ADVANCING SOYBEAN BREEDING THROUGH DISEASE RESISTANCE, CLIMATE RESILIENCE, AND HIGH-THROUGHPUT PHENOTYPING

by

ETHAN JOHN MENKE

(Under the Direction of Zenglu Li)

ABSTRACT

The development of soybean cultivars with disease resistance and abiotic stress tolerance is an important breeding goal. Disease resistance provided by PI 398469 to Southern Soybean Stem Canker (SSC) caused by *Diaporthe aspalathi* was evaluated in a mapping population. Bulked segregant analysis using Infinium SoySNP50k BeadChips was used to identify putative genomic regions on the soybean chromosomes for SSC resistance. KASP markers were developed for these regions and then used to genotype the recombinant inbred lines (RILs) derived from G81-2057 × PI 398469 and perform quantitative trait locus (QTL) analysis. A single QTL on chromosome (Chr) 14 was identified in the RILs and confirmed in the F2:3 population derived from the same cross between G81-2057 and PI 398469. This QTL was confirmed to be the Rdm3 locus using the RILs derived from a cross of G81-2057 × Crockett and both PI 398469 and Crockett were found to carry resistance alleles at the Rdm3 locus for resistance to SSC.

Slow canopy wilting is a drought tolerance mechanism that is critical for developing climate-resilient cultivars. PI 471938, a slow canopy wilting accession, was crossed with Hutcheson to create a RIL mapping population that was evaluated for canopy wilting under rain-fed conditions in seven environments. Three QTLs on Chrs 2, 8, and 9 were identified for slow

canopy wilting that accounted for 11, 10, and % of the phenotypic variation, respectively. These QTLs could be exploited by breeders to develop new drought-tolerant cultivars.

Unmanned aerial vehicles (UAVs) can allow plant breeders to collect measurements for traits that are not or cannot be collected in large-scale breeding programs. Canopy coverage (CC) has been studied in early maturity group soybean, but this trait has not been evaluated in elite southern soybean germplasm. In this study, UAVs were used to estimate canopy coverage (CC) in elite southern soybean breeding lines from the University of Georgia Soybean Breeding Program's Preliminary and Advance yield trials. CC estimates were collected using a UAV at 4, 5, 6, and 7 weeks post-planting between the 2018 and 2021 growing seasons. CC estimates at 5 to 6 weeks after planting were shown to be most informative in elite southern soybean germplasm. Average CC estimates for the three datasets ranged from 23-32%, 36-49%, 49-55 %, and 43-48% for Weeks 4, 5, 6, and 7, respectively. Heritability of estimated CC ranged from 0.05 in 2018-19 for Week 4 to 0.52 in 2020-21 for Week 7. The relationships of CC with yield, 100-seed weight, maturity, lodging, and plant height were investigated, which could be used by breeders to make informed selections in their breeding programs.

INDEX WORDS: Soybean, *Glycine max*, Single nucleotide polymorphism (SNP), Disease resistance, Southern stem canker, Quantitative trait locus (QTL), *Diaporthe aspalathi*, Drought tolerance, Canopy wilting, Unmanned aerial vehicle (UAV), Canopy coverage, High-throughput phenotyping

ADVANCING SOYBEAN BREEDING THROUGH DISEASE RESISTANCE, CLIMATE RESILIENCE, AND HIGH THROUGHPUT PHENOTYPING

by

ETHAN JOHN MENKE

B.S., Newman University, 2014

M.S., Kansas State University, 2017

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2022

ADVANCING SOYBEAN BREEDING THROUGH DISEASE RESISTANCE, CLIMATE RESILIENCE, AND HIGH THROUGHPUT PHENOTYPING

by

ETHAN JOHN MENKE

Major Professor:	Zenglu Li
Committee:	James Buck
	Josh Clevenger
	Ai-Ping Hu
	Scott Jackson

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2022

DEDICATION

To my parents, Gary and Judy Menke, thank you for supporting me in all that I have done in my life. To my daughter Charlotte, I hope this inspires you to follow your dreams. Finally, to my wife Dr. Jenna Menke, Ph.D., without your support and encouragement, none of this would be possible. You have been with me every step of this journey and I could have had a better companion.

ACKNOWLEDGEMENTS

I am thankful for all the people that I have met and who have helped me to develop into the professional that I am today, and for their help in completing this dissertation research. To Dr. Zenglu Li, thank you for your support as I worked to complete this project. Thank you also for your dedication to helping your students develop the skill necessary to be successful in and outside of the lab and classroom. To my committee members, Dr. Scott Jackson, Dr. James Buck, Dr. Josh Clevenger, and Dr. Ai-Ping Hu thank you gentlemen for your giving your valuable time to provide input on my research. I would like to thank my fellow member of the Soybean Breeding and Genetics Lab for their dedication, insight, and technical support: Brice Wilson, Greg Gokalp, Brian Little, Dale Wood, Earl Baxter, Ricky Zoller, Dr. Jeff Boehm, Dr. Miles Ingwers, and Dr. Rustem Ustun. A Special thank you to Nicole Bachleda, Tatyana Nienow, and Carol Picard, you have provided immense support in my projects and I cannot express how applicative I am. Thank you to the current and former Li lab graduate students: Mark Miller, Sam McDonald, Alexandra Ostezan, Renan Souza, Muhammad "Habib" Widyawan, Nathaniel Burner, Dr. Ivy Tran, Dr. Clint Steketee, Dr. Ben Stewart-Brown, Dr. Zack King, Liz Prenger, and Brooks Arnolds, you were great friends and colleges during my studies. This research was funded by the United Soybean Board, the Glen and Helen Burton Feeding the Hungry Scholarship, and the John Ingle Innovation in Plant Breeding Award. I sincerely apologize to anyone that I failed to recognize- I cannot express the recognition that everyone deserves.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Soybean History, Production, and Improvement

Soybean History

Soybean (*Glycine max* (L.) Merr.) is an oil seed crop that is native to East Asia. It was believed to have been domesticated between 1700-1100 BCE in central China and then spread throughout the rest of China and Korea (Hymowitz & Newell, 1981). It then spread to Japan and other south Asian countries, by sea and land trade routes, in the 1st century CE. By the 1730s soybean found their way to various spots across Europe where they were cultivated in botanical gardens (Hymowitz & Newell, 1981).

Soybeans were introduced into North America by Samuel Bowen, who had been an English seafarer who traveled to China in the late 1750s and was believed to have acquired soybean seeds (Hymowitz & Harlan, 1983). After returning from China, Bowen moved to the Colony of Georgia, and in 1765 asked Henry Young, the Surveyor-General of Georgia, to plant the soybean seeds that he had brought from China (Hymowitz & Harlan, 1983). Bowen then used the crop to produce soy sauce and other products for export back to England, as well as forage for livestock (Hymowitz & Harlan, 1983). By 1804 soybean was being cultivated in Pennsylvania. By 1851 it had moved west into Illinois and then quickly spread throughout the Midwest (Hymowitz, 1987). The first report of soybean being introduced to South America was not seen until 1892 (Shurtleff & Aoyagi, 2004).

Soybean Production

In 2021, US farmers harvested approximately 35 million hectares of soybeans, and with increased demand, this number is projected to rise in 2022 (USDA NASS Acreage Report, 2021). Currently, strong pricing of soybean (~$600/ metric ton) in 2022 leads to increased production to meet current demand. This demand made soybean one of the most economically important row crops grown in the United States of America, worth almost $30 billion in 2020 (SoyStats, 2021). The average yield across the US was 3.4 metric tons per hectare during the 2020 growing season leading to a total production of 112.5 million metric tons of soybean which was a 15% increase in production from 2019 (SoyStats, 2021).

Soybean is the largest oil seed crop in the world, accounting for approximately 30% of the world's vegetable oil production (SoyStats, 2021). After the oil is extracted from the bean the remaining meal is a protein-rich supplement that is mainly used as a feedstock in animal production. Soybean meal accounts for approximately 70% of the total protein meal that is used in the world (SoyStats, 2021). The U.S. is the 2nd largest producer, accounting for 28% of total soybean production, Brazil is the number one producer, accounting for about 38% of world production, and Argentina is the number 3 producer, accounting for about 14% of production (SoyStats, 2021). The U.S. exported approximately 45.8 million metric tons in 2020, making it the second-largest exporter of soybean in the world following Brazil (SoyStats, 2021).

Soybean Physiology

Soybean is a dichotomous annual legume plant. It has an average above-ground vegetative growth between 75-125 cm, containing 14-26 nodes (Lersten & Carlson, 2004). The first nodes contain cotyledons, followed by two opposite, primary leaves on the second node. All remaining

leaves that are produced on the plant's main stem and lateral branches are alternating trifoliate on long petioles (Lersten & Carlson, 2004). Root development of the plant is primarily branched, lateral, secondary roots. Eighty to 90% of the root dry weight is in the top 15 cm from the soil surface, with root growth continuing until the termination of top growth of the soybean (Shibles et al., 1975). Root growth is most favorable at a soil temperature between 22-27° C, with top growth to root ratio being driven more by sunlight than by temperature (Lersten & Carlson, 2004). *Rhizobium japonicum,* which is nitrogen-fixing bacteria, shares a symbiotic relationship with soybean root (Lersten & Carlson, 2004).

Soybean is diploid, containing 20 pairs of chromosomes ($2n = 2x = 40$), and monoecious, meaning the flowers contain both male and female reproductive organs and are self-fertilizing. There is less than a 1% chance of outcrossing to another plant and it is thought that most fertilization occurs before the bud of the plant opens (Carlson & Lersten, 2004). The plants are short-day photoperiod sensitive and flowering is initiated due to changes in daylight (Carlson & Lersten, 2004). Soybean express determinant or indeterminate growth habits, in which determinate lines have a cluster of pods on the terminal raceme, while indeterminate lines do not express this growth (Carlson & Lersten, 2004). Flowering takes place for approximately four to six weeks but is dependent on the growth habits of the line. Indeterminate lines flower sequentially up the stem of the plant, whereas determinate plants flower within a few days of one another on the entire plant. Pod numbers can vary from two to more than 20 pods on a single node, and up to 400 pods on a single plant and pods usually contain between one and five seeds (Kato et al., 1954).

Soybean Breeding

The single most important trait when developing new soybean varieties is the yield potential because this determines the worth of the varieties to farmers. Other agronomically important traits, such as biotic and abiotic stress resistance, are important in maintaining high yields across different environments. For soybean production to meet predicted future demands yield will need to increase by 2.4% annually across the globe (Ray et al., 2012). Between 1989 and 2020, the US has seen on average a ~35 kg per hectare increase in soybean yield, which is approximately a 1% increase in yield per year (SoyStats, 2021). This puts the U.S. behind the needed increase in production to supply the global soybean demand.

Other key agronomic traits in soybean include lodging and shattering resistance and maturity. Soybean plants that fail to maintain an erect growth habit during the growing season as described as being lodged. Lodging causes significant yield reduction when lodging occurred at or after the R5 growth stage (Fehr & Caviness, 1977; Woods & Swearingin, 1977). Shattering is when the soybean seeds do not remain in the pod until harvesting can occur and is the largest source of pre-harvest yield loss in soybean (Philbrook & Oplinger, 1989). Shattering is greatly affected by weather, with early maturing lines experiencing less shattering than later-maturing lines (Philbrook & Oplinger, 1989). Breeding efforts have selected lines with resistance to shattering and in most elite cultivars it is only observed in cases where harvesting is delayed several weeks after maturity.

There are currently 13 different maturity groups in North America (MG 000 to X) based on the latitudes, at which lines can optimally grow. Soybean flowering is controlled by photoperiod and temperature (Major et al., 1975) and it will begin flowering under reduced light exposure. The cool temperature later in the growing season can also delay flowering (Carlson & Lersten, 2004).

Earlier maturity soybeans are bred to flower under longer daylight conditions that are experienced in the northern growing regions, and moving soybeans out of the region of adaptation will greatly reduce the ability to produce seeds (Heatherly & Elmore, 2004).

Soybean Southern Stem Canker (*Diaporthe aspalathi*)

Disease

Soybean stem canker is caused by the fungal complex *Diaporthe phaseolorum* (Morgan-Jones, 1989): Northern stem canker (NSC) is caused by *D. phaseolorum f. sp. caulivora,* and Southern stem canker (SSC) is caused by *D. aspalathi (syn. Diaporthe phaseolorum var. meridionalis*). The northern and southern designation for stem canker was derived based on the geographic regions in the U.S. where they were observed (Rupe, 2015). Soybean stem canker causes significant yield loss in most soybean-producing regions in North and South America. From 1996 to 1998 in the southern U.S., stem canker caused yield losses of 12,000 to 18,000 metric tons (Wrather et al., 2001). Between 2006 and 2009, both northern and southern soybean stem canker caused yield losses of 130,000 to 180,000 metric tons (Koenning & Wrather, 2010).

Taxonomy

The *Diaporthe*/*Phomopsis* disease complex is a wide-ranging plant pathogen and is responsible for stem canker, seed decay, and stem blight, which cause significant economic losses in soybeans (Santos et al., 2011). The lifecycle of this fungus is a complex containing both a teleomorph, sexually reproducing *Diaporthe phaseolorum,* and anamorph *Phomopsis phaseoli,* which is asexual in its reproduction (Morgan-Jones, 1989). The classification of the fungi causing stem canker has changed since the first observation of the disease. Welch and Gilman (1948)

differentiated the pathogen that causes stem canker from *Diaporthe phaseolorum* var. *soja* (Lehaman) Wehm, which is responsible for pod and stem blight. However, the researchers falsely identified the pathogen *Diaporthe phaseolorum var. batatatis* (Harter &Fields) Wehm as the causative agent of stem canker (Welch & Gilman, 1948). Athow and Caldwell (1954) correctly identified the *Diaporthe phaseolorum* (Cke.,&Ell.) Sacc. var. *caulivora* Athow & Caldwell as the causative agent of soybean stem canker from *Diaporthe phaseolorum var. soja.* The stem canker pathogen *D. phaseolorum* var. *caulivora* was differentiated from *D. phaseolorum* var. *soja* by the lack of an associated anamorph, the clusters and shape of its perithecia, and smaller asci and ascospores (Athow & Caldwell, 1954).

By the late 1950 and 60's the differentiation of two separate species of *D. phaseolorum* was in question. Threinen et al. (1959) used radiation to create mutant lines of *D. phaseolorum* var. *soja* and *D. phaseolorum* var. *caulivora*. Mutants of *D. phaseolorum* var. *caulivora* were shown to cause symptoms of stem canker, while other mutants caused symptoms of pod and stem bight consistent with *D. phaseolorum* var. *soja* (Threinen et al., 1959). Whitehead (1966) observed *D. phaseolorum* var. *soja* on Birds-foot trefoil (*Lotus corniculatus*) and showed that these strains caused symptoms similar to stem canker on soybean. These studies invalidated the claim that *D. caulivora* and *D. soja* are separate species in the *D. phaseolorum* complex (Kulik, 1984). But strains of the fungi that cause stem canker are easily identifiable from other *D. phaseolorum* strains (Kmetz et al., 1978). Kulik (1984) proposed that the varieties be labeled as *formae speciales* on their abilities to induce symptoms of stem canker on soybean. Backman et al. (1985) showed that northern and southern biotypes of the *D. phaseolorum* differed enough to justify the introduction of *formae speciales* for the separate geographical regions, based on heat tolerance, morphology, and epidemiology. Biotypes differences lead to the creation of two *formae speciales*, with northern

stem canker retaining the name *D. phaseolorum* f. sp. *caulivoria,* while southern stem canker was named *D. phaseolorum* f. sp. *meridionalis* (Morgan-Jones, 1989).

The development of DNA sequencing technologies has allowed the differentiation of the causal agent of stem canker by the use of EFI1-α and ITS regions (Fernández and Hanlin, 1996; Lu et al., 2010). This is a more efficient and precise way to identify species in the *Diaporthe/Phomopsis* complex, compared to host association and the morphological differences in the pathogen as was done in prior studies (Santos et al., 2011). Using PCR-RFLP, Zang et al. (1998) indicated that northern and southern stem canker were varieties of *D. phaseolorum* and should be classified as *D. phaseolorm* var. *caulivoria* and *D. phaseolorm* var. *meridionalis*, respectively. It was reported that *D. phaseolorm* var. *meridionalis* was not a variety of *D. phaseolorum*, and thus proposed renaming the pathogen *Diaporthe aspalathi* (Janse Van Rensburg et al., 2006). Similarly, Santo et al. (2011) concluded that *D. phaseolorm* var. *caulivoria* was also not a variety *of D. phaseolorm* but was a separate species of *Diaporthe caulivoria.*

Diaporthe aspalathi is an ascomycetous fungus sexually through the perithecia (Zhang et al., 1998). On acidified potato dextrose agar, *D. aspalathi* produced tufted mycelium and white colonies that became tan with age, stroma were infrequent, alpha conidia were produced in pycnidia after 25-30 days and solitary perithecia with long beaks were observed at 35 days (Pioli et al., 2003).

It should be noted that both the northern and southern stem canker has been referred to as *D. phaseolorum* var. *caulivoria* in many previous publications before the differentiation of formal species. In this review, we assume that when describing southern stem canker produced by southern isolates, it will refer to *Diaporthe aspalathi* (syn. *Diaporthe phaseolorum var. meridionalis*). In prior research, the resistance gene symbol *Rdc* was used. Similar to the pathogen

name changes, after the creation of two the *formae speciales*, a resistance gene symbol for southern stem canker was renamed to *Rdm*. This review will assume that *Rdm* denotes the resistance genes to *D. aspalathi* when describing experiments where southern isolates were used for disease resistance screening.

Disease History

Northern stem canker was first observed in the late 1940s in the corn and soybean belt of the U.S. and was caused by the fungus *Diaporthe caulivora* (syn. *Diaporthe phaseolorum var. caulivora*). Northern stem canker was a major disease in the Midwest in the late 1940s and 50s due to the highly susceptible cultivars, mainly Blackhawk and Hawkeye, being used in production (Hildebrand, 1952). Once these cultivars were removed from production the disease incidence decreased and outbreaks became sporadic causing limited yield loss (Weaver et al., 1984; Backman et al., 1985).

Southern stem canker was first observed in the southern U.S. in Mississippi in 1973, and spread to Alabama in 1977, Tennessee in 1981, Georgia and South Carolina in 1982, followed by Florida, Louisiana, Arkansas, and Texas in 1984 (Krauz & Fortnum, 1983; Backman et al., 1985). During this time the disease caused a loss of up to 80% in susceptible fields (Backman et al., 1985). This disease caused economic losses of $37 million ($102 million in 2022) to producers in 1983 (Backman et al., 1985). Southern stem canker was identified in Brazil in 1989, Paraguay in 1992 (Sato et al., 1993), and Argentina in 1993 (Pioli et al., 1997). It has also been observed in the northern states of the U.S. during the 1999 and 2000 growing seasons in Illinois (Gravert et al., 2001), and Wisconsin from 2000 to 2003 (Li et al., 2004).

Like the case of northern stem canker, many of the popular varieties in production were susceptible to SSC when it emerged as a major problem in the Southern US (Backman et al., 1985). SSC is still a concern for producers in the southern US, but continued breeding efforts have developed many resistant cultivars and crop management practices have helped reduce major disease losses that were observed in the 1980s.

Disease Symptoms

Soybean stem canker derives its name from the lesions, or cankers, that are produced by causal organisms. *D. aspalathi* infects the soybean plants by entering through wounds on the stems, petioles, and petioles bases (Plotez & Shokes, 1987). Infection can occur through the leaves, but this leading to the development of stem canker is rare. The pathogen is introduced by the splashing of soil particles containing the pathogen onto the plant by rainfall or irrigation (Rupe, 2015).

Disease symptoms of southern stem canker begin during the early reproductive growth stage (R1) of the plant. The disease remains dormant for most of the vegetative growth of the plant (V1-V7) even though it has infected the plant during this period (Fehr & Caviness, 1977; Rupe, 2015). Cankers are initially small and slightly sunken with red-brown color with a greyish brown center and they first appear on a lower node on the plant (Rupe, 2015). As the disease progresses the canker size increases longitudinally and becomes more sunken, with the edges of the canker maintaining the reddish-brown color, while the center stays greyish brown color (Rupe, 2015). The disease causes the pith and vascular tissue to turn brown and become discolored (Rupe, 2015). The stem around the canker will remain green (Grau et al., 2004). Southern stem canker rarely girdles the stem, which helps identify the disease from Phytophthora root and stem rot (Rupe, 2015). Multiple infections on the same stem are possible with southern stem canker, which is not

generally the case with northern stem canker. The stem tissue that is killed by the pathogen will eventually block the upward flow of water to the plant, causing water stress and eventually the premature death of the plant (Hildebrand, 1952).

Foliar symptoms appear most commonly as interveinal chlorosis and necrosis, but tip dieback can occur. Eventually, leaf death will occur and the leaves will remain attached to the stem until harvest. The foliar symptoms that are caused by southern stem canker are very similar to other foliar diseases such as soybean sudden death syndrome, brown stem rot, and red crown rot, which can make identifying the disease through visual evaluation of the leaves difficult (Rupe, 2015). The foliar symptoms are observed as the result of phytotoxins produced by *D. aspalathi* (Lalitha et al., 1989). Lalitha et al. (1989) isolated phytotoxins from four different isolates of *D. aspalathi,* each having a different level of phytotoxin production. The increased production of these phytotoxins was correlated to increase canker size and disease symptoms. These phytotoxins caused foliar leaf symptoms on excised soybean trifoliates (Lalitha et al., 1989).

Disease Management

Predicting if and when southern stem canker will occur is a major challenge for producers. For this reason, producers should focus on the prevention of the disease. Rothrock et al. (1985) reported that the mean disease incidence could go from 1.6% of the field showing symptoms of southern stem canker in 1983 to over 74% of the field showing symptoms in 1984. This study demonstrated the ability of *D. aspalathi* to go from a negligible problem to an epidemic in a single growing season (Rothrock et al., 1985).

Crop debris of infected plants is the primary source of infection between fields. Farmers should avoid the spread of the disease between infected and non-infected fields. This can be done

through the cleaning of tractors, tillage equipment, and planters between fields, and working in non-infected fields first (Backman et al., 1985). Infected soybean should not be used as a seed source and this is a more important fact for seed producers than farmers (Backman et al., 1985). Another method to limit southern stem canker is to reduce the level of inoculum that accumulates in the field. Tyler et al. (1983) reported that infection of southern stem canker was greater in the no-till plot as compared to any of the other tillage practices. Tillage is thought to limit the severity of the southern stem canker from one year to the next by the reduction of soybean residue that is left in the field to cause infestation next season (Tyler et al., 1983; Backman et al., 1985). The author observed that Brown Spot (*Septoria glycines*) in the tilled plots of the experiment and these plots had a large amount of leaf loss on lower stems. The authors speculated that this lower leaf loss may have caused some of the difference in disease severity between tillage methods. No-till plots had more lower leaves where *D. aspalathi* could collect and infect the plants (Tyler et al., 1983). Rothrock et al. (1985) indicated that conventional tillage reduced the disease severity in soybean fields compared to no-till fields but conventional till fields yielded significantly lower than no-till fields. Fretais et al. (2002) showed that no-till plots had reduced southern stem canker symptoms when compared to minimum-till plots, as well as increased yields. The authors proposed that minimum tillage was not effective at completely reducing the inoculum in the field. This lack of complete disease control from minimum tillage combined with heavy rainfall caused the disease to be more prevalent (Freitas et al., 2002). The results from this study contradicted those of Rothrock et al. (1985) where tillage reduced the disease severity. Both Fretais et al. (2002) and Tyler et al. (1983) observed that additional environmental factors such as soil type, weather events, and additional disease pressure could lead to the difference in the effects of tillage on the control of SSC.

The planting date of the infected field should be delayed to reduce seedling exposure to the initial release of inoculum in May and mid-June (Backman et al., 1985). This limits the total number of plants that are infected, as well as decreases the growing season in which the disease infects the plants causing less yield loss (Backman et al., 1985). Multiple studies have shown that infections of soybean plants during the early vegetative stage between V3 and V7 had the greatest impact on the development of southern stem canker disease symptoms (Smith & Backman, 1989; Rupe et al., 1999).

Many environmental conditions affect the development and severity such as temperature, rainfall and irrigation, and inoculum density. Levels of inoculum are directly related to the incidence of southern stem canker (Ploetz & Shokes, 1985, 1987) with conidia and ascospores both able to cause southern stem canker symptoms (Ploetz & Shokes, 1985). Ploetz and Shokes (1987) showed that southern stem canker infection occurs most frequently between 28 and 34 °C and it also can cause infection between 10 and 22°C. However, at 40°C and above the pathogen was unable to cause infection. Diamicone et al. (1987) indicated that free moisture increases the disease incidence and dead plants from inoculation in greenhouse studies over plants that were inoculated and kept in a dry environment (Damicone et al., 1987). Subbarao et al. (1992) reported similar results that plants that were grown under irrigation had higher levels of southern stem canker as compared to plots grown without irrigation. These symptoms were even consistent across resistant and susceptible cultivars (Subbarao et al., 1992). Under irrigated conditions, *D. aspalathi* was able to reproduce and have a second cycle of inoculation in a season (Subbarao et al., 1992). The symptoms of this second inoculation with *D. aspalathi* were irregular due to the long incubation period of the pathogen (Subbarao et al., 1992). These reactions to temperature and moisture might also influence why delayed planting is beneficial to help prevent southern stem

canker. In most locations in the southern U.S., June is a warmer and drier climate than May, making it less conducive for the development of southern stem canker (Smith & Backman, 1989; Rupe et al., 1996).

Backman et al. (1985) stated that both pre and post-emergence fungicides were effective in limiting the spread of the disease. For effective application of these fungicides, the application should occur during vegetative growth, this makes application decisions difficult due to visible disease symptoms not occurring until later in the season (Backman et al., 1985).

Several weed and crop species have been identified as alternative hosts that create inoculum that can infect soybeans in later seasons (Roy & Miller, 1983; Black et al., 1996). Roy and Miller (1983) isolated *D. aspalathi* isolates for cotton (*Gossypium hirsutum*) that were able to infect soybean and produce stem canker symptoms on soybeans. Seventeen weed species were tested for their compatibility as alternative hosts to *D. aspalathi*, allowing the fungi to survive, reproduce and infect soybean (Black et al., 1996). Fourteen of 17 weeds, black nightshades (*Solanum nigrum*), entire leaf morning-glory (*Ipomoea hederacea var. interfriudcula*), hairy indigo (*Indigofera hirsute*), hemp sesbania (*Sesbania exaltata*), ivy-leaf morning-glory (*Ipomoea hederacea*), northern joint-vetch (*Aeschynomene virgnica*), pitted morning-glory (*Ipomoea lacunose*), prickly sida (*Sida spinose*), redweed (*Melochia corchorifolia*), sicklepod (*Cassia obtisifolia*), smallflower morning-glory (*Jacquemontia tamnifolia*), spiny amaranth (*Amaranthus spinosus*), tall mooring-glory (*Ipomoea purpurea*), and wild poinsettia (*Euphorbia heterophylla*) were shown to be alternative hosts to *D. aspalathi*, while barnyard grass (*Echinochloa crus-galli*), Johnsongrass (*Sorghum lacunosa*) and curly dock (*Rumex crispus*) did not serve as alternative host to the fungi (Black et al., 1996). All weed species tested were asymptomatic, except for hemp sesbania and hairy indigo which developed stem lesions similar to symptoms of stem canker in

soybean (Black et al., 1996). Lalitha et al. (1989) tested 12 plant species for the reactions to *D. aspalathi* and only observed reactions on soybean and lima bean (*Phaseolus lunatus*), making lima bean an alternative host to *D. aspalathi*.

Crop rotation between soybean and non-susceptible hosts such as corn or grain sorghum is recommended to reduce inoculum build-up (Roy & Miller, 1983). With weed species serving as alternative hosts and an inoculant source for *D. aspalathi*, herbicide applications or tillage are key factors in preventing southern stem canker (Black et al., 1996).

By far the most effective and useful method of prevention of southern stem canker has been through the development of resistant varieties (Backman et al., 1985). During the late 1970 and 80's most of the soybean cultivars that were in production in the southern U.S. were susceptible to *D. aspalathi* (Backman et al., 1985). Weaver et al. (1984) screened 41 adapted cultivars in maturity groups V-VIII for resistance to southern stem canker, and observed a wide range of reactions to the disease, suggesting that resistance was controlled by multiple resistance sources. Both authors advocated for the development of resistant cultivars by removal of susceptible cultivars from breeding programs and production (Weaver et al., 1984; Backman et al., 1985). For this reason, the Georgia soybean production handbook advocates for the selection of resistant cultivars as the primary method of controlling southern stem canker (Bryant, 2021).

Disease Resistance

Currently, five major resistance genes, *Rdm1*, *Rdm2*, *Rdm3*, *Rdm4,* and *Rdm5* have been named or identified from different sources to prevent southern stem canker (Kilen et al., 1985; Kilen and Hartwig, 1987; Bowers et al., 1993; Tyler, 1996; Chiesa et al., 2009). However, these *Rdm* genes do not provide equal resistance to all strains of *D. aspalathi* (Keeling, 1982). Keeling

(1982) reported that some lines with resistance to northern soybean stem canker didn't have resistance to southern stem canker. Ploetz and Shokes (1987) demonstrated that plant resistance to southern stem canker was due to host-pathogen interaction, with resistant plants having the same frequency of infections as susceptible plants. Keeling (1982) first reported that the cultivars Tracy and Centennial had no natural infections of *D. aspalathi* in 1979 and 1980. Both cultivars out-yielded the susceptible cultivar J77-339 when all three cultivars were grown under disease pressure (Keeling, 1982). For the resistance in both Tracy and Centennial, the source of resistance to *D. aspalathi* was proposed to be the cultivar CNS which is their common soybean ancestor (Keeling, 1982). In that study, Keeling (1982) also developed a greenhouse inoculation protocol to evaluate lines for stem canker resistance using *D. aspalathi* infested toothpicks.

Tracy-M, a reselection for metribuzin tolerance out of Tracy (resistant to SSC), was crossed with J77-339 (susceptible to SSC) to create a segregating population to determine the inheritance of resistance of the *Rdm* gene (s) in Tray-M (Kilen et al., 1985). Seedlings were inoculated using infected toothpicks in the greenhouse and rated 30, 35, and 40 days after inoculation. Plants at day 40 that failed to develop symptoms were determined to be resistant to SSC (Kilen et al., 1985). Forty days after inoculation, all J77-339 plants had died, while all Tracy-M and F_1 plants were healthy and showed no disease symptoms and this demonstrated that the *Rdm* gene was dominantly inherited (Kilen et al., 1985). F_3 plants segregated at a ratio of 7 resistant, 8 segregating, and 1 susceptible, which indicated that two dominant genes were present in Tracy-M that were conferring resistance to *D. aspalathi* (Kilen et al., 1985). Keeling and Hartwig (1987) took the F_3 plants that were identified in Kilen et al. (1985) with a segregation ratio equal to 3:1 for resistance and advanced them to F_4 generation to select uniformly disease-free F_4 lines. Two experimental lines X404 and X412 were chosen and then backcrossed to J77-339, respectively, which produced

BC_1F_2 populations. Both populations segregated in a 3:1 ratio for resistance that indicated a single gene conferring resistance in each line of two lines (Kilen & Hartwig, 1987). Both X404 and X412 were then crossed to produce F_2 progeny that segregated in a 15:1 ratio, indicating two genes for resistance being present (Kilen & Hartwig, 1987). This confirmed that two independent genes (*Rdm1* and *Rdm2*) were present in Tracy-M that were responsible for providing resistance *to D. aspalathi* (Kilen et al., 1985; Kilen & Hartwig, 1987).

Two soybean cultivars Crockett and Dowling were observed to have resistance to southern stem canker (Bowers, 1990; Bowers et al., 1993). To determine if these were new resistance sources, Dowling, Crockett, and Tracy-M (*Rdm1* and *Rdm2*) were crossed to two susceptible parents Coker 338 and Johnson, respectively, and resulting progeny were screened for resistance using the toothpick inoculation method (Keeling, 1985; Bowers et al., 1993). F_2 populations derived from either Dowling or Crockett segregated in a 3:1 ratio (resistant to susceptible), indicating that a single dominant gene conferred the resistance (Bowers et al., 1993). However, the populations derived from Tracy-M, segregated in a 15:1 ratio for resistance, which confirmed the previous report that it possesses two independent genes *Rdm1* and *Rdm2* for resistance (Kilen & Hartwig, 1987; Bowers et al., 1993). To determine if Dowling and Crockett carried the same resistance gene, the two cultivars were crossed and the F_2 progeny segregated in a 15:1 ratio for resistance, indicating that the two cultivars carried different resistance genes (Bowers et al., 1993). Crockett and Dowling were then each crossed to Tracy-M to form two segregating populations, the resulting $F_{2:3}$ populations were each observed to be segregating for resistance in a ratio of 37 resistant, 26 segregating, and 1 susceptible (Bowers et al., 1993). It indicated that both cultivars carried different genes that confer resistance to *D. aspalathi* as compared to Tracy-M. The authors proposed that these loci be named *Rdm3* for Crockett and *Rdm4* for Dowling (Bowers et al., 1993)

Soybean cultivar Hutcheson was released in 1988, and due to its high yields and disease resistance, it was widely used as a parent to many of the southern soybean lines that were developed in the late 1980s and early 1990s (Tyler, 1996). It was reported that Hutcheson is resistant to southern stem canker, but its source of resistance was unknown (Tyler, 1996). To determine the resistance source Tyler (1996) crossed Hutcheson to two susceptible parents Hartwig and J77-339, with F_2 progeny from both crosses segregating in a 3:1 ratio, indicating a single dominant gene responsible for resistance. Hutcheson was then crossed to each of the known resistance sources D85-10404 (*Rdm1*), D85-10412 (*Rdm2*), Crockett (*Rdm3*), and Dowling (*Rdm4*) (Tyler, 1996). In three crosses between D85-10404 (*Rdm1*), D85-10412 (*Rdm2*), and Crockett (*Rdm3*) and Hutcheson, a segregation ratio of 15:1 were observed, indicating that Hutcheson has a different source of resistance than *Rdm1*, *Rdm2*, and *Rdm3*. The cross between Hutcheson and Dowling (*Rdm4*) produced no susceptible lines, which indicated that the resistance in Hutcheson is allelic to *Rdm4* (Tyler, 1996). The author proposed that the resistance gene *Rdm4* in Hutcheson was inherited from the cultivars York that was a grandparent of Hutcheson (Tyler, 1996).

Chiesa et al. (2009) crossed Hutcheson to J77-339 and developed $F_{2:3}$ families, these families were then inoculated with two different isolates of *D. aspalathi*. The authors indicated that Hutcheson carries two independent genes that control resistance and these two genes are not tightly linked, a recombination fraction of 29%, but are close enough to segregate together in most crosses. The authors proposed that these two genes confer race-specific resistance to the two-different isolated of *D. aspalathi* (Chiesa et al., 2009). In addition, Chiesa et al (2013) confirmed two independent dominant resistance genes in Hutcheson and observed isolate specific interactions in the progeny of Hutcheson with susceptible cultivars J77-339 and RA702. The authors also

showed that *Rdm2* possessed incomplete penetrance, where lines with the resistance allele failed to provide SSC resistance when interacting with specific isolates of *D. aspalathi* (Chiesa et al., 2013). Chiesa et al. (2017) mapped the *Rdm4* and *Rdm5* in the Hutcheson × J77-339 population developed by Chiesa et al. (2009) to a 44 cM region on Chr 8 between Satt_162 and Satt_233. *Rdm4* and *Rdm5* were located approximately 17.6 cM apart, which was constant with the genetic distance that was previously reported by Chiesa et al. (2009).

Understanding the genomic locations of the genes for resistance to southern stem canker can aid breeders to determine the heritability of resistance and improve their ability to make selections. Shearin (2011) mapped resistance in the crosses between J77-339 (susceptible) and D85-10412 (*Rdm2*), Crockett (*Rdm3*), Dowling (*Rdm4*), and PI398469 (*Rdm?*) using $F_{2:3}$ populations and SSR markers. All populations segregated in a 3:1 ratio for resistance. *Rdm2* and *Rdm4* from D85-10412 and Dowling, respectively, were both located on the linkage group (LG) D1b (Chr 2) in different chromosomal regions. *Rdm3* from Crockett and *Rdm?* from PI 398469 were mapped to a similar genomic region on LG-B2 (Chr 14) and were clustered near a region with several known disease resistance motifs (Shearin, 2011).

Currently, genome-wide association studies have also been used to understand the resistance to southern stem canker and to identify the novel loci for resistance. Using 112 accessions with SSC phenotypes available in the USDA Soybean Germplasm Collection, Chang et al. (2016) identified a significant SNP, ss715617951, for resistance to southern stem canker on Chr 14. ss715617951 was located approximately 200 Kb from the significant SNP, ss715617869, that was identified as providing resistance to northern stem canker (Chang et al., 2016). The authors suggested that the same resistance locus might be responsible for resistance to both diseases.

Maldonado Dos Santos et al., (2019) performed a GWAS using a panel of 295 soybean accessions from a variety of maturity groups and different regions of origin, including the known SSC resistance sources. These accessions were phenotyped using the toothpick inoculation method and disease ratings were taken 60 days after inoculation. The % of dead plants was calculated by the equation [$\%DP = ((DP + \text{IP}/2)/\text{TP})*100$], where DP is the number of dead plants, IP is the number of infected plants, and TP is the total number of inoculated plants. A genotype by sequencing approach was used to generate genotypic data on the panel, with about 30,000 SNPs being identified after data processing and filtering. The genome-wide association study identified a genomic region on Chr 14 that was significantly associated with resistance to southern stem canker, with the peak SNP located at 1,744,370 bp (Maldonado Dos Santos et al., 2019). The authors then developed KASP markers based on the most significant SNPs identified in the GWAS and used them to screen a subset of the GWAS panel. They found that the KASP marker designed off the SNP at 1,744,370 bp on Chr 14 was able to separate resistant from susceptible lines with an accuracy of 98%. This study found that Hutcheson had the same haplotype for resistance that was seen in PI 398469, and proposed that Hutcheson possesses a resistance gene on Chr 14, which conflicts with the reports of Chiesa et al. (2017) that *Rdm4* and *Rdm*5 were located on Chr 8. The authors stated that this difference in mapping could be due to different isolates of *D. aspalathi* being used in each experiment (Maldonado Dos Santos et al., 2019).

High Throughput Phenotyping and Canopy Closure in Soybean

Worldwide the rate of genetic gain for crop yield is currently 1.3% which is insufficient to meet the predicted needs which is the gain of 2.4% in crop yield a year (Ray et al., 2012). In the last 80 years breeding for high-yielding soybean has led to an increase of 29 kg ha^{-1} per year in

maturity group II, III, and IV soybeans (Rincker et al., 2014). For soybean cultivars in MG V, VI, VII released between 1928-2008 a 13.7 kg ha^{-1} yr^{-1} was observed, this rate of gain is lower than earlier maturity group soybean but is in line with previous reports (Boehm et al., 2019). Specht et al. (2015) looked at how the genetic gains were partitioned between improved cultivars and better agronomic practices. The authors indicated that approximately two-thirds of the yield gain that has been observed on farms has been through the continual development of new high-yielding cultivars and their adoption by farmers. The remaining one-third of the improvement in yield was due to better agronomic practices such as row spacing, herbicide use, along with a synergistic effect between new cultivars and better environmental practices and the adoption of earlier maturity groups (Specht et al., 2014)

High-throughput Genotyping

Plant breeders have made great advances in the last 25 years in the efficiency and rate of genetic gain that they can accomplish using genetic and genomic data. These advances continue to become cheaper and faster using new technologies such as next-generation sequencing and genomic selection (White et al., 2012). The reduction in sequencing cost has allowed for the development of whole-genome sequences for soybean and many other crops and the sequence in the case of soybean will allow researchers to identify and characterize the genetic basis of many soybean traits (Schmutz et al., 2010). In addition, this reduction in the cost of genotyping has allowed the development of techniques such as genomic section by breeders to predict the performance of lines in breeding populations and increase the rate of genetic gain that plant breeders can achieve in a single cycle (Poland & Rife, 2012). This knowledge can then be used by breeders in the faster development of new cultivars to meet future soybean demands.

High-throughput Phenotyping

These new genetic technologies have created a bottleneck in the ability of researchers to generate high-quality phenotypic data (Furbank & Tester, 2011). This is due to the number of lines in these genomic studies as well as the need for skilled personnel to collect data at key physiological times during the growing seasons. This causes traditional phenotyping to be slow and costly. For this reason, many researchers have been looking at increasing the efficiency of phenomics, the study of phenotyping, by using new methods to increase the precision and throughput of traditional phenotyping methods (Tardieu et al., 2017). High throughput phenotyping research has looked at many different levels of phenotyping, ranging from automatic measurements taken in greenhouses to field level measurements (Araus & Cairns, 2014). Field-based high-throughput phenotyping will allow breeders to collect measurements on plant health, disease resistance, abiotic stress, plant characteristics, and architecture such as height, and potentially predict yield and other key agronomic traits noninvasively with less labor (Furbank & Tester, 2011; Araus & Cairns, 2014). It will enable breeders to perform selection more efficiently by allowing them to observe and evaluate more lines.

Most high-throughput phenotyping methods rely on some type of remote sensing, which usually is to use photography or digital imaging to collect data on plants without physically touching them. These images can be collected from standard RGB cameras, hyperspectral radiometers, or thermal cameras. Both RGB pictures and hyperspectral measurements have been used to approximate leaf chlorophyll content (Chappelle et al., 1992; Gitelson et al., 2005). Depending on the traits that researchers wish to estimate, vegetative indices can derive from values acquired from images or measurements. One of the most commonly used indices is the normalized vegetative index (NDVI) which was used to estimate green biomass (Deering, 1978; Tucker et al.,

1979) and is calculated by the relationship between the reflectance of visible red and near-infrared light (NIR), $NDVI = \frac{(Red-NIR)}{(Red=NIR)}$.

Researchers have evaluated many different methods for accurate and quick data collection (Araus and Cairns, 2014). These methods have involved the use of manned and unmanned ground-based or aerial vehicles or handheld sensors operated by researchers on the plot level (Araus & Cairns, 2014). For most of these phenotyping platforms, such as handheld and ground-based vehicles, implementation in a large-scale plant breeding program is limited by the inability of these systems to cover a large number of plots or multiple locations in breeding programs (Makanza et al., 2018). Due to these limitations, much of the recent research has focused on the use of unmanned aerial vehicles (UAV) as a platform for the collection of phenotypic data (Xie & Yang, 2020). UAVs can produce high-resolution data on the plot level and cover large-scale breeding trials at a lower cost than other collection methods such as satellites and manned aircraft (Araus & Cairns, 2014; Shi et al., 2016; Makanza et al., 2018; di Gennaro et al., 2018). High-throughput phenotyping has been used in many crops to evaluate the traits such as yield estimation, disease resistance, abiotic stress tolerance, and agronomic and physiological traits (Araus & Cairns, 2014; Shi et al., 2016; Makanza et al., 2018; di Gennaro et al., 2018)

High-throughput Phenotyping in Breeding

Ma et al. (2001) showed that soybean spectral reflectance at the R4 to R5 growth stage could be used to predict high yielding from low yielding soybean varieties using the normalized vegetative index (NDVI). Using UAVs to collect canopy area measurements on approximately 3000 soybean lines, Yu et al. (2016) found a significant correlation between the canopy area of a plot and the seed yield (r =0.56). However, with yield being a highly quantitative trait and

influenced by environmental variation, using high-throughput phenotyping to estimate is difficult to reproduce accurately. Yu et al. (2016) also examined the use of drones for estimating the maturity of soybean plants as opposed to visually evaluating maturity. The estimation model was able to predict 63% of ratings being an exact match with visually maturity evaluations. The authors stated that the lack of accuracy was due to maturity being traditionally evaluated based on the maturity of the pod, not the leaves. Therefore, the model would not be able to correctly estimate the maturity of the plants which still have foliar material on the plants. If a sample difference of plus or minus a day was allowed, the prediction accuracy increased to 95% with visual evaluation, which is sufficient accuracy for most breeding programs (Yu et al., 2016).

Nutter et al. (2002) used three different collection methods satellite, aerial, and ground-based imaging to collect phenotypic data on soybean plants under soybean cyst nematode (SCN) pressure. The authors showed that with an increase in the initial SCN population, a decrease in reflectance at 810 nm was observed, whereas yield increased with increased reflectance at 810nm (Nutter et al., 2002). This is consistent with the previous research that decreased reflectance near-infrared region is a sign of plant stress and overall health (Filella & Penuelas, 1994; Peñuelas et al., 1997). This study also indicated that all three methods were similarly effective to estimate the effects of the SCN population on yield. Bajwa et al. (2017) used handheld reflectance measurements to estimate the resistance of soybean cultivars for soybean sudden death syndrome (SDS) and SCN. The authors were able to characterize plants using vegetative indices as healthy approximately 95% of the time, and diseased 58% of the time (Bajwa et al., 2017). However, the prediction model could neither identify the difference between slightly diseased and healthy plants nor distinguish between the SCN and SDS (Bajwa et al., 2017).

Makanza et al. (2018) used aerial digital imaging to estimate canopy coverage and senescence, or maturity, in maize breeding trials. The current method of collection for these traits is through visual evaluation of individual plots by skilled researchers, which is time-consuming and can be subjective based on the evaluators' judgment (Makanza et al., 2018). To create a more accurate estimate of senescence and canopy coverage with less physical labor, the researchers used RGB images from a UAV platform. The authors used some automated extraction processes to increase the speed at which images were analyzed (Makanza et al., 2018). These processes increased the speed and number of lines that were phenotyped, but they did require precise GPS data to allow for automated extraction by ArcGIS. From each plot, three measurement values were extracted from the UAV images and the pixels in the image were classified as soil cover, yellow, green, or dry (senesced or brown leaves) canopy. Once the image pixels were classified, plot color classifications were exported as numeric values that then were used to estimate canopy coverage and senescence in these breeding populations. Canopy coverage was estimated by using green cover and a ratio of green cover divided by the plot area; both measurements had a significant genetic correlation to grain yield. Using these image classification measurements, the authors formed a senescence index, with which they compared to ground-truth data (Makanza et al., 2018). No consistent relationship between aerial and visual evaluations was found and the authors suspected that was due to differences between researchers' visual evaluations (Makanza et al., 2018).

di Gennaro et al. (2017) phenotyped barley for vigor using both visible and near-infrared vegetative indices from a UAV. In the meantime, the authors also used ground-based sensors to collect images to compare the effectiveness and accuracy of aerial data acquisition (di Gennaro et al., 2018). This experiment used similar methods as Makanza et al. (2018) to process and extract

the drone images and to extract the data from GPS referenced images using computer programs. In this study, vegetative indices created from UAV measurements were significantly correlated with those collected with ground-based sensors (di Gennaro et al., 2018). Compared to ground-based sensors, UAVs tended to have less power in detecting vigor in the barley genotypes. The major benefit of the use of UAVs can be seen in the breakdown of time to collect these measurements. With a UAV the experiment can be evaluated in a single day, whereas to evaluate by hand it would take researchers three days to complete the same task (di Gennaro et al., 2018). The reduction in time would allow for greater numbers of genotypes to be evaluated and selected (di Gennaro et al., 2018).

Canopy Closure in Soybeans

Increased early crop growth rate, or canopy closure, has shown to have a beneficial impact on yield in soybeans (Bullock et al., 1998). In most cases, this canopy closure or crop growth rate has been evaluated in terms of how row spacing affects growth rate and subsequently grain yield. In the Midwest U.S., it has been shown that narrow row width of 76 cm or less increased the yield of soybeans (Beatty et al., 1982). However, in the southern U.S., the effects of the row spacing have a less consistent reaction to the narrowing row spacing (Beatty et al., 1982; Parvez et al., 1989; Board et al., 1990; Bullock et al., 1998). This lack of consistency in the southern U.S. is probably due to the production of determinate soybeans, while in the Midwest indeterminate soybeans are predominantly grown (Bullock et al., 1998). Determinate soybean cultivars tend to be larger and when grown in a full season production they tend to achieve row closure even when planted in rows with spacing greater than 76 cm (Board et al., 1990; Bullock et al., 1998). When these determinate soybeans are double-cropped or planted following winter wheat harvest, narrow

row spacing leads to a significant increase in yield (Parvez et al., 1989; Bullock et al., 1998). Bullock et al. (1998) showed that soybean yield was not affected by the leaf area index during grain fill (R5). The lack of effect of increased leaf index on the grain yield during the grain fill stage (R5) supports the theory that rapid early season canopy growth rate and an increased number of pods led to the increased grain yields that are observed (Bullock et al., 1998). Early canopy closure in soybeans leads to increased water use efficiency by decreasing the water that is lost by soil evaporation (Purcell & Specht, 2004). Soybean yields are driven by the ability of plants to acquire photosynthetic radiation, with earlier canopy closer leading to more photosynthetic radiation being absorbed by the plant (Edwards et al., 2005). The efficiency of the crop canopy to intercept solar radiation and convert that solar energy into a storable form is the driving force behind increasing yield (Gifford et al., 1984).

To produce high-yielding cultivars, selecting varieties with the ability to acquire the maximum solar radiation during the growing season is important. This could be accomplished through plants with faster canopy coverage to intercept more light earlier in the growing season (Gifford et al., 1984; Edwards et al., 2005; Kaler et al., 2018). With the increased use of herbicides in soybean production, there has been an increase in the number of resistant weeds (Benbrook, 2012). Weed pressure during the growing season has a significant negative impact on the yields of crops. To combat these resistant weeds many different management methods have been evaluated including tillage, row spacing, rapid canopy coverage, and cultivar selection. Jha and Norsworthy (2009) reported that spring tillage did not affect the emergence of the weed *Palmer amaranth* in soybeans fields due to a limited effect on soil temperature. The authors found that while tillage had no effect, canopy closure had a beneficial effect on the reduction of the emergence of weeds during the growing season (Jha & Norsworthy, 2009). Jha et al. (2008) indicated that

narrow row spacing decreased the biomass production of two common weeds, Palmer amaranth, and Pusley species, in southern soybean production. The faster soybean canopy closure led to increased light competition between the weeds and soybeans leading to decreased weed biomass. While fast canopy closure reduced weed biomass, there was a need for multiple glyphosate applications to effectively manage both weed populations and reduce yield loss (Jha et al., 2008). Busan et al. (1997) evaluated the effects of growth rate and competitiveness of 16 different soybean genotypes against 12 weed species. The authors indicated that increased weed populations have decreased soybean yields (Bussan et al., 1997). The authors reported that there were differences between genotypes in their competitiveness, but no consistent interaction was observed between genotypic completeness and increased yields (Bussan et al., 1997). This poses a challenge for breeders in selecting genotypes that are competitive against weeds and high yielding (Bussan et al., 1997)

Canopy coverage is a highly quantitative trait that is controlled by many genetic and environmental factors and their interactions (Xavier et al., 2017; Kaler et al., 2018). Xavier et al. (2017) examined the genetic architecture of the canopy coverage in soybeans by evaluating approximately 5600 lines from 40 SoyNAM populations over three years. Canopy coverage was measured using both a handheld camera based on the method described by Purcell, (2000) in 2013-2014 and a UAV in 2014 and 2015. After evaluating both handheld and aerial imaging data from 2014, the authors found no difference between the two platforms and then elected to collect data via the UAV only in 2015. These measurements were collected from 2 to 8 weeks after planting which covered the physiological growth stages V2 through R5. These images were processed and data was extracted to create average canopy coverage, which was a function-based individual canopy coverage measurement taken throughout the season. They then performed a genome-wide

association study and identified six genomic regions on Chrs 1, 5, 6, 9, 10, and 19 that were significantly associated with canopy coverage measurements taken from 14 to 64 d after planting (Xavier et al., 2017). The effect of the SNP located on Chr 19 was significant on all the days that canopy coverage was collected. The second-largest SNP was observed on Chr 10 but was only significantly associated with canopy closure measurements 14 to 26 days after planting. The SNPs on Chrs 1, 5, and 9 tended to be associated with measurements that were collected 53 - 56 days after planting when the canopy was closed, and these SNPs caused smaller canopy closure estimates. SNPs on chromosome 10 were only associated with measurements from 14-26 days after planting. The SNP on Chr 9 was significant on days 14-35 and 59-64 after planting, like the SNP on Chr 19 (Xavier et al., 2017). The QTL on Chr 19 appeared to be a positive association with an increase of 47.3 kg ha^{-1} grain yield and none of these QTLs seemed to be associated with a significant increase in days to maturity (Xavier et al., 2017).

Kaler et al. (2018) performed a GWAS using 373 soybean accessions from the USDA Soybean Germplasm Collection. Using the method developed by Purcell (2000) two canopy coverage estimates CC1 23-28 d (V2-V3) and CC2 8-16 days later (V4-V7) were taken at five locations across Kansas and Arkansas. Canopy coverage estimates were created from the ratio of green pixels to total pixels in the image. Correlations between the canopy coverage measurements at all five locations were significant and positive (r = 0.10 - 0.39). Thirty-six SNPs were significantly associated with the CC1 in at least one environment, with six of the SNPs being significant in at least two of the environments. Fifty-six SNPs were associated with CC2 in at least one environment, with 11 SNPs being significant in two environments. These significant SNPs were located across all 20 soybean chromosomes, with only 12 SNPs being identified that were significant with both CC1 and CC2 (Kaler et al., 2018). All of the SNPs identified had minor allelic

effects, with the SNP on Chr 18 having the highest allelic effect of 0.08 (Kaler et al., 2018). The authors also identified four SNPs located in similar genomic regions on Chrs 1, 5, 6, and 9 that were identified by Xavier et al. (2017) and nine regions on the chromosomes where canopy coverage and seed weight are associated together. This is consistent with a previous report on a positive association between canopy coverage, seed weight, early-season plant height, and weed suppression (Place et al., 2011).

Using RGB UAV images Moreira et al. (2019) estimated canopy coverage vs traditional selection methods for the selection of plant rows to be advanced to yield testing. The authors selected plant rows for the 1st-year yield test based on three selection methods. The first one was the traditional selection method of using BLUPs for yield as the selection criteria. The second method was to select lines based on the largest BLUPs for canopy coverage and the final group was determined by calculating the BLUPs for yield with canopy coverage used as a covariate and selecting the top lines(Moreira et al., 2019). After repeating the experiment for two years, the authors found no difference in the three selection methods and stated that selecting based on canopy coverage could be used by breeders to increase efficiency when selecting new lines. They also stated that using CC estimates could improve selection in environments, where traditional phenotypic data is of low quality or missing (Moreira et al., 2019).

Drought Tolerance and Breeding for Drought Tolerance

Effects of drought on soybean

Over the next 30 years, climate change is expected to cause a significant impact on agricultural production across the globe. An average of 0.16-0.18°C per decade increase in global surface temperature has been observed over the past 50 years (Ray et al., 2019). During this period,

the top 10 global crops experienced an increased average growing season temperature, and more variability in precipitations (Ray et al., 2019). In soybean, drought is the largest abiotic stress and causes up to 40% yield loss (Specht et al., 1999; Purcell and Specht, 2004). Specht et al. (1999) investigated the effects of drought on soybean by observing differences between irrigated and rain-fed soybean production from 1972 to 1997 across Nebraska. Over this time a difference in yield of 800 kg ha^{-1} was observed between irrigated and rain-fed production. The authors indicated that not all the yield difference was due to water stress and yield differences existed even in the years with high rainfall. They stated that this difference in yield was due to more intensive management of the irrigated production areas as compared to rain-fed production fields (Specht et al., 1999). Between 1980 and 2003, 10 major drought events in the U.S. resulted in at least one billion dollars of loss (Ross et al., 2003). While hurricanes and tropical storms were more prevalent, drought caused 43% of the monetary losses that were caused by extreme weather (Ross et al., 2003). Approximately only 10% of US soybean production is grown under irrigation, and thus soybean production is dependent on timely rainfall throughout the growing season.

Drought Tolerance Mechanisms and Slow Canopy Wilting

Multiple drought tolerance and avoidance mechanisms have been observed in soybean which could be incorporated into modern cultivars to reduce the effects of drought. These traits include high transpiration efficiency, leaf osmotic adjustment, partial stromal closure, low epidermal conductance, high relative water content, leaf pubescence, lower accumulation of ureide in petioles, high transpiration efficiency, early vigor, reactive oxygen species scavenging, and more expansive or deeper rooting (Manavalan et al., 2009). One of the traits that have been well studied by researchers is the slow canopy wilting trait. This is due to the relative ease of

phenotyping compared to other drought tolerance traits and the observed benefits that lines possessing the trait have (Kunert & Vorster, 2020). Slow canopy wilting is a highly complex trait that relies on multiple water conservation methods throughout the plant to provide drought tolerance (Ye et al., 2020).

Three methods have been proposed as the physiological mechanisms responsible for the slow wilting trait (Ries et al., 2012). The first mechanism was high water use efficiency (WUE), high radiation use efficiency (REU), and conservation of soil moisture. The lines with this mechanism would more efficiently convert transpired water into biomass and have relatively high yield potential in both drought and non-drought environments (Ries et al., 2012). The second mechanism is low stomatal conductance, low RUE, low WUE, and conserved soil moisture (Ries et al., 2012). This mechanism would produce lines that have a relatively low photosynthetic capacity (Ries et al., 2012). Lines with this mechanism would perform well in drought environments but would have low yield potential in optimum environments (Ries et al., 2012). The third mechanism is for slow canopy wilting and deeper rooting, but this mechanism has yet to be identified in any slow wilting cultivars (Ries et al., 2012).

Sources of Slow Canopy Wilting

PI 416937, a MG V accession from Japan, was observed to have a tolerance to drought conditions, the primary phenotype of the slow canopy wilting (Sloane et al., 1990). It was evaluated under drought stress and irrigation conditions in a split-plot design with Forrest, a leading MG 5 commercial cultivar at the time. Both cultivars were evaluated for leaf water potential, relative water content, soil moisture, and transpiration, and at the end of the season for plot yield under rain-out shelters. Under a drought stress condition, PI 416937 has higher leaf water

potential as compared to Forrest, but no difference between genotypes was observed under irrigation. Similarly, relative water content was significantly high in PI 416937 as compared to Forrest under a stressed condition, with the 6.3% difference being observed in the afternoon. The PI showed significantly more transpiration and increased water use efficiency when compared to Forrest under drought stress. However, no differences were observed in soil water potential between the two cultivars. Compared to PI 416937, Forrest had a significantly higher yield under non-stress conditions. Under drought conditions, seed yield was reduced by ~50% in Forrest, while in the PI only a 30% yield loss was observed. PI 416937 had a fibrous and extensive lateral root system which allows it to extract more water from the soil under drought stress (Goldman et al., 1989; Sloane et al., 1990; Hudak & Patterson, 1996).

Sinclair et al. (2008) examined the effects of atmospheric vapor pressure deficit (VPD) on PI 416937 and two conventional soybean genotypes. Vapor pressure deficit is the difference between the moisture in the air and the total moisture that the air can hold. Increased VPD indicates water stress in the environment. Compared to the two commercial cultivars, the authors found that PI 416937 maintained constant transpiration. As VPD increased to 1.9 kPa, all three cultivars had consistent linear increases in transpiration and after 1.9 kPa, unlike the two commercial cultivars, transpiration in PI 416937 did not increase (Sinclair et al., 2008). Based on pedigree records in the USDA Uniform Soybean Trials (Soybase.org), PI 416937 has been incorporated into the southern soybean germplasm. One example is USDA 8002, a high-yielding drought-tolerant MG VIII soybean cultivar that was developed by the USDA-ARS (Carter et al., 2016).

Three additional plant introductions PI 471938, a MG V accession from Nepal, and PI 567690, and PI 567731, MG III soybean accessions from central China, were also reported that possess the slow canopy wilting phenotype (Sadok et al., 2012; Pathan et al., 2014). Although PI

471938 possesses the slow wilting trait, based on experimental evidence it does not possess any of the four methods of slow canopy wilting that were demonstrated by PI 416937 (Sadok et al., 2012). PI 567690 and PI 567731 were identified to be the most tolerant lines in a screening of 250 MG III soybean accessions for drought tolerance by Pathan et al. (2014). Both of these PIs were shown to maintain their yield even under drought stress (Pathan et al., 2014).

GWAS and QTL mapping for the slow canopy wilting trait

The highly quantitative nature of the slow wilting trait leads to the need for understanding the underlying genetics by QTL mapping and genome-wide association studies. Using 79 RILs derived from a cross of KS4895 by Jackson that were evaluated for drought tolerance in three environments in Arkansas and North Carolina and genotyped using SSR markers, Charlson et al. (2009) identified four QTLs on Chrs 8, 14, 17, and 13. These QTLs explained 5 to 10 % of the phenotypic variation that was observed in RILs for tolerance to drought. Similarly, Abdel-Haleem et al. (2012) found seven QTLs related to canopy wilting in a population derived from 'Benning' × PI 416937 that was evaluated in five rain-fed environments for canopy wilting. Two of the seven QTLs identified were from the fast wilting parent Benning, and the remaining five QTLs were contributed from PI 416937. A locus identified on Chr 12 accounted for over 25 % of the phenotypic variation for slow canopy wilting that was observed in the population. Hwang et al. (2015) used three slow wilting genotypes Jackson, PI 424140, and PI 416937 to form five RIL populations that were used for mapping slow canopy wilting QTLs. RILs were evaluated for drought tolerance in three environments in North Carolina, Kansas, and Arkansas. The authors identified nine QTLs on Chrs 2, 5, 11, 14, 17, and 19 in at least two of the populations. The QTL on Chr 11 explained up to 39% of the phenotypic variation (Hwang et al., 2015). Using the same

populations, the authors performed a meta-analysis to narrow the regions that were identified in the previous mapping study and found eight meta- QTL clusters on Chrs 2, 5, 11, 19. These clusters on average explained 13% of the phenotypic variation (Hwang et al., 2016). Two major QTLs were identified on Chrs 11 and 19 but they were environmentally unstable (Hwang et al., 2016).

Ye et al. (2020) mapped the slow canopy wilting trait in two RIL populations developed from crosses of Pana × PI 567690 and Magellan × PI 567731. These populations were evaluated in two environments in Kansas and Missouri between 2013 and 2015. Four QTLs were identified on Chrs 5, 9, 12, and 19 in the Pana × PI 567690 population which were present in at least two of the environments and were in the region reported in previous mapping work (Ye et al., 2020). Two additional QTLs were identified on Chrs 6 and 10 in the Magellan × PI 567731 population that were present in at least two environments. Both QTLs were not previously reported and each QTL explained 20-30% of the phenotypic variation.

Kaler et al. (2017) performed a GWAS using a panel of 373 maturity group IV accessions from the USDA Soybean Germplasm Collection and the SoySNP50K iSelect BeadChip. These lines were phenotyped in two environments in Arkansas and Kansas for two years for slow canopy wilting. Sixty-one SNPs were identified to be significantly associated with canopy wilting across four environments and 21 of those SNPs were present in at least two environments (Kaler et al., 2017). The 61 SNPs in this study were located in 51 putative loci on 19 of 20 soybean chromosomes (Kaler et al., 2017). Using a panel of 162 soybean genotypes ranging from MG VI-VIII from 30 countries around the world and the SoySNP50K iSelect BeadChip, Steketee et al. (2020) identified 45 unique SNPs at 44 loci for canopy wilting on 19 chromosomes. Of these QTLs, 12 significant loci identified by Steketee et al. (2020) were in similar genomic regions as the loci that were identified by Kaler et al. (2017).

Breeding for drought tolerance

Previous studies have shown that drought tolerance traits such as slow canopy wilting can have a beneficial impact on soybean yields (Sloane et al., 1990; Devi et al., 2014). These studies also indicated that sources of slow canopy wilting tend to have lower yield potential than more advanced commercial cultivars. For breeders to exploit the slow canopy wilting traits it would need to be introgressed into elite breeding materials. But the introgression of complex traits such as slow canopy wilting can be difficult and time-consuming. Both PI 416937 and PI 471936 have been used extensively by southern breeding programs to breed for drought tolerance based on the pedigree records from lines in the USDA Uniform Soybean test and variety registrations (Soybase.org, 2022). The development of markers based on the mapping work could be used to more efficiently select for drought tolerance using marker-assisted selection and genomic prediction.

Objectives

The dissertation is divided into three chapters focusing on advancing soybean breeding through disease resistance, climate resilience, and high throughput phenotyping. The objectives of Chapter 2 are to: i) understand the resistance to SSC provided by PI 398469, ii) develop markers that can be used to select for resistance to SSC provided by PI 398469, and iii) determine if Crockett and PI 398469 carry the same QTL for resistance to SSC. The objectives of Chapter 3 are to: i) evaluate a RIL population in repeated field experiments for canopy wilting, and ii) elucidate genomic regions responsible for canopy wilting in PI 471938. The objectives of Chapter 4 are to: i) develop a workflow to phenotype canopy coverage, ii) understand the canopy coverage trait per se in elite southern breeding materials to see if it can be selected as an additional breeding

trait, and iii) determine the relationship between rapid canopy coverage and the agronomically important traits 100-seed weight, plant height, lodging, and maturity.

References

Abdel-Haleem, H., Carter, T. E., Purcell, L. C., King, C. A., Ries, L. L., Chen, P., Schapaugh, W., Sinclair, T. R., & Boerma, H. R. (2012). Mapping of quantitative trait loci for canopy-wilting trait in soybean (Glycine max L. Merr). *Theoretical and Applied Genetics. 125*(5), 837–846. doi.org/10.1007/s00122-012-1876-9

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, *19*(1), 52–61. doi.org/10.1016/j.tplants.2013.09.008

Athow, K. L., & Caldwell, R. m. (1954). A comparative study of *Diaporthie* stem canker and pod and stem blight of soybean. *Phytopathology*, *44*, 319–325.

Backman, P., Weaver, D., & Morgan-Jones, G. (1985). Soybean Stem Canker: An Emerging Disease Problem. *Plant Disease*, *69*(8), 641–647.

Bajwa, S. G., Rupe, J. C., & Mason, J. (2017). Soybean disease monitoring with leaf reflectance. *Remote Sensing*, *9*(2). doi.org/10.3390/rs9020127

Bandara, A. Y., Weerasooriya, D. K., Bradley, C. A., Allen, T. W., & Esker, P. D. (2020). Dissecting the economic impact of soybean diseases in the United States over two decades. *PLoS ONE*, *15*(4). doi.org/10.1371/journal.pone.0231141

Beatty, K. D., Eldridge, I. L., & Simpson, A. M. (1982). Soybean Response to Different Planting Patterns and Dates. *Agronomy Journal*, *74*(5), 859–862. doi.org/10.2134/agronj1982.00021962007400050021x

Benbrook, C. (2012). Impacts of genetically engineered crops on pesticide use in the U.S.–the first sixteen years. *Environmental Sciences Europe*, *24*, 1–13.

Black, B. D., Padgett, G. B., Russin, J. S., Griffin, J. L., Snow, J. P., & Berggren Jr.G.T. (1996). Potential Weed Hosts for *Diaporthe phaseolorum* var. *caulivora*, Causal Agent for Soybean Stem Canker. *Plant Disease*, *80*, 763–765.

Board, J. E., Harville, B. G., & Saxton, A. M. (1990). Branch Dry Weight in Relation to Yield Increases in Narrow-Row Soybean. *Agronomy Journal*, *82*(3), 540–544. doi.org/10.2134/agronj1990.00021962008200030021x

Boehm, J. D., Abdel-Haleem, H., Schapaugh, W. T., Rainey, K., Pantalone, V. R., Shannon, G., Klein, J., Carter, T. E., Cardinal, A. J., Shipe, E. R., Gillen, A. M., Smith, J. R., Chen, P., Weaver, D. B., Boerma, H. R., & Li, Z. (2019). Genetic improvement of us soybean in maturity groups V, VI, and VII. *Crop Science*, *59*(5), 1838–1852. doi.org/10.2135/cropsci2018.10.0627

Bowers, G. R. (1990). Registration of 'Crockett' Soybean. *Crop Science*, *30*(2), 427. doi.org/10.2135/cropsci1990.0011183x003000020049x

Bowers, G. R., Ngeleka, K., & Smith, O. (1993). Inheritance of Stem Canker Resistance in Soybean Cultivars Crockett and Dowling. *Crop Science*, *33*(1), 67–70. doi.org/10.2135/cropsci1993.0011183x003300010010x

Bryant, C. (2021). *Soybean Production in Georgia*. Athens, GA

Bullock, D., Khan, S., & Rayburn, A. (1998). Soybean yield response to narrow rows is largely due to enhanced early growth. *Crop Science*, *38*(4), 1011–1016. doi.org/10.2135/cropsci1998.0011183X003800040021x

Bussan, A. J., Burnside, O. C., Orf, J. H., & Puettmann, K. J. (1997). Field Evaluation of Soybean (*Glycine max*) Genotypes for Weed Competitiveness. *Weed Science*, *45*(1), 31–37. about.jstor.org/terms

Carlson, J. B., & Lersten, N. R. (2004). Reproductive Morphology. In *Soybean: Improvement, Production, and Uses* (3rd ed., Vol. 3, pp. 59–95). American Society of Agronomy.

Carter, T. E., Todd, S. M., & Gillen, A. M. (2016). Registration of 'USDA-N8002' Soybean Cultivar with High Yield and Abiotic Stress Resistance Traits. *Journal of Plant Registrations*, *10*(3), 238–245. doi.org/10.3198/jpr2015.09.0057crc

Chang, H. X., Lipka, A. E., Domier, L. L., & Hartman, G. L. (2016). Characterization of disease resistance loci in the USDA soybean germplasm collection using genome-wide association studies. *Phytopathology*, *106*(10), 1139–1151. doi.org/10.1094/PHYTO-01-16-0042-FI

Chappelle, E. W., Kim, M. S., & Mcmurtrey, J. E. (1992). Ratio Analysis of Reflectance Spectra (RARS): An Algorithm for the Remote Estimation of the Concentrations of Chlorophyll A, Chlorophyll B, and Carotenoids in Soybean Leaves. *Remote Sensing of Environment*, *39*, 239–247.

Charlson, D. v, Bhatnagar, S., King, C. A., Ray, J. D., Sneller, C. H., Carter, T. E., & Purcell, L. C. (2009). Polygenic inheritance of canopy wilting in soybean [*Glycine max* (L.) Merr.]. *Theoretical and Applied Genetics.*, *119*, 587–594. doi.org/10.1007/s00122-009-1068-4

Chiesa, M. A., Cambursano, M. v., Pioli, R. N., & Morandi, E. N. (2017). Molecular mapping of the genomic region conferring resistance to soybean stem canker in Hutcheson soybean. *Molecular Breeding*, *37*(5), 65. doi.org/10.1007/s11032-017-0660-6

Chiesa, M. A., Pioli, R. N., Cambursano, M. v., & Morandi, E. N. (2013). Differential expression of distinct soybean resistance genes interacting with Argentinean isolates of *Diaporthe phaseolorum* var. *meridionalis*. *European Journal of Plant Pathology*, *135*(2), 351–362. doi.org/10.1007/s10658-012-0091-5

Chiesa, M. A., Pioli, R. N., & Morandi, E. N. (2009). Specific resistance to soybean stem canker conferred by the *Rdm4* locus. *Plant Pathology*, *58*(6), 1032–1038. doi.org/10.1111/j.1365-3059.2009.02145.x

Damicone, J. P., Berggren G.T., & Snow, J. P. (1987). Effect of Free Moisture on Soybean Stem Canker Development. *Phytopathology*, *77*(11), 1568–1572.

Deering, D. W. (1978). *Rangeland Reflectance Characteristics Measured by Aircraft and Spacecraft Sensors*. Ph.D. Dissertation. Texas A&M University.

Devi, J. M., Sinclair, T. R., Chen, P., & Carter, T. E. (2014). Evaluation of elite southern maturity soybean breeding lines for drought-tolerant traits. *Agronomy Journal*, *106*(6), 1947–1954. doi.org/10.2134/agronj14.0242

di Gennaro, S. F., Rizza, F., Badeck, F. W., Berton, A., Delbono, S., Gioli, B., Toscano, P., Zaldei, A., & Matese, A. (2018). UAV-based high-throughput phenotyping to discriminate barley vigour with visible and near-infrared vegetation indices. *International Journal of Remote Sensing*, *39*(15–16), 5330–5344. doi.org/10.1080/01431161.2017.1395974

Edwards, J. T., Purcell, L. C., & Karcher, D. E. (2005). Soybean yield and biomass responses to increasing plant population among diverse maturity groups: II. Light interception and utilization. *Crop Science*, *45*(5), 1778–1785. doi.org/10.2135/cropsci2004.0570

Fehr, W. R., & Caviness, C. E. (1977). *Stages of Soybean Development*. http://lib.dr.iastate.edu/specialreports/87

Fernández, F. A., & Hanlin, R. T. (1996). Morphological and RAPD Analyses of *Diaporthe phaseolorum* from Soybean. *Mycologia*, *88*(3), 425–440.

Filella, I., & Penuelas, J. (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass, and hydric status. *International Journal of Remote Sensing*, *15*(7), 1459–1470.

Freitas, M. A., Cafe Filho, A. C., & Nasser, L. C. B. (2002). Cultural practices and genetic resistance as factors affecting soybean stem canker and plant yield in the Cerrado. *Fitopatologia Brasileira*, *27*(1), 5–11.

Furbank, R. T., & Tester, M. (2011). Phenomics - technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, *16*(12), 635–644. doi.org/10.1016/j.tplants.2011.09.005

Gifford, R. M., Thorne, J. H., Hitz, W. D., & Giaquinta, R. T. (1984). Crop Productivity and Photoassimilate Partitioning. *Science*, *225*(4664), 801–808.

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, *32*(8), 1–4. doi.org/10.1029/2005GL022688

Goldman, I. L., Carter, T. E., & Patterson, R. P. (1989). Differential genotypic response to drought stress and subsoil aluminum in soybean. *Crop Science*, *29*(2), 330–334. doi.org/10.2135/cropsci1989.0011183X002900020020x

Grau, C. R., Dorrance, A. E., Bond, J., & Russin, J. S. (2004). Fungal Diseases. In *Soybeans: Improvement, Production, and Uses,* (3rd ed., Vol. 16, pp. 679–763). American Society of Agronomy.

Gravert, C. E., Li, S., & Hartman, G. L. (2001). Occurrence of *Diaporthe phaseolorum* var. *meridionalis* on Soybean in Illinois. *Plant Disease*, *85*(11), 1211. doi.org/10.1094/PDIS.2001.85.11.1211A

Heatherly, L. G., & Elmore, R. W. (2004). Managing Inputs for Peak Production. In *Soybeans: Improvement, Production, and Uses* (3rd ed., Vol. 16, pp. 451–536). American Society of Agronomy.

Hildebrand, A. A. (1952). Stem canker. A disease of increasing importance on Soybeans in Ontario. *Soybean Digest*, *12*(9), 12–15.

Hudak, C. M., & Patterson, R. P. (1996). Root distribution and soil moisture depletion pattern of a drought-resistant soybean plant introduction. *Agronomy Journal*, *88*(3), 478–485. doi.org/10.2134/agronj1996.00021962008800030020x

Hwang, S., King, C. A., Chen, P., Ray, J. D., Cregan, P. B., Carter, T. E., Li, Z., Abdel-Haleem, H., Matson, K. W., Schapaugh, W., & Purcell, L. C. (2016). Meta-analysis to refine map position and reduce confidence intervals for delayed-canopy-wilting QTLs in soybean. *Molecular Breeding*, *36*(7), 91. doi.org/10.1007/s11032-016-0516-5

Hwang, S., King, C. A., Ray, J. D., Cregan, P. B., Chen, P., Carter, T. E., Li, Z., Abdel-Haleem, H., Matson, K. W., Schapaugh, W., & Purcell, L. C. (2015). Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. *Theoretical and Applied Genetics*, *128*(10), 2047–2065. doi.org/10.1007/s00122-015-2566-1

Hymowitz, T. (1987). Introduction of the Soybean to Illinois. *Economic Botany*, *41*(1), 28–32.

Hymowitz, T., & Harlan, J. R. (1983). Introduction of Soybean to North America by Samuel Bowen in 1765. *Economic Botany*, *37*(4), 371–379.

Hymowitz, T., & Newell, C. A. (1981). Taxonomy of the Genus *Glycine*, Domestication and Uses of Soybeans. *Economic Botany*, *35*(3), 272–288.

Janse Van Rensburg, J. C., Lamprecht, S. C., Groenewald, J. Z., Castlebury, L. A., Crous, P. W., & Lamprecht, S. (2006). Characterisation of *Phomopsis* spp. associated with die-back of rooibos (*Aspalathus linearis*) in South Africa. *Studies of Mycology*, *55*, 65–74.

Jha, P., & Norsworthy, J. K. (2009). Soybean Canopy and Tillage Effects on Emergence of Palmer Amaranth (*Amaranthus palmeri*) from a Natural Seed Bank. *Weed Science*, *57*(6), 644–651. doi.org/10.1614/WS-09-074

Jha, P., Norsworthy, J. K., Bridges, W., & Riley, M. B. (2008). Influence of Glyphosate Timing and Row Width on Palmer Amaranth *(Amaranthus palmeri*) and Pusley (*Richardia* spp.) Demographics in Glyphosate-Resistant Soybean. *Weed Science*, *56*(3), 408–415. doi.org/10.1614/ws-07-174.1

Kaler, A. S., Ray, J. D., Schapaugh, W. T., Davies, M. K., King, C. A., & Purcell, L. C. (2018). Association mapping identifies loci for canopy coverage in diverse soybean genotypes. *Molecular Breeding*, *38*(5), 50. doi.org/10.1007/s11032-018-0810-5

Kaler, A. S., Ray, J. D., Schapaugh, W. T., King, C. A., & Purcell, L. C. (2017). Genome-wide association mapping of canopy wilting in diverse soybean genotypes. *Theoretical and Applied Genetics*, *130*(10), 2203–2217. doi.org/10.1007/s00122-017-2951-z

Kato, I., Sakaguchi, S., & Naito, Y. (1954). Development of flower parts and seed in the soybean plant, Glycine max M. *Tokai Kinki National Agriculture Experimental Station (Japan) Bulletin*, *1*, 96–114.

Keeling, B. L. (1982). A Seedling Test for Resistance to Soybean Stem Canker Caused by *Diaporthe phaseolorum* var. *caulivora*. *Phytopathology*, *72*(7), 807–809.

Kilen, T. C., & Hartwig, E. E. (1987). Identification of Single Genes Controlling Resistance to Stem Canker in Soybean. *Crop Science*, *27*(5), 863–864. doi.org/10.2135/cropsci1987.0011183x002700050005x

Kilen, T. C., Keeling, B. L., & Hartwig, E. E. (1985). Inheritance of Reaction to Stem Canker in Soybean. *Crop Science*, *25*(1), 50–51. doi.org/10.2135/cropsci1985.0011183x002500010014x

Kmetz, K. T., Schmitthenner, A. F., & Ellett, C. W. (1978). Soybean Seed Decay: Prevalence of Infection and Symptom Expression Caused by *Phomopsis* sp., *Diaporthe phaseolorum* var. *sojae*, and *D. phaseolorum* var. *caulivora*. *Phytopathology*, *68*, 863–840.

Koenning, S. R., & Wrather, J. A. (2010). Suppression of Soybean Yield Potential in the Continental United States by Plant Diseases from 2006 to 2009. *Plant Health Progress*, *11*(1). doi.org/10.1094/php-2010-1122-01-rs

Krauz, J. P., & Fortnum, B. A. (1983). An Epiphytotic of *Diaporthe* Stem Canker of Soybean in South Carolina. *Plant Disease*, *67*(10), 1128–1129.

Kulik, M. M. (1984). Symptomless Infection, Persistence, and Production of Pycnidia in Host and Non-Host Plants by *Phomopsis batatae*, *Phomopsis phaseoli*, and *Phomopsis sojae*, and the Taxonomic. *Source: Mycologia*, *76*(2), 274–291.

Kunert, K., & Vorster, B. J. (2020). In search for drought-tolerant soybean: Is the slow-wilting phenotype more than just a curiosity? *Journal of Experimental Botany*, *71*(2), 457–460. doi.org/10.1093/jxb/erz235

Lalitha, B., Snow, J. P., & Berggren, G. T. (1989). Phytotoxin production by *Diaporthe phaseolorum* var. *caulivora*, the causal organism of stem canker of soybean. *Phytopathology*, *79*(4), 499–504.

Lersten, N. R., & Carlson, J. B. (2004). Vegetative Morphology. In *Soybeans : Improvement, Production, and Uses*. 3rd ed., Vol. 16, pp. 15-57). American Society of Agronomy.

Li, S., Kurtzweil, N. C., Grau, C. R., & Hartman, G. L. (2004). Occurrence of Soybean Stem Canker (*Diaporthe phaseolorum* var. *meridionalis*) in Wisconsin. *Plant Disease*, *88*(5), 576. doi.org/10.1094/PDIS.2004.88.5.576B

Ma, B. L., Dwyer, L. M., Costa, C., Cober, E. R., Morrison, M. J., Dwyer, L. M., Cober, E. R., & Morrison, M. J. (2001). Early Prediction of Soybean Yield from Canopy Reflectance Measurements. *Agronomy Journal*, *93*, 1227–1234.

Major, D. J., Johnson, D. R., & Luedders, V. D. (1975). Evaluation of Eleven Thermal Unit Methods for Predicting Soybean Development. *Crop Science*, *15*(2), 172–174. doi.org/10.2135/cropsci1975.0011183x001500020008x

Makanza, R., Zaman-Allah, M., Cairns, J. E., Magorokosho, C., Tarekegne, A., Olsen, M., & Prasanna, B. M. (2018). High-throughput phenotyping of canopy cover and senescence in maize field trials using aerial digital canopy imaging. *Remote Sensing*, *10*(2), 330. doi.org/10.3390/rs10020330

Maldonado Dos Santos, J. V., Ferreira, E. G. C., Passianotto, A. L. D. L., Brumer, B. B., Santos, A. B. dos, Soares, R. M., Torkamaneh, D., Arias, C. A. A., Belzile, F., Abdelnoor, R. V., & Marcelino-Guimarães, F. C. (2019). Association mapping of a locus that confers southern stem canker resistance in soybean and SNP marker development. *BMC Genomics*, *20*(1), 798. doi.org/10.1186/s12864-019-6139-6

Manavalan, L. P., Guttikonda, S. K., Tran, L.-S., & Nguyen, H. T. (2009). Physiological and molecular approaches to improve drought resistance in soybean. *Plant & Cell Physiology*, *50*(7), 1260–1276. doi.org/10.1093/pcp/pcp082

Moreira, F. F., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2019). Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. *Plant Methods, 15*(1), 139. doi.org/10.1186/s13007-019-0519-4

Morgan-Jones, G. (1989). The *Diaporthe/Phomopsis* complex: taxonomic considerations. *World Soybean Research Conference, 4*, 1699–1706.

Nutter, F. W., Tylka, G. L., Guan, J., Moreira, A. J. D., Marett, C. C., Rosburg, T. R., Basart, J. P., & Chong, C. S. (2002). Use of Remote Sensing to Detect Soybean Cyst Nematode-Induced Plant Stress. *Journal of Nematology, 34*(3), 222–231.

Parvez, A. Q., Gardner, F. P., & Boote, K. J. (1989). Determinate- and indeterminate-type soybean cultivar responses to pattern, density, and planting date. *Crop Science, 29*(1), 150–157. doi.org/10.2135/cropsci1989.0011183X002900010034x

Pathan, S. M., Lee, J.-D., Sleper, D. A., Fritschi, F. B., Sharp, R. E., Carter, T. E., Nelson, R. L., King, C. A., Schapaugh, W. T., Ellersieck, M. R., Nguyen, H. T., & Shannon, J. G. (2014). Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought. *Journal of Agronomy and Crop Science*, *200*(3), 231–236. doi.org/10.1111/jac.12053

Peñuelas, J., Isla, R., Filella, I., & Araus, J. L. (1997). Visible and near-infrared reflectance assessment of salinity effects on barley. *Crop Science*, *37*(1), 198–202. doi.org/10.2135/cropsci1997.0011183X003700010033x

Philbrook, B. D., & Oplinger, E. S. (1989). Soybean Field Losses as Influenced by Harvest Delays. *Agronomy Journal*, *81*(2), 251–258. doi.org/10.2134/agronj1989.00021962008100020023x

Pioli, R., Gattuso, S., Prado, D., & Borghi, A. (1997). Recent Outbreak of Stem Canker (*Diaporthe phaseolorum* var. *meridionalis*) of Soybean in Santa Fe, Argentina. *Plant Disease*, *81*(10), 1215.

Pioli, R. N., Morandi, E. N., Martìnez, M. C., Lucca, F., Tozzini, A., Bisaro, V., & Hopp, H. E. (2003). Morphologic, molecular, and pathogenic characterization of *Diaporthe phaseolorum* variability in the core soybean-producing area of Argentina. *Phytopathology*, *93*(2), 136–146. doi.org/10.1094/PHYTO.2003.93.2.136

Place, G. T., Reberg-Horton, S. C., Carter, T. E., & Smith, A. N. (2011). Effects of soybean seed size on weed competition. *Agronomy Journal*, *103*(1), 175–181. doi.org/10.2134/agronj2010.0195

Ploetz, R. C., & Shokes, F. M. (1985). Soybean stem canker incited by ascospores and conidia of the fungus causing the disease in the southeastern United States. *Plant Disease*, *69*(11), 990–992.

Ploetz, R. C., & Shokes, F. M. (1987). Factors Influencing Infection of Soybean Seedlings by Southern *Diaporthe phaseolorum*. *Phytopathology*, *77*(6), 786–790.

Poland, J. A., & Rife, T. W. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. *The Plant Genome*, *5*(3). doi.org/10.3835/plantgenome2012.05.0005

Purcell, L. C. (2000). Soybean canopy coverage and light interception measurements using digital imagery. *Crop Science*, *40*(3), 834–837. doi.org/10.2135/cropsci2000.403834x

Purcell, L. C., & Specht, J. E. (2004). Physiological Traits for Ameliorating Drought Stress. In *Soybeans: improvement, production, and uses* (Vol. 3, pp. 569–620). American Society of Agronomy.

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., & Chatterjee, S. (2019). Climate change has likely already affected global food production. *PLoS ONE*, *14*(5) e0217148. doi.org/10.1371/journal.pone.0217148

Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., & Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nature Communications*, *3*. 1293. doi.org/10.1038/ncomms2296

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. v., & Chatterjee, S. (2019). Climate change has likely already affected global food production. *PLoS ONE*, *14*(5). doi.org/10.1371/journal.pone.0217148

Ries, L. L., Purcell, L. C., Carter, T. E., Edwards, J. T., & King, C. A. (2012). Physiological traits contributing to differential canopy wilting in soybean under drought. *Crop Science*, *52*(1), 272–281. doi.org/10.2135/cropsci2011.05.0278

Rincker, K., Nelson, R., Specht, J., Sleper, D., Cary, T., Cianzio, S. R., Casteel, S., Conley, S., Chen, P., Davis, V., Fox, C., Graef, G., Godsey, C., Holshouser, D., Jiang, G.-L., Kantartzi, S. K., Kenworthy, W., Lee, C., Mian, R., … Diers, B. (2014). Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. *Crop Science*, *54*(4), 1419–1432. doi.org/10.2135/cropsci2013.10.0665

Ross, T. & Lott, N. (2003). *A Climatology of 1980-2003 Extreme Weather and Climate Events*. http://www.ncdc.noaa.gov/ol/reports/billionz.html

Rothrock, C. S., Hobbs, T., & Phillips, D. (1985). Effects of Tillage and Cropping System on Incidence and Severity of Southern Stem Canker of Soybean. *Phytopathology*, *75*(10), 1156. doi.org/10.1094/phyto-75-1156

Roy, K. W., & Miller, W. A. (1983). Soybean Stem Canker Incited by Isolates of *Diaporthe* and *Phomopsis* spp. from Cotton in Mississippi. *Plant Disease*, *67*(2), 135–137.

Rupe, C. J., Sutton, E. A., Becton, C. M., & Gbur Jr, E. E. (1996). Effect of Temperature and Wetness Period on Recovery of the Southern Biotype of *Diaporthe phaseolorum* var. *caulivora* from Soybean. *Plant Disease*, *80*(3), 255–257.

Rupe, J. C. (2015). Stem Canker. In G. L. Hartman, J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis, & K. L. Steffey (Eds.), *Compendium of Soybean Disease and Pests* (5th ed., pp. 85–88). The American Phytopathological Society.

Rupe, J. C., Sutton, E. A., & Gbur, E. E. (1999). Effect of Soybean Growth Stage at the Time of Inoculation with *Diaporthe phaseolorum* var. *meridionales* on Stem Canker Development and Yield. *Plant Disease*, *83*(6), 582–586.

Sadok, W., Gilbert, M. E., Raza, M. A. S., & Sinclair, T. R. (2012). Basis of slow-wilting phenotype in soybean PI 471938. *Crop Science*, *52*(3), 1261–1269. doi.org/10.2135/cropsci2011.11.0622

Santos, J. M., Vrandečić, K., Ćosić, J., Duvnjak, T., & Phillips, A. J. L. (2011). Resolving the *diaporthe* species occurring on soybean in Croatia. *Persoonia: Molecular Phylogeny and Evolution of Fungi*, *27*, 9–19. doi.org/10.3767/003158511X603719

Sato, T., de Viedma, L. Q., Alvarez, E., Romero, M. I., & Morel, W. (1993). First Occurrence of Soybean Southern Stem Canker in Paraguay. *Japan Agricultural Research Quarterly*, *27*, 20–26.

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., Song, Q., Thelen, J. J., Cheng, J., Xu, D., Hellsten, U., May, G. D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M. K., Sandhu, D., Valliyodan, B., … Jackson, S. A. (2010). Genome

sequence of the palaeopolyploid soybean. *Nature*, *463*(7278), 178–183. doi.org/10.1038/nature08670

Shearin, Z. (2011). *Facilitating Breeding for Resistance to Southern Stem Canker and Elevated Seed Oleic Acid Content in Soybean*. Ph.D. Dissertation. The University of Georgia.

Shibles, R., Anderson, I. C., & Gibson, A. H. (1975). Soybean. In *Crop physiology : some case histories* (pp. 151–189). Cambridge University Press.

Shi, Y., Alex Thomasson, J., Murray, S. C., Ace Pugh, N., Rooney, W. L., Shafian, S., Rajan, N., Rouze, G., Morgan, C. L. S., Neely, H. L., Rana, A., Bagavathiannan, M. v., Henrickson, J., Bowden, E., Valasek, J., Olsenholler, J., Bishop, M. P., Sheridan, R., Putman, E. B., … Yang, C. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE*, *11*(7). doi.org/10.1371/journal.pone.0159781

Shurtleff, W., & Aoyagi, A. (2004). *History of Soy in Latin America*. A Chapter from the Unpublished Manuscript, History of Soybeans and Soyfoods, 1100 B.C. to the 1980s.

Sinclair, T. R., Zwieniecki, M. A., & Holbrook, N. M. (2008). Low leaf hydraulic conductance associated with drought tolerance in soybean. *Physiologia Plantarum*, *132*(4), 446–451. doi.org/10.1111/j.1399-3054.2007.01028.x

Sloane, R. J., Patterson, R. P., & Carter, T. E. (1990). Field drought tolerance of a soybean plant introduction. *Crop Science*, *30*, 118–123. doi.org/10.2135/cropsci1990.0011183X003000010027x

Smith, E. F., & Backman, P. A. (1989). Epidemiology of Soybean Stem Canker in the Southeastern United States: Relationship Between Time of Exposure to Inoculum and Disease Severity. *Plant Disease*, *73*(6), 464–468.

Soybase (2022). Soybean Pedigree Database, Retrieved from www.soybase.org

SoyStats (2021). Planting, production, and yield data. Retrieved from www.soystats.com/

Specht, J. E., Diers, B. W., Nelson, R. L., de Toledo, J. F. F., Torrion, J. A., & Grassini, P. (2014). Soybean. In *Yield Gains in Major U.S. Field Crops* (Vol. 33, pp. 311–355). doi.org/10.2135/cssaspecpub33.c12

Specht, J., Hume, D., & Kumudini, S. (1999). Soybean yield potential—A genetic and physiological perspective. *Crop Science*, *39*, 1560–1570.

Steketee, C. J., Schapaugh, W. T., Carter, T. E., & Li, Z. (2020). Genome-wide association analyses reveal genomic regions controlling canopy wilting in soybean. *G3: Genes, Genomes, Genetics*, *10*(4), 1413–1425. doi.org/10.1534/g3.119.401016

Subbarao, K. v., Snow, J. P., Berggren, G. T., Damicone, J. P., & Padgett, G. B. (1992). Analysis of Stem Canker Epidemics in Irrigated and Nonirrigated Conditions on Differentially Susceptible Soybean Cultivars. *Phytopathology*, *82*(12), 1251–1256.

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., & Bennett, M. (2017). Plant Phenomics, From Sensors to Knowledge. *Current Biology*, 27, R770–R783. doi.org/10.1016/j.cub.2017.05.055

Threinen, J. T., Kommedahl, T., & Klug, R. J. (1959). Hybridization between radiation-induced mutants of two varieties of *Diaporthe phaseolorum*. *Phytopathology*, *49*, 797–801.

Tucker, C. J., Elgin, J. H., Mcmurtrey, J. E., & Fan, C. J. (1979). Monitoring Corn and Soybean Crop Development with Hand-Held Radiometer Spectral Data. *Remote Sensing of Environment*, *8*, 237–248.

Tyler, D. D., Overton, J. R., & Chambers, A. Y. (1983). Tillage effects on soil properties, diseases, cyst nematodes, and soybean yields. *Journal of Soil and Water Conservation*, *38*(4), 374–376.

Tyler, J. M. (1996). Characterization of stem canker resistance in “Hutcheson” soybean. *Crop Science*, *36*(3), 591–593. doi.org/10.2135/cropsci1996.0011183X003600030011x

USDA NASS. (2021, June). *Acreage Report* (ISSN: 1949-1522). Retrieved from usda.library.cornell.edu/concern/publications/j098zb09z?locale=en

Weaver, D. B., Cosper, B. H., Backman, P. A., & Crawford, M. A. (1984). Cultivar Resistance to Field Infestations of Soybean Stem Canker. *Plant Disease*, *68*(10), 877–879.

Welch, A. W., & Gilman, J. C. (1948). Hetero-and homo-thallic types of *Diaporthe* on soybeans. *Phytopathology*, *38*, 628–637.

Whitehead, M. D. (1966). Stem Canker and Blight of Birdsfoot-Trefoil and Soybeans Incited by *Diaporthe phaseolorum* var *sojae*. *Phytopathology*, *56*, 396–400.

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., Feldmann, K. A., French, A. N., Heun, J. T., Hunsaker, D. J., Jenks, M. A., Kimball, B. A., Roth, R. L., Strand, R. J., Thorp, K. R., Wall, G. W., & Wang, G. (2012). Field-based phenomics for plant genetics research. *Field Crops Research*, *133*, 101–112. doi.org/10.1016/j.fcr.2012.04.003

Woods, S. J., & Swearingin, M. L. (1977). Influence of Simulated Early Lodging upon Soybean Seed Yield and its Components. *Agronomy Journal*, *69*(2), 239–242. doi.org/10.2134/agronj1977.00021962006900020011x

Wrather, J. A., Stienstra, W. C., & Koenning, S. R. (2001). Soybean disease loss estimates for the United States from 1996 to 1998. *Canadian Journal of Plant Pathology*, *23*(2), 122–131. doi.org/10.1080/07060660109506919

Xavier, A., Hall, B., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2017). Genetic Architecture of Phenomic-Enabled Canopy. *Genetics*, *206*(June), 1081–1089. doi.org/10.1534/genetics.116.198713/-/DC1.1

Xie, C., & Yang, C. (2020). A review on plant high-throughput phenotyping traits using UAV-based sensors. *Computers and Electronics in Agriculture*, *178*. doi.org/10.1016/j.compag.2020.105731

Ye, H., Song, L., Schapaugh, W. T., Ali, M. L., Sinclair, T. R., Riar, M. K., Mutava, R. N., Li, Y., Vuong, T., Valliyodan, B., Pizolato Neto, A., Klepadlo, M., Song, Q., Grover Shannon, J., Chen, P., Nguyen, H. T., & Foyer, C. (2020). The importance of slow canopy wilting in drought tolerance in soybean. *Journal of Experimental Botany*, *71*(2), 642–652. doi.org/10.1093/jxb/erz150

Yu, N., Li, L., Schmitz, N., Tian, L. F., Greenberg, J. A., & Diers, B. W. (2016). Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sensing of Environment*, *187*, 91–101. doi.org/10.1016/j.rse.2016.10.005

Zhang, A. W., Riccioni, L., Pedersen, W. L., Kollipara, K. P., & Hartman, G. L. (1998). Molecular identification and phylogenetic grouping of *Diaporthe phaseolorum* and *Phomopsis longicolla* isolates from soybean. *Phytopathology*, *88*(12), 1306–1314. doi.org/10.1094/PHYTO.1998.88.12.1306

.

CHAPTER 2

GENETIC MAPPING AND CONFIRMATION OF THE *RDM3* LOCUS UNDERLYING RESISTANCE TO SOUTHERN STEM CANKER IN SOYBEAN

[1]Ethan Menke, M. Habib Widyawan, Nicole Bachleda, James W. Buck, Zack Shearin, H.R. Boerma, and Zenglu Li. To be submitted to *Crop Science*

Abstract

Southern stem canker (SSC) in soybean, caused by the fungal pathogen *Diaporthe aspalathi*, is a major disease in the southern United States. It can cause up to 80% yield loss in severely infected fields. PI 398469, a maturity group VI soybean accession from South Korea, is one of six known resistance sources to SSC. To determine the genetics of the resistance provided by PI 398469, $F_{2:3}$ and F_5-derived recombinant inbred line (RIL) populations were developed from a cross of susceptible G81-2057 × PI 398469. $F_{2:3}$ and RIL populations were phenotyped in the greenhouse by inoculating individual plants with toothpicks colonized with *D. aspalathi*. SSC-resistant and susceptible bulks were created from the $F_{2:3}$ population. The bulks consisted of 20 resistant and 18 susceptible families, respectively, and were genotyped using the Infinium SoySNP50K BeadChips for a bulked segregant analysis (BSA). Based on BSA results, 22 KASP markers were designed to genotype the RIL population. Quantitative trait locus (QTL) analysis identified *qRdm_14*, on chromosome (Chr) 14 with an R^2 of 62%. The $F_{2:3}$ population confirmed the presence of the QTL on Chr 14. RILs derived from G81-2057 × Crockett were phenotyped for SSC resistance and genotyped with our KASP markers to examine the relationship between *Rdm3* from Crockett and *qRdm_14* from PI 398469. PI 398469 and Crockett were found to carry alleles at the *Rdm3* locus for resistance to SSC. The QTL and KASP marker information could be used to screen breeding lines for resistance to southern stem canker.

Keywords: Soybean, disease resistance, southern stem canker, quantitative trait locus, *Diaporthe aspalathi*, *Rdm3*

Introduction

Southern Stem Canker (SSC) of soybean [*Glycine max* (L.) Merrill] is caused by the fungus *Diaporthe aspalathi* (syn. *Diaporthe phaseolorum var. meridionalis*). *D. aspalathi* is an ascomycetous fungus that reproduces sexually (Zhang et al., 1998). Historically, *D. aspalathi* was referred to as *D. phaseolorum var. meridionalis* until recently when it was differentiated from other *Diaporthe* species and renamed (Janse Van Rensburg et al. 2006).

SSC was first observed in southern U.S. soybean production in 1973 in Mississippi (Backman et al., 1985). By 1983, the pathogen had moved across the southern U.S. and was present in most soybean-producing areas. Stem canker across the U.S. was responsible for yield losses of 1.6 million metric tons of soybean between 2015 and 2019 (Bradley et al., 2021). Significant crop damage was observed with susceptible cultivars having up to 80% yield loss (Krauz &Fortnum, 1983). SSC has also become a major problem for soybean producers in South America (Wrather et al., 1997). In 1998 1.4 million metric tons of soybean yield loss was experienced in Argentinian, Brazil, and Bolivia (Wrather et al., 2001).

Stem canker is named after the lesions or canker caused by *D. aspalathi*. The pathogen overwinters on infected soybean debris and new infections occur in spring when inoculum moves from debris to wounded stems or petioles during irrigation or rainfall events. The fungus remains quiescent for most of the vegetative growth stages (VE-V7) of the plant and symptoms don't appear until the early reproductive stage R1 (Fehr et al., 1971). Cankers appear as small slightly sunken areas with a reddish-brown color on the lower nodes of the plant (Rupe, 2015). As disease progression continues, the lesion increases in size longitudinally and becomes more sunken. The pith and vascular tissues turn brown and discolored, but the tissue around the canker typically

remains green and the stem is rarely girdled. Eventually, the pathogen will block the upward flow of water to the plant causing water stress and premature plant death (Hildebrand, 1952).

Foliar symptoms of SSC are most commonly observed as interveinal chlorosis and eventual necrosis. These symptoms are caused by the production of phytotoxins by *D. aspalathi* (Lalitha et al., 1989). These symptoms are similar to Fusarium wilt and other soybean diseases and can result in misidentification of SSC in the field. The detrimental effects of SSC on soybean yields have been limited in the southern U.S. by deploying cultivars with several sources of genetic resistance in addition to changes in production techniques such as the increased use of foliar fungicide and seed treatments (Backman et al., 1985). 'Tracy-M' was the first cultivar identified as being resistant to SSC under field conditions across the southern U.S. (Kilen et al., 1985).

Kilen et al. (1985) developed a population from a cross of Tracy-M × J77-339, a susceptible breeding line. Based on the phenotyping results, both the F_2 generation and F_3 families segregated in a manner that indicated the Tracy-M contained two major dominant genes for resistance to SSC. D85-10404 and D85-10412 are experimental lines derived from Tracy-M and each was believed to contain a single resistance gene. By crossing D85-10404 and D85-10412 to J77-339, Kilen and Hartwig (1987) developed two F_2 populations, confirming that both lines possess a single resistance gene. D85-10404 and D85-10412 were also crossed with each other to produce an $F_{2\ 3}$ population, which segregated in a 15:1 ratio, indicating that each experimental line contained a different resistance gene that was named *Rdm1* (D85-10404) and *Rdm2* (D85-10412) (Kilen & Hartwig, 1987).

Two additional sources of SSC resistance were identified in the cultivars 'Crockett' and 'Dowling' through field evaluations (Backman et al., 1985; Bowers, 1990). After examining the relationship between the genes controlling SSC in Crockett, Dowling, and Tracy-M, Bowers et al.

(1993) concluded that Crockett and Dowling each possessed a single locus that provided resistance to SSC. These loci were different from *Rdm1* and *Rdm2* present in Tracy-M and were then named *Rdm3* from Crockett and *Rdm4* from Dowling (Bowers et al., 1993).

Pioli et al. (2003) investigated the differential response of soybean cultivars to isolates from the *Diaporthe*/Phomopsis complex collected across North and South America. They identified isolates that caused disease in the Dowling (*Rdm4*) background, but not on 'Hutcheson', suggesting that an additional resistance gene may be present in Hutcheson (Pioli et al., 2003). Chiesa et al. (2017) confirmed the presence of two independent resistance genes in the $F_{2:3}$ population derived from J77-339 × Hutcheson by using a *D. aspalathi* isolate that was virulent on *Rdm4*. The results indicated that Hutcheson carries two non-allelic genes that provide resistance to SSC and the authors named the second gene as *Rdm5* (Chiesa et al., 2017). Two additional sources of resistance to *D. aspalathi*, PI 398469 and PI 230976 were found in the USDA Soybean Germplasm Collection (Tyler, 1995). Crosses between PI 398469 or PI 230976 and the susceptible breeding line J77-330 were made to produce two segregating populations. Segregation analyses indicated that both plant introductions contain a single dominant resistance gene for resistance to *D. aspalathi*. PI 398469 and PI 230976 were then crossed with D85-10404, D85-10412, Crockett, and Dowling and the results of this study indicated that PI 398469 and PI 230976 carry genes that are not allelic with *Rdm1, 2, 3*, and *4* (Tyler, 1995). Crockett (*Rdm*3) and PI 398469 (*Rdm?*) were then crossed to a susceptible line J77-339 to produce $F_{2:3}$ populations and both QTLs were roughly mapped using SSR markers to a similar region on Chr 14 (Shearin, 2011). However, this study could not determine if *Rdm3* and *Rdm?* were at the same locus.

To identify resistance loci to SSC, Chang et al. (2016) performed a genome-wide association study (GWAS) using 112 accessions from USDA Soybean Germplasm Collection that

were phenotyped by Ferrari de Novoa and Scandiani (2009) and genotyped with the Infinium SoySNP50K BeadChip. A genomic region on Chr 14 was identified that provided resistance to SSC with a peak at the SNP marker ss71561795 (Gm14_1942681_T_C). Using 295 soybean accessions from Asia, and North and South America, Maldonado Dos Santos et al. (2019) also performed a GWAS for resistance to SSC. These accessions were phenotyped with a Brazilian strain of *D. aspalathi*, using a toothpick inoculation method. The plants were rated for resistance based on the percentage of dead plants 60 days after the inoculation. These accessions were genotyped using a genotyping by sequencing method. GWAS along with the haplotype analysis identified one significant genomic region located at 1,744,370 bp on Chr 14 that provided resistance to SSC.

The objectives of this research were to: i) understand the resistance to SSC provided by PI 398469; ii) develop markers that can be used by breeders to select for resistance to SSC provided by PI 398469, and iii) determine if Crockett and PI 398469 carry the same QTL for resistance to SSC.

Materials and Methods

Plant materials and population development

PI 398469 is a maturity group VI plant introduction (PI) that was collected from South Korea in 1975. It was identified as resistance to SSC through evaluation of the USDA Soybean Germplasm Collection (Tyler, 1995). PI 398469 was crossed with G81-2057 to develop mapping populations. ‘G81-2057’ is a breeding line developed at the University of Georgia (UGA) that was derived from a cross of R75-120 × J74-35. It was identified as being susceptible to SSC through

multiple disease screenings (Campbell et al., 2017). G81-2057 has been also used by the UGA breeding program as a susceptible check for evaluation of resistance to SCC.

The cross between G81-2057 and PI 398469 was made in the summer of 2017 at the UGA J. Phil Campbell Sr. Research and Education Center in Watkinsville, Ga. F_1 seeds were advanced in the Illinois Crop Improvement Winter Nursery in Puerto Rico during the winter of 2017-2018. During the summer of 2018, the F_2 seeds were planted at the UGA Soybean Breeding Program's nursery at the Iron Horse Plant Science Farm, Watkinsville, GA. In the fall of 2018, 175 individual F_2 plants were harvested to create an $F_{2:3}$ population. In addition to single F_2 plants, the population was also advanced for the F_3 and F_4 generations using a modified single-seed descent method in the Illinois Crop Improvement Winter Nursery in Puerto Rico during the winter of 2018 - 2019. In the 2019 growing season, the F_5 seed was planted at the UGA Soybean Breeding Program's nursery and individual plants were harvested to generate 224 F_5-derived recombinant inbred lines (RILs) for QTL mapping.

Similarly, to map the *Rdm3* locus from Crockett, a set of 153 RILs was developed from a cross of G81-2057 × Crockett at UGA in 2019. Using a similar method as described above the population was advanced through single seed descent until the F_5 where individual plants were harvested and threshed individually to produce RILs. Crockett was developed from a cross of 'Hampton 266' × PI 171451 by Texas A&M University in 1978 (Bowers, 1990).

Phenotyping for resistance to southern stem canker in the greenhouse

The 175 $F_{2:3}$ families derived from G81-2057 × PI 398469 and 224 F_5-derived RILs derived from the same parents were evaluated in a randomized complete block design with six replicates each containing two plants. The 153 RILs derived from the G81-2057 × Crockett cross were also

evaluated in a randomized complete block design with four replications each containing two plants. Experiments included resistant checks Tracy-M (*Rdm1/2*), Crockett (*Rdm*3), Dowling (*Rdm4*), and Hutcheson (*Rdm4/5*), and a susceptible check, J77-339 along with the parents G81-2057 and PI 398469 of the population. The three populations were evaluated at the UGA Plant Pathology Greenhouses in Griffin, GA. Lines from the $F_{2:3}$ families and both RIL populations were grown in 15 10 × 10 cm-cell flats (Kord Presto, Riverhead, NY) filled with SunGro Metro-Mix 830 potting mix (Agawam, Ma). The 12 outside cells of each flat were used, with the three middle cells being left vacant to promote better airflow and light penetration (Harris et al., 2015). Three seeds of each line were planted in a pot. After three weeks each plot was thinned down to two seedlings, and then the seedlings were supported by wooden stakes to prevent mechanical damage caused by inoculation. Seedlings grew under 12 hr of supplemental light and were watered 3 to 4 times each week and fertilizer was supplied to the plants weekly at a rate of 200 µg mL^{-1} from a stock of 20-20-20 Scotts water-soluble fertilizer (The Scotts Company, Marysville, OH).

Diaporthe aspalathi BL 22-14 was isolated from susceptible soybean grown in the UGA soybean stem canker nursery in Bledsoe, Ga in 2014 and identified by ITS sequence (Campbell et al. 2017). The isolate was prepared for long-term storage by lyophilizing infected stem tissue, and short-term storage was accomplished by storing inoculant on PDA slants at 4°C, a working stock of inoculant was prepared by transferring the isolate onto fresh PDA every 2-3 weeks.

A modified toothpick assay was used to inoculate soybean seedlings (Campbell et al. 2017, Keeling, 1982) Briefly, toothpicks were halved and autoclaved twice in fresh distilled water and placed in petit dishes containing potato dextrose agar (PDA). The plates were seeded in the middle of the toothpicks with a mycelium plug of *D. aspalathi* grown on PDA for two weeks and incubated at room temperature (~22°C) on a laboratory bench for three weeks to allow for the colonization

of the toothpicks by the fungi. Approximately three weeks after planting, seedlings were wounded with a dissecting needle above the cotyledons, a colonized toothpick was placed in the wound and it was sealed with petroleum jelly (Campbell et al. 2017). After three weeks plants were rated for the disease on a 1 to 5 scale (1 = no disease, 2 = discoloration of cortex, 3 = visible stem canker or lesions, 4 = stem lesions, foliar interveinal chlorosis and necrosis, and 5 = dead plant) (Fig. 2.1).

Bulked Segregant Analysis

Leaf tissue was collected from the parents and $F_{2:3}$ plants during phenotyping. A resistant bulk, consisting of 20 families which were highly resistant to SSC based on greenhouse screenings, was created. Similarly, a susceptible bulk was formed by combining leaf tissue of 18 extremely susceptible families from greenhouse screening. A modified CTAB protocol optimized for genotyping with the Infimum SoySNP50K BeadChip was used to extract genomic DNA from parents and bulks (Keim et al., 1988; Song et al., 2013). The resistant and susceptible bulks along with two parents were genotyped with the SoySNP50K iSelect BeadChip at the Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD. SNP genotypes were called using GenomeStudio 2.0 (Illumina, San Diego, U.S.). Bulked segregant analysis (BSA) was used to identify the putative genomic regions that were responsible for resistance to SSC. These genomic regions were identified by comparing the two parental genotypes to identify the polymorphic SNPs between G81-2057 and PI 398469 and by examining differences in the resistant and susceptible bulks.

KASP marker design and genotyping

Bulked leaf tissue was collected from each of the $F_{2:3}$ families and RILs during the greenhouse evaluation. Leaf tissue was lyophilized for 24 h in 50 mL falcon tubes and ground into a fine powder using a Spex Sample Prep Geno/Grinder (SPEX, Metuchen, NJ). Genomic DNA was extracted using the protocol that was previously mentioned. Polymorphic SNPs between G81-2057 and PI 398469 that were identified in the BSA were selected to design KASP markers for mapping QTLs for resistance to SSC.

KASP markers were designed with FAM dye-labeled forward primers for the selection of alleles from PI 398469 and VIC dye-labeled primers were designed for the selection of alleles from G81-2057 based on SoySNP50K data. The assays were performed by following the protocol of Pham et al. (2015) for genotyping RILs derived from G81-2057 × PI 398469. Briefly, a 4 μL reaction consisted of 2 μL of KASP master mix, 0.05 μL of primer mix, and 2 μL of genomic DNA (>10ng/ μL). A KASP PCR protocol was used as follows: 94°C for 15 min, followed by 10 cycles of touch down PCR from 65°C to 57°C with a 0.8°C decrease per cycle, followed by 30 cycles of 94°C for 20 s and 57°C for 1 min. A Roche LightCycler 480 (Roche Diagnostic Corporation, Basel, Switzerland) was used for endpoint reading and performing allele calls for each of the KASP markers. The same set of markers with the same protocol mentioned above was used to genotype $F_{2:3}$ families derived from G81-2057 × PI 398469 for QTL confirmation. Only SNPs polymorphic between G81-2057 and Crockett were selected to genotype the RILs derived from G81-2057 × Crockett.

QTL analysis

Disease scores were analyzed using the JMP 16 Pro to fit a mixed model where both genotype and replication were treated as a random variable. The broad-sense heritability was estimated using the formula $H^2 = \frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2}$ where H^2 represents the broad-sense heritability, σ_g^2 represents the genotypic variance, and σ_e^2 represents the error.

Single marker analysis was performed using the Lme4 function in R (Bates et al., 2015). The significance of SNP markers was tested by performing a linear regression with genotype treated as a fixed effect. Based on the genotyping results, a linkage map of the RILs derived from G81-2057 × PI 398469 was constructed using JoinMap 4.1 software (van Ooijen, 2006). The linkage between markers was analyzed using the regression function and the recombination distances were calculated using the Kosambi mapping function. Using the genotypic least-square means, QTL mapping was performed using the composite interval mapping function of Windows QTL Cartographer 2.5 with a walking speed of 0.5 cM, two control markers, and a window size of 5 cM (Wang et al., 2012). The same methods were used to construct linkage maps and analyze the QTLs in the $F_{2:3}$ population of G81-2057 × PI 398469 and the RIL population derived from G81-2057 × Crockett.

Candidate gene analysis within a QTL region

The peak SNPs from the QTL region identified from the RIL and $F_{2:3}$ populations were used for the identification of candidate genes responsible for resistance to SSC. Candidate genes and their functions were identified in SoyBase by using the Glyma2.1 gene models that were located within 10 kb of the peak SNP's physical location (Soybase.org). Genome annotation and

gene ontological (GO) terms related to disease resistance, stress tolerance, pathogen tolerance, and cell signaling were used to determine the candidate genes that were examined in this study.

Results

Phenotypic reactions to D. aspalathi

Resistant and susceptible checks showed consistently, expected phenotypic reactions across tests and reps, indicating the test data was good for further mapping work. The mean SSC score of the RILs derived from G81-2057 × PI 398469 and the $F_{2:3}$ population was 2.56 with a range of 1.1 to 4.8 and 2.17 with a range from 1.2 to 4.5, respectively (Table S2.1 and Fig. S2.1). For both populations, genotype, but not replication, explained a significant portion of the variance that was observed in the response to SSC at a p-value of 0.001 (Table 2.1). Any differences that were observed in replications to SSC are likely due to the differences in raters in a given replication. The broad-sense heritability was estimated at 0.69 and 0.45 for the RIL and $F_{2:3}$ populations, respectively (Table 2.1). These results indicated that resistance to SSC is a highly heritable trait, disease evaluations are stable across replicates, and our phenotypes are reliable to map QTLs for resistance to SSC.

Bulked segregant analysis

Through BSA, eight informative genomic regions were identified on Chrs 2, 5, 9, 10, 11, 12, 14, and 20, respectively, and 1-2 KASP markers were developed for each of these genomic regions to map the QTLs for resistance to SSC. SNP markers that were tested on Chrs 2, 5, 11, 12, and 20 through single marker analysis were found not to be significantly associated with resistance to SSC. Initial testing of SNP markers on Chrs 9, 10, and 14 were found to be significantly

associated with resistance to SSC (Table 2.3), and then based on SoySNP50K data, seven additional KASP SNP marker assays were designed in the QTL regions to saturate the linkage map for further QTL analysis

QTL mapping in the RIL population

Based on the genomic region identified in the BSA, 12 KASP markers that were polymorphic between G81-2057 and PI 398469 were designed to elucidate the underlying QTL on Chr 14 associated with SSC resistance. Single marker analysis indicated that these markers were significantly associated with resistance to SSC ($p = 0.05$) (Table 2.3). A linkage map with a length of 23.3 cM was generated (Fig 2.2a). QTL analysis revealed a significant association with resistance to SSC at a significance level of $\alpha=0.05$ based on 1000 permutation tests. A single major QTL named *qRdm-14*, was located in the region between markers GSM_933 and GSM_921 on Chr 14 spanned 2.7 Mb (Fig. 2.2a). The peak of this QTL was identified at marker GSM_975 with a LOD score of 47.2. The QTL *qRdm_14* has an additive allele effect of -0.97, indicating that the allele from PI 398469 contributed to the SSC resistance (Table 2.2). The 62% of the phenotypic variation in the SSC score in the RIL population was accounted for by *qRdm_14*, indicating that it is a major QTL that provided resistance to SSC (Table 2.2). RILs that possess the GG alleles of SNP marker GSM_975 from PI 398469 lowered the SSC score by 2.3, or more than 2.5 times lower than lines with the AA alleles from G81-2057 (Fig. 2.3a).

Four KASP markers were designed in the putative genomic region on Chr 9 identified in the BSA to create a linkage map. Of these four SNP markers, GSM_929 and GSM_930 were significantly associated with resistance to SSC through single marker analysis (Table 2.1). The linkage group generated by the four KASP markers spanned 44.3 cM, however, no significant

QTLs were detected on Chr 9 based on interval mapping. Similarly, five KASP markers were developed on Chr 10 to span the region that was identified through BSA. Based on the single-marker analysis, both GSM_923 and GSM_1006 were significantly associated with SSC resistance (p=0.05) (Table 2.3). The map spanned 22.4 cM on Chr 10 based on linkage analysis of the five markers. Similarly, interval mapping did not detect any QTL in this region.

QTL confirmation with the $F_{2:3}$ population

KASP markers used in QTL mapping in the RIL population were also used to genotype $F_{2:3}$ families to confirm the QTL, *qRdm-14*, identified in the RIL population. Single-marker analysis indicated that all 12 KASP markers were significantly associated with SSC resistance in the $F_{2:3}$ population (p= 0.05). KASP markers on Chr 14 formed a linkage group that spanned 26.2 cM. QTL analysis with the composite interval mapping detected the QTL with a LOD score of 9.1. In this $F_{2:3}$ population, the QTL was mapped between GSM_942 and GSM_932 on Chr 14, which spanned 3.9 cM (Fig. 2.2b). The peak of the QTL was located at GSM_943 (Fig. 2.2b), which is located downstream of 0.6 cM from GSM_975, the peak SNP mapped in the G81-5057 x PI398469 RIL population. This genomic location overlaps the region that was identified in the RILs, confirming the presence of a QTL on Chr 14. The QTL accounted for 21% of the phenotypic variation observed in the $F_{2:3}$ families and had an additive effect of -0.36, indicating that $F_{2:3}$ families possess alleles from PI 398469 that provide resistance to SSC (Table 2.2). Similar to the RIL population, possessing the GG alleles from PI 398469 at SNP GSM_943 reduced the SSC score by half when compared to the susceptible alleles AA from G81-2057 (Fig. 2.3b).

Determination of the Rdm 3 locus in Crockett and PI 398469

The disease ratings of the RILs derived from G81-2057 × Crockett ranged from 1.0 to 4.75 with a population mean of 2.29. Of the 12 KASP markers on Chr 14 that were used for *qRdm-14* mapping in the G81-2057 × PI 398469 RIL population, seven SNPs were found to be polymorphic between G81-2057 and Crockett. These seven markers were used to genotype the G81-2057 × Crockett RIL population and generate a linkage group that spanned 10.2 cM (Fig. 2.2c). QTL analysis detected a single QTL located in the regions between GSM_943 and GSM_975, spanning 1.1 cM (Fig. 2.2c). The peak marker of this QTL was GSM_943 which had a LOD score of 26.6 (Table 2.2 and Fig. 2.2c). This QTL explained 57% of phenotypic variation and had an additive allelic effect of -0.74, indicating that the resistance was coming from Crockett. CIM mapping in the RIL population derived from G81-2057 × Crockett indicated that both GSM_943 and GSM_975 were significantly associated with resistance to SSC. These markers were also the peak markers in the G81-2057 × PI 398469 RIL and $F_{2:3}$ populations, which confirms that PI 398469 and Crockett both possess the resistance alleles at the *Rdm3* locus.

Candidate genes at the QTL region for resistance to southern stem canker

A total of 64 gene models were annotated on the Williams 82.a2 reference genome between the SNP markers GSM_942 (1,316,725 bp) and GSM_932 (1,767,144 bp) based on a current SoyBase search. Five gene models were located in this region within 10 Kb of the peak SNPs GSM_943 and GSM_975 identified in three mapping populations and were related to disease resistance (Table 2.4). These five gene models were selected based on annotations related to disease resistance. The gene model *Glyma.14g020100* was annotated as an Arabidopsis Ethylene-Response factor 3 involved in DNA binding, while the gene model *Glyma.14g022600* was

annotated as a DNA-binding bromodomain-containing protein. *Glyma.14g022900* was annotated as an RNI – like superfamily protein go term related to leucine-rich repeats (LRR). The remaining two gene models were related to cell wall modification (*Glyma.14g020300*) and ribosomal regulations (*Glyma.14g020500*).

Discussion

Identification and confirmation of the QTL from PI 398469

SSC resistance inherited from PI 398469 was highly heritable, with an H^2 = 0.70 in the RILs, and H^2 = 0.56 in the $F_{2:3,}$ indicating that the SSC resistance can easily be selected for in a breeding program. Using the BSA, a putative genomic region on Chr 14 was identified that was associated with resistance to SSC. Based on SNPs that were polymorphic between the parents of the population, 12 KASP markers from that genomic region on Chr 14 were developed to map and confirm the QTL *qRdm-14* for resistance to SSC using both RIL and $F_{2:3}$ populations. The QTLs detected on Chr 14 for the QTL underlying SSC resistance in both RIL and $F_{2:3}$ populations were overlapping and the peak SNP markers are separated by ~170 Kb based on the physical position of the SNPs on the William82.a2 soybean reference genome. Based on the recombination observed in RILs the peak marker GSM_975 was located approximately 2.1 cM from GSM_943, which was identified as the peak marker in the $F_{2:3}$ population. The recombination between these two markers was at 0.6 cM based on the $F_{2:3}$ population linkage map. In addition, the SNP markers GSM_975 and GSM_943 in both RIL and $F_{2:3}$ populations met the threshold LOD scores, indicating that they were all significantly associated with resistance to SSC. With the linkage maps being made based on the recombination of two separate populations developed from the same initial cross, a slight difference in the peak markers for the QTLs is most likely due to the difference in recombination

that was observed in the RIL and $F_{2:3}$ populations. The same phenomenon was also observed in the study by Yi et al. (2018) for tassel-related traits in maize where the QTL mapping was performed in both RIL and $F_{2:3}$ populations. The authors stated that the discrepancies in QTL locations in this study were due to different recombination in the two populations (Yi et al., 2018).

Additionally, a larger number of heterozygous plants in the $F_{2:3}$ population were phenotyped and their mixed phenotypic reactions could skew the mapping results slightly. While there are slight discrepancies in the linkage maps that were generated in this study they have a similar recombination pattern. In both RIL and $F_{2:3}$ populations, multiple additional markers, GSM_921, GSM_932, and GSM_973 were located in the QTL interval. The commonality of the results for both RIL and $F_{2:3}$ populations derived from the cross of G81-2057 × PI 398469 can be used to conclude that a single major QTL, named *qRdm-14*, from PI 398469 on Chr 14 was providing resistance to SSC.

Rdm3 resistance alleles present in both Crockett and PI 398469

Using an F_2 population derived from Crockett (*Rdm3*) × PI 398469 (*Rdm?*) that segregated in a 15:1 ratio (resistant vs susceptible) Tyler (1995) indicated that two major genes were segregating in this population and each resistance source might carry an independent gene for resistance. Previous mapping work performed by the UGA soybean breeding program indicated that the *Rdm*3 gene from Crockett was located on Chr 14, which was mapped to the similar genomic location of the major QTL from PI 398469 (data unpublished). In this study, KASP markers designed for the G81-2057 × PI 398469 populations were used to genotype the G81-2057 × Crockett RIL population to understand the relationship between the two resistance sources. A single QTL was identified in both populations in a common genomic region on Chr 14. This

confirmed that the *Rdm3* QTL from Crockett is located on Chr 14 and PI 398469 carries a resistance allele at the *Rdm3* locus that provides resistance to SSC. However, from this study, we could not conclude if both Crockett and PI 398469 carry the same resistance allele at the *Rdm3* locus on Chr 14. Further studies by sequencing the locus could provide evidence.

Comparisons with other reported QTLs

The mapping results from this study confirm the presence of a QTL, *Rdm3* on Chr 14 of soybean from the resistance sources PI 398469 and Crockett. Chang et al. (2016) performed a GWAS for SSC resistance using 112 accessions from USDA Soybean Germplasm Collection that have previously been phenotyped for SSC and had genotypic information from the SoySNP50k BeadChip (Chang et al., 2016). The GWAS identified two SNPs, ss7156178 (1731256 bp) and ss715617951 (1938019 bp) on Chr 14 that were associated with resistance to northern and southern stem canker, respectively. SNP marker ss715617951 was located approximately 3.3 Kb downstream of GSM_975 based on the physical position on Williams 82.a2 reference genome. The SNP marker ss715617951 was not used in this study due to the region being unfavorable for the KASP marker design.

Maldonado Dos Santos et al. (2019) also found a single locus on Chr 14 that was significantly associated with resistance to SSC based on their GWAS results. This GWAS was performed using 295 accessions from around the world which were phenotyped with *D. aspalathi* from Brazil (Maldonado Dos Santos et al., 2019). GSM_921 in our study was designed at the position of 1,744,370 bp on Chr 14 based on the Williams 82 assembly a2, which was a significant marker from the GWAS results reported by Maldonado Dos Santos et al. (2019). Based on the haplotype analysis, the authors also stated that the QTL on Chr 14 accounted for the resistance to

SSC from PI 398469 (*Rdm?*), Crockett (*Rdm3*), and Hutcheson (*Rdm5*). However, Cheisa et al. (2017) mapped both *Rdm4* and *Rdm5* in Hutcheson to Chr 8 and the loci are separated by 17.6 cM. The result from Maldonado Dos Santos et al. (2019) indicated that the *Rd*m5 locus might not be located on Chr 14. The *QTL Rdm3* identified in both G81-2057 × PI 398469 RIL and $F_{2:3}$ populations and the G81-2057 × Crockett RIL population from this study are in alignment with the results from Chang et al. (2016) and Maldonado Dos Santos et al. (2019). In addition, a GWAS using the UGA advanced breeding lines trials (unpublished data) identified a significant association with resistance to SSC in the same genomic region of the *Rdm3* on Chr 14 that was mapped in this study. This indicates the prevalence of the *Rdm3* as a major resistance source in elite southern soybean germplasm in the US and many soybean accessions from around the world.

Candidate genes at the QTL region for resistance to southern stem canker

A large number of candidate genes were observed in the QTL region, but the focus was placed on the gene models with annotations related to disease resistance and cell signaling which would most likely be related to SSC resistance. Within 10 kb up or down stream of GSM_943 (1,444,397 bp), a peak marker at the *Rdm3* locus, three gene models of interest were identified based on the gene annotations of the William82.a2 reference genome on SoyBase (Soybase.org, 2022). *Glyma.14g020100* is annotated as an ethylene response transcription factor (ERF) which plays a role in a plant's response to both abiotic and biotic stresses (Thirugnanasambantham et al., 2015). In multiple other plant species such as tobacco, rice, and Arabidopsis *(Arabidopsis thaliana)*, ERF is involved in the regulation of plant defenses against fungal pathogens (Thirugnanasambantham et al., 2015). The second gene model *Glyma.14g020300* is annotated as pectinesterase, which is involved in cell wall modification. In soybean, it has been shown to

potentially play a role in resistance to southern root-knot nematode (Pham et al., 2013). *Glyma.14g020500* is the third gene model annotated as a Ribosomal S17 family protein, which plays a key role in protein translation and assembly. In rice, Ribosomal S17 family proteins were shown to have increased expression in response to abiotic and biotic stress (Saha et al., 2017).

In the 20 kb region surrounding the marker GSM_975, two gene models with annotations related to disease resistance were present. A DNA-binding bromodomain-containing protein is the predicted gene model for *Glyma.14g022600*. DNA binding bromodomain proteins work in conjunction with nucleotide-binding leucine-rich repeats (NLR) to control immune signaling into potato pathogens (Sukarta et al., 2020). *Glyma.14g022900* is annotated as RNI-like superfamily protein/F-Box/Leucine-Rich Repeat Protein, and these proteins have been shown to have increased expression in plants under stress (An et al., 2016). Genes in the region surrounding GSM_943 and GSM_975 identified in this study would be good targets for further gene expression studies. Further study of these gene models could provide the underlying molecular mechanism that is providing resistance to SSC and help develop functional markers for effective selection of SSC resistance.

KASP markers for marker-assisted selection

Phenotyping for SSC is a tedious process, requiring individual toothpicks to be inserted into each of the plants being screened and care to ensure that the stems of plants are not broken and the toothpicks do not fall out before disease symptoms develop. Phenotyping for SSC also requires significant greenhouse space for 6-8 weeks per experiment depending on disease progression. The KASP markers identified for the *Rdm3* locus are effective in separating SCC phenotypes and could be used to select breeding lines for resistance to SSC. As an example,

GSM_975 and GSM_943 separate the resistant and susceptible RILs with 94% accuracy (Fig. 2.3a,b). In each population, possessing the homozygous alleles at GSM_975 and GSM_943 loci from PI 398469 reduced the mean SSC score by 50% as compared to the heterozygous or homozygous alleles from G81-2057. By analyzing the UGA elite soybean germplasm in the breeding pipeline using a GWAS approach, the same highly significant major QTL on Chr 14 was identified (unpublished data), indicating that the QTL, *Rdm3* provides an important role in SSC resistance. However, to confirm the effectiveness of these markers for selection in non-related material, additional testing might be needed. These markers identified in this research will provide a simple and effective way to screen a large number of breeding lines for potential resistance to SSC and select lines that contain a desired resistant genotype for advancement. This would make the breeding process more efficient by reducing the number of lines in greenhouse evaluation each year to just the lines that carry the resistance alleles.

Conclusion

Genetic mapping using RILs identified a major QTL, *qRdm-14*, on Chr 14 from PI 398469 that is responsible for resistance to SSC. The QTL was confirmed with the $F_{2:3}$ population derived from the same parents. In both RIL and $F_{2:3}$ populations, possessing alleles from PI 398469 significantly reduced SSC disease. To better understand the relationship between SSC resistance from PI 39869 and Crockett, SNP markers developed from the G81-2057 × PI 398469 RILs were used to map *Rdm3* in Crockett. Both PI 398469 and Crockett possess resistance alleles at the *Rdm3* locus for resistance to SSC. The peak KASP markers GSM_975 and GSM_943 tightly linked to the *Rdm3* locus will be useful for screening large numbers of lines quickly with simple KASP genotyping.

References

An, J. P., Li, R., Qu, F. J., You, C. X., Wang, X. F., & Hao, Y. J. (2016). Apple F-box protein MdMAX2 regulates plant photomorphogenesis and stress response. *Frontiers in Plant Science, 7*, 1685.doi.org/10.3389/fpls.2016.01685

Backman, P., Weaver, D., & Morgan-Jones, G. (1985). Soybean Stem Canker: An Emerging Disease Problem. *Plant Disease*, *69*(8), 641–647.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi.org/10.18637/jss.v067.i01

Bowers, G. R. (1990). Registration of 'Crockett' Soybean. *Crop Science*, *30*(2), 427. doi.org/10.2135/cropsci1990.0011183x003000020049x

Bowers, G. R., Ngeleka, K., & Smith, O. (1993). Inheritance of Stem Canker Resistance in Soybean Cultivars Crockett and Dowling. *Crop Science*, *33*(1), 67–70. doi.org/10.2135/cropsci1993.0011183x003300010010x

Bradley, C. A., Allen, T. W., Sisson, A. J., Bergstrom, G. C., Bissonnette, K. M., Bond, J., Byamukama, E., Chilvers, M. I., Collins, A. A., Damicone, J. P., Dorrance, A. E., Dufault, N. S., Esker, P. D., Faske, T. R., Fiorellino, N. M., Giesler, L. J., Hartman, G. L., Hollier, C. A., Isakeit, T., … Wise, K. A. (2021). Soybean Yield Loss Estimates Due to Diseases in the United States and Ontario, Canada, from 2015 to 2019. *Plant Health Progress*, *22*(4), 483–495. doi.org/10.1094/PHP-01-21-0013-RS

Campbell, M. A., Li, Z., & Buck, J. W. (2017). Development of southern stem canker disease on soybean seedlings in the greenhouse using a modified toothpick inoculation assay. *Crop Protection*, *100*, 57–64. doi.org/10.1016/j.cropro.2017.05.026

Chang, H. X., Lipka, A. E., Domier, L. L., & Hartman, G. L. (2016). Characterization of disease resistance loci in the USDA soybean germplasm collection using genome-wide association studies. *Phytopathology*, *106*(10), 1139–1151. doi.org/10.1094/PHYTO-01-16-0042-FI

Chiesa, M. A., Cambursano, M. v., Pioli, R. N., & Morandi, E. N. (2017). Molecular mapping of the genomic region conferring resistance to soybean stem canker in Hutcheson soybean. *Molecular Breeding*, *37*(5), 1-12. doi.org/10.1007/s11032-017-0660-6

Fehr, W. R., Caviness, C. E., Burmood, D. T., & Pennington, J. S. (1971). Stage of Development Descriptions for Soybeans, *Glycine max* (L.) Merrill. *Crop Science*, *11*(6), 929–931. doi.org/10.2135/cropsci1971.0011183x001100060051x

Harris, D. K., Kendrick, M. D., King, Z. R., Pedley, K. F., Walker, D. R., Cregan, P. B., Buck, J. W., Phillips, D. v., Li, Z., & Boerma, H. R. (2015). Identification of Unique Genetic Sources of Soybean Rust Resistance from the USDA Soybean Germplasm Collection. *Crop Science*, *55*(5), 2161–2176. doi.org/10.2135/cropsci2014.09.0671

Hildebrand, A. A. (1952). Stem canker. A disease of increasing importance on Soybeans in Ontario. *Soybean Digest*, *12*(9), 12–15.

Janse Van Rensburg, J. C., Lamprecht, S. C., Groenewald, J. Z., Castlebury, L. A., Crous, P. W., & Lamprecht, S. (2006). Characterisation of *Phomopsis* spp. associated with die-back of rooibos (*Aspalathus linearis*) in South Africa. *Studies of Mycology*, *55*, 65–74.

Keeling, B. L. (1982). A Seedling Test for Resistance to Soybean Stem Canker Caused by *Diaporthe phaseolorum* var. *caulivora*. *Phytopathology*, *72*(7), 807–809.

Keim, P., Olson, T. C., & Shoemaker, R. C. (1988). A rapid protocol for isolating soybean DNA. *Soybean Genetics Newsletter*, *18*, 150–152.

Kilen, T. C., & Hartwig, E. E. (1987). Identification of Single Genes Controlling Resistance to Stem Canker in Soybean. *Crop Science*, *27*(5), 863–864. doi.org/10.2135/cropsci1987.0011183x002700050005x

Kilen, T. C., Keeling, B. L., & Hartwig, E. E. (1985). Inheritance of Reaction to Stem Canker in Soybean. *Crop Science*, *25*(1), 50–51. doi.org/10.2135/cropsci1985.0011183x002500010014x

Krauz, J. P., & Fortnum, B. A. (1983). An Epiphytotic of *Diaporthe* Stem Canker of Soybean in South Carolina. *Plant Disease*, *67*(10), 1128–1129.

Lalitha, B., Snow, J. P., & Berggren, G. T. (1989). Phytotoxin production by *Diaporthe phaseolorum* var. *caulivora*, the causal organism of stem canker of soybean. *Phytopathology*, *79*(4), 499–504.

Maldonado Dos Santos, J. V., Ferreira, E. G. C., Passianotto, A. L. D. L., Brumer, B. B., Santos, A. B. dos, Soares, R. M., Torkamaneh, D., Arias, C. A. A., Belzile, F., Abdelnoor, R. V., & Marcelino-Guimarães, F. C. (2019). Association mapping of a locus that confers southern stem canker resistance in soybean and SNP marker development. *BMC Genomics*, *20*(1), 1–13. doi.org/10.1186/s12864-019-6139-6

Pham, A. T., McNally, K., Abdel-Haleem, H., Roger Boerma, H., & Li, Z. (2013). Fine mapping and identification of candidate genes controlling the resistance to southern root-knot nematode in PI 96354. *Theoretical and Applied Genetics*, *126*(7), 1825–1838. doi.org/10.1007/s00122-013-2095-8

Pioli, R. N., Morandi, E. N., Martìnez, M. C., Lucca, F., Tozzini, A., Bisaro, V., & Hopp, H. E. (2003). Morphologic, molecular, and pathogenic characterization of *Diaporthe*

phaseolorum variability in the core soybean-producing area of Argentina. *Phytopathology*, *93*(2), 136–146. doi.org/10.1094/PHYTO.2003.93.2.136

Rupe, J. C. (2015). Stem Canker. In G. L. Hartman, J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis, & K. L. Steffey (Eds.), *Compendium of Soybean Disease and Pests* (5th ed., pp. 85–88). The American Phytopathological Society.

Saha, A., Das, S., Moin, M., Dutta, M., Bakshi, A., Madhav, M. S., & Kirti, P. B. (2017). Genome-wide identification and comprehensive expression profiling of ribosomal protein small subunit (RPS) genes and their comparative analysis with the large subunit (RPL) genes in rice. *Frontiers in Plant Science*, *8*, 1553. doi.org/10.3389/fpls.2017.01553

Shearin, Z. (2011). *Facilitating Breeding for Resistance to Southern Stem Canker and Elevated Seed Oleic Acid Content in Soybean.* Ph.D. Dissertation. The University of Georgia.

Song, Q., Hyten, D. L., Jia, G., Quigley, C. v, Fickus, E. W., Nelson, R. L., & Cregan, P. B. (2013). Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PloS One*, *8*(1), e54985. doi.org/10.1371/journal.pone.0054985

Soybase (2022). William82.a2 reference genome, Retrieved from www.soybase.org

Sukarta, O. C. A., Townsend, P. D., Llewelyn, A., Dixon, C. H., Slootweg, E. J., Pålsson, L. O., Takken, F. L. W., Goverse, A., & Cann, M. J. (2020). A DNA-Binding Bromodomain-Containing Protein Interacts with and Reduces Rx1-Mediated Immune Response to Potato Virus X. *Plant Communications*, *1*(4), 100086. doi.org/10.1016/j.xplc.2020.100086

Thirugnanasambantham, K., Durairaj, S., Saravanan, S., Karikalan, K., Muralidaran, S., & Islam, V. I. H. (2015). Role of Ethylene Response Transcription Factor (ERF) and Its

Regulation in Response to Stress Encountered by Plants. *Plant Molecular Biology Reporter, 33*(3),347-357. doi.org/10.1007/s11105-014-0799-9

Tyler, J. M. (1995). Additional sources of stem canker resistance in soybean plant introductions. *Crop Science*, *35*(2), 376–377. doi.org/10.2135/cropsci1995.0011183X003500020015x

van Ooijen, J. W. (2006). JoinMap® 4: Software for the calculation of genetic linkage maps in experimental populations. *Kyazma BV, Wageningen*.

Wang, S. C., Basten, J., & Zeng, Z. B. (2012). *Windows QTL Cartographer 2.5* (2.5). Department of Statistics, North Carolina State University.

Wrather, J. A., Anderson, T. R., Arsyad, D. M., Gai, J., Ploper, L. D., Porta-Puglia, A., Ram, H. H., & Yorinori, J. T. (1997). Soybean disease loss estimates for the top 10 soybean producing countries in 1994. *Plant Disease*, *81*(1), 107–110. doi.org/10.1094/PDIS.1997.81.1.107

Wrather, J. A., Stienstra, W. C., & Koenning, S. R. (2001). Soybean disease loss estimates for the United States from 1996 to 1998. *Canadian Journal of Plant Pathology*, *23*(2), 122–131. doi.org/10.1080/07060660109506919

Yi, Q., Liu, Y., Zhang, X., Hou, X., Zhang, J., Liu, H., Hu, Y., Yu, G., & Huang, Y. (2018). Comparative mapping of quantitative trait loci for tassel-related traits of maize in F 2 : 3 and RIL populations. *Journal of Genetics*, *97*(1), 253–266. doi.org/10.1007/s12041-018-0908-x

Zhang, A. W., Riccioni, L., Pedersen, W. L., Kollipara, K. P., & Hartman, G. L. (1998). Molecular identification and phylogenetic grouping of *Diaporthe phaseolorum* and *Phomopsis longicolla* isolates from soybean. *Phytopathology*, *88*(12), 1306–1314. doi.org/10.1094/PHYTO.1998.88.12.1306

Figures and Tables

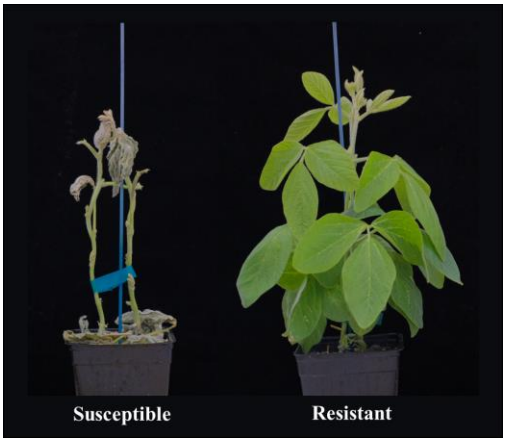


Figure. 2.1. Phenotypic response of soybean plants G81-2057 (susceptible) with a SSC score of 5 and Crockett (resistant) with a SSC score of 1, four weeks post-inoculation with *Diaporthe aspalathi* that causes southern stem canker.

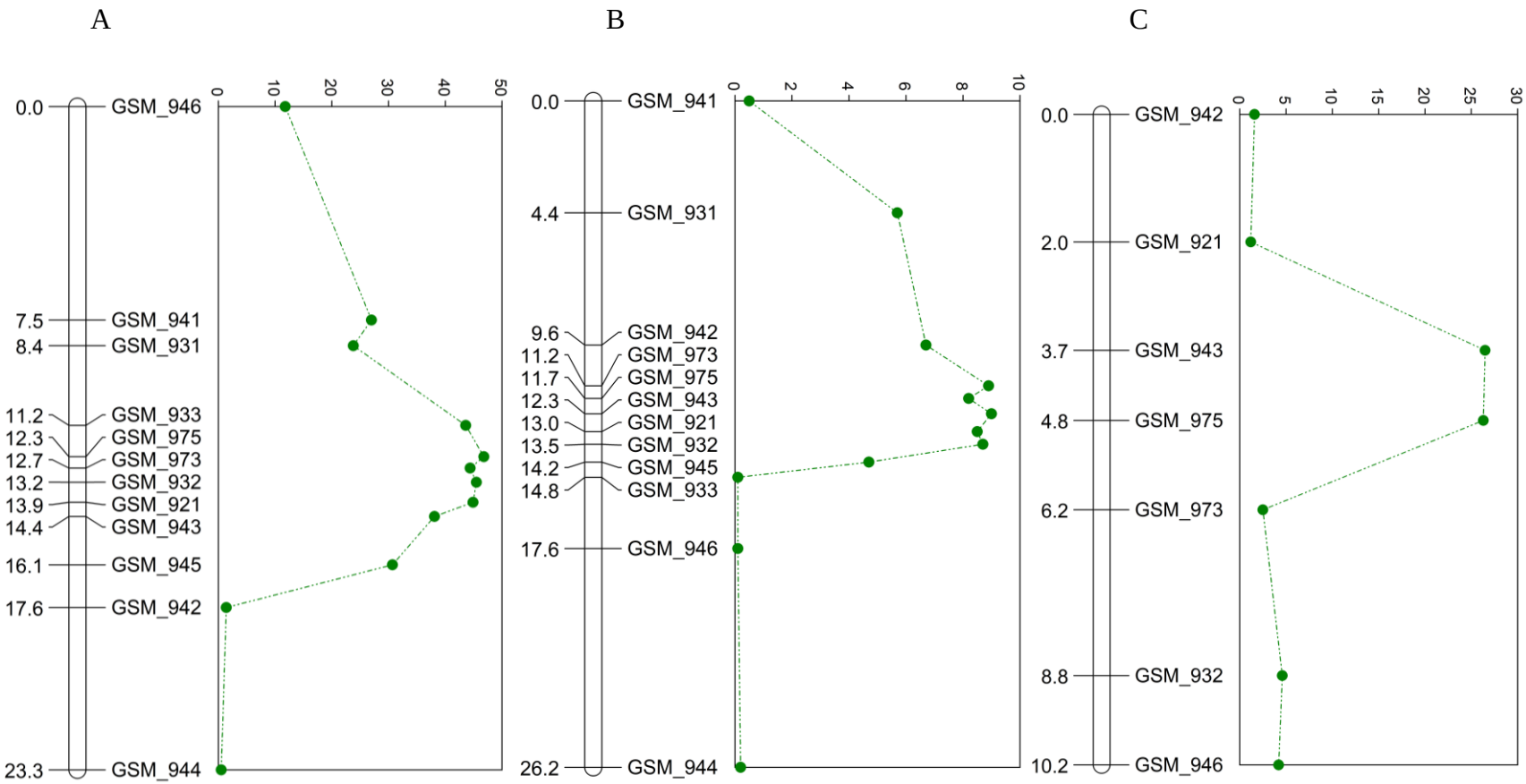


Figure 2.2. Linkage maps and QTLs identified for resistance to southern stem canker on chromosome 14. A) QTL detected in the recombinant inbred line (RIL) population derived from G81-2057 × PI 398469; B) QTL detected in the $F_{2:3}$ population derived from G81-2057 × PI 398469; C) QTL detected from the RIL population derived from G81-2057 × Crockett

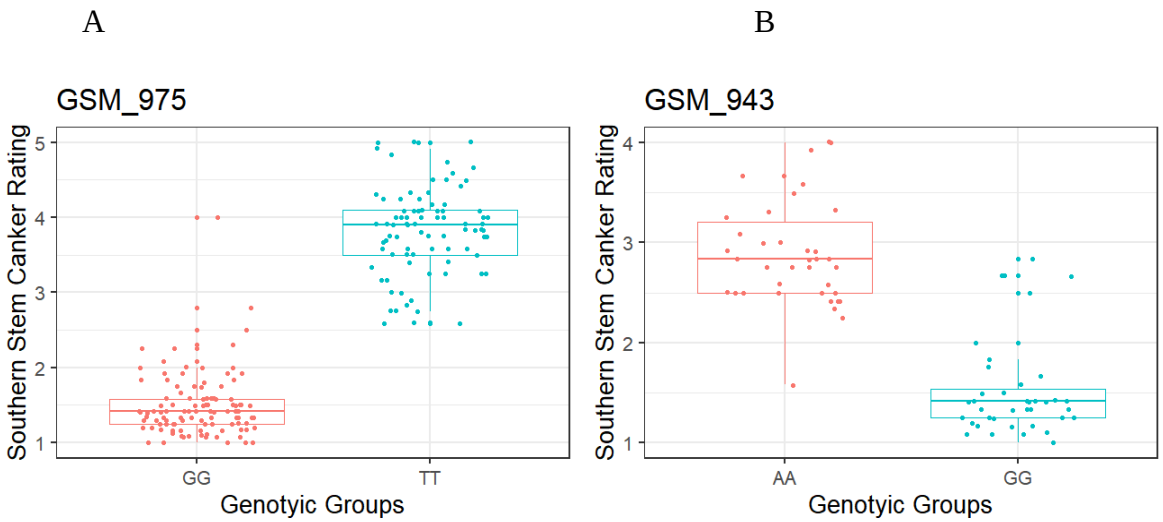


Figure 2.3. Box plot showing the genotype-phenotype association at two peak SNP markers from the mapping populations. (A) GSM_975, a peak marker for the QTL detected in the recombinant inbred line (RIL) population derived from G81-2057 × PI 398469 where the GG genotype had a means SSC score of 1.4 and the TT genotype had a mean SSC score of 3.8; (B) GSM_943, the peak marker for the QTL detected in the $F_{2:3}$ population derived from G81-2057 × PI 398469 with the AA genotype having a mean SSC score of 2.9 and the GG genotype had a mean SSC score of 1.5.

Table 2.1. Analysis of variance (ANOVA) of SSC scores for the recombinant inbred line (RIL) and $F_{2:3}$ populations derived from G81-2057 × PI 398469

Population	Effect	df	Variance Component	Standard Error	p-Value	Broad Sense Heritability (H^2)
RILs	Genotype	223	1.45	0.15	<.0001	0.68
	Replication	5	0.04	0.03	0.14	
	Residual	1114	0.64	0.03		
	Total	1343	2.13	0.15		
$F_{2:3}$ Population	Genotype	174	0.54	0.07	<.0001	0.45
	Replication	5	0.01	0.02	0.48	
	Residual	869	0.64	0.03		
	Total	1049	1.18	0.07		

Table 2.2. QTL identified on chromosome 14 associated with resistance to southern stem canker in both RIL and $F_{2:3}$ populations derived from a cross of G81-2057 × PI 398469 and the RIL population derived from the cross of G81-2057 × Crockett.

Population	Chr	Most significant marker	LOD **-1** interval (cM)	LOD score	R^2	Additive effect	Source of Favorable Allele
G81-2057 × PI 398469 RILs	14	GSM_975	2.7	46.7	0.62	0.97	PI 398469
G81-2057 × PI 398469 $F_{2:3}$	14	GSM_943	3.9	9.1	0.21	0.36	PI 398469
G81-2057 × Crockett RILs	14	GSM_943	1.1	26.6	0.57	0.74	Crockett

Table 2.3. Single marker analysis for the KASP markers that were designed based on the bulked segregant analysis in the $F_{2:3}$ population derived from G81-2057 × PI 398469.

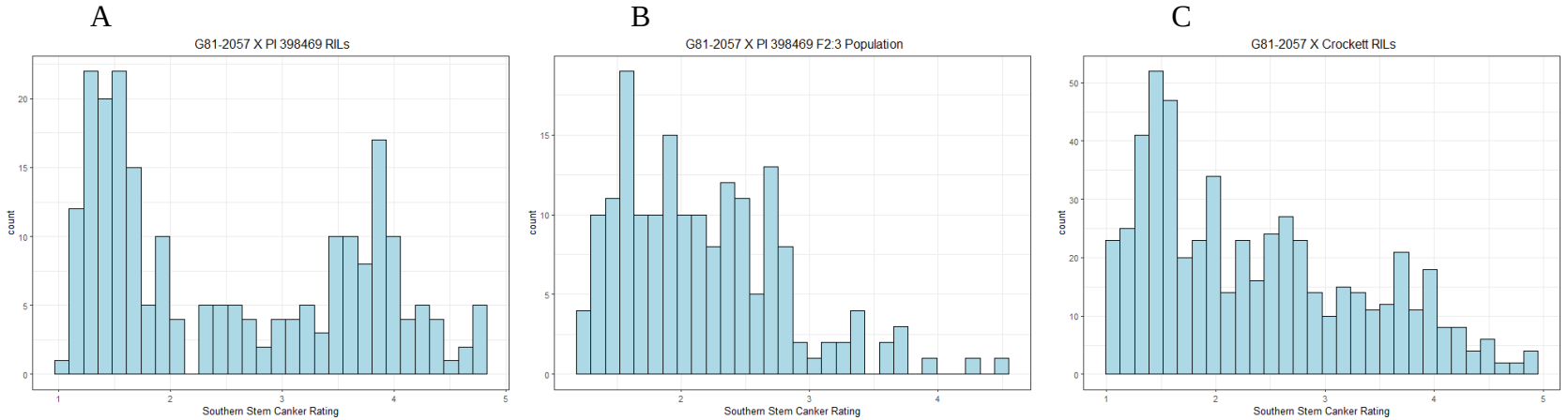
Marker Name	SNP ID	Chr	Physical position (bp) [a]	Favorable Allele	F value
GSM_931	Gm14_701101_A_C	14	707,505	C	56.4***†
GSM_941	Gm14_937811_T_C	14	939,553	C	49.9***
GSM_942	Gm14_1315167_A_G	14	1,316,725	G	49.9***
GSM_973	Gm14_1423892_A_G	14	1,425,590	G	110.1***
GSM_943	Gm14_1442735_A_G	14	1,444,397	G	97.4***
GSM_975	Gm14_1610874_T_G	14	1,612,516	G	190.8***
GSM_933	Gm14_1657815_T_G	14	1,658,941	T	213.7***
GSM_944	Gm14_1723025_T_C	14	1,724,237	T	147.2***
GSM_945	Gm14_1735755_T_C	14	1,736,967	T	136.4***
GSM_921	Gm14_1,744,370	14	1,744,370	C	146.1***
GSM_932	Gm14_1765933_A_G	14	1,767,144	A	199.9***
GSM_946	Gm14_1879468_A_C	14	1,874,810	A	31.4***
GSM_1007	Gm10_4638769_C_T	10	4,655,396	C	1.8
GSM_1010	Gm10_4668552_G_A	10	4,685,288	G	0.1
GSM_923	Gm10_4670275_G_A	10	4,687,011	A	2.5*
GSM_1006	Gm10_5175579_C_T	10	5,197,746	C	3.4*
GSM_1005	Gm09_40354246_C_T	9	43,031,363	T	2.4
GSM_930	Gm09_41070222_G_A	9	43,784,512	A	2.7
GSM_1009	Gm09_42586932_T_G	9	45,783,045	G	1.98
GSM_929	Gm09_43821505_C_T	9	47,026,346	C	4.6**

†***, **, and * indicates significance at a p-value of 0.001, 0.01 and 0.05, respectively
[a] Physical position is based on the William 82.a2 reference genome

Table 2.4. Gene models that were identified in the QTL region on Chr 14 in both G81-2057 × PI 398469 RIL and $F_{2:3}$ populations based on gene annotations of the Williams 82.a2 reference genome from SoyBase (http://soybase.org).

Gene Model	Gene Annotation	GO Term
Glyma.14g020100	*ATERF3, ERF3*	DNA-Binding Transcription Factor Activity
Glyma.14g020300	*Pectinesterase*	Cell Wall Modification
Glyma.14g020500	Ribosomal S17 Family Protein	Translation, Negative Regulation of Catalytic Activity, Ribosome, Structural Constituent of Ribosome, Enzyme Inhibitor Activity
Glyma.14g022600	DNA-Binding Bromodomain-Containing Protein	Protein Binding
Glyma.14g022900	RNI-Like Superfamily Protein	F-Box/Leucine-Rich Repeat Protein

Appendix A: Supplemental Figure



Supplementary Figure 2.S1. Distribution of mean southern stem canker scores. A) G81-2057 × PI 398468 RILs; B) G81-2057 × PI 398468 $F_{2:3}$; and C) G81-2057 × Crockett RILs. Scale of southern stem canker rating on a 1 to 5 scale (1 = no disease, 2 = discoloration of cortex, 3 = visible stem canker or lesions, 4 = stem lesions, foliar interveinal chlorosis and necrosis, 5 = dead plant).

Appendix B: Supplemental Tables

Supplementary Table 2.S1. KASP marker sequences that were designed for the G81-2057 × PI 398469 recombinant inbred line population.

Marker Name	SNP ID	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
GSM_931	Gm14_701101_A_C	ACTGAAACACACCAACCTCGA[A/C]	TGGCATATGCTGACGGAGTT
GSM_941	Gm14_937811_T_C	TTCAATGACATATTTCAATTTGACC[C/T]	GTCAGGCCCTTAAGGTAGGC
GSM_942	Gm14_1315167_A_G	CGAAGAAGAAGAAGGATCGCT[G/A]	TTCACGAACTCACGCTCTG
GSM_973	Gm14_1423892_A_G	GTGCATTCAGTCGTGCGG[G/A]	GCTCTTGCATCATATATCTCCTTCC
GSM_943	Gm14_1442735_A_G	GTAGGAATAAATAGTACTGGACCAG[G/A]	TGACTCTCGTCCTTGAGCTG
GSM_975	Gm14_1610874_T_G	CGGTTGCCTTGATATTGATTTCA[G/T]	AGTGATGCCTTGCCCATCAA
GSM_933	Gm14_1657815_T_G	AGACTTGGCCTGGTTGGA[T/G]	TTAAAGAGGGTGTCTCGTTTGGA
GSM_944	Gm14_1723025_T_C	AATTCCCTCCATTTTCCCC[T/C]	ATCGGGAGAAAGAGGTGGAT
GSM_945	Gm14_1735755_T_C	TCCTCACCCTCTCATTCAAAC[T/C]	AAGGAGAGACATGGGTGTGG
GSM_921	Gm14_1,744,370	GTGTCAAAATCAATTTTAGTTGAAC[C/T]	CATCCTTGAAAGCAGCTGTG
GSM_932	Gm14_1765933_A_G	GTGGACCTAGAAGGATCACTTC[A/G]	TCTCTTCCTAAGAAAGTTCAAGCAG
GSM_946	Gm14_1879468_A_C	TCACTGCCATAGCCAATGA[A/C]	TAAGCATGATCGGTGGGATT
GSM_1007	Gm10_4638769_C_T	GTATTCCGCCTCTGCACTAGAT[C/T]	TTCTGGGCATTGTGTATTTCTTG
GSM_1010	Gm10_4668552_G_A	GCTACCCCAGTGAAGGTGC[G/A]	ATGAGGCCCTTCTGAAGCATG

GSM_923	Gm10_4670275_G_A	TTCTCAAAGGGTATCCACCT[G/A]	GAGACGAAGCTGTGTTCCACT
GSM_1006	Gm10_5175579_C_T	GCGGAGCTGCAATTTCCTT[C/T]	GCACACCACCACAACCACTGTA
GSM_1005	Gm09_40354246_C_T	TTAGACCCGATCCCATCTGTT[T/C]	CAGAAACTGCATGGCCCTATC
GSM_930	Gm09_41070222_G_A	ATGGCAATGAGTAGGATGAC[A/G}	GCAACCATTTCTTCTTTCTGGA
GSM_1009	Gm09_42586932_T_G	ATTTTTCCGTACAAACAGGCAACTA[G/T]	
GSM_929	Gm09_43821505_C_T	AAATAAAGAGGAATGAATGTATTG[C/T]	CAGATTGGAGTACCGTGGAT

CHAPTER 3

GENETIC MAPPING REVEALS THE COMPLEX GENETIC CONTROLLING OF SLOW CANOPY WILTING IN SOYBEAN PI 471938

[1]Ethan Menke, Clinton J. Steketee, William T. Schapaugh, Thomas E. Carter Jr, Benjamin Fallen, and Zenglu Li. To be summitted to *Theoretical and Applied Genetics*.

Abstract

In soybean [*Glycine max* (L.) Merr.] drought stress is the leading cause of yield loss from abiotic stress in rain-fed U.S. growing areas. With only 10% of the US soybean production under irrigation, plants must possess physiological mechanisms to tolerate drought to combat drought stress. Slow or delayed canopy wilting is a physiological trait that is observed in a few exotic plant introductions (PIs) and may lead to yield improvement under drought stress. In this study, canopy wilting on 130 recombinant inbred lines (RILs) derived from Hutcheson × PI 471938 under drought stress was visually evaluated and RILs were genotyped with the SoySNP6K iSelect BeadChip. Over four years, field evaluations of canopy wilting were conducted under rain-fed conditions at three locations across the US (Georgia, Kansas, and North Carolina). Due to the variation in weather between locations and years, the phenotypic data were collected from seven environments. Substantial variation in canopy wilting was observed among the genotypes in the RIL population across environments. Three QTLs were identified for canopy wilting from the RIL population using composite interval mapping on chromosomes 2, 8, and 9 based on combined environmental analyses. These QTLs inherited the favorable alleles from PI 471938 and accounted for 11, 10, and 14% of phenotypic variation. The QTLs identified through this research can be used as targets for further investigation to understand the mechanisms of slow canopy wilting. These QTLs could be deployed to improve drought tolerance by allowing for the targeted selection of the genomic regions in PI 471938 that provide slow canopy wilting in new soybean cultivars.

Keywords: Soybean, Drought Tolerance, Canopy Wilting, Quantitative Trait Locus

Introduction

Soybean [*Glycine max* (L.) Merr.] is the largest oilseed crop globally, providing over a quarter of the vegetable oil and almost ~70% of the plant protein meal used worldwide. Global demand for soybean has led it to be the second most cultivated row crop in the United States, with an estimated 33.6 million hectares of soybean planted in 2020 (SoyStats, 2020). Even with its economic importance, only less than 10% of soybean hectares in the United States are under irrigation (Specht et al. 2014). Lack of irrigation leaves soybean extremely vulnerable to drought stress, which can cause more than a 40% reduction in yield (Specht et al. 1999; Purcell & Specht 2004).

Canopy wilting is caused by a decrease in turgor pressure in the soybean leaves and is a trait commonly used by soybean breeders to identify differential responses to stress. Slow or delayed canopy wilting has been observed in exotic soybean germplasm and is controlled by multiple plant mechanisms. A maturity group (MG) VI, plant introduction (PI) 416937 from Japan, has been observed to have slower canopy wilting under drought conditions than other existing cultivars (Sloane et al. 1990). PI 416937 also has an extensive lateral root system, with a large root surface area (Hudak &Patterson 1996; Pantalone & Rebetzke 1996) combined with low stomatal conductance (Tanaka et al. 2010). Fletcher et al. (2007) showed that under high vapor pressure deficit (VPD), PI 416937 reached a maximum transpiration rate at 2.0 kPa, while commercial cultivars showed increased transpiration rates at VPD greater than 2.0 kPa. This decreased transpiration at high VPD allows for the conservation of moisture, thus increasing the water use efficiency of the plants (Fletcher et al. 2007). The above results indicate that PI 416937 uses water conservation as its mechanism of the slow wilting phenotype that may protect yield under drought conditions.

PI 471938, an accession from Nepal is a MG V introduction that exhibits slow canopy wilting as well, but the mechanism for this response to drought stress is unknown (Sadok et al. 2012; Bagherzadi et al. 2017). PI 471938 has shown normal nitrogen fixation under soil drying conditions (Sinclair et al. 2000; Devi & Sinclair 2013; Riar et al. 2018). It has been used by multiple southern breeding programs to develop cultivars in the Southeastern US (Devi et al. 2014; Carter et al. 2016). Based on pedigree data of lines that appeared in the USDA Uniform Tests, cultivar registrations, and plant variety protection applications, PI 471938 is a parent of six varieties developed from the population used in this experiment and is in the ancestry of at least 25 other breeding lines that reached the Uniform Yield Tests (Soybase.org, 2021). In addition, PI 567690 and PI 567731, both MG III, have been identified as two new sources of slow canopy wilting for early maturity group soybean breeders (Pathan et al. 2014; Ye et al. 2020).

Several studies of canopy wilting have been performed to understand the underlying genetics using bi-parental populations and genome-wide association mapping studies. Kaler et al. (2017) used 373 MG IV soybean genotypes as a genome-wide association panel to identify genomic regions associated with slow canopy wilting. In this study, the authors found 61 single nucleotide polymorphisms (SNPs) that tagged 51 loci on 19 of the 20 soybean chromosomes (Kaler et al. 2017). Steketee et al. (2020) used a panel of 162 MG VI-VIII accessions and cultivars to identify genomic regions associated with canopy wilting. The study identified 45 unique SNPs related to differential canopy wilting at 44 loci in this population (Steketee et al. 2020). Twenty genomic regions on chromosomes (Chrs) 1, 4, 6, 9, 12, 15, 18, and 19 from Steketee et al. (2020) were also identified by Kaler et al. (2017). Abdel-Haleem et al. (2012) used a recombinant inbred line (RIL) population derived from a cross of ‘Benning’ (a fast wilting MG VII cultivar) × PI 416937 to identify seven quantitative trait loci (QTLs). These QTLs explained 75% of the

phenotypic variation observed in canopy wilting using multiple interval mapping (Abdel-Haleem et al. 2012). Of the seven QTLs identified in multiple locations, five QTLs on Chrs 2, 4, 5, 12, and 19 inherited the favorable alleles from PI 416937 for the slow canopy witling trait. The two remaining QTL identified in this population on Chrs 14 and 17, inherited the favorable alleles from the fast wilting cultivar Benning.

Charlson et al. (2009) investigated the effects of drought stress on a RIL population developed from a cross of 'KS4895', a fast wilting cultivar, and Jackson, a slow wilting cultivar. Four QTLs that explained 47% of the phenotypic variation in canopy wilting were identified on Chrs 8, 13, 14, and 17. The slow canopy wilting cultivar, Jackson, is present in the pedigree of the fast wilting cultivar Benning, which could explain the beneficial drought tolerance alleles identified from Benning in the Benning × PI 416937 RIL population (Charlson et al. 2009; Abdel-Haleem et al. 2012). Using five RIL populations involving three slow wilting genotypes Jackson, PI 424140, and PI 416937, Hwang et al. (2015) found seven QTL clusters on Chrs 2, 5, 8, 11, 17, and 19 based on 95% confidence intervals from at least two mapping populations (Hwang et al. 2015). The populations from the previous study used by Hwang et al. (2015) were used to perform a meta-QTL analysis, which identify nine meta-QTLs, in eight QTL clusters on Chrs 2, 5, 11, 17, and 19 with a reduced confidence interval (Hwang et al. 2016). Ye et al. (2020) mapped the QTLs in two RIL populations derived from PI 567690 and PI 567731, two MG III exotic landraces. In the 'Pana' × PI 567690 RIL population, eight QTLs were identified, which were located at similar chromosomal positions to the QTLs identified in both Abdel-Haleem et al. (2012) and Hwang et al. (2016). Two QTLs on Chrs 6 and 10 that were identified in the 'Magellan' × PI 567731 were not identified in previous QTL mapping studies. The objectives of this study were to i) evaluate a

RIL population in repeated field experiments for canopy wilting, and ii) elucidate genomic regions responsible for canopy wilting in PI 471938.

Materials and methods

Plant materials

A cross between 'Hutcheson' (PI 518664) and PI 471938 was made in 1998 in Raleigh, NC, USA. Hutcheson is an MG V cultivar developed by Virginia Tech (Buss et al. 1988). PI 471938 is an MG V plant introduction characterized previously as a slow wilting soybean accession (Carter et al. 1999; Hufstetler et al. 2007; Sadok et al. 2012). The F_1 seed from this cross was grown at the USDA Tropical Agricultural Research Station in Isabela, Puerto Rico. The F_2 to F_4 generations were advanced by single seed descent (Brim 1966) throughout the inbreeding process. The F_4 plants were harvested individually and used to develop the 130 F_4-derived recombinant inbred lines (RILs) used in this study.

Evaluation of canopy wilting

The Hutcheson × PI 471938 RIL population was evaluated in Athens, GA (2016_GA) and Salina, KS (2016_KS) in 2016. Two-row plots were planted at both locations with three replications in GA and two replications in KS using a randomized complete block design. In 2018 the RIL population was evaluated in Midville, Ga (2018_GA), Salina, Ks (2018_KS), and Sandhills, North Carolina (2018_NC). In 2019 the population was evaluated in Midville, Ga (2019_GA), Salina, KS (2019_KS), and Sandhills, North Carolina (2019_NC). The 2018 and 2019 experiments were planted as two-row plots in a randomized complete block design, with three

replications. All environments were planted with 0.76 m row spacing at a seeding density of 32 seed m^{-2}.

Canopy wilting was rated in increments of five on a scale from 0 to 100: 0 = no wilting present; 20 = slight wilting and some rolling in the top of the canopy; 40 = somewhat severe leaf rolling at the top of the canopy, moderate wilting of leaves throughout the rest of the canopy, and some loss of petiole turgidity; 60 = severe wilting of leaves throughout the entire canopy, with advanced loss of petiole turgidity; 80 = plants with petioles severely wilted and dead leaves throughout much of the canopy; and 100 = plant death.

The RILs were evaluated for canopy wilting by taking the mean of three ratings taken between 26 August and 16 September (during pod filling) by three raters as the phenotypic score for the 2016_GA environment, and the average of the rating taken by three raters on 16 Sept 2016 was used for mapping. A single rating taken on 27 July 2016 (during flowering) by one rater was used as the phenotypic score for the 2016_KS environment. In 2017, an evaluation of this RIL population at three locations (Athens, GA, Salina, KS, and Sandhills, NC) was attempted, but no canopy wilting scores were recorded because of minimal water stress. In 2018 no canopy wilting evaluations were performed in Sandhills, NC, due to a lack of drought stress during the growing season. One wilting rating was taken on 17 Sept. 2018 (during pod filling) at the 2018_GA environment by one rater. Three wilting ratings were collected in the 2018_KS environment on 1, 17, and 18 Sept. 2018; the rating from 1 Sept. was used for QTL mapping. The 2019_GA environment was rated four times during the growing season, with one rater rating on 31 July and 13 Aug 2019 (flowering). The ratings on 12 Sept. 2019 (pod fill) were collected by two raters, with an average of the two on 12 Sept. 2019 used for mapping. A single observer scored the 2019

ratings at the 2019_KS environment on 23 Sept 2019 (pod fill). The ratings at the 2019_NC environment were collected on 17 Sept. 2019 by a single rater (pod fill).

Genotype data and quality control

DNA was extracted from leaf tissue and genotyped with the SoySNP6K iSelect BeadChip (Song et al. 2014). The leaf tissue collection and DNA extraction procedures were the same as described in Steketee et al. (2020). These genotyping efforts generated 5,403 genome-wide SNPs that were analyzed using GenomeStudio software (Illumina Inc., San Diego, CA, USA) to perform SNP quality control for segregation distortion and compression of genotype calls. Monomorphic markers between the two parents were removed, leaving 1,258 polymorphic SNP markers available to create a genetic map. Forty-six additional markers were removed that did not meet requirements for joining a linkage group during the genetic map construction, leaving a total of 1,212 polymorphic SNP markers to be used in QTL mapping.

Statistical analyses

Analysis of variance (ANOVA) was conducted using PROC MIXED in SAS version 9.4 (SAS Institute 2014). The model for the combined analysis was built by treating genotype, environment, and genotype by environment interaction and replication within the environment as random variables using the Standard Least Squares personality and REML method. Genotype means were separated by Fisher's least significant difference (LSD) test at the $\alpha = 0.05$ probability level. Broad-sense heritability was calculated on an entry-mean basis according to Holland et al. (2002), with the variance components being calculated using a model where all variables were treated as random. Correlations of genotype means were calculated using PROC CORR in SAS

version 9.4. Best linear unbiased predictions (BLUPs) were calculated for canopy wilting scores across all environments using SAS version 9.4. For individual environments only, genotype and replication were used and treated as random variables to calculate BLUPs. Using BLUP values for each genotype across and within environments helped to account for variation caused by environmental factors and missing data. BLUPs were used as the phenotypic values for subsequent QTL analyses.

Genetic map construction and QTL analyses

The 1,212 polymorphic SNP markers for the Hutcheson × PI 471938 RIL population were used to construct a genetic map in JoinMap 4.1 (Van Ooijen 2006). The logarithm of odds (LOD) criterion of greater than six was used to establish linkage groups. As necessary, some groups were then forced together to form 20 linkage groups based on the known chromosomes and physical positions of the SNP markers. Maximum likelihood (ML) mapping with the default settings was used to convert recombination frequencies into map distances in centiMorgans (cM). These cM positions were then used in subsequent QTL mapping.

The software package Windows QTL Cartographer (WinQTLCart) 2.5 (Wang et al. 2012) was used for composite interval mapping (CIM) using Model 6 of the Zmapqtl program module. The genome was scanned with a walking speed of 0.5 cM and a window size of 10 cM, and the forward-backward regression method was used to choose cofactors. The significance LOD threshold was determined by 1,000 permutations, with a significance level of $\alpha = 0.05$. The significance threshold, LOD = 3.3 was used for the combined analysis. Significant QTLs were identified by the peak of the QTL meeting or exceeding the LOD score. QTL peaks positions were

determined by the highest score on a specific chromosome. MapChart 2.32 (Voorrips, 2002) was used to visualize the genetic maps and QTL mapping results.

Candidate gene identification

The peak SNPs of the QTLs on Chr 2, 8, and 9 identified through composite interval mapping were used to identify nearby candidate genes. Steketee et al. (2020) found a median distance of 9 kb and a mean distance of 26 kb between SNP markers used in their GWAS of drought-tolerant cultivars. For this study, a 10 kb region before or after the peak of the significant QTL was used to identify potential gene models.

Results

Genetic variation of canopy wilting for the RIL population

The RIL population exhibited a wide range of canopy wilting among the RILs, and the wilting was more severe in the Georgia environments (Fig. 3.1). Genotypes, environments, and their interactions were statistically significant ($p < 0.05$) for canopy wilting scores (Table 3.1). Correlation between the environments canopy wilting scores ranged from r = -0.17 to 0.44 (Table 3.2). The broad-sense heritability of canopy wilting on an entry-mean basis for the combined environments was 0.29. No RILs in the combined environments had a lower mean wilting score than the slow wilting parent, PI 471938. Twenty-three of the RILs had higher canopy wilting based on mean performance in the combined environments as compared to the fast wilting parent, Hutcheson.

QTL mapping of canopy wilting trait for the RIL population

In the combined environments three QTLs were identified on Chr 2, 8, and 9. These QTLs accounted for between 11% and 14% of the phenotypic variation that was observed for slow canopy wilting. On Chr 2 a QTL, *qWilt_Gm2*, was identified and explained 11% of the phenotypic variation that was observed in canopy wilting (Table 3.3). The peak marker for the qWilt_Gm2 was Gm02_15067760_G_A, which was located at 15,271,225 bp with a confidence interval (CI) of this QTL spanned 4.7 Mb (14,220,378-18,913,725 bp) (Fig. 3.2a). A QTL was identified on Chr 8 in the combined environments (*qWilt_Gm8*) and accounted for 10% of the phenotypic variation observed in slow canopy wilting (Table 3.3). The QTL *qWilt_Gm8* spanned a CI of 1.5 Mb (44,267,551-45,913,059 bp) on Chr 8 (Fig. 3.2b, Table 3.3). The peak marker for *qWilt_Gm8* was Gm08_44368268_A_G (45,403,652 bp). The QTL on Chr 9 (*qWilt_Gm9*) was identified in the combined analysis, which accounted for 14% of the observed variance in the wilting score (Table 3.3). The peak of qWilt_Gm9 was at marker Gm09_36486860_T_C (39,047,264 bp) with a CI of 6.3 Mb (36,455,035-42,790,738 bp) (Fig. 3.1c, Table 3.3). From the combined analysis, QTLs on Chrs 2, 8, and 9 all had a positive allelic effect (Table 3.3). Positive additive effects indicate that the mean canopy wilting score for the RILs possessing the allele from PI 471938 was lower than those possessing the alleles from Hutcheson.

Discussion

Physiological mechanisms for canopy wilting and relationship to other traits

Slow canopy wilting could lead to less yield reduction during drought stress for soybean (Sloane et al. 1990). A previous study proposed three different combinations of physiological mechanisms that could lead to delayed canopy wilting (Ries et al. 2012). One is a combination of

increased water use efficiency (WUE), high radiation use efficiency (RUE) which is an increase in above-ground dry weight per unit of intercepted radiation, and soil moisture conservation. Genotypes in this group would utilize transpired water for biomass production, and higher RUE would have increased photosynthesis and would be expected in drought-stressed and optimal growing conditions. The second combination is low stomatal conductance, low RUE, low WUE, and soil moisture conservation. The genotypes in this group would have low transpiration that would reduce potential photosynthetic capacity and would be better at conserving water during drought stress conditions as might be expected in desert flora. However, this second combination of physiological attributes could reduce overall yield potential, especially in well-watered environments. Deeper rooting is a third mechanism proposed that could delay canopy wilting in soybean (Ries et al. 2012). Given these advantages and trade-offs for different physiological traits, identifying soybean germplasm with the optimal combination to reduce canopy wilting during drought stress will be different depending on the target environment.

Much like canopy wilting in soybean, evaluation of other crops for drought tolerance commonly uses secondary traits for indirect yield selection under stressed conditions. Leaf rolling reduces exposed leaf area, thereby decreasing transpiration and reducing light interception, and can be observed in crops such as maize, rice, and wheat (Kadioglu et al. 2012). The earlier leaf rolling occurs on a given day, or the longer duration of rolling indicates the plant is experiencing more stress (Rauf et al. 2016). Therefore, ratings and selections can be made to identify plants with reduced leaf rolling during drought periods to improve drought tolerance (Rauf et al. 2016). In maize, another trait that is evaluated to improve drought adaptation is the anthesis-silking interval (ASI). This trait is negatively correlated with grain yield under drought conditions and has been a breeding target due to the ease of measurement and moderate heritability (Tuberosa 2012).

Additional traits such as the stay-green, root architecture, and canopy temperature depression can impact a plant's ability to tolerate drought stress and have been evaluated in several crop species (Winterhalter et al. 2011; Tuberosa 2012; Mace et al. 2012).

PI 471938 has been used extensively in southern breeding programs and it has been used in the pedigree of over 30 lines that have been evaluated for potential commercial use. These lines were advanced to regional testing, indicating that they possessed potential as commercial cultivars for use in the Southern United States. An example is USDA-N8002 which was derived from both PI 471937 (25% by pedigree) and PI 416937 (12.5% by pedigree) and possesses drought-resistant traits with high yield potential (Carter et al. 2016). This is a source for the slow wilting trait and demonstrates that breeding for drought tolerance doesn't have to result in the incorporation of deleterious traits from plant introductions.

Slow canopy wilting in soybean is potentially related to many possible physiological mechanisms, which are highly complex. This study identified three QTLs on Chrs 2, 8, and 9 that each explains a relatively small portion of phenotypic variation. This indicated that the genetic control of the slow canopy wilting could be governed by many minor QTLs.

Genotype by environment interaction and heritability

Although the genotype by environment interactions was significant ($p < 0.05$) with the population (Table 3.1) and the severity of wilting experienced in the seven environments varied (Fig. 3.1). The correlations between wilting scores between locations in a single year were relatively high (r =0.22- 0.44), except for the 2016_GA and 2016_KS locations (r=0.07), indicating that the genotypes tested wilted similarly across environments in a given year (Table 3.2). Using CIM, eight QTLs were identified for canopy wilting in six of the seven individual environments

tested (Table S3.1). No significant QTL was detected in the 2019_NC environment at a LOD= 3.4. These QTL accounted for 8 to 20% of the phenotypic variation that was observed in an individual environment. Heritability across environments was also moderate to high, with heritability comparable to those observed in previously canopy wilting QTL mapping and GWAS studies (Charlson et al. 2009; Abdel-Haleem et al. 2012; Hwang et al. 2015; Kaler et al. 2017). The QTLs *qWilt_Gm2* and *qWilt_Gm8* were located in similar regions as QTLs identified in the individual environments (Table 3.3 and S3.1). The remaining six QTLs were identified in only one environment, this is most likely due to the highly complex nature of the slow canopy wilting and the environmental variance.

Comparisons of genetic mapping for canopy wilting to previous studies

In this study, three QTLs were associated with canopy wilting using the composite interval mapping method in the combined environmental analysis. Common QTLs were identified on Chrs 2, 8, and 9. The QTL *qWilt_Gm2* is located approximately 1.9 Mb upstream of the peak of the meta QTL *mqCanopywilt-003* identified by Hwang et al. (2016). *qWilt_Gm2* was located in a similar genomic region to the significant QTL *qSW-Gm02* that was identified by Abdel-Haleem et al. (2012) in a RIL population derived from Benning × PI 416937. The peak of *qWilt_Gm8* on Chr 8 was located 199 kb from the significant SNP Gm08_44751317_C_T identified in the GWAS for canopy wilting (Kaler et al. 2017). The Chr 9 QTL, *qWilt_Gm9* was located 2.3 Mb from a reported significant SNP identified in the GWAS for canopy wilting done by Kaler et al. (2017) and 2.1 Mb from the GWAS performed by Steketee et al. (2020). In this study, three QTLs identified in the combined environments explained a relatively small portion of the phenotypic variation that was observed in canopy wilting. The QTLs identified in this study are in genomic regions as in

previous mapping and association studies (Abdel-Haleem et al. 2012; Hwang et al. 2016; Kaler et al. 2017). These QTLs only provide a small portion of the genetic control of the slow canopy wilting trait from PI 471938 and this is most likely due to the highly complex nature of the slow canopy wilting trait.

Candidate genes at canopy wilting significant genomic regions

Three gene models, *Glyma.02g147900, Glyma.02g148000,* and *Glyma.02g148100* within the region 10 Kb from the peak of the QTL *qWilt_Gm2* were identified. This included a gene model for Aluminum-Activated Malate Transporters, which play a role in cell signaling of abiotic stress (Ramesh et al., 2018; Shelp et al., 2012). In the 10 kb region on either side of the peak marker for *qWilt_Gm8*, three gene models, *Glyma.08g337000*, *Glyma.08g337100,* and *Glyma.08g337200* were present. These gene models have gene ontologies such as calcium ion transport, transmembrane transport, and protein binding. *Glyma.08g337000* encodes $Ca2^{+}$ exchange proteins in Arabidopsis and these proteins help regulate stomatal movements (Cho et al. 2009). The *Glyma.08g337100* gene model encodes a Chaperone DnaJ-domain superfamily protein and in rice grown under drought stress, proteins in this family showed elevated transcription under drought stress (Luo et al. 2019). *qWilt_Gm9* which is located 10 kb from two gene models *Glyma.09g166400* and *Glyma.09g166500* that have gene ontologies terms that include RNA modification, an integral component of membrane, intracellular membrane-bounded organelle. The third gene model, *Glyma.09g166400,* encodes an organic cation/carnitine transporter. In Arabidopsis, organic cation/carnitine transporters were involved in lateral root formation, and plants containing knockouts of this gene lead to increased root growth (Lelandais-Brière et al. 2007). The gene models that were identified around each of the three QTLs that were present in

the common environment are related to stress tolerance in soybean or other crops. These gene models could provide targets for gene expression studies to identify the mechanism that underlies drought tolerance in PI 471938.

Implications for soybean breeding

Breeding for drought tolerance is a complex process. It requires adequate drought stress every year to effectively evaluate large numbers of genotypes. As shown in this experiment, relying on natural drought can be difficult. In two out of four years during this study, at least one location did not experience any drought symptoms. Lack of consistent phenotyping, combined with the labor and resources needed to conduct drought field evaluations of many soybean breeding lines becomes challenging to accomplish. For these reasons, the results from this study could allow for use of QTLs and marker information to aid in the selection of lines for the slow canopy wilting trait. In the Hutcheson × PI 471938 mapping population, three QTLs on Chrs 2, 8, and 9 have been significantly associated with slow canopy wilting in the combined environments. In all the genomic regions surrounding the peak SNPs, gene models indicate potential mechanisms related to drought tolerance. These three regions are good targets for further research to narrow down the underlying drought tolerance mechanism in PI 471938. Based on this and previous research, drought tolerance appears to be a highly quantitative trait with a large number of underlying mechanisms that lead to the slow canopy wilting phenotype. Using the markers identified in this study to select for drought tolerance through marker-assisted selection would allow for the selection of drought tolerance without having to completely rely on favorable phenotyping conditions. This could prove difficult due to a large number of QTLs with little phenotypic control that would need to be tracked in a breeding program. Additionally, the QTLs

from PI 471938 could be used when training genomic selection models to help predict the performance of lines for drought tolerance without having to phenotype them under drought stress in the early generations. This would relieve the need to have consistent environmental conditions to effectively phenotype breeding populations. This would allow for the deployment of these QTLs in new cultivars more efficiently where the slow wilting trait could help soybean producers limit some of the significant effects of drought in soybean production.

Conclusions

Using the 137 RILs derived from Hutcheson × PI 471938, three QTLs on Chrs 2, 8, and 9 for slow canopy wilting were identified using CIM and accounted for 10-14 % of the phenotypic variation that is observed in drought tolerance. The genomic locations of these QTLs identified in this study are in proximity to previously reported mapping and GWAS results. The candidate genes located near all three QTLs are targets for expression studies to understand the functions of these genes that control slow canopy wilting in PI 471938. The QTLs discovered in this study will allow for improved efficiency in breeding drought-tolerant soybeans, through marker-assisted selection and genomic selection. These improved drought-tolerant cultivars can then be used by soybean producers to meet the climate challenges that they face due to drought.

References

Abdel-Haleem, H., Carter, T. E., Purcell, L. C., King, C. A., Ries, L. L., Chen, P., Schapaugh, W., Sinclair, T. R., & Boerma, H. R. (2012). Mapping of quantitative trait loci for canopy-wilting trait in soybean (*Glycine max* L. Merr). *Theoretical and Applied Genetics*, *125*(5), 837–846. doi.org/10.1007/s00122-012-1876-9

Bagherzadi, L., Sinclair, T. R., Zwieniecki, M., Secchi, F., Hoffmann, W., Carter, T. E., & Rufty, T. W. (2017). Assessing water-related plant traits to explain slow-wilting in soybean PI 471938. *Journal of Crop Improvement*, *31*(3), 400–417. doi.org/10.1080/15427528.2017.1309609

Brim, C. A. (1966). A modified pedigree method of selection in soybeans. *Crop Science*, *6*, 220.

Buss, G. R., Camper Jr., H. M., & Roane, C. W. (1988). Registration of "Hutcheson" soybean. *Crop Science*, *28*(6), 1024–1025.

Carter, T. E., de Souza, P. I., & Purcell, L. C. (1999). Recent advances in breeding for drought and aluminum resistance in soybean. In H. Kauffman (Ed.), *World Soybean Conference VI* (pp. 106–125).

Carter, T. E., Todd, S. M., & Gillen, A. M. (2016). Registration of 'USDA-N8002' Soybean Cultivar with High Yield and Abiotic Stress Resistance Traits. *Journal of Plant Registrations*, *10*(3), 238–245. doi.org/10.3198/jpr2015.09.0057crc

Charlson, D. V., Bhatnagar, S., King, C. A., Ray, J. D., Sneller, C. H., Carter, T. E., & Purcell, L. C. (2009). Polygenic inheritance of canopy wilting in soybean [*Glycine max* (L.) Merr.]. *Theoretical and Applied Genetics.*, *119*, 587–594. doi.org/10.1007/s00122-009-1068-4

Cho, D., Kim, S. A., Murata, Y., Lee, S., Jae, S. K., Nam, H. G., & Kwak, J. M. (2009). De-regulated expression of the plant glutamate receptor homolog AtGLR3.1 impairs long-

term Ca2+-programmed stomatal closure. *Plant Journal*, *58*(3), 437–449. doi.org/10.1111/j.1365-313X.2009.03789.x

Devi, J. M., Sinclair, T. R., Chen, P., & Carter, T. E. (2014). Evaluation of elite southern maturity soybean breeding lines for drought-tolerant traits. *Agronomy Journal*, *106*(6), 1947–1954. doi.org/10.2134/agronj14.0242

Devi, M. J., & Sinclair, T. R. (2013). Nitrogen fixation drought tolerance of the slow-wilting soybean PI 471938. *Crop Science*, *53*(5), 2072–2078. doi.org/10.2135/cropsci2013.02.0095

Fletcher, A. L., Sinclair, T. R., & Allen, L. H. (2007). Transpiration responses to vapor pressure deficit in well watered 'slow-wilting' and commercial soybean. *Environmental and Experimental Botany*, *61*(2), 145–151. doi.org/10.1016/j.envexpbot.2007.05.004

Holland, J. B., Nyquist, W. E., & Cervantes-Martínez, C. T. (2003). Estimating and Interpreting Heritability for Plant Breeding: An Update. In J. Janick (Ed.), *Plant Breeding Reviews* (Vol. 22, pp. 9–112). John Wiley and Sons, Inc. doi.org/10.1002/9780470650202.ch2

Hudak, C. M., & Patterson, R. P. (1996). Root distribution and soil moisture depletion pattern of a drought-resistant soybean plant introduction. *Agronomy Journal*, *88*(3), 478–485. doi.org/10.2134/agronj1996.00021962008800030020x

Hufstetler, E. V., Boerma, H. R., Carter, T. E., & Earl, H. J. (2007). Genotypic variation for three physiological traits affecting drought tolerance in soybean. *Crop Science*, *47*, 25–35.

Hwang, S., King, C. A., Chen, P., Ray, J. D., Cregan, P. B., Carter, T. E., Li, Z., Abdel-Haleem, H., Matson, K. W., Schapaugh, W., & Purcell, L. C. (2016). Meta-analysis to refine map position and reduce confidence intervals for delayed-canopy-wilting QTLs in soybean. *Molecular Breeding*, *36*(7), 91. doi.org/10.1007/s11032-016-0516-5

Hwang, S., King, C. A., Ray, J. D., Cregan, P. B., Chen, P., Carter, T. E., Li, Z., Abdel-Haleem, H., Matson, K. W., Schapaugh, W., & Purcell, L. C. (2015). Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. *Theoretical and Applied Genetics*, *128*(10), 2047–2065. doi.org/10.1007/s00122-015-2566-1

Kadioglu, A., Terzi, R., Saruhan, N., & Saglam, A. (2012). Current advances in the investigation of leaf rolling caused by biotic and abiotic stress factors. *Plant Science, 182*, 42–48. doi.org/10.1016/j.plantsci.2011.01.013

Kaler, A. S., Ray, J. D., Schapaugh, W. T., King, C. A., & Purcell, L. C. (2017). Genome-wide association mapping of canopy wilting in diverse soybean genotypes. *Theoretical and Applied Genetics, 130*(10), 2203–2217. doi.org/10.1007/s00122-017-2951-z

Lelandais-Brière, C., Jovanovic, M., Torres, G. A. M., Perrin, Y., Lemoine, R., Corre-Menguy, F., & Hartmann, C. (2007). Disruption of AtOCT1, an organic cation transporter gene, affects root development and carnitine-related responses in Arabidopsis. *Plant Journal*, *51*(2), 154–164. doi.org/10.1111/j.1365-313X.2007.03131.x

Luo, Y., Fang, B., Wang, W., Yang, Y., Rao, L., & Zhang, C. (2019). Genome-wide analysis of the rice J-protein family: identification, genomic organization, and expression profiles under multiple stresses. *Biotech, 9, 358*- doi.org/10.1007/s13205-019-1880-8

Mace, E. S., Singh, V., van Oosterom, E. J., Hammer, G. L., Hunt, C. H., & Jordan, D. R. (2012). QTL for nodal root angle in sorghum (Sorghum bicolor L. Moench) co-locate with QTL for traits associated with drought adaptation. *Theoretical and Applied Genetics*, *124*(1), 97–109. doi.org/10.1007/s00122-011-1690-9

Pantalone, V., & Rebetzke, G. (1996). Phenotypic evaluation of root traits in soybean and applicability to plant breeding. *Crop Science*, *36*, 456–459.

Pathan, S. M., Lee, J.-D., Sleper, D. A., Fritschi, F. B., Sharp, R. E., Carter, T. E., Nelson, R. L., King, C. A., Schapaugh, W. T., Ellersieck, M. R., Nguyen, H. T., & Shannon, J. G. (2014). Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought. *Journal of Agronomy and Crop Science*, *200*(3), 231–236. doi.org/10.1111/jac.12053

Purcell, L. C., & Specht, J. E. (2004). Physiological Traits for Ameliorating Drought Stress. In *Soybeans: improvement, production, and uses* (Vol. 3, pp. 569–620). American Society of Agronomy.

Ramesh, S. A., Kamran, M., Sullivan, W., Chirkova, L., Okamoto, M., Degryse, F., McLaughlin, M., Gilliham, M., & Tyerman, S. D. (2018). Aluminum-activated malate transporters can facilitate GABA transport. *Plant Cell*, *30*(5), 1147–1164. doi.org/10.1105/tpc.17.00864

Rauf, S., Al-khayri, J. M., Zaharieva, M., Monneveux, P., & Khalil, F. (2016). Breeding strategies to enhance drought tolerance in crops. In *Advances in Plant Breeding Strategies: Agronomic, Abiotic and Biotic Stress Traits* (Vol. 2, pp. 397-445). Springer, Cham

Riar, M. K., Cerezini, P., Manandhar, A., Sinclair, T. R., Li, Z., & Carter, T. E. (2018). Expression of drought-tolerant N fixation in heterogeneous inbred families derived from PI471938 and Hutcheson soybean. *Crop Science*, *58*, 364–369. doi.org/10.2135/cropsci2017.02.0089

Ries, L. L., Purcell, L. C., Carter, T. E., Edwards, J. T., & King, C. A. (2012). Physiological traits contributing to differential canopy wilting in soybean under drought. *Crop Science*, *52*(1), 272–281. doi.org/10.2135/cropsci2011.05.0278

Sadok, W., Gilbert, M. E., Raza, M. A. S., & Sinclair, T. R. (2012). Basis of slow-wilting phenotype in soybean PI 471938. *Crop Science*, *52*(3), 1261–1269. doi.org/10.2135/cropsci2011.11.0622

SAS Institute. (2014). *The SAS system for Windows. Release 9.4*.

Shelp, B. J., Bozzo, G. G., Zarei, A., Simpson, J. P., Trobacher, C. P., & Allan, W. L. (2012). Strategies and tools for studying the metabolism and function of γ-aminobutyrate in plants. II. integrated analysis. *Botany*, *90*(9), 781–793. doi.org/10.1139/b2012-041

Sinclair, T. R., Purcell, L. C., Vadez, V., Serraj, R., King, C. A., & Nelson, R. (2000). Identification of soybean genotypes with N fixation tolerance to water deficits. *Crop Science*, *40*, 1803–1809. doi.org/10.2135/cropsci2000.4061803x

Sloane, R. J., Patterson, R. P., & Carter, T. E. (1990). Field drought tolerance of a soybean plant introduction. *Crop Science*, *30*, 118–123. doi.org/10.2135/cropsci1990.0011183X003000010027x

Soybase (2022). Soybean Pedigree Database, Retrieved from www.soybase.org/uniformtrial/index.php?page=lines

SoyStats (2021), Retrieved from http://soystats.com/

Specht, J. E., Diers, B. W., Nelson, R. L., de Toledo, J. F. F., Torrion, J. A., & Grassini, P. (2014). Soybean. In *Yield Gains in Major U.S. Field Crops* (Vol. 33, pp. 311–355). doi.org/10.2135/cssaspecpub33.c12

Specht, J., Hume, D., & Kumudini, S. (1999). Soybean yield potential—A genetic and physiological perspective. *Crop Science*, *39*, 1560–1570.

Steketee, C. J., Schapaugh, W. T., Carter, T. E., & Li, Z. (2020). Genome-wide association analyses reveal genomic regions controlling canopy wilting in soybean. *G3: Genes, Genomes, Genetics*, *10*(4), 1413–1425. doi.org/10.1534/g3.119.401016

Tanaka, Y., Fujii, K., & Shiraiwa, T. (2010). Variability of leaf morphology and stomatal conductance in soybean [Glycine max (L.) Merr.] cultivars. *Crop Science*, *50*(6), 2525–2532. doi.org/10.2135/cropsci2010.02.0058

Tuberosa, R. (2012). Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology*, *3*, 347. doi.org/10.3389/fphys.2012.00347

van Ooijen, J. W. (2006). JoinMap® 4: Software for the calculation of genetic linkage maps in experimental populations. *Kyazma BV, Wageningen*.

Voorrips, R. E. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity*, *93*(1), 77–78. doi.org/10.1093/jhered/93.1.77

Winterhalter, L., Mistele, B., Jampatong, S., & Schmidhalter, U. (2011). High throughput phenotyping of canopy water mass and canopy temperature in well-watered and drought stressed tropical maize hybrids in the vegetative stage. *European Journal of Agronomy*, *35*(1), 22–32. doi.org/10.1016/j.eja.2011.03.004

Ye, H., Song, L., Schapaugh, W. T., Ali, M. L., Sinclair, T. R., Riar, M. K., Mutava, R. N., Li, Y., Vuong, T., Valliyodan, B., Pizolato Neto, A., Klepadlo, M., Song, Q., Grover Shannon, J., Chen, P., Nguyen, H. T., & Foyer, C. (2020). The importance of slow canopy wilting in drought tolerance in soybean. *Journal of Experimental Botany*, *71*(2), 642–652. doi.org/10.1093/jxb/erz150

Figures and tables

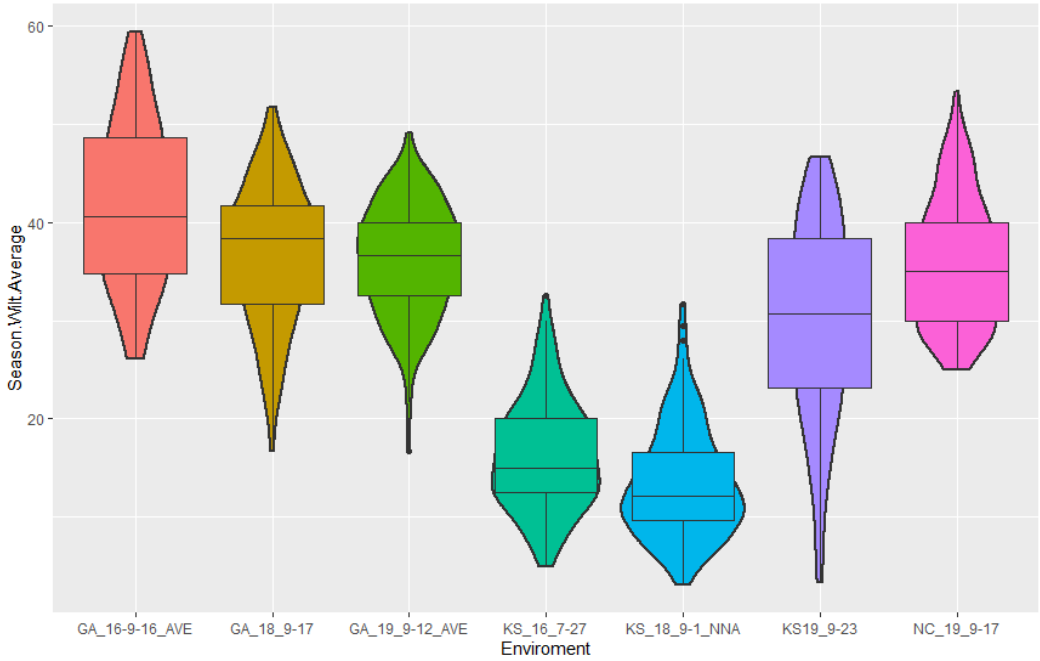


Figure 3.1. Distribution of canopy wilting scores for the recombinant inbred line population across the environments. Environments are named as Location-Year-Date, with Georgia (GA), Kansas (KS), and North Carolina (NC) as locations.

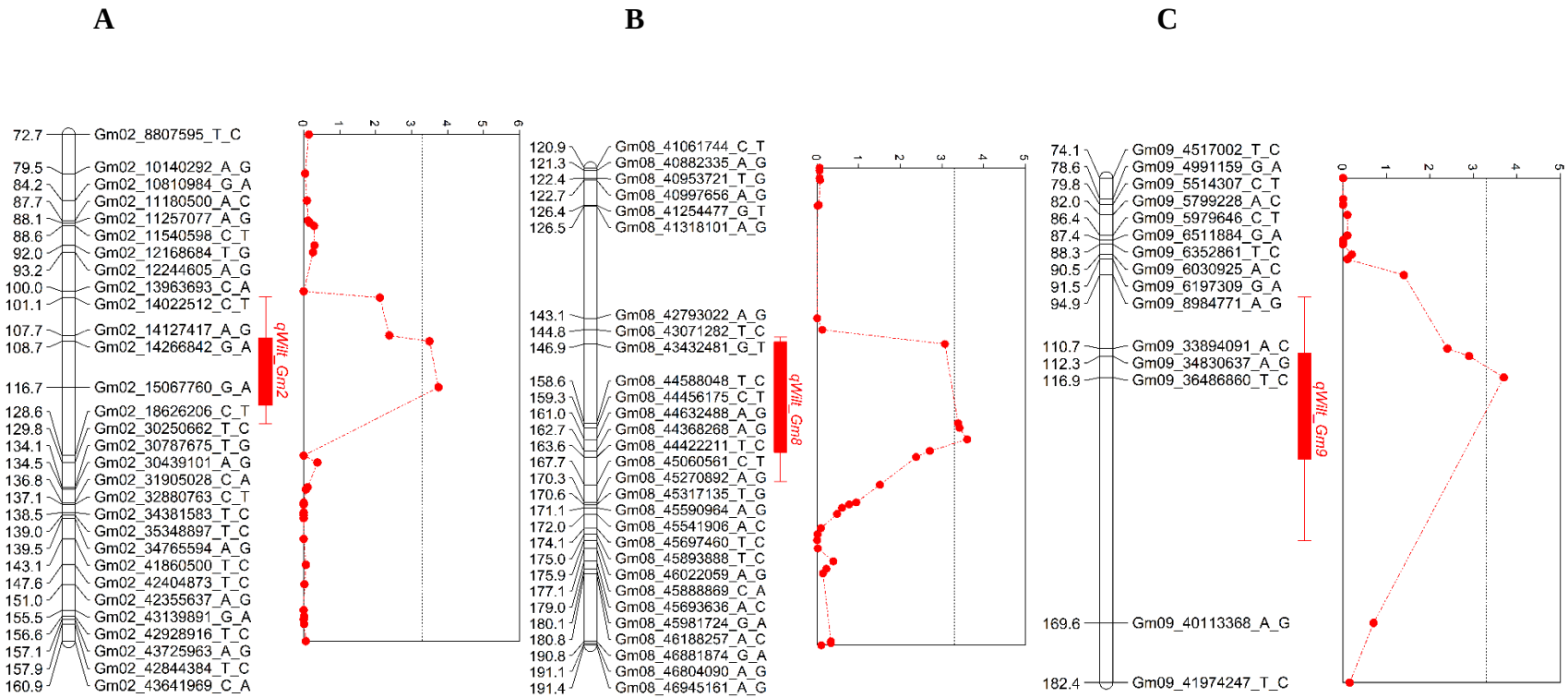


Figure 3.2. Composite interval mapping for canopy wilting in the recombinant inbred line population derived from Hutcheson × PI 471938 in the combined environments. Genetic maps with cM positions for chromosomes with QTLs meeting logarithm of odds (LOD) significance thresholds of 3.3 which is indicated by the dotted black lines. A) Segment of chromosome 2 harboring the *qWilt_Gm2*; B) Segment of chromosome 8 harboring the *qWilt_Gm8*; *C*) Segment of chromosome 9 harboring the *qWilt_Gm9*

Table 3.1. Analysis of variance (ANOVA) for effects of genotype (G), environment (E), and their interaction based on the RIL population canopy wilting scores. The G × E MS was used as the denominator of the F Value for significance testing.

Source	DF	F Value	P > F
Genotype (G)	129	2.38	<0.0001
Environment (E)	6	32.5	<0.0001
G × E	774	1.68	<0.0001

Table 3.2**.** Pearson correlations of canopy wilting scores among the environments.

	2016_KS	2018_GA	2018_KS	2019_GA	2019_KS	2019_NC
2016_GA	0.07	0.20*	0.15	0.21*	0.25**	0.26**
2016_KS		-0.08	0.15	0.19*	-0.09	-0.05
2018_GA			0.22*	-0.17	0.22*	0.13
2018_KS				0.44***	0.38***	0.25**
2019_GA						0.29***
2019_KS						0.38***

***, ***, and * indicate significance at <0.0001, 0.01, and 0.05, respectively

Table 3.3. QTLs for canopy wilting that were identified by composite interval mapping (CIM) for the Hutcheson × PI 471938 RIL population in the combined environments.

QTL Name	Chr [a]	Peak Marker	Pos (cM) [b]	CI (cM) [c]	Pos (bp) [d]	CI (bp) [e]	LOD [f]	Additive Effect [g]	R^2 (%)	Source of favorable Allele
qWilt_Gm2	2	Gm02_15067760_G_A	114.8	102.5-127.7	15271225	14220378-18913725	3.9	1.00	0.11	PI 471938
qWilt_Gm8	8	Gm08_44368268_A_G	162	146.9-167.9	45403652	44267551-45913059	3.6	0.95	0.10	PI 471938
qWilt_Gm9	9	Gm09_36486860_T_C	116.9	105.6-162.7	39047264	36455035-42790738	3.7	0.98	0.14	PI 471938

[a] Chromosome
[b] Position in centiMorgans based on the genetic map
[c] Confidence interval in centiMorgans which includes all SNPs that met logarithm of the odds (LOD) threshold
[d] Glyma.Wm82.a2 physical position of the peak SNP marker
[e] Confidence interval based on Glyma.Wm82.a2 physical positions of all SNPs that met logarithm of the odds (LOD) threshold
[f] Logarithm of the odds (LOD) of peak SNP marker
[g] Additive allelic effect

Table 3.4. Candidate genes in the QTL regions and their functional annotation using the Glyma2.1 gene models in SoyBase (www.soybase.org) within plus or minus 10 kb of the SNP with the highest LOD peak for canopy wilting QTLs from composite interval mapping.

QTL Name	SNP	Gene Name	Annotation
qWilt_Gm2	Gm02_15067760_G_A	*Glyma.02g147900*	Aluminum Activated Malate Transporter Family Protein
		Glyma.02g148000	Salt Stress Response/Antifungal
		Glyma.02g148100	F-Box Family Protein
qWilt_Gm8	Gm08_44368268_A_G	*Glyma.08g337000*	Cation Exchanger 5
		Glyma.08g337100	Chaperone Dna j-Domain Superfamily Protein
		Glyma.08g337200	Transducin/WD40 Repeat-Like Superfamily Protein
qWilt_Gm9	Gm09_36486860_T_C	*Glyma.09g166400*	Organic Cation/Carnitine Transporter1
		Glyma.09g166500	Integral Component of Membrane
		Glyma.09g166600	Nucleotide-Diphospho-Sugar Transferase Family Protein

Appendix A: Supplemental Tables

Supplementary Table 3.S1. QTLs for canopy wilting that were identified with composite interval mapping (CIM) for the Hutcheson × PI 471938 RIL population in the individual environments.

QTL Name	Chr[a]	Peak Marker	Pos (cM)[b]	CI (cM)[c]	Pos (bp)[d]	CI (bp)[e]	LOD[f]	Effect[g]	R^2 (%)	Env[h]
qWilt_Gm2.1	2	Gm02_12,244,605_A_G	97.2	88.3-107.3	12582331	11349126-14470563	5.1	1.6	0.13	Salina, KS
qWilt_Gm7	7	Gm07_6,781,309_T_C	45.3	43.3-50.9	6819959	6898693-7409086	4.2	1.7	0.11	Plains GA
qWilt_Gm8	8	Gm08_44,632,488_A_G	160.3	146.9-169.7	45140042	44267551-45913059	3.6	1.8	0.10	Salina, KS
qWilt_Gm13_1.1	13	Gm13_29,481,243_C_A	125.2	123.7-127.2	30681387	30117998-30875555	6.0	-1.7	0.13	Athens, GA
qWilt_Gm13_1.2	13	Gm13_39,978,113_C_A	244.2	230.5-251.4	41157215	39311323-42650588	4.0	-1.4	0.13	Athens, GA
qWilt_Gm13.3	13	Gm13_38,133,840_A_C	229.9	222.6-238.8	39311323	38159424-39797053	7.9	2.1	0.2	Midville, GA
qWilt_Gm15	15	Gm15_6,925,513_T_C	39.7	38.2-42.3	6939367	6841404-7358153	3.8	-1.4	0.08	Midville, GA
qWilt_Gm16	16	Gm16_7,851,145_G_A	58.9	51.9-63.4	8008387	6858504-8197236	4.5	-1.0	0.11	Salina, KS

[a] Chromosome
[b] Position in centiMorgans based on the genetic map
[c] Confidence interval in centiMorgans which includes all SNPs that met logarithm of the odds (LOD) threshold
[d] Glyma.Wm82.a2 physical position of peak SNP marker
[e] Confidence interval based on Glyma.Wm82.a2 physical positions of all SNPs that met logarithm of the odds (LOD) threshold
[f] Logarithm of the odds (LOD) of peak SNP marker
[g] Additive allelic effect
[h] Environment
[I] Date of canopy wilting rating

CHAPTER 4

PHENOTYPING CANOPY COVERAGE OF ELITE SOUTHERN SOYBEAN BREEDING LINES USING UNMANNED AERIAL VEHICLES[1]

[1] Ethan Menke, Zhiwu Zhang, Ai-Ping Hu, Zenglu Li. To be submitted to Remote Sensing.

Abstract

To increase breeding selection efficiency, plant breeders have had to adopt new high throughput phenotyping methods such as aerial imaging with unmanned aerial vehicles (UAVs). UAVs enable plant breeders to take measurements that are traditionally not collected due to the difficulty of phenotyping on large-scale breeding populations. Canopy coverage (CC) effects have been reported for the indeterminate soybean in early maturity groups, but this trait has not been evaluated in elite southern soybean germplasm. The objectives of this study were to: 1) develop a phenotyping method to evaluate canopy coverage in elite southern soybean germplasm; 2) understand the genetic variation of canopy coverage per se in the elite southern soybean populations; and 3) determine the effects of canopy coverage on agronomically important traits in a breeding program. Advanced and preliminary yield trials from the University of Georgia breeding program were phenotyped for CC using a UAV from 2018 to 2021. Flights were conducted in Weeks 4, 5, 6, and 7 after planting, and images were then used to generate CC estimates to evaluate the effects of CC on agronomic traits. The results indicated that flight images collected in Weeks 5 and 6 appear to be the most informative and heritable estimates for canopy coverage. Average CC estimates for the three datasets ranged from 23-32%, 36-49%, 49-55 %, and 43-48% for Weeks 4, 5, 6, and 7, respectively. Heritability of estimated CC ranged from 0.05 in 2018-19 for Week 4 to 0.52 in 2020-21 for Week 7. After establishing the heritability of the CC, the relationship between CC, yield, 100-seed weight, maturity, lodging, and plant height was determined. A significant negative correlation was observed between increased CC at week 4 for yield in three data sets, while a positive correlation was observed between 100-seed weights and CC estimates from Week 5 in 2018-19, and 2019-20. Plant height had a significant positive association with CC estimates in Weeks 4, 5, and 6 in 2019-20 and 2020-21. Lodging tended to be

positively associated with CC in the 2018-19 and 2019-20 data sets. Maturity was significantly negatively associated with CC in weeks 4 and 7 in the 2018-19 data set. The methodology and workflow generated in this study can enable plant breeders to rapidly phenotype canopy coverage in their breeding programs for selection.

Keywords: Unmanned aerial vehicle, Canopy coverage, High throughput phenotyping, Agronomic traits, Soybean

Introduction

The development of high throughput precision phenotyping is needed to advance the development of new varieties. In the last 25-years, the rapid deployment of new high-throughput genotyping had created a need for large amounts of high-quality phenotypic data. Generating the phenotypic data could greatly help the full utilization of these genetic advancements. New phenotyping methods have been rapidly developed in the last 10 years to address the bottleneck that has been created by these genetic technologies (Furbank &Tester, 2011; Fahlgren et al., 2015). These high throughput phenotyping technologies have been used in laboratories, greenhouses, and field-based experiments (Yang et al., 2020). Multiple methods of phenotyping have been deployed for field-based research, including handheld sensors, carts, and unmanned aerial vehicles (UAVs) (Yang et al., 2020). For field-based phenotyping to be used by plant breeders, the ability to phenotype a large number of relatively small plots with speed and accuracy is required. UAVs have become one of the major tools in high-throughput phenotyping because of their decreasing cost and ease of use, which can be used in phenotyping many plots in a matter of minutes as compared to traditional methods of phenotyping (Araus et al., 2018; Yang et al., 2020). UAVs

have been used to estimate abiotic stress, plant disease, and basic phenotypic traits such as plant height and maturity to increase the efficiency of plant breeding (Singh et al., 2016; Dobbels & Lorenz, 2019; Moreira et al., 2019; Zhou et al., 2021).

Canopy coverage (CC) in soybean is measured as the rate at which the soybean canopy closes the inter-row space. This trait has predominantly been studied in terms of how row spacing affects the rate at which the soybean canopy closes. These studies have shown that decreased row spacing increases the rate of CC and tends to increase yield (Boerma & Ashley, 1982). Faster CC could be selected for and was shown to decrease weed pressure in soybean production, without affecting grain yield (Jannink et al., 2001). This could make it useful in combating weed resistance, which has become a major problem in soybean production (Peterson et al., 2018). Traditionally, soybean breeders have not evaluated or selected lines for rapid CC, because it is labor-intensive to phenotype, but high throughput/precision phenotyping (HTPP) technologies such as UAVs are making it possible to phenotype this trait.

Xavier et al. (2017) performed a genome-wide association study (GWAS) on the SoyNAM population (www.soybase.org/SoyNAM), which consisted of 140 recombinant inbred lines (RILs) derived from each of 40 soybean lines that were crossed to a common parent, IA3023. Lines were phenotyped for CC using UAVs as well as digital images from 14-56 days after planting (Xavier et al., 2017). Both UAV and ground-based images were segmented in a binary classification scheme to estimate CC. Ground-based images were batch processed using the SigmaScanPro software to generate CC estimates in the method described by Karcher and Richardson (2005), while for UAV images orthomosaics were generated using Pix4Dmapper software. These orthomosaic images were then processed using ENVI 5.0 (L3Harris Geospatial Solutions, Inc., Broomfield, CO, USA) to generate binomial estimates of CC. The authors used individual CC

measurements from each observation date to perform a genome-wide association study (GWAS) (Xavier et al., 2015, 2017). Six genomic regions associated with CC were identified on six chromosomes. One QTL on Chromosome (Chr) 19 was positively associated with increased CC across the entire season and provided a 47.30 kg ha^{-1} increase in yield without increasing the maturity of the lines. The authors stated that CC was more heritable (H^2=0.76) compared to yield (H^2=0.58). While this population gives a wide range of diversity for mapping traits, it may not provide an accurate representation of the usefulness of using CC estimates in an elite breeding program. In addition, the lines were evaluated in an augmented design without replication, erroneous decisions could be made based on a single observation.

Kaler et al. (2018) examined a diverse panel of 373 MG IV soybean accessions from the USDA Soybean Germplasm Collection. Two groups of lines were selected: the first group had a yield greater than 2,500 kg/ha with good agronomic characteristics, and the second group was selected for diversity in addition to agronomic traits. Accessions were evaluated for the rate of CC over two years in five environments. Canopy coverage was estimated using images from a digital camera as described by Purcell (2000), at two time points approximately 4 and 6 weeks after emergence. Similar to Xavier et al (2017), SigmaScanPro was used to generate CC estimates using hue and saturation of images to select and classify pixels as soybean canopy or other using a batch process (Karcher & Richardson, 2005). CC estimates were then used to perform GWAS to identify genomic regions that were associated with CC (Kaler et al., 2018). This study found 15 genomic regions on 10 chromosomes associated with CC. These 15 regions have multiple SNPs in common between the two CC measurements and four of these regions were located within the six regions identified by Xavier et al. (2017). With the size of the population and the number of locations used in this study, collecting ground-based measurements took five days for the 1st evaluation and eight

days for the 2nd collection date. The large time differences between each of the measurements could add additional error to the mapping for loci associated with CC.

Moreira et al. (2019) evaluated the use of UAVs to estimate the CC of plant rows for the selection of lines to advance into yield trials in MG III. Red-Green-Blue (RGB) color images were collected in 2015 and 2016 in Indiana. In each season $F_{4:5}$ progeny rows in a non-replicated trial, with no common lines between years, were evaluated. UAV images were processed using Pix4D Mapper to generate orthomosaic images and then MATLAB (MathWorks, Natick, MA, USA) was used to create multilayer orthoimages for each plot, and finally, ENVI was used to generate CC estimates by using maximum likelihood classification. For each progeny row, best linear unbiased predictions (BLUPs) were calculated for yield, average canopy coverage (ACC), and yield with ACC being used as a covariate (Yield|ACC) (Moreira et al., 2019). Three groups each year were formed by selecting progeny rows based on the highest BLUPs for yield, ACC, and Yield|ACC (Moreira et al., 2019). These selections were then evaluated in replicated yield trials to compare the yield potential among these groups. The maturity of lines in the yield trials was collected and used as a covariate to produce BLUPs for the replicated yield trials to account for the effects of maturity on yield. The authors found that selecting for CC alone or in combination with yield (Yield|ACC) in the progeny row provided similar accuracy as using yield alone when selecting lines for advancement to yield testing for both years. This could allow breeders to make selections without having yield data, or in cases where yield data is of poor quality (Moreira et al., 2019). However, in most soybean breeding programs, yield testing is not performed at the progeny row stage due to resource limitations. When calculating BLUPs for the progeny rows maturity was not used as a covariate which could lead to bias in the selection of lines for later maturity.

Herrero-Huerta et al. (2020) examined the possibility of using UAVs to estimate the yield of 382 soybean RILs selected as a subset of the SoyNAM population. Twelve RILs were selected from 32 subpopulations of the SoyNAM panel to create a panel with similar maturity while keeping phenotypic variance for yield and other agronomic traits (Lopez et al., 2019). Plots were phenotyped at 61 and 70 days after planting during the 2018 growing season using both multispectral and RGB images on UAVs (Herrero-Huerta et al., 2020). The authors used Pix4Dmapper for producing orthomosaic images and 3D point clouds and CC estimates were then extracted for individual plots using the 'Triple S' pipeline. Triple S is an open-source python image processing pipeline that generates CC estimates using K-mean clustering to classify soybean canopy from the ground in images (Herrero-Huerta et al., 2019). The data from these images were then used to train two machine learning models, Random Forest and XGBoost, to predict the yield of these RILs. The data was divided into sets by withholding 15% of the plots being used for testing, and the remaining plots were used to train the machine learning models. After training the models the testing data was used to predict yield with an accuracy of R^2=0.38 and 0.42 for Random Forest and XGBoost, respectively. These studies demonstrated the relationship between soybean yield and rapid CC in northern soybean germplasm. Additionally, these experiments showed the potential for breeders to exploit the rapid CC trait and select for increased soybean yield during early generation testing. However, the authors stated that yield prediction models need to be retrained in different environments and when new lines are added. This would limit the usefulness of this method in a breeding program where there is a constant change in the locations and lines being evaluated from year to year.

The objectives of this study are to i) develop a workflow to phenotype CC; ii) understand the CC trait per se in elite southern breeding materials to determine if it can be selected as an

additional breeding trait; and iii) determine the relationship between rapid CC and agronomically important traits such as yield, 100-seed weight, plant height, lodging, and maturity.

Materials and Methods

Plant Materials and Experiment Design

In this study, elite breeding lines developed at the University of Georgia were evaluated over four growing seasons, from 2018 through 2021. Each year, the trials consist of lines in different stages of evaluations for potential release as new cultivars for the Southern United States. The progeny derived from these elite × elite crosses were selected and advanced over multiple seasons to generate stable inbred lines with sufficient seed for evaluation of yield. Each year, lines selected from plant rows are evaluated for yield and agronomic traits in UGA preliminary yield trials (PYTs) in replicated multi-location yield trials. Based on their performance in the PYTs, selected lines are advanced into UGA advanced yield trails (AYT) for further testing the following year. After selection again, the lines are advanced into regional yield trials across the Southeastern United States. Due to the fluctuating nature of the forward breeding process, comparisons were made between years by evaluating the breeding lines in their PYTs and AYTs occurring between consecutive years in this study. The 2018-19 data set consisted of 136 breeding lines; the 2019-20 data set was 170 breeding lines, and the 2020-21 data set was 186 breeding lines (Table 4.1).

In addition, 42 breeding lines that were observed to have fast and slow CC were selected from the lines grown in 2019 PYTs. These 42 breeding lines were planted in the 2020 growing season and were evaluated for CC to confirm their slow and fast CC phenotype. In 2021, 34 of these 42 lines with fast or slow CC were selected and planted as a validation set in a randomized complete block design with three replications to evaluate CC, agronomic traits, and yield. The

validation set along with the AYT experiments in 2021 were visually evaluated for CC by a single rater, using a 0 to 100% scale to estimate the canopy closure of each plot on each date that flight images were collected.

The lines were evaluated at the University of Georgia Iron Horse Plant Sciences Farm (33° 44' ~18", -83°18' ~10") near Watkinsville, GA. During the breeding process as lines are advanced from PYTs to AYTs, the number of rows per plot and replications of each line is increased to have a better evaluation of the yield across environments. Plots were planted on 9th June 2018, 15th and 16th May 2019, 1st and 2nd June 2020, and 17th and 18th May 2021. While the number of rows per plot varied between two-row plots for the PYTs and four-row plots for AYTs, all plots had a planted length of 4.8 m with a row spacing of 76 cm using an Almaco cone planter fitted with Almaco SkyTrip GPS tripping (Almaco, Nevada, Iowa) (Table 4.1). During the late growing season, the plots were end-trimmed to a length of 3.6 m to limit the edge effect of plots by removing equal portions of each side of the plots. Throughout the growing season, multiple applications of herbicide, fungicide, and insecticide were applied as needed following recommendations from the Georgia Soybean Production Guide (Bryant, 2021). Agronomic traits including plant height, lodging, maturity, and seed size were collected on yield plots in each of the growing seasons. Plant height was collected by measuring three representative plants from the middle two rows of each plot. Lodging of each plot was scored on a scale of one to five with one being completely erect and five being completely prostate. Maturity (R8) of each plot was noted, when 95% of the pods in a plot reached their mature colors (Fehr et al., 1971). Plots were harvested using research plot combines with weight systems to record plot seed weight and moisture at the time of harvest. Plot yields were converted into kg/ha^{-1} at 13% moisture for further analysis. Seed size was measured as 100-seed weight (g).

UAV Phenotyping

Aerial images were collected using a DJI Phantom 4pro with a DJI 20-megapixel RGB camera by placing the camera in a nadir view (DJI, Shenzhen, China). The Phantom 4Pro was controlled using a flight control app DroneDeploy (DroneDeploy, San Francisco, California) that controlled flight path, elevation, and speed. Images were collected at 36.5 m above the ground level with a ground sample distance of approximately 10 mm per pixel at a flight speed of approximately 3 m $s^{-1,}$ and a minimum of 65% front and side overlap between images. The aerial images were acquired on near cloud-free days between 10:00 am and 2:00 pm (EDT) weekly from emergence (VE) to flowering (R1).

Image Processing and Extraction

RGB images were downloaded from the UAV via the onboard SD card, on which the images were saved during the flight and uploaded to the image processing software Pix4D Mapper (Pix4D, Prilly, Switzerland) to generate orthomosaic field images. Pix4D automatically extracted the GPS metadata from each drone image (Fig. 4.1). The "3d maps" function processing template in Pix4D was chosen for this processing. When the process was complete orthomosaic models were created from each flight and saved for further image processing. Orthomosaics were then imported into QGIS (QGIS Association, 2022) to subdivide orthomosaics into sets of individual tests for the input into the software GReenfield Image Decoder (GRID) (Chen & Zhang, 2020). GRID is a python-based image analysis software that segments images based on pixel-wise K-means cluster analysis and provides a graphical user interface (GUI) to implement plot data extraction from large-scale field data (Fig. 4.1). After the center of plots is indicated, GRID then extracts features from each plot, such as the total number of pixels of the plot area, an average of

original channels, and default vegetative indices. It then exports the feature values for further evaluation of the CC trait (Fig. 4.1). After exporting image features from GRID, percent vegetation estimates were calculated and CC estimates for each week were analyzed along with agronomic data that were collected during and after the growing seasons (Fig. 4.1).

Statistical Analysis

A generalized linear model (GLM) was used to estimate covariance components for the CC trait. The PROC GLM in SAS 9.4 (SAS Institute, Cary, NC) was used to perform an ANOVA using the model suggested by Bondari, (2003): $y = 1\mu + G_i + E_j + (GE)_{ij} + \epsilon_{ijk}$ where μ is the grand mean; G_i is the genotypic effect of the ith genotype; E_j is the environmental effect of the jth environment; $(GE)_{ij}$ is the interaction term between the ith line in the jth environment. and ϵ_{ijk} is the residual error in the model. ANOVAs were fitted independently for each of the traits of interest. PROC MIXED was used to estimate the variance components of plot yield, plant height, 100-seed weight, lodging, and maturity using a similar model as above for CC estimates, but in these instances genotype, environment, and genotype by environment were treated as random effects.

Broad sense heritability on an entry mean basis was calculated as $H^2 = \sigma_G^2/[\sigma_g^2 + \left(\frac{\sigma_{GE}^2}{y}\right) + \left(\frac{\sigma_e^2}{ry}\right)]$ where σ_G^2 is the genotypic variance; σ_{GE}^2 is the genotype by environment variance; and σ_e^2 is the residual variance; y is the number of years; and r is the number of replications. To account for the variation in the number of replicates, the harmonic mean of replication within each environment was calculated each year (Holland et al., 2003; Schmidt et al., 2019).

Pearson correlation between the LS means of each CC estimate and BLUPs for yield, 100-seed weight, plant height, lodging, and maturity for each of the years of the three data sets were

calculated using the multivariate function in JMP Pro 16 (SAS Institute, Cary, NC). Additional correlation calculations were made of individual years in each data set to evaluate the relationship of CC and agronomic traits between years in each data set.

Results

Phenotyping canopy coverage

The implementation of a method for phenotyping a large-scale breeding program was assessed in this research throughout the experiment. With a commercially available drone and flight control software, phenotyping the UGA Soybean Breeding yield trials at a single location could be easily completed in around 1-2 hours. This is significantly less time than is needed to evaluate even a small portion of the field for a single trait with visual evaluations. Orthomosaic image generation in Pix4D requires additional time but could be completed overnight. GRID was designed to meet three challenges that make using orthomosaic UAV images for high-throughput phenotyping difficult (Chen & Zhang, 2020). The 1st challenge that GRID was developed to overcome was the lack of geographical information in most orthomaic images. Software such as Quantum Geographical Information System (QGIS) (QGIS Project,QGIS.org) can be used to overcome this problem by manually defining the areas of interest in an image, but this can be time-consuming which creates a second challenge (Chen & Zhang, 2020). GRID addresses both of these challenges by using the layout of most field-based experiments in a consistent grid with repeating rectangular plots (Chen & Zhang, 2020). By identifying plot centers based on the layout of the rows and columns in a field experiment, GRID can automatically be segmented into individual plots. Additionally, the software allowed for slight adjustments in areas of interest to accurately extract plot information. The final challenge in using orthomosaic images usually requires the

extraction of image features by training a model to classify pixels, this can be time-consuming (Chen & Zhang, 2020). GRID was developed to be an easy-to-operate software that could be used to extract usable data from UAV images. The use of the GRID image processing system made plot images extraction feasible on a large scale for a breeding program. Processing a single flight for CC estimates for the entire UGA breeding program with GRID took approximately 4 hours. This is a considerable time saving compared to the collections of CC based on visual evaluations, which would have taken multiple days to complete the ratings.

Variation of canopy coverage

In the 2018-19 data set, CC ranged from 16-46%, 19-61%, 30-74%, and 30-66% in Weeks 4, 5, 6, and 7, respectively. CC estimates from Week 5 in 2018 were negatively correlated with Week 5 CC estimates from 2019, but no significant correlation was observed between 2018 and 2019 for Weeks 4, 6, and 7(Table 4.2). For the 2019-20 data set, Week 4 had CC ranging from 11-41% with a mean of 23%; Week 5 had a mean of 36% and ranged from 14-51%; for Week 6 a range of 21-67% was observed with a mean of 40; and for Week 7, CC estimates ranged from 33%-61% with a mean of 47%. Significant positive correlations were observed between 2019 and 2020 for CC estimates from Weeks 4 and 6, while no correlations were observed between 2019 and 2020 for Weeks 5 and 7(Table 4.2). In the 2020-21 data set, the mean CC measurements ranged from 21-48% with a mean of 32% for Week 4; Week 5 ranged from 39 -67% with a mean of 49%; the mean of Week 6 was 49% with a range of 35-63%, and Week 7 CC estimates ranged from 21-57% with a mean of 43%. The CC estimates from 2020 and 2021 in Week 7 were significantly positively correlated, however, the CC estimates from Weeks 4,5, and 6 were not significantly correlated between 2020 and 2021 (Table 4.2).

To evaluate the effectiveness of using UAVs to estimate CC, correlations between CC estimates from UAV images and visual evaluation collected on the 2021 AYT and validation set were calculated. It was observed that there was a significant positive correlation between visual evaluation and UAV CC estimates ($r=0.09$ for Week 4 and 0.52 for Weeks 5 and 6) at $p \leq 0.05$ (Fig. 4.2). However, no significant correlation was observed between CC estimates in Week 7 between UAV images and visual evaluations in the 2021 AYT and validation set (Fig. 4.2). No differences in ground-based CC estimates were observed between the fast and slow CC groups of the validation set grown in 2021. In contrast, for UAV data, a significant difference was observed only in Week 4 between the fast and slow CC groups in the validation set. The slow CC group in the validation set had a mean CC estimate of 26% for Week 4, whereas the fast CC group had a mean CC estimate of 29%. For both collection methods, there was no difference between the fast and slow CC groups for Weeks 5, 6, and 7 in the validation set for 2021.

Analysis of Variance and Heritability of Canopy Coverage

In all three data sets, ANOVAs revealed that genotype, year, and genotype by year all explained a significant portion of the variance in CC estimates from flights in Weeks 4, 5, 6, and 7 in 2018-19 and 2020-21. However, for the 2019-20 dataset Week 7 CC estimates neither genotype nor genotype by year interaction was significant and only the year was significant in explaining the observed variation in CC. For the lines included in the 2018-19 dataset, the heritability of CC ranged from 0.05 in Week 4 to 0.43 in Week 7, with CC estimated from Weeks 5 and 7 flights being the most heritable (Table 4.3). For the 2019-20 dataset, the heritability of CC was 0.08 in Week 7 to 0.30 in Week 6 (Table 4.3). In this dataset flights from Weeks 4 and 6 had

the highest heritability. The heritability of the 2020-21 dataset ranged from 0.12 for Week 4 to 0.52 in Week 7 CC estimates, with Week 7 being the most heritable (Table 4.3).

Correlation of Agronomic Traits with Canopy Coverage

After examining the heritability of the CC trait in each of the data sets, the correlation between agronomically important traits and CC was considered. Plot yield was significantly positively correlated with the Weeks 5 and 6 CC estimate in the 2018-19 data set ($p \leq 0.01$), however, the Week 4 CC estimate was significantly negatively correlated to yield ($p \leq 0.01$) (Table 4.4). In 2018-19, Week 7 CC estimates were not significantly associated with soybean yields (Table 4.4). Plot yield was significantly negatively correlated with Weeks 4 and 5 CC estimates for the 2019-20 data set ($p \leq 0.01$), while Weeks 6, and 7 estimates were not significant (Table 4.4). CC estimates and yield for Weeks 5, 6, and 7 were not correlated in the 2020-21 dataset ($p \leq 0.05$), while yield and CC estimates from Week 4 were observed to have a significant negative correlation in the 2020-21 dataset ($p \leq 0.05$) (Table. 3).

For 100-seed weight, significant positive correlations with Weeks 5 and 6 CC estimates were observed in the 2018-19 dataset ($P \leq 0.01$), and CC estimates of Weeks 4 and 7 were not significantly correlated to 100-seed weight (Table 4. 3). In the 2019-20 data set, CC estimates from Week 5 were positively associated ($p \leq 0.01$) with 100-seed weight, while Week 6 was negatively associated. ($p \leq 0.001$) (Table 4.4). No significant correlation was observed between CC estimates in Weeks 4 and 7 and 100-seed weight in the 2019-20 data set. In the 2020-21 data set no significant association was present between all CC estimates and 100-seed weight.

Maturity was significantly negatively associated with CC estimates for Weeks 4 and 7 in the 2018-19 dataset (Table 4.4), and no association between maturity and CC was observed in

Weeks 5 and 6. Maturity and CC from all weeks in both the 2019-20 and 2020-21 data sets were not significantly correlated.

Plant height and CC estimates from Weeks 4, 5, 6, and 7 were not significantly associated with plant height in 2018-19 (Table 4.4). In the 2019-20 data set CC estimates from Weeks 4, 5, and 6 and plant heights were significantly positively correlated ($p \leq 0.01$), and Week 7 was not correlated to plant height (Table 4.4). CC estimates from all weeks in the 2020-2021 data set were significantly correlated to plant height ($p < 0.05$). For 2020-21 CC estimates from Weeks 4,5, and 6 there was a positive correlation to plant height, while a negative correlation of CC estimate from Week 7 was observed (Table 4.4).

CC estimates for Week 5 in the 2018-19 dataset were positively associated with lodging ($p \leq 0.001$) (Table 4.4). No significant correlations were observed for CC estimates from Weeks 4, 6, and 7 in 2018-19 (Table 4.4). Only Week 5 and 7 CC estimates from 2019-20 were significantly positively associated with lodging ($p \leq 0.01$), and Week 4 and 6 were not correlated with lodging. No correlations were observed between any CC estimate and lodging for the 2020-21 dataset (Table 4.4).

Discussion

Phenotyping Canopy Coverage of Soybean

Creating a streamlined methodology is a key part of implementing high throughput phenotyping. In this experiment phenotyping of plots was efficiently accomplished by using commercially available UAVs and flight planning software. An entire large breeding research field was easily phenotyped in one or two 15 min drone missions for each CC estimate. This is significantly less time than visually scoring canopy coverage in the 2021 AYT and validation set

which required a single rater several hours to complete although additional time is needed to post-process and extract image values. Each of the previous studies using UAVs also required post-image processing to generate orthomosaic images. Herrero-Huerta et al. (2020), Moreira et al. (2019), and Xavier et al. (2017) all used Pix4Dmapper to generate orthomosaic images, which is the same software that was used in this study. Both Moeira et al. (2019) and Xavier et al. (2017) used the geospatial software ENVI to perform image classification and feature extraction. Both Xavier et al. (2017) and Kaler et al. (2018) phenotyped CC using a digital camera, and extracted CC estimates using the SigmaScan Pro software (SPSS, Inc, Chicago, Ill) to classify pixels as soybean canopy or not by batch processing images based on pixel hue and saturation (Karcher and Richardson, 2005). Herrero-Huerta et al. (2020) used the open-source python package Triple S (Statistical computing of Segmented Soybean multispectral imagery) to generate CC estimates using k-means clustering. Both ENVI and Triple S create a binary classification of pixels as either soybean canopy or non-canopy. Triple S used a K-means clustering similar to GRID, while ENVI classification used a different classification scheme. In the case of Xavier et al (2017) a supervised classification, where training images with labeled pixels were given to the software to train the classification model to generate CC estimates from UAV images. Unlike the method used above by Kaler et al. (2017), GRID used unsupervised classification which didn't require model training to make CC estimates and it also allowed for easy processing of UAV images to generate CC estimates of the line under yield testing in a soybean breeding program.

These estimates were then used to examine the relationship between CC and agronomic traits in southern soybeans. It was observed that CC estimates collected during Weeks 5 (V3-4) and 6 (V5-6) were the most stable relationship with other agronomic traits. These UAV estimates for Weeks 5 and 6 were significantly correlated with ground-based CC estimates in the 2021 AYT

and validation set. This indicates that Week 5 and 6 estimates are most likely representative of actual CC measurements based on the correlation between UAV CC estimates and ground-truthed data. This aligns with the results from Kaler et al (2018) where the CC estimate from the 2nd collection period identified 12 SNPs that were significantly associated with CC. The peak SNP identified by Xavier et al (2017) on Chr 19 was significantly associated with CC coverage measurements that were collected during a similar time as this study. Moreira et al. (2019) used a season average CC estimate from all flight dates, which included CC estimates from 5 and 6 weeks post-planting. This method requires significant resources to collect and process images to generate a single measurement. Additionally, averaging the CC estimates across dates could hide some of the effects that individual CC estimates had on yield. For this reason, breeders could focus on completing flights between 5 and 6 weeks post-planting in the southern germplasm, which would provide useful information without generating a large amount of unused data.

Validation of CC estimates in elite breeding lines

In the 2021 AYT and validation set of the CC estimates from UAVs vs. visual estimation of CC, no significant correlation was observed between Weeks 7 from the UAV and ground-truthed data. The CC estimates from the Week 4 UAV ($p \leq 0.05$) and ground-truthed data were less correlated than estimates from Weeks 5 and 6 ($p \leq 0.001$). Flights that were collected before four weeks post-planting tended to be uninformative and this is most likely due to GRID's inability to distinguish plants in these images due to plants being smaller than the pixel resolution. In flights after Week 7, image segmentation became difficult for the software due to the lack of variation in the images of the canopy. In the studies by Moreira et al. (2019) and Xavier et al. (2017), UAV flights were collected approximately 2-8 weeks after planting, and in this study, UAV flights ended

approximately one week earlier. The difference in the ending of the collection period is due to the later planting date of southern breeding programs and that flights were stopped at the onset of flowering (R1) due to limited vegetate growth after this stage in determinant soybeans. Herrero-Huerta et al (2020) stated that they used the flight from 15 days post-planting as a reference to build a digital surface model of the field that was later phenotyped using a multispectral camera, indicating that little vegetative growth was present in those images. Moreira et al. (2019) collected only one CC estimate before 4 weeks post-planting in both years of the study. Kaler et al. (2018) collected images with a hand-held digital camera at approximately 4 and 6-weeks post-planting. These collection times agree with our results that phenotyping for CC is feasible between 4 and 7 weeks post-planting. Additionally, these data collection dates appear to have consistent heritability, indicating that UAV phenotyping can provide reliable estimates of canopy coverage.

The validation set of 34 lines that were selected as having fast and slow CC only exhibited differences in mean CC estimates at Week 4. This could indicate that even cultivars with slow canopy growth can catch up with lines that have fast CC in a relatively short time after planting. This also indicates the large environmental variance that is observed in the CC trait between years. The validation set shows the difficulty of selecting lines for advancement based only on a single trait such as canopy coverage which has a large genotype by environmental component. Additionally, initial phenotypes were confirmed in non-replicated single-row plots in 2020 and this could have led to an error in selecting lines for the validation set of Slow and Fast CC groups grown in 2021. Moreira et al. (2019) stated that abnormal CC growth in one study year had a significant impact on the results of the study, indicating the difficulty of non-replicated trials when phenotyping traits with large environmental variability like CC.

Evaluation of canopy coverage estimates and relationship with agronomic traits

Most previous studies that sought to evaluate the effects of rapid CC focused on early maturing soybean genotypes in MG II through IV. In most of the northern US soybean production, the MG 0-IV cultivars are predominantly indeterminate, whereas the southern United States tends to grow MG V-VIII cultivars which tend to have a determinant growth type. In these cases, indeterminate soybean cultivars continue growing in stem length and leaf production through and after flowering, while determinate cultivars cease stem growth at the onset of flowing or shortly thereafter (Bernard, 1972). This trend still holds in the UGA Soybean Breeding Program where determinant lines are predominant. For southern breeding programs to utilize the fast CC trait, the relationship between fast CC and later maturity groups needed to be further explored. Additionally, Xavier et al. (2017) and Herrero-Huerta et al (2020) used the SoyNAM mapping population and Kaler et al. (2018) used accessions from the USDA Soybean Germplasm Collection. The material used in these studies was not elite lines and thus the results from these studies may not be applied in elite breeding programs.

100-seed weight or seed size is an important trait in soybean which is commonly phenotyped by breeders. Seed size can play a role in seed composition. In most studies seed size is positively associated with increased protein in the soybean seed, but no association between oil and seed size was observed (Panthee et al., 2005). Commercial MG II-IV cultivars developed between 1928 and 2008 were used to examine the genetic gain that was achieved in soybeans for yield and other agronomic traits (Rincker et al., 2014). No relationship between year release and 100-seed weight was observed, indicating that yield increases over this time were due to increased seed per-unit area in new cultivars over older cultivars (Rincker et al., 2014). Seeds from three soybeans cultivars were sorted using mesh screens into multiple groups based on seed size (Place

et al., 2011). These sorted seeds were then planted and increased seed size was shown to increase the early season biomass in soybean (Place et al., 2011). Place et al. (2011) showed that this increase in biomass helped reduce the weed pressure in test plots, which could allow for reduced use of herbicides in soybean production. QTL mapping results indicated that 100-seed weight is controlled by a large number of genes that contribute a portion of the variation of the 100-seed weight trait (Ikram et al., 2020). In this study, CC estimates were positively associated with increased 100-seed weight, indicating that larger seed size led to larger canopy in four flight dates in Weeks 5 and 6 in 2018-19 and Week 5 and 7 in 2019-20. However, in Week 6 from 2019-20, a significant negative correlation was observed, and when examining the data in 2019 and 2020 independently only 2019 PYT data was significantly negatively associated with 100-seed weight. In the 2020-21 data set, no significant correlation was observed in 100-seed weight and CC estimates, which may be due to 100-seed weight not being collected in the 2020 PYTs. A positive correlation between 100-seed weight and CC is in agreement with previous reports that 100-seed weight and canopy development.

Yield is the most agronomically important trait in soybean. Since soybean cultivation increased in the late 1920's breeders have been working to increase the soybean yield across the United States (Boehm et al., 2019). Between 1928 and 2008, the yield of soybean was increased by 13.7 kg ha^{-1} per year across maturity groups V, VI, and VII (Boehm et al., 2019). Lines released between 1928 and 2008 in maturity groups III and IV had an increase of 23 and 25 kg ha^{-1} per year, respectively (Christenson et al., 2014). Between 1924 and 2012 in the United States, in MG 00-VII soybean, the rate of genetic gain for yield was 23.3 kg ha^{-1} per year (Specht et al., 2014). This indicates that breeders are continuing to make genetic improvements in selecting for higher soybean yields. In this study, positive correlations between increased CC estimates and soybean

yields were observed for Weeks 5 and 6 from 2018-19, while in 2019-20 a significant negative correlation was seen. The fast CC estimates from Weeks 5 and 6 in the 2018-19 data set had heritability that was comparable with the yield in the three data sets that were observed (Table 4.3 and 4.4). However, Week 4 estimates were significantly negatively correlated with yield in the three data sets, indicating that rapid early growth before Week 4 may not be beneficial to yield in southern soybean cultivars (Table 4.4).

Herrero-Huerta et al. (2020) found that utilizing UAV estimates can accurately predict yield, but required training the prediction model in every season to provide accurate predictions. But this method would require continued development of training models to estimate yield accurately inside of individual years. Moreira et al. (2019) showed that the selection of lines for advancement in early maturity group soybeans could be accomplished by selecting for CC with similar effectiveness as selecting for yield alone. Unlike the study by Moreira et al. (2019), yield testing of non-replicated progeny rows is not performed in most breeding programs due to the cost and time involved in harvesting so many yield plots and the limited reliability of these yield trials. The heritability of the CC trait in southern cultivars shows the possibility that CC estimates could be used for selection for increased CC. This means that breeders could potentially exploit this trait to make more informed selections of plant row materials that are advanced into yield trials.

Maturity is a key agronomic trait in soybeans due to the major role that it plays in soybean yield. When breeders are evaluating new varieties for potential release they are constantly evaluating cultivars for maturity to ensure that they are making selections for their target environments. For Weeks 4 and 7 CC estimates a negative association was observed with maturity in 2018-19. For the remaining dates in 2018-19 and all dates in 2019-20 and 2020-21, no significant correlation between CC and maturity was observed. These differences could be due to

differences in planting dates between 2018 and 2019. In 2018 planting was significantly delayed due to weather and could be responsible for the negative correlation between CC and maturity. Xavier et al. (2017) indicated that CC could be selected without increasing the days to maturity in the northern germplasm. This could be different in southern germplasm, where delayed planting may cause increased CC to be related to early maturity. This indicates that if selecting lines for CC, breeders would need to pay close attention to the changes in maturity.

2018-19 Week 5 and 2019-20 Weeks 5 and 7 estimates were positively associated with lodging and CC estimates. Lodging tends to be positively associated with CC in two of the three data sets. Lodging is collected at maturity in soybean breeding programs, this could cause differences in the relationship between CC and lodging. Lodging that is present early in the season could be interpreted by the software system as increased coverage. While lines the lodge later in the season may not show a relationship with CC. Most previous studies do not mention lodging when observing the CC trait. Moreira et al. (2019) did mention that in one year of study data increased plant height did lead to an increase in lodging and caused a problem when generating accurate canopy estimates on single row replicated plots.

Plant height in the 2019-20 and 2020-21 data sets tended to be positively associated with increased Weeks 4, 5, and 6 CC estimates and this is expected because an increase in plant height should be accompanied by an increase in canopy width (Table 4.4). Additionally, a rapid increase in plant height early in the growing season would allow plants to out-compete weeds in soybean fields, allowing for decreased yield loss from weed pressure (Jannink et al., 2000, 2001; Place et al., 2011). For the Week 7 measurement in 2020-21, a negative association between plant height and increased CC could indicate that shorter plants increase their canopy width later in the vegetative growth period than taller plants. Plant height and lodging were positively correlated in

both the 2018-19 and 2019-20 data sets and in the 2020-21 data set a negative correlation was observed. A similar trend was observed between plant height and yield wherein the 2018-19 and 2019-20 data sets there was a significantly negative correlation, which was not observed in the 2020-21 data set. Due to personnel limitations in 2020 plant height was not collected in PYTs. Herrero-Huerta (2020) used UAV estimated plant height as a trait in training their model for predicting yield in soybeans. Most other studies did not examine the relationship between plant height and canopy coverage. In the southern determinate varieties that were examined in this study, most plant height will be completed before the onset of flowering. The positive relationship observed between CC and plant height may indicate that selecting for early season canopy coverage could lead to an increase in plant height.

Variability in measurements for CC was observed between years, in the 2018-19 data set for all CC estimates except for Week 5, where 2018 had higher mean CC estimates as compared to 2019 estimates. Similar results were observed in the other two data sets where 2019 and 2021 had higher means of CC estimates as compared to 2020 CC estimates for both the 2019-20 and 2020-2021 data sets. This is expected due to the different environmental conditions over the years. Planting dates can vary by as much as three weeks between years. In 2019 and 2021 planting was done in mid-May, while in 2018 and 2020 planting occurred in early June due to unfavorable weather conditions in mid-May. In 2021 a significant increase in yield was observed across all experiments. This could be due to the earlier planting in the 2021 season, which possibly provided an optimal growing environment. Additional environmental variance such as the location where an individual test was planted in the field was observed to cause differences in the CC estimates. Variability in CC estimates is observed in most UAV phenotyping methods, Herrero-Huerta et al.

(2020) stated that variability in individual flights and environments could skew results in yield prediction based on CC estimates.

In the UGA soybean breeding program, it has been observed that populations with the glufosinate tolerant trait have progeny with more indeterminate growth habits than other populations in the breeding programs. This could explain why in the 2019-20 Week 6 flight and 2020-21 data the lines tested from the glufosinate tolerance program caused significant distortion in the data and prevented accurate heritability estimates from being calculated. This trend was observed in the 2019-20 CC estimate from Week 6.

Conclusion

Estimating CC can be accomplished fast and accurately using UAVs. This study established the feasibility of evaluating elite southern breeding lines for CC. Moderate heritability was observed for CC estimates from Weeks 5 and 6 in elite southern soybean breeding lines in the UGA Soybean breeding program between 2018 and 2021. CC estimates could be collected using a UAV to select lines with more rapid CC in elite southern breeding programs. In this experiment flights collected in Weeks 5 and 6 seems to be the most informative for the selection of CC. Breeders are currently limited in their ability to yield test early generation material. The correlations observed between CC estimates and traits such as yield and 100-seed weight could be exploited by breeders to select early generation material for advancements to yield trials. Using UAVs to select for increased CC could make breeding programs more efficient in evaluating only the most promising lines in yield testing.

References

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, *19*(1), 52–61. doi.org/10.1016/j.tplants.2013.09.008

Bernard, R. L. (1972). Two Genes Affecting Stem Termination in Soybeans. *Crop Science*, *12*(2), 235–239. doi.org/10.2135/cropsci1972.0011183x001200020028x

Boehm, J. D., Abdel-Haleem, H., Schapaugh, W. T., Rainey, K., Pantalone, V. R., Shannon, G., Klein, J., Carter, T. E., Cardinal, A. J., Shipe, E. R., Gillen, A. M., Smith, J. R., Chen, P., Weaver, D. B., Boerma, H. R., & Li, Z. (2019). Genetic improvement of us soybean in maturity groups V, VI, and VII. *Crop Science*, *59*(5), 1838–1852. doi.org/10.2135/cropsci2018.10.0627

Boerma, H. R., & Ashley, D. A. (1982). Irrigation, Row Spacing, and Genotype Effects on Late and Ultra-Late Planted Soybeans. *Agronomy Journal*, *74*(6), 995–999. doi.org/10.2134/agronj1982.00021962007400060015x

Bondari, K. (2003). Statistical Analysis of Genotype X Environment Interaction in Agricultural Research. *The Proceedings of the Southeast SAS Users Group*.

Bryant, C. (2021). *Soybean Production in Georgia*.

Chen, C. J., & Zhang, Z. (2020). GRID: A python package for field plot phenotyping using aerial images. *Remote Sensing*, *12*(11). doi.org/10.3390/rs12111697

Christenson, B. S., Schapaugh, W. T., An, N., Price, K. P., & Fritz, A. K. (2014). Characterizing changes in soybean spectral response curves with breeding advancements. *Crop Science*, *54*(4), 1585–1597. doi.org/10.2135/cropsci2013.08.0575

Dobbels, A. A., & Lorenz, A. J. (2019). Soybean iron deficiency chlorosis high throughput phenotyping using an unmanned aircraft system. *Plant Methods*, *15*(1). doi.org/10.1186/s13007-019-0478-9

Fahlgren, N., Gehan, M. A., & Baxter, I. (2015). Lights, camera, action: High-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology*, *24*, 93–99. doi.org/10.1016/j.pbi.2015.02.006

Fehr, W. R., Caviness, C. E., Burmood, D. T., & Pennington, J. S. (1971). Stage of Development Descriptions for Soybeans, Glycine Max (L.) Merrill. *Crop Science*, *11*(6), 929–931. doi.org/10.2135/cropsci1971.0011183x001100060051x

Furbank, R. T., & Tester, M. (2011). Phenomics - technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, *16*(12), 635–644. doi.org/10.1016/j.tplants.2011.09.005

Herrero-Huerta, M., Rodriguez-Gonzalvez, P., & Rainey, K. M. (2020). Yield prediction by machine learning from UAS-based mulit-sensor data fusion in soybean. *Plant Methods*, *16*(1). doi.org/10.1186/s13007-020-00620-6

Holland, J. B., Nyquist, W. E., & Cervantes-Martínez, C. T. (2003). Estimating and Interpreting Heritability for Plant Breeding: An Update. In J. Janick (Ed.), *Plant Breeding Reviews* (Vol. 22, pp. 9–112). John Wiley and Sons, Inc. doi.org/10.1002/9780470650202.ch2

Ikram, M., Han, X., Zuo, J. F., Song, J., Han, C. Y., Zhang, Y. W., & Zhang, Y. M. (2020). Identification of QTNs and their candidate genes for 100-seed weight in soybean (Glycine max L.) using multi-locus genome-wide association studies. *Genes*, *11*(7), 1–22. doi.org/10.3390/genes11070714

Jannink, J. L., Jordan, N. R., & Orf, J. H. (2001). Feasibility of selection for high weed suppressive ability in soybean: Absence of tradeoffs between rapid initial growth and sustained later growth. *Euphytica*, *120*(2), 291–300. doi.org/10.1023/A:1017540800854

Jannink, J. L., Orf, J. H., Jordan, N. R., & Shaw, R. G. (2000). Index selection for weed suppressive ability in soybean. *Crop Science*, *40*(4), 1087–1094. doi.org/10.2135/cropsci2000.4041087x

Kaler, A. S., Ray, J. D., Schapaugh, W. T., Davies, M. K., King, C. A., & Purcell, L. C. (2018). Association mapping identifies loci for canopy coverage in diverse soybean genotypes. *Molecular Breeding*, *38*(5). doi.org/10.1007/s11032-018-0810-5

Karcher, D. E., & Richardson, M. D. (2005). Batch analysis of digital images to evaluate turfgrass characteristics. *Crop Science*, *45*(4), 1536–1539. doi.org/10.2135/cropsci2004.0562

Lopez, M. A., Xavier, A., & Rainey, K. M. (2019). Phenotypic variation and genetic architecture for photosynthesis and water use efficiency in soybean (Glycine max l. Merr). *Frontiers in Plant Science*, *10*. doi.org/10.3389/fpls.2019.00680

Moreira, F. F., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2019). Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. *Plant Methods*, *15*(1). doi.org/10.1186/s13007-019-0519-4

Panthee, D. R., Pantalone, V. R., West, D. R., Saxton, A. M., & Sams, C. E. (2005). Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. *Crop Science*, *45*(5), 2015–2022. doi.org/10.2135/cropsci2004.0720

Peterson, M. A., Collavo, A., Ovejero, R., Shivrain, V., & Walsh, M. J. (2018). The challenge of herbicide resistance around the world: a current summary. *Pest Management Science*, *74*(10), 2246–2259. doi.org/10.1002/ps.4821

Place, G. T., Reberg-Horton, S. C., Carter, T. E., & Smith, A. N. (2011). Effects of soybean seed size on weed competition. *Agronomy Journal*, *103*(1), 175–181. doi.org/10.2134/agronj2010.0195

Purcell, L. C. (2000). Soybean canopy coverage and light interception measurements using digital imagery. *Crop Science*, *40*(3), 834–837. doi.org/10.2135/cropsci2000.403834x

Rincker, K., Nelson, R., Specht, J., Sleper, D., Cary, T., Cianzio, S. R., Casteel, S., Conley, S., Chen, P., Davis, V., Fox, C., Graef, G., Godsey, C., Holshouser, D., Jiang, G.-L., Kantartzi, S. K., Kenworthy, W., Lee, C., Mian, R., … Diers, B. (2014). Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. *Crop Science*, *54*(4), 1419–1432. doi.org/10.2135/cropsci2013.10.0665

Schmidt, P., Hartung, J., Rath, J., & Piepho, H. P. (2019). Estimating broad-sense heritability with unbalanced data from agricultural cultivar trials. *Crop Science*, *59*(2), 525–536. doi.org/10.2135/cropsci2018.06.0376

Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine Learning for High-Throughput Stress Phenotyping in Plants. In *Trends in Plant Science* (Vol. 21, Issue 2, pp. 110–124). Elsevier Ltd. doi.org/10.1016/j.tplants.2015.10.015

SoyStats (2021).Yield data. Retrieved from www.soystats.com/

Specht, J. E., Diers, B. W., Nelson, R. L., de Toledo, J. F. F., Torrion, J. A., & Grassini, P. (2014). Soybean. In *Yield Gains in Major U.S. Field Crops* (Vol. 33, pp. 311–355). doi.org/10.2135/cssaspecpub33.c12

Xavier, A., Hall, B., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2017). Genetic Architecture of Phenomic-Enabled Canopy. *Genetics*, *206*, 1081–1089. doi.org/10.1534/genetics.116.198713/-/DC1.1

Xavier, A., Xu, S., Muir, W. M., & Rainey, K. M. (2015). NAM: Association studies in multiple populations. *Bioinformatics*, *31*(23), 3862–3864. doi.org/10.1093/bioinformatics/btv448

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., & Yan, J. (2020). Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. *Molecular Plant*, *13*(2), 187–214. doi.org/10.1016/j.molp.2020.01.008

Zhou, J., Zhou, J., Ye, H., Ali, M. L., Chen, P., & Nguyen, H. T. (2021). Yield estimation of soybean breeding lines under drought stress using unmanned aerial vehicle-based imagery and convolutional neural network. *Biosystems Engineering*, *204*, 90–103. doi.org/10.1016/j.biosystemseng.2021.01.017

Figures and Tables

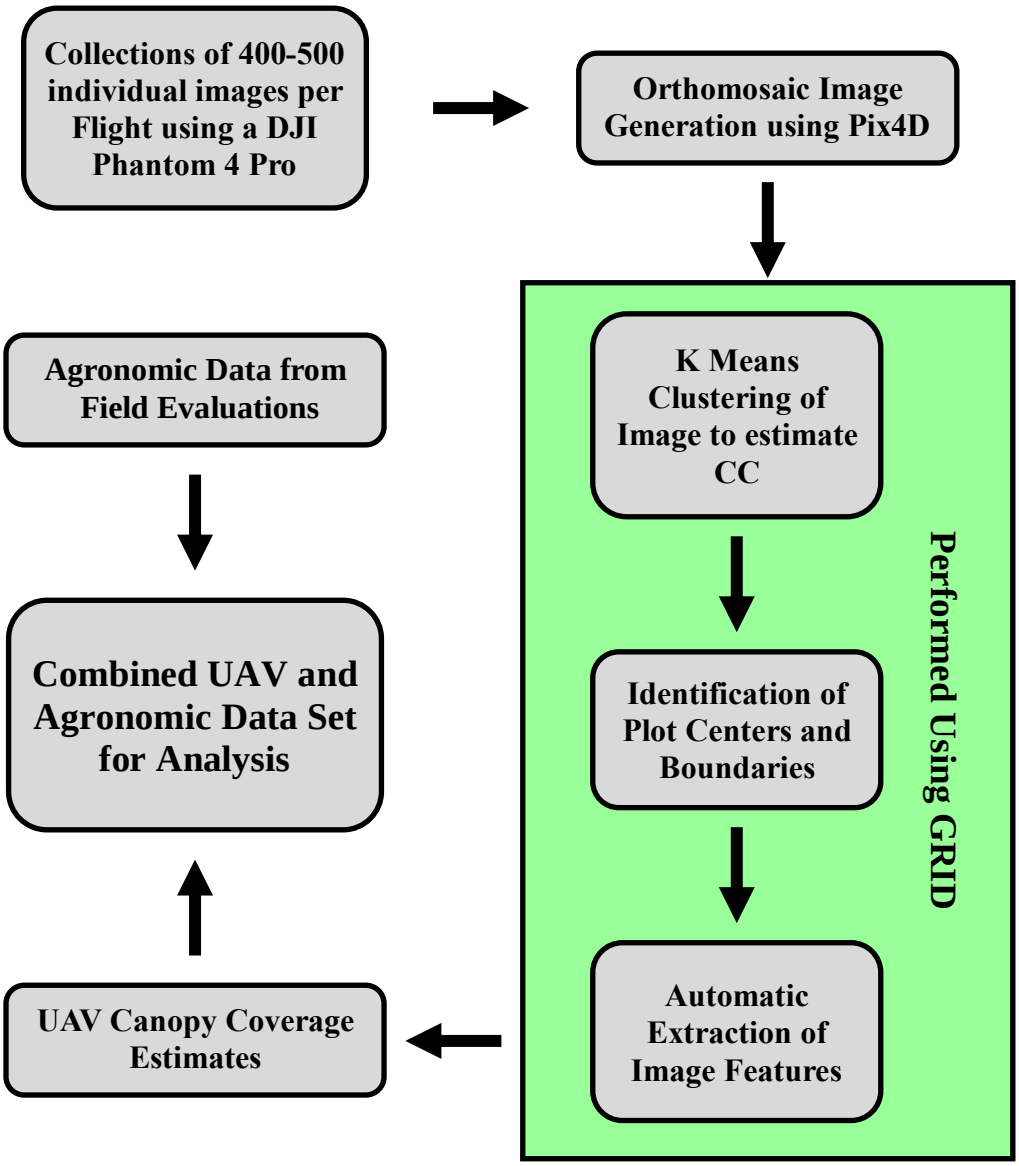
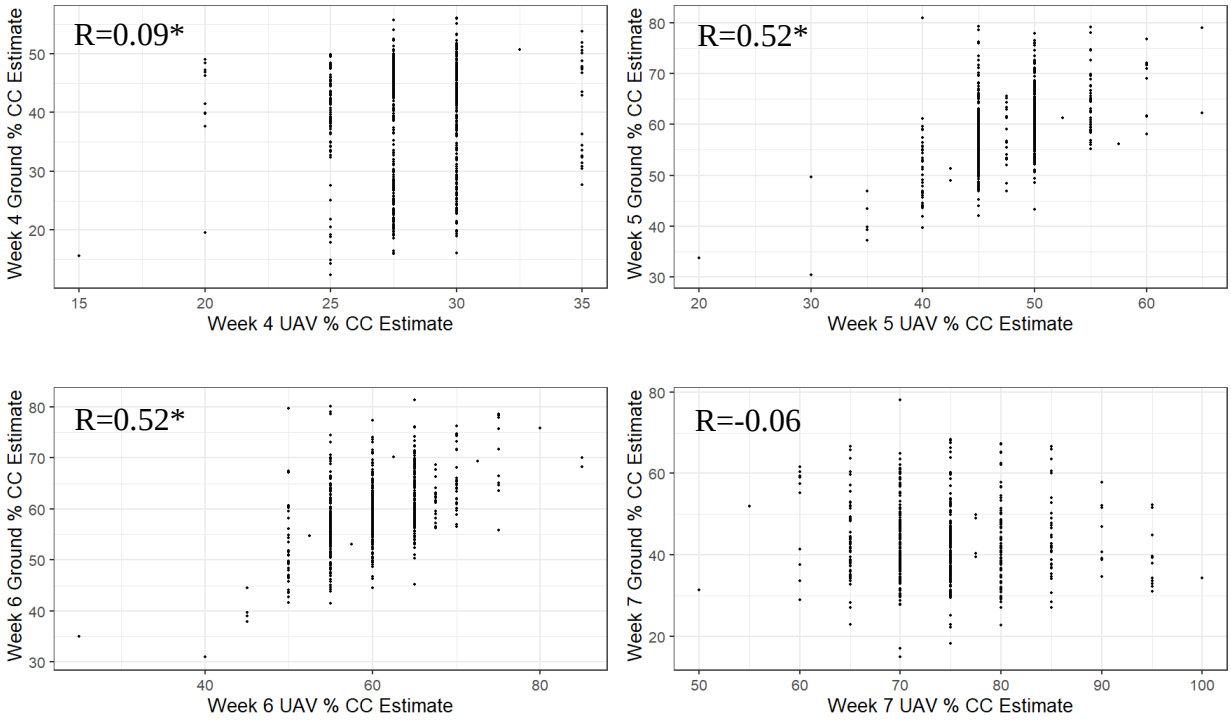


Figure 4.1. Overview of UAV image acquisition and processing to generate canopy coverage estimates.



** and ** indicate significance at p ≤ 0.05 and 0 .00, respectively*

Figure 4.2. Scatter plots of visual canopy coverage estimates compared with canopy coverage estimated from UAV images for the 2021 Canopy coverage validation set in the 2nd-year yield trials. Pearson correlation between each ground-truthed and UAV estimate for each week is indicated in the individual plots for each week.

Table 4.1. The number of genotypes evaluated in the 2018-19, 2019-20, and 2020-21 data sets.

Data Set	Number of Genotypes Evaluated	Replications		Planted Rows per Plot		Row length (m)	Row Spacing (cm)
		PYT*	AYT†	PYT*	AYT†		
2018-19	136	2	3	2	4	4.8	76
2019-20	170	2	3	2	4	4.8	76
2020-21	186	2	3	2	4	4.8	76
Validation Set-2021	34		3		2	4.8	76

*AYT -Advance yield trials
†PYT - preliminary yield trial

Table 4.2. Correlations between canopy coverage estimates from each flight collected between the years in each data set.

Year	Week 4	Week 5	Week 6	Week 7
2018-19	0.01	0.2*	-0.06	0.17
2019-20	0.16*	-0.04	0.24**	-0.13
2020-21	0.11	0.12	0.08	0.

, **, and * indicate significance at $p \leq 0.05$, 0.01, and 0.001 respectively*

Table 4.3. Broad sense heritability (H^2) of canopy coverage for each UAV flight based on the weeks post-planting in each of the three data sets.

Year	Week 4	Week 5	Week 6	Week 7
2018-19	0.05	0.41	0.31	0.43
2019-20	0.29	0.18	0.30†	0.08
2020-21	0.12†	0.38†	0.29†	0.52†

† indicates that lines in the data set from glufosinate tolerance breeding populations were excluded from the evaluations

Table 4.4. Correlations between canopy coverage estimates from each flight collected post-planting and key agronomically important traits across each of the three data sets

Year	Trait	Week 4	Week 5	Week 6	Week 7
2018-2019	Yield	-0.19**	0.16**	0.19**	-0.14
	100 Seed Weight	0.004	0.2**	0.42***	0.01
	Plant height	0.12	0.10	-0.06	0.07
	Lodging	-0.02	0.38***	0.1	-0.13
	Maturity	-0.21***	0.08	0.05	-0.18**
2019-2020	Yield	-0.16**	-0.28***	0.06	-0.03
	100 Seed Weight	-0.05	0.16***	-0.26***	0.14*
	Plant height	0.33***	0.16**	0.18**	-0.11
	Lodging	0.05	0.2***	-0.03	0.15**
	Maturity	0.003	0.04	-0.01	0.02
2020-2021	Yield	-0.13*	0.04	-0.04	0.10
	100 Seed Weight [a]	0.14	0.13	0.11	0.06
	Plant height [a]	0.42***	0.17*	0.17*	-0.29***
	Lodging	-0.05	0.07	0.04	0.09
	Maturity	-0.05	-0.07	-0.01	0.02

, **, and * indicate significance at $p \leq 0.05$, 0.01, and 0.001 respectively*

[a] Only the lines from 202 1AYT were used for analysis due to a lack of data for either trait from the 2020 preliminary yield trials.

CHAPTER 5

SUMMARY

Soybean [*Glycine max* (L.) Merr.] is one of the most cultivated oilseed crops in the United States and the world. The demand for soybean is set to increase as the world population increases. To meet this demand, soybean breeders must improve genetic gains of soybean through increased yield, disease resistance, abiotic stress tolerance, and more efficient breeding technologies.

Southern stem canker (SSC) in soybean, caused by the fungal pathogen *Diaporthe aspalathi,* is a major disease in the southern United States. It can cause up to 80% yield loss in severely infected fields. In Chapter 2, QTL mapping was performed to understand the resistance to SSC provided by PI 398469 and Crockett. RIL and $F_{2:3}$ populations were developed from a cross between G81-2057 and PI 398469 and an additional RIL population was also developed from a cross between G81-2057 and Crockett. Both RIL and the $F_{2:3}$ populations were phenotyped for resistance to SSC using the toothpick inoculation method. Phenotypes from the $F_{2:3}$ population derived from G81-2057 by PI 398469 were used to perform a bulked segregant analysis (BSA). The resistant and susceptible bulks along with the parents G81-2057 and PI 398469 were genotyped using the SoySNP50K BeadChip. BSA identified putative genomic regions that were related to SSC resistance. Two KASP markers were developed for each of the regions identified in the BSA and then used to genotype the G81-2057 × PI 39846 RIL population. Based on the single-marker analysis, genomic regions on Chrs 9, 10, and 14 were identified that were significantly associated with resistance to SSC. Additional KASP markers were developed to saturate the genomic regions and perform composite interval mapping. QTL analysis of the G81-

2057 × PI 398469 RIL population identified a QTL, *qRdm_14*, on Chr 14 with an R^2 of 62%, which confirmed the BSA result. The presence of minor QTLs for resistance to southern stem canker on Chrs 9 and 10 was examined, but no QTL for resistance to SSC was found. The G81-2057 × PI 398469 $F_{2:3}$ population was then genotyped using the KASP markers, confirming the presence of *qRdm_14* on Chr 14. Previous studies indicated that Crockett carries the *Rdm*3 resistance gene. To map the *Rdm*3 from Crockett and understand the relationship between *Rdm3* and *qRdm_14*, RILs derived from G81-2057 × Crockett were phenotyped for SSC resistance and genotyped with the KASP markers designed for the previous populations. The mapping results of the G81-2057 × Crockett RIL population confirmed that both PI 398469 and Crockett carry resistance alleles at the *Rdm3* locus for resistance to SSC. These KASP markers and QTL can be used by breeders to select for SSC resistance in breeding populations.

Slow or delayed canopy wilting is a physiological trait that is observed in a few exotic plant introductions (PIs) and may lead to yield improvement under drought stress. PI 471938, an accession from Nepal is a MG V introduction that exhibits slow canopy wilting, but the mechanism for this response to drought stress is unknown. In chapter 3, the RIL population derived from Hutcheson × PI 471938 was evaluated for canopy wilting under drought stress and genotyped with the SoySNP6K iSelect BeadChip. Field evaluations were conducted across the US in Georgia, Kansas, and North Carolina over four years to evaluate canopy wilting under rain-fed conditions, and data from a total of seven environments were collected due to the variability of weather and each location and year and used for QTL mapping. Significant variation among the genotypes for canopy wilting was observed among the RILs across environments. QTL analysis identified three QTLs for slow canopy wilting from the RIL population on Chrs 2, 8, and 9 based on the combined environmental analysis. These QTLs accounted for 11, 10, and 14% of phenotypic variation,

respectively. These QTLs can be used as targets to understand the physiological mechanisms of slow canopy wilting traits in PI 471938. Additionally, they could be deployed in breeding programs to incorporate the slow canopy wilting QTLs from PI 471938 into breeding material through targeted selection to develop drought-tolerant soybean cultivars.

Chapter 4 focused on the use of UAVs to estimate canopy coverage and understand the canopy coverage trait in the elite southern soybean germplasm. UAVs enable plant breeders to take measurements that are traditionally not collected due to the difficulty of phenotyping on large-scale breeding populations. Canopy coverage (CC) has been studied in the early maturity groups of soybean, but this trait has not been evaluated in elite southern soybean germplasm. The objectives of this study were to 1) develop a phenotyping method to evaluate canopy coverage in elite southern soybean cultivars; 2) understand the genetic variation of canopy coverage per se in the elite southern soybean populations; and 3) determine the effects of canopy coverage on agronomically important traits in a breeding program. The University of Georgia breeding programs advanced and preliminary yield trials were phenotyped for CC using UAVs between 2018 and 2021. UAVs were flown at 4, 5, 6, and 7 weeks after planting, and images were used to generate CC estimates. These CC estimates were then used to better understand the CC trait in elite southern cultivars and determine the relationship between CC and other agronomically important traits. Flights in Weeks 5 and 6 appeared to be the most informative and heritable estimates for canopy coverage. Average CC estimates in Week 4 ranged from 23-34% and increased throughout the growing season until Weeks 6 and 7 which ranged from 49-55 %. Canopy coverage had a range of heritability from 0.05 in 2018-19 Week 4 to 0.52 in 2020-21 Week 7. After the CC trait per se was evaluated, the relationships between CC, yield, 100-seed weight, maturity, lodging, and plant height were explored. A significant positive correlation was observed

between increased CC estimates at weeks 5 and 6 for yield in 2018-19, while a negative correlation was observed between yield and CC estimates from Week 4 in all three data sets that were evaluated in this study. Plant height has a significant positive association with CC estimates from Weeks 4, 5, and 6 in the 2019-20 and 2020-21 data sets. Early maturity was only significantly associated with increased CC in Weeks 4 and 7 of the 2018-29 data set. The methodology and workflow generated in this study can enable plant breeders to rapidly phenotype canopy coverage in their breeding program for selection.