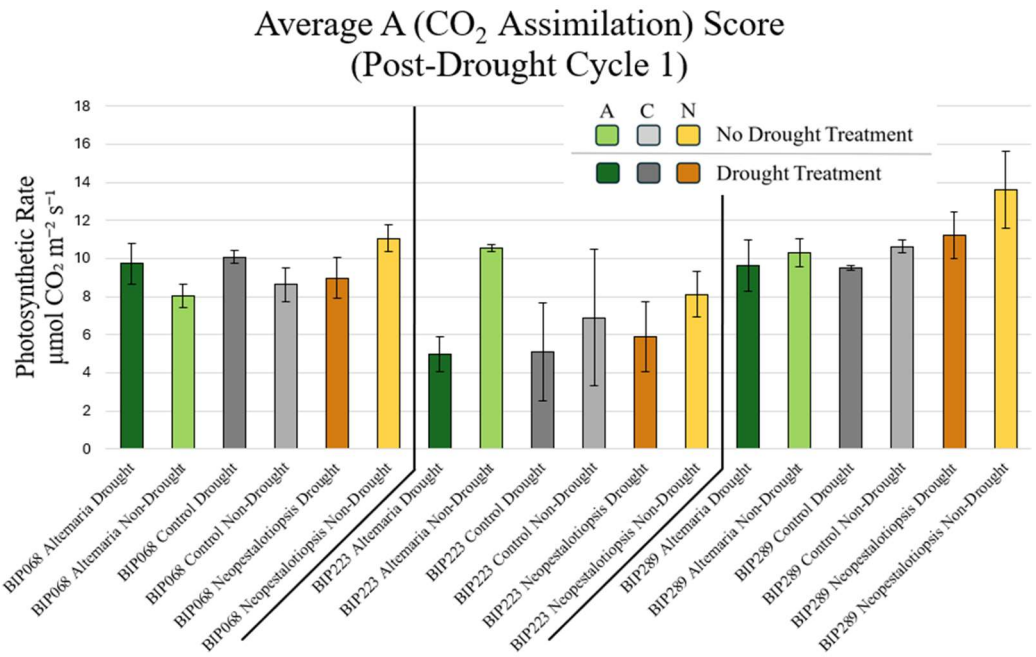

Figure 3.21. The Average A (CO_2 Assimilation) score obtained from LI-COR measurements on plants at the end of the first drought cycle. Each bar represents the average of two replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The green bars represent the Alternaria alternata (A) inoculated plants, gray bars represent the controls (C), and the orange bars represent Neopestalotiopsis (N) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

The second time point was taken after drought cycle 3 and 1-2 days prior to harvest (Table 3.4I). Unfortunately, none of the parameters were significantly linked to CO_2 assimilation. However, it should be noted that while the weather was mostly sunny during the first LI-COR time point, for the second time point the sky was cloudy during both data collection days. Additionally, plants were not showing visible signs of wilting at the time of the second LI-COR time point.

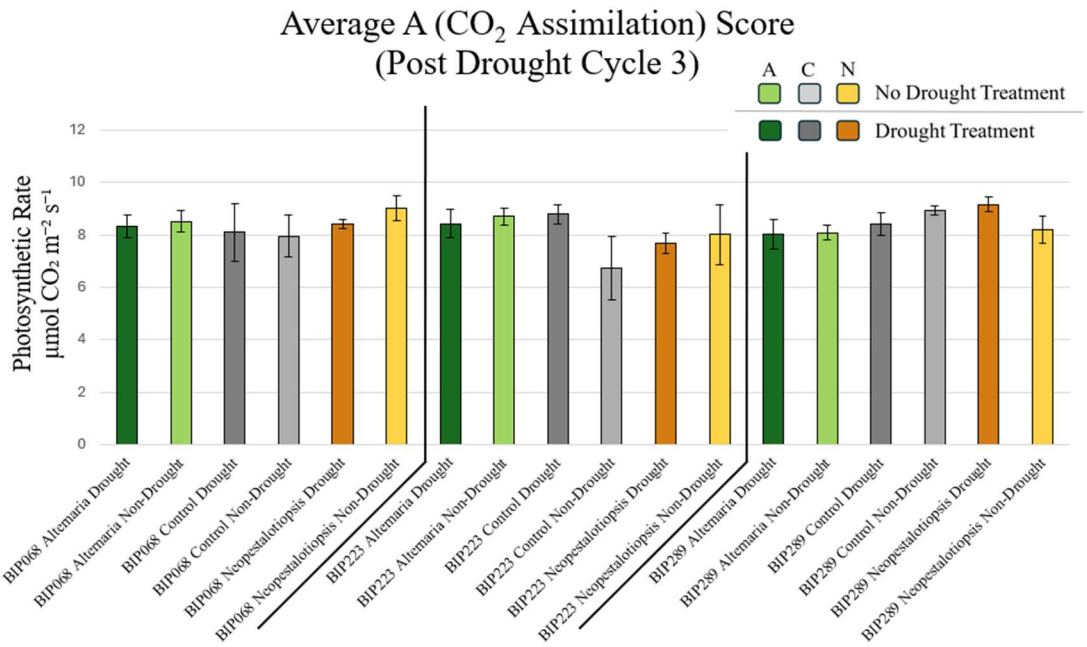

Figure 3.22. The Average A (CO_2 Assimilation) score obtained from doing LI-COR measurements on plants at the end of the third (final) drought cycle at the end of the experiment. Each bar represents the average of two replicates within a given genotype/treatment/drought combination. The vertical lines separate the different genotypes. Error bars represent the standard error of the mean. The green bars represent the Alternaria alternata (A) inoculated plants, gray bars represent the controls (C), and the orange bars represent Neopestalotiopsis (N) inoculated plants. The darker/more saturated colored bars represent the drought treatment plants and the lighter/less saturated bars represent the non-drought plants.

Correlation Analysis

CO2The following correlation analysis was conducted on the full dataset (all treatments). The correlation matrix between height and stalk diameter measurements pre-drought (Figure 3.23) has a moderate negative correlation (R = -0.6). This suggests that increased plant height is correlated with thinner stalks, however the relationship is not statistically significant.

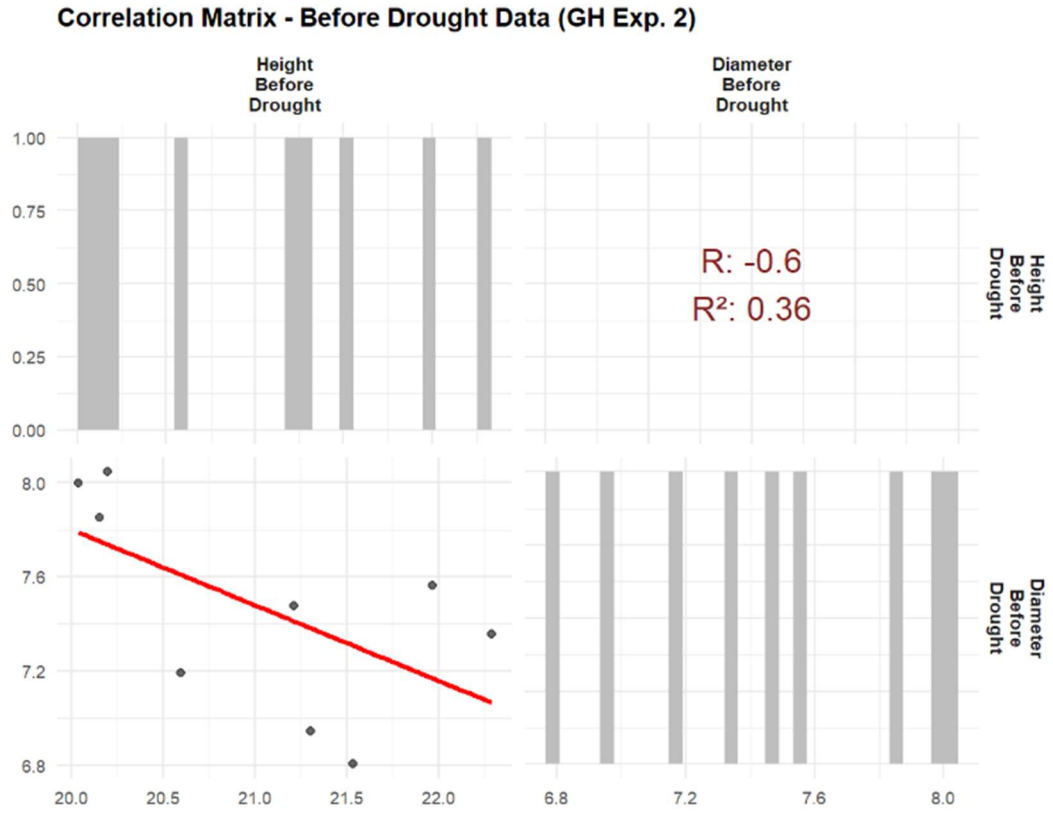

Figure 3.23. The correlation matrix comparing the traits (height and stalk diameter) before the drought treatment. For each pairwise correlation between traits, Pearson's correlation coefficient (r) and coefficients of determination (R^2) were calculated to gauge the strength of the relationship (r) and the proportion of variance explained (R^2). Statistical significance thresholds for Pearson's correlation coefficient were defined as p-values < 0.05 (*), 0.01 (**), and 0.001 (***).

Figure 3.24 shows the correlation matrix comparing height, stalk diameter, biomass, anthracnose severity, and wilting score after the drought treatment. There was a moderate negative correlation between plant height and wilting score (R = -0.57, P > 0.05), however it wasn't statistically significant. In contrast, there was a strong positive correlation between stalk diameter and wilting score (R = 0.69, P < 0.05), suggesting that plants with thinner stalks wilted less. Similarly, there was a positive correlation between biomass and wilting score (R =0.71, P < 0.05), indicating that plants with more biomass

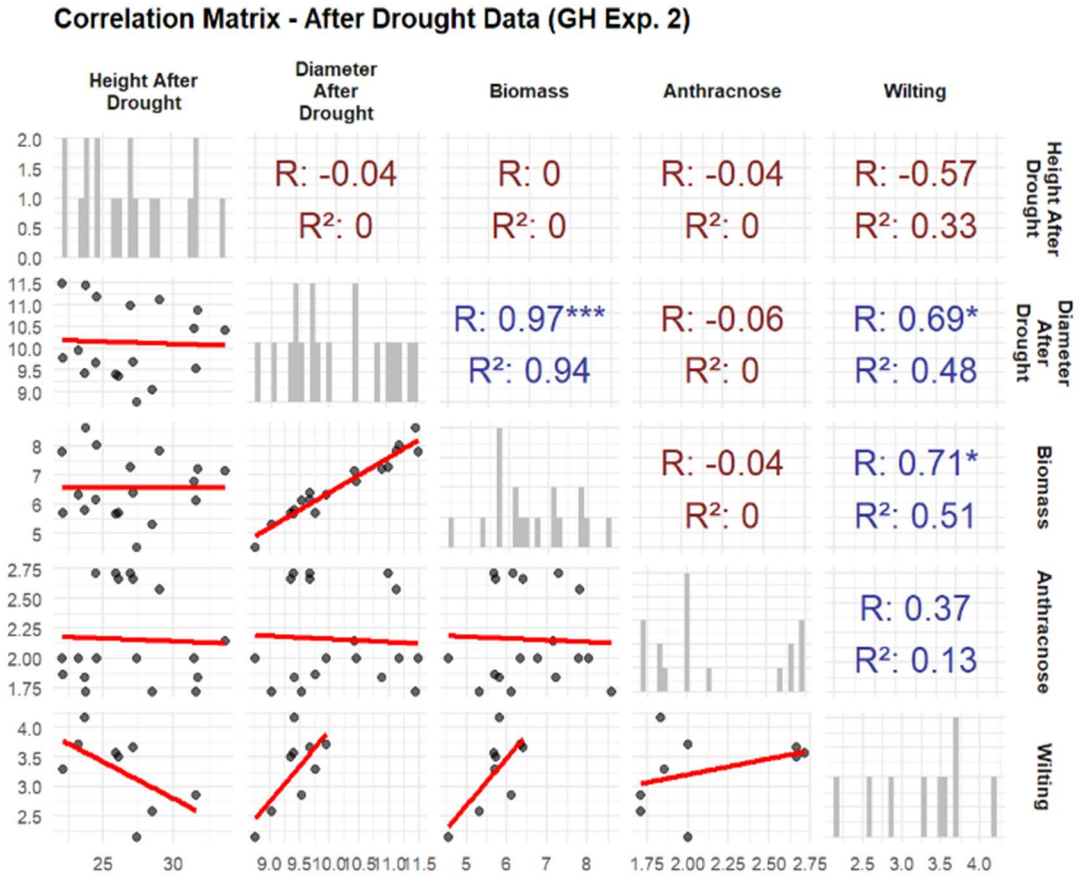

Figure 3.24. The correlation matrix comparing the traits (height, stalk diameter, biomass, anthracnose severity, and wilting score) after the drought treatment. For each pairwise correlation between traits, Pearson's correlation coefficient (r) and coefficients of determination (R^2) were calculated to gauge the strength of the relationship (r) and the proportion of variance explained (R^2). Statistical significance thresholds for Pearson's correlation coefficient were defined as p-values < 0.05 (*), 0.01 (**), and 0.001 (***).

exhibited more wilting. Biomass and stalk diameter also had a strong positive correlation ($R = 0.97$, $P < 0.001$), which suggests that plants with thicker stalks tend to produce more biomass. Anthracnose was not significantly correlated with any of the other traits, likely because the experiment was concluded before disease symptoms became severe. This lack of anthracnose effect is ideal as it was not meant to be a trait of interest.

Microscopy

Microscopy was done on Greenhouse Experiment 2 roots to determine if fungal inoculation with *A. alternata* and *Neopestalotiopsis* was successful. While septate endophyte structures were identified in the inoculated roots (Figure 3.25 A, B, and C), the controls (Figure 3.25 E and F) also showed significant colonization by fungi. Additionally, some roots, including *A. alternata* and *Neopestalotiopsis* inoculated roots, showed colonization by non-septate fungi (Figure 3.25 D). Additionally, visual

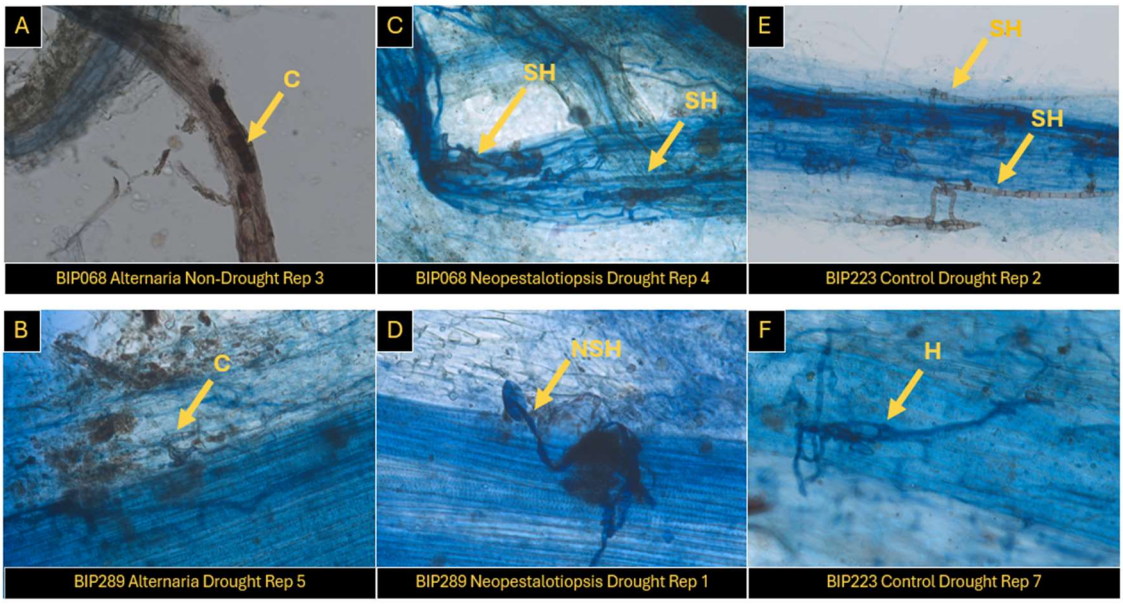

Figure 3.25. Representative microscopy images taken of Greenhouse Experiment 2 roots. A and B show A. alternata-inoculated roots, C and D show Neopestalotiopsis-inoculated roots, and E and F show control roots. Arrows indicate fungal structures of interest such as hyphae or conidia (hyphae: H; septate hyphae: SH; non-septate hyphae: NSH; conidia: C).

quantification of fungal structures did not reveal a statistically significant increase in endophyte presence in roots of *A. alternata*- and *Neopestalotiopsis*-inoculated plants compared to the controls. This suggests that either fungal inoculation was not successful, or that soil or seed contamination introduced other fungal contaminants (Supplemental S3.2).

PCR Analysis

One extra plant per genotype and treatment in Greenhouse Experiment 1 from those that were removed three weeks after planting were assessed by PCR amplification with fungal-specific ITS primers for fungal presence. However, not all the *A. alternata* inoculated plants showed fungal presence (Figure 3.26). Additionally, there was fungal presence in some of the control samples that were not inoculated with endophytes (Figure 3.26). This suggests fungal contamination. However, it should be noted that the control for BIP068 only has weak amplification of the *EIF4α* sorghum amplicon, so the absence of a band for the corresponding control ITS amplicon is not indicative of fungal absence. At the end of Greenhouse Experiment 1, DNA from harvested roots was used to again amplify the ITS region in a subset of samples. Samples were then sent for Sanger sequencing. The results for the *A. alternata* inoculated plants were inconclusive, potentially due to other fungal contaminants. Additionally, the ITS sequence from one

sequenced control sample had high similarity (93.59% identity, 89% query cover) to *Saccharomyces cerevisiae* (Yeast).

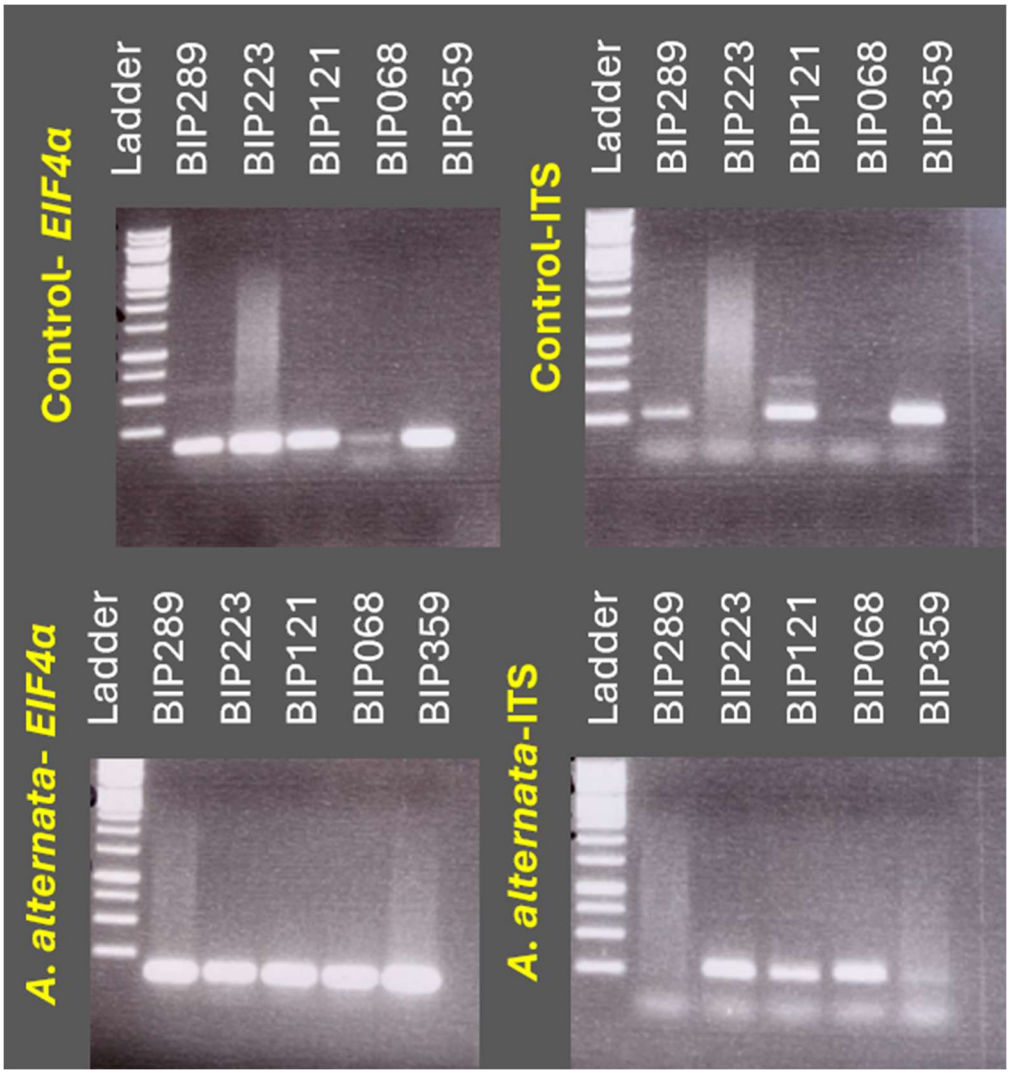

Figure 3.26. PCR gel image for Greenhouse Experiment 1 extra plant sampling analysis. The top left quadrant shows bands for amplicons generated from control plants using the EIF4α sorghum primers. The top right quadrant shows bands for amplicons generated from control plants using the ITS primers. The bottom left quadrant shows bands for amplicons generated from A. alternata-inoculated plants using the EIF4α sorghum primers. The bottom right quadrant shows bands for amplicons generated from A. alternata-inoculated plants using the ITS primers.

Due to the failure of Sanger sequencing at identifying *A. alternata* samples in Greenhouse Experiment 1, samples were not sequenced for Greenhouse Experiment 2.

The microscopy data already indicate significant fungal contamination, so sequencing would likely be inconclusive again. However, in an attempt to confirm *A. alternata* and *Neopestalotiopsis*, specific primers were used. The specific primers were run alongside the sorghum primers (*EIF4α)* to ensure quality DNA. Unfortunately, neither the *A. alternata* nor the *Neopestalotiopsis* specific primers were able to successfully amplify the isolates used in this experiment.

Discussion

This experiment compared the effects of septate endophytes across different sorghum genotypes. Plants were grown in a mixture of sand and field soil to create nutrient limiting conditions to better assess potential endophyte advantages. Plants were also subjected to drought treatments. Greenhouse Experiment 1 only had three replicates per genotype/endophyte (*A. alternata*)/drought treatment group, limiting its statistical power. This contrasts with Greenhouse Experiment 2, which not only had seven replicates per treatment group and two different endophyte inoculation treatments (*A. alternata* and *Neopestalotiopsis*), but also imposed a different drought treatment with set watering intervals. Additionally, plants were ordered by replicate and treatment in Greenhouse Experiment 1 while plants from Greenhouse Experiment 2 were randomized. As such, while results for each experiment can be viewed in parallel, they may not be completely comparable due to their different experimental designs.

Both experiments demonstrated, as expected, significant genotypic effects influencing various traits such as plant height, stalk diameter, biomass, and drought

tolerance. This suggests that genotypic variation is linked to variation in plant growth, including under drought conditions. This is consistent with Chen et al. (2016), which looked at the effect of drought on maize seedlings and found that there was significant genotypic variation in terms of the level of drought resistance and the efficiency of recovery from water deficit. Different sorghum genotypes have also been found to have variable drought tolerances as well (Luquet et al., 2019).

The effects of drought across the two greenhouse experiments also were relatively consistent, with drought affecting plant height, stalk diameter, and biomass accumulation. However, drought effects in Greenhouse Experiment 1 were much more significant than in Greenhouse Experiment 2. This difference in magnitude of drought effects may be a result of experimental design. In other words, withholding water until plants became very wilted (Greenhouse Experiment 1) may have caused a larger impact on plant performance than prolonged drought through intermittent watering (Greenhouse Experiment 2). In Greenhouse Experiment 2, drought intervals were set so that water was provided once ~3-5 plants showed distinct signs of wilting (scores of 3-4 based on the 1-5 scale described previously). This meant that some treatment groups did not experience enough water deficit to induce wilting before water was again provided. For example, BIP223 determined the length of the drought cycles as BIP223 plants (regardless of endophyte inoculation) had earlier severe wilting compared to the other genotypes. That means watering resumed when BIP289 and BIP068 showed either mild or no wilting. While this design was necessary to ensure plants survived, it could be contributing to the smaller drought effects in Greenhouse Experiment 2. Greenhouse Experiment 2 may have also not been long enough to stress the plants to the same degree as in Greenhouse

Experiment 1. Achten et al. (2010) compared how varying levels of drought severity affect growth and biomass production in *Jatropha curcas*. They found that while well-watered plants exhibited the highest biomass yields, plants under moderate drought stress (40% soil moisture) didn't produce as much biomass, but were still able to continue growing, albeit at a slower rate than non-droughted plants. In contrast, the extreme drought plants (no irrigation) halted shoot growth altogether and started shedding leaves. Biomass allocations also shifted towards roots instead of aboveground biomass. This shift into halted growth, leaf wilting, and shedding began 12 days after drought treatment. These results may provide insights into the varying drought treatments in Greenhouse Experiment 1 and 2. Greenhouse Experiment 1 more closely mimicked the extreme drought treatment, withholding water completely until plants began wilting. Greenhouse Experiment 2 more closely follows the trend of the moderate drought stress treatment, where plants continued growing and accumulating biomass, though with reduced biomass yields compared to controls. However, it should be noted that the severe aphid infestation that Greenhouse Experiment 1 suffered may have influenced plant biomass measurements, meaning patterns may have varied slightly under non-infected conditions.

Another interesting aspect of the drought was the effect that wilting and wilting progression had on plant traits. In Greenhouse Experiment 1, stalk diameter and biomass were both negatively correlated with wilting progression. In other words, plants that had thinner stalks and/or lower biomass were able to survive longer without water before reaching a wilting score of 5. Similarly, in Greenhouse Experiment 2, plants with thinner stalks and/or lower biomass were correlated with less wilting, meaning they were less affected by the drought conditions. This follows previously established patterns for

drought stressed plants, where plant size is decreased as a result of things such as changes in turgor pressure and the availability of photosynthates (Yang et al., 2021). Luquet et al. (2019) found that stem size was reduced under drought conditions in sorghum, however, they also noted significant genotypic variation in terms of drought effects on stems. The general pattern is that drought-stressed plants invest more energy towards root growth and reduce resource allocations to the stalks and leaves, as smaller aboveground tissues help reduce transpiration and water loss (Eziz et al., 2017). Given this, the greenhouse experiment results seem reasonable in terms of the observed reduction in stalk diameter and biomass conferring better drought tolerance as seen through reduced wilting.

Soil tests revealed that the soils for both Greenhouse Experiments 1 and 2 are likely deficient in nitrogen, phosphorus, and potassium based on recommended soil nutrient levels for optimal agricultural crop production (E.S. Marx, 1999; Wortmann et al., 2013). This is ideal for the current study, as the utility of inoculating biofuel sorghum with potentially beneficial endophytes is being explored. As biofuel sorghum production will likely be largely allocated to marginal lands (Mullet et al., 2014; Olson et al., 2012; Rooney, 2014; Rooney et al., 2007), determining endophytes that could optimize growth on deficient soils would be invaluable.

Unfortunately, while there were some trends in terms of endophyte effect on plant performance, especially in Greenhouse Experiment 1, the PCR and microscopy results cast doubt on the reliability of the results. In Greenhouse Experiment 2, fungi were found across endophyte inoculated and non-inoculated plants. This suggests either seed or soil contamination. While seeds were surface sterilized, seed endophytes could have been present. However, the diversity of fungal endophytes colonizing the sorghum roots

suggest the contamination was more likely of soil origin. Additionally, some of the microscopy images (Figure 3.25 E) show a structure that resembles *Rhizopus*. The *Rhizopus* genus represents a group of saprotrophic fungi that is widespread in soil and can act as a post-harvest pathogen, feeding on dead plant tissues (Gryganskyi et al., 2018). Since *Rhizopus spp*. would therefore be expected to be external colonizers that target dead or dying plant tissues, its presence on the roots is likely a result of soil contamination rather than being transmitted through seeds. That means that the soil autoclaving was not thorough enough. While autoclave tape was buried in the soil during autoclaving to assess internal temperature, if the tape wasn't perfectly in the center of the soil, it may not have been an accurate representation of the temperature of the entire soil bag. For future similar studies, autoclaving time may need to be increased and soil volume in bags decreased to ensure complete sterilization.

Additionally, while there have been instances of *A. alternata* being isolated from sorghum roots (Janet, 1983), there isn't research on whether *Neopestalotiopsis* can colonize sorghum roots. The *Neopestalotiopsis* isolate used in this experiment was identified from switchgrass. That means we have no direct evidence on its efficiency in colonizing sorghum. Additionally, while microscopy results indicate the presence of septate endophytes, we don't have enough evidence to conclusively identify them as *Neopestalotiopsis.* As such, it's important to consider the possibility that *Neopestalotiopsis* may not be able to successfully colonize sorghum roots or may only colonize at very low rates.

Microscopy data was not gathered on Greenhouse Experiment 1 to assess the level of contamination in the inoculated plants compared to the controls. The higher

significance of *A. alternata* inoculation in Greenhouse Experiment 1 may suggest that conditions were less contaminated, allowing for *A. alternata* effects to be more clearly seen, or that the effects of *A. alternata* were maintained under severe drought stress even in the presence of other endophytes. However, PCR results and Sanger sequencing suggest the presence of at least some fungal contaminants since direct sequencing of the ITS amplicons generated for *A. alternata* inoculated plants did not yield good sequence, suggesting a mixed origin of the amplicons. Additionally, one of the control samples from Greenhouse Experiment 1 that was sequenced shows high sequence similarity to *Saccharomyces cerevisiae* (Yeast). In other words, both greenhouse experiments suffered from contamination.

Given the clear contamination observed in Greenhouse Experiment 2 as well as likely contamination in Greenhouse Experiment 1, reliable conclusions can't be drawn about the effect of *A. alternata* and *Neopestalotiopsis* on sorghum performance. So, while some trends were observed, such as the increased stalk diameters and delayed wilting in *A. alternata* inoculated plants in Greenhouse Experiment 1, these patterns would need to be validated. So, while the possibility of *A. alternata* and *Neopestalotiopsis* providing a fitness advantage to their host plants is still possible, this study was not able to conclusively demonstrate this.

Conclusion

This study represents a preliminary analysis of the effects of *Alternaria alternata* and *Neopestalotiopsis* on different sorghum genotypes under drought stress. This experiment

was able to successfully isolate *Neopestalotiopsis* from switchgrass roots. *Neopestalotiopsis* as well as *A. alternata* were then cultured on plates and used to generate inoculum for applying to sorghum seeds. After plants were well-established, drought was imposed to assess the effects that endophytes had on drought tolerance. Results revealed that sorghum genotype had a significant effect on various traits such as plant height, stalk diameter, biomass, and drought tolerance. Unfortunately, PCR and microscopy results revealed the presence of microbial contamination. As such it is difficult to accurately determine any potential endophytic effects. So, while this represents the first analysis of endophytic *Neopestalotiopsis* and *A. alternata* across sorghum genotypes, future research would need to be conducted under more sterile conditions to elucidate benefits, or drawbacks, of inoculating with these endophytes.

References

Achten, W. M., Maes, W. H., Reubens, B., Mathijs, E., Singh, V. P., Verchot, L., & Muys, B. (2010). Biomass production and allocation in *Jatropha curcas L.* seedlings under different levels of drought stress. *Biomass and Bioenergy*, *34*(5), 667-676.

Arnold, A. E., & Lutzoni, F. (2007). Diversity and host range of foliar fungal endophytes: are tropical leaves biodiversity hotspots? *Ecology*, *88*(3), 541-549.

Baggio, J. S., Forcelini, B. B., Wang, N.-Y., Ruschel, R. G., Mertely, J. C., & Peres, N. A. (2021). Outbreak of leaf spot and fruit rot in Florida strawberry caused by *Neopestalotiopsis* spp. *Plant Disease*, *105*(2), 305-315.

Brenton, Z. W., Cooper, E. A., Myers, M. T., Boyles, R. E., Shakoor, N., Zielinski, K. J., Rauh, B. L., Bridges, W. C., Morris, G. P., & Kresovich, S. (2016). A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. *Genetics*, *204*(1), 21-33.

Chen, D., Wang, S., Cao, B., Cao, D., Leng, G., Li, H., Yin, L., Shan, L., & Deng, X. (2016). Genotypic variation in growth and physiological response to drought stress and re-watering reveals the critical role of recovery in drought adaptation in maize seedlings. *Frontiers in Plant Science*, *6*, 1241.

Cobb, A. B., Wilson, G. W., Goad, C. L., Bean, S. R., Kaufman, R. C., Herald, T. J., & Wilson, J. D. (2016). The role of arbuscular mycorrhizal fungi in grain production and nutrition of sorghum genotypes: Enhancing sustainability through plant-microbial partnership. *Agriculture, Ecosystems & Environment*, *233*, 432-440.

Darapanit, A., Boonyuen, N., Leesutthiphonchai, W., Nuankaew, S., & Piasai, O. (2021). Identification, pathogenicity and effects of plant extracts on *Neopestalotiopsis* and *Pseudopestalotiopsis* causing fruit diseases. *Scientific Reports*, *11*(1), 22606.

DeMers, M. (2022). *Alternaria alternata* as endophyte and pathogen. *Microbiology*, *168*(3), 001153.

E.S. Marx, J. H., and R.G. Stevens. (1999). Soil Test Interpretation Guide is snot ehdivorce don't rhoguth stones when you live in a glass house -riemer. file:///C:/Users/Rachel%20Rackers/Downloads/ec1478.pdf

Eziz, A., Yan, Z., Tian, D., Han, W., Tang, Z., & Fang, J. (2017). Drought effect on plant biomass allocation: A meta-analysis. *Ecology and evolution*, *7*(24), 11002-11010.

Ferreira, A. P., dos Santos Oliveira, J. A., Polonio, J. C., Pamphile, J. A., & Azevedo, J. L. (2023). Recombinants of *Alternaria alternata* endophytes enhance inorganic phosphate solubilization and plant growth hormone production. *Biocatalysis and Agricultural Biotechnology*, *51*, 102784.

Gryganskyi, A. P., Golan, J., Dolatabadi, S., Mondo, S., Robb, S., Idnurm, A., Muszewska, A., Steczkiewicz, K., Masonjones, S., & Liao, H.-L. (2018). Phylogenetic and phylogenomic definition of *Rhizopus* species. *G3: Genes, Genomes, Genetics*, *8*(6), 2007-2018.

Hu, H., Jeewon, R., Zhou, D., Zhou, T., & Hyde, K. D. (2007). Phylogenetic diversity of endophytic *Pestalotiopsis* species in *Pinus armandii* and *Ribes spp.*: evidence from rDNA and β-tubulin gene phylogenies. *Fungal diversity*.

Jacob, J. K. S., Witzel, K., & dela Cruz, T. E. E. (2023). Comparative diversity and functional traits of fungal endophytes in response to elevated mineral content in a Mangrove ecosystem. *Journal of Fungi*, *9*(12), 1186.

Janet, E. (1983). Fungal colonization of stalks and roots of grain sorghum during the growing season. *Plant Disease*, *417*.

Khanna, M., Chen, L., Basso, B., Cai, X., Field, J. L., Guan, K., Jiang, C., Lark, T. J., Richard, T. L., & Spawn-Lee, S. A. (2021). Redefining marginal land for bioenergy crop production. *Gcb Bioenergy*, *13*(10), 1590-1609.

Konstantinova, P., Bonants, P. J., Van Gent-Pelzer, M. P., & Van Der Zouwen, P. (2002). Development of specific primers for detection and identification of *Alternaria*

spp. in carrot material by PCR and comparison with blotter and plating assays. *Mycological Research*, *106*(1), 23-33.

Koske, R., & Gemma, J. (1989). A modified procedure for staining roots to detect VA mycorrhizas.

Kumar, M., & Shukla, P. (2005). Use of PCR targeting of internal transcribed spacer regions and single-stranded conformation polymorphism analysis of sequence variation in different regions of rRNA genes in fungi for rapid diagnosis of mycotic keratitis. *Journal of Clinical Microbiology*, *43*(2), 662-668.

Lee, J. C., Yang, X., Schwartz, M., Strobel, G., & Clardy, J. (1995). The relationship between an endangered North American tree and an endophytic fungus. *Chemistry & Biology*, *2*(11), 721-727.

LI-COR Biosciences. (2017). *Using the LI-6800 Portable Photosynthesis System*. LI-COR Biosciences. https://www.licor.com/env/products/photosynthesis/LI-6800

Luo, W., Ping, X., Zhou, J., Gao, S., Huang, X., Song, S., Xu, J., & He, W. (2025). *Alternaria alternata* JTF001 Metabolites Recruit Beneficial Microorganisms to Reduce the Parasitism of *Orobanche aegyptiaca* in Tomato. *Biology*, *14*(2), 116.

Luquet, D., Perrier, L., Clément-Vidal, A., Jaffuel, S., Verdeil, J. L., Roques, S., Soutiras, A., Baptiste, C., Fabre, D., & Bastianelli, D. (2019). Genotypic covariations of traits underlying sorghum stem biomass production and quality and their regulations by water availability: Insight from studies at organ and tissue levels. *Gcb Bioenergy*, *11*(2), 444-462.

Maharachchikumbura, S. S., Guo, L.-D., Chukeatirote, E., Bahkali, A. H., & Hyde, K. D. (2011). *Pestalotiopsis*—morphology, phylogeny, biochemistry and diversity. *Fungal diversity*, *50*, 167-187.

Maharachchikumbura, S. S., Hyde, K. D., Groenewald, J. Z., Xu, J., & Crous, P. W. (2014). *Pestalotiopsis* revisited. *Studies in Mycology*, *79*(1), 121-186.

Maw, M. J., Houx III, J. H., & Fritschi, F. B. (2017). Maize, sweet sorghum, and high biomass sorghum ethanol yield comparison on marginal soils in Midwest USA. *Biomass and Bioenergy*, *107*, 164-171.

McGonigle, T. P., Miller, M. H., Evans, D., Fairchild, G., & Swan, J. A. (1990). A new method which gives an objective measure of colonization of roots by vesicular—arbuscular mycorrhizal fungi. *New Phytologist*, *115*(3), 495-501.

Mullet, J., Morishige, D., McCormick, R., Truong, S., Hilley, J., McKinley, B., Anderson, R., Olson, S. N., & Rooney, W. (2014). Energy Sorghum—a genetic model for the design of C4 grass bioenergy crops. *Journal of experimental botany*, *65*(13), 3479-3489.

Munir, E., Yurnaliza, Y., Lutfia, A., & Hartanto, A. (2022). Isolation and characterization of phosphate solubilizing activity of endophytic fungi from *Zingiberaceous* species. *OnLine J. Biol. Sci*, *22*, 149-156.

Olson, S. N., Ritter, K., Rooney, W., Kemanian, A., McCarl, B. A., Zhang, Y., Hall, S., Packer, D., & Mullet, J. (2012). High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops. *Biofuels, Bioproducts and Biorefining*, *6*(6), 640-655.

Op De Beeck, M., Lievens, B., Busschaert, P., Declerck, S., Vangronsveld, J., & Colpaert, J. V. (2014). Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PloS one*, *9*(6), e97629.

Ortas, I. (2012). The effect of mycorrhizal fungal inoculation on plant yield, nutrient uptake and inoculation effectiveness under long-term field conditions. *Field crops research*, *125*, 35-48.

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rebello, C. S., Wang, N.-Y., Marin, M. V., Baggio, J. S., & Peres, N. A. (2023). Detection and species differentiation of *Neopestalotiopsis spp.* from strawberry (*Fragaria× ananassa*) in Florida using a high-resolution melting analysis. *PhytoFrontiers™*, *3*(1), 156-163.

Rodriguez, R., White Jr, J., Arnold, A. E., & Redman, a. R. a. (2009). Fungal endophytes: diversity and functional roles. *New Phytologist*, *182*(2), 314-330.

Rooney, W. L. (2014). Cellulosic energy cropping systems. In (pp. 110-123). John Wiley & Sons.

Rooney, W. L., Blumenthal, J., Bean, B., & Mullet, J. E. (2007). Designing sorghum as a dedicated bioenergy feedstock. *Biofuels, Bioproducts and Biorefining*, *1*(2), 147-157.

Sabia, E., Claps, S., Morone, G., Bruno, A., Sepe, L., & Aleandri, R. (2015). Field inoculation of arbuscular mycorrhiza on maize (*Zea mays L.*) under low inputs: preliminary study on quantitative and qualitative aspects. *Italian Journal of Agronomy*, *10*(1), 30-33.

SAS. (2023). *SAS® 9.4M8 Global Statements: Reference*. SAS Institute Inc.

Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A., Toomet, O., Crowley, J., & Hofmann, H. (2021). GGally: Extension to'ggplot2'.

Shi, J., Li, B., Wang, S., Zhang, W., Shang, M., Wang, Y., & Liu, B. (2024). Occurrence of Neopestalotiopsis clavispora causing Apple leaf spot in China. *Agronomy*, *14*(8), 1658.

Smith, S. E., & Read, D. J. (2010). *Mycorrhizal symbiosis*. Academic press.

Stamenković, O. S., Siliveru, K., Veljković, V. B., Banković-Ilić, I. B., Tasić, M. B., Ciampitti, I. A., Đalović, I. G., Mitrović, P. M., Sikora, V. Š., & Prasad, P. V. (2020). Production of biofuels from sorghum. *Renewable and Sustainable Energy Reviews*, *124*, 109769.

Sudhakar Reddy, P., Srinivas Reddy, D., Sivasakthi, K., Bhatnagar-Mathur, P., Vadez, V., & Sharma, K. K. (2016). Evaluation of sorghum [*Sorghum bicolor (L.)*] reference genes in various tissues and under abiotic stress conditions for quantitative real-time PCR data normalization. *Frontiers in Plant Science*, *7*, 529.

Tralamazza, S. M., Piacentini, K. C., Iwase, C. H. T., & de Oliveira Rocha, L. (2018). Toxigenic *Alternaria* species: impact in cereals worldwide. *Current opinion in food science*, *23*, 57-63.

Wei, J.-G., Xu, T., Guo, L.-D., Liu, A.-R., Zhang, Y., & Pan, X.-H. (2007). Endophytic Pestalotiopsis species associated with plants of Podocarpaceae, Theaceae and Taxaceae in southern China. *Fungal Divers*, *24*(1), 55-74.

Wortmann, C. S., Ferguson, R. B., Hergert, G. W., Shapiro, C. A., & Shaver, T. M. (2013). Nutrient management suggestions for grain sorghum. *Bulletin of Nebraska University*, 206-215.

Xin, Z., & Chen, J. (2012). A high throughput DNA extraction method with high yield and quality. *Plant methods*, *8*, 1-7.

Yang, X., Lu, M., Wang, Y., Wang, Y., Liu, Z., & Chen, S. (2021). Response mechanism of plants to drought stress. *Horticulturae*, *7*(3), 50.

Zhou, L. S., Tang, K., & Guo, S. X. (2018). The plant growth-promoting fungus (PGPF) *Alternaria sp.* A13 markedly enhances *Salvia miltiorrhiza* root growth and active ingredient accumulation under greenhouse and field conditions. *International journal of molecular sciences*, *19*(1), 270.

CHAPTER 4

CONCLUSIONS

MiRNA Analysis

Highlights

The goals of this chapter were to (1) identify miRNAs from sorghum roots harvested from a bioenergy association panel (BAP) in Georgia as well as Arizona, (2) Determine which miRNAs were differentially expressed between Georgia and Arizona, (3) identify potential miRNA targets, and (4) identify miRNA eQTL via a miRNA expression Genome Wide Association Study (eGWAS). The experiment successfully identified 98 miRNAs, with 20 representing novel miRNAs not previously identified. Of the miRNAs identified in Georgia, the expression of 14 miRNAs was determined to vary significantly by genotype. Differential expression analysis yielded 37 miRNAs that varied significantly in expression between Georgia and Arizona. And lastly, the miRNA eGWAS identified many trans eQTL that could be useful in elucidating complete miRNA regulatory pathways. In other words, the additional information provided by eQTL could improve understanding of what environmental conditions may influence regulatory eQTL gene expression, which then modulates miRNA expression levels, impacting expression of target genes, and ultimately influencing a trait.

Literature searches for miRNAs of interest revealed the potential links between miRNA expression and AMF regulation. For example, miR396 targets GROWTH REGULATING FACTOR (GRF) transcription factors, which are involved in growth and stress responses. But of note is the fact that they've been linked to AMF symbiosis, specifically as a negative regulator (Ledford et al., 2024). As this was one of the miRNAs determined to be significantly influenced by sorghum genotype. This could indicate that miR396 could be one mechanism through which sorghum is able to regulate its interactions with AMF. While the specific AMF abundance as a result of variable miR396 expression has not yet been analyzed, target gene prediction did at least corroborate mi396's function in sorghum by identifying GRF transcription factor genes as the targets.

Limitations

Functions of miRNAs implied from the literature, while useful as a preliminary analysis of miRNA roles, may vary from actual patterns observed in context. As such, one of the main limitations of this study is the lack of AMF phenotypic data to compare miRNA expression against. Another limitation is the fact that the small RNA reads used as inputs for miRNA identification came from only a subset of the entire dataset. For example, only 227 (out of 1,011) samples from the Georgia field experiment and 24 (out of 915) from the Arizona field experiment were used. As such, preliminary conclusions may not be accurate in the context of the full dataset. It should also be noted that the genotypes represented for each of the replicates was uneven. As a result, eGWAS was only conducted separately on two of the replicates that had enough samples. However, the genotypes represented in each replicate were not equivalent. Additionally, the inputs

for each of the GWAS runs represented fewer than 100 samples. This means the likelihood of false positives or negatives is much higher than if the entire dataset were represented (Ko et al., 2024).

Future Directions

The next steps following this analysis would be to sequence the rest of the samples from both locations. Then similar analyses as described here could be repeated with the full dataset. Additionally, miRNA results could be cross-analyzed against data from collaborators. This includes AMF abundance, AMF species composition, mRNA expression, and mRNA eGWAS. AMF data could be useful in making more direct conclusions on which miRNAs influence AMF colonization. The mRNA expression data could be useful in confirming specific miRNA targets by seeing if increased miRNA expression is correlated with decreased target mRNA expression. Any overlapping mRNA eQTL could also be interesting as they may indicate that miRNA expression is strongly linked to variable target mRNA expression. Overlapping miRNA-mRNA eQTL for genes linked to favorable traits such as increased AMF colonization or favorable AMF species compositions, could become invaluable breeding targets for producing future varieties with ideal miRNA expression patterns.

Endophyte Greenhouse Experiment

Highlights

Chapter 3 sought to (1) determine if endophytic *A. alternata* or *Neopestalotiopsis* could provide fitness advantages to sorghum, (2) determine if potential endophyte benefits, or drawbacks, were influenced by drought, and (3) determine if the benefits or detriments of a given endophyte were genotype specific in a greenhouse setting. *Neopestalotiopsis* was successfully isolated from switchgrass roots and both *Neopestalotiopsis* and *A. alternata* were cultured on plates. During the experiment, drought was determined to show significant variation across the different sorghum genotypes. Unfortunately, while some plant traits showed certain patterns in response to endophyte, likely contamination made it difficult to support significant conclusions.

Limitations

Unfortunately, despite efforts to sterilize soil, PCR analysis and microscopy conducted after the experiment revealed other fungal contaminants. This ultimately makes it difficult to confidently declare a specific effect of *A. alternata* or *Neopestalotiopsis* on different sorghum genotypes under variable drought treatments. Additionally, while there is evidence that *A. alternata* is capable of colonizing sorghum roots (Janet, 1983), there is no direct support that *Neopestalotiopsis* can successfully interact with sorghum roots. Given the fungal contamination, it's difficult to determine if *Neopestalotiopsis* even successfully colonized the sorghum roots. So, while structures

resembling the morphology characteristic of *Neopestalotiopsis* were identified microscopically, colonization can't be determined with 100% certainty.

Future Directions

While this study was ultimately unable to determine significant endophyte effects on sorghum performance, it could still be useful for informing future experimentation. Lessons learned from sterilization and inoculation methods used could be invaluable for repeating a similar, but optimized, experiment in the future.

Overall Conclusions

Both chapters discussed here represent efforts in optimizing sorghum biofuel production on marginal lands by exploring mechanisms that may help mediate harsh environmental conditions. Exploration of miRNA mediated modulation of important genes, such as those linked to AMF compositions and abundances, could be useful in identifying sorghum genotypes with favorable miRNA expression patterns. By better understanding the genetic regulatory pathways linked to traits of interest, future sorghum varieties can be bred to enhance sorghum biomass production on marginal lands. The role of non-mycorrhizal endophytes represents a largely unexplored avenue for enhancing plant performance. Determining which endophytes can provide the best fitness advantages as well as identifying genotypes that are able to maximize endophyte benefits could be another interesting dimension to add to the already ongoing efforts of utilizing

AMF for better biofuel production. Ultimately, optimization of the biofuel sorghum production system on marginal lands could help produce more sustainable aviation fuel.

References

Janet, E. (1983). Fungal colonization of stalks and roots of grain sorghum during the growing season. *Plant Disease*, *417*.

Ko, B. S., Lee, S. B., & Kim, T.-K. (2024). A brief guide to analyzing expression quantitative trait loci. *Molecules and Cells*, *47*(11), 100139.

Ledford, W. C., Silvestri, A., Fiorilli, V., Roth, R., Rubio-Somoza, I., & Lanfranco, L. (2024). A journey into the world of small RNAs in the arbuscular mycorrhizal symbiosis. *New Phytologist*, *242*(4), 1534-1544.