DISCOVERY OF GERMPLASM AND GENOMIC REGIONS TO IMPROVE SOYBEAN DROUGHT TOLERANCE

by

CLINTON JAMES STEKETEE

(Under the Direction of Zenglu Li)

ABSTRACT

Drought stress is the most important abiotic constraint affecting soybean [*Glycine max* (L.) Merr.] yield in rain-fed production areas. Combating this stress requires soybean plants which possess physiological mechanisms to tolerate episodic drought stress, because less than 10% of U.S. soybean production areas are irrigated. Evaluation of physiological traits that relate to drought tolerance can be used in breeding programs to identify genomic regions associated with the traits and genotypes with favorable combinations of alleles. The objective of this research was to evaluate drought tolerance related traits in two populations in order to identify germplasm and genetic loci to improve soybean drought tolerance. A panel of over 200 genetically diverse soybean accessions genotyped with the SoySNP50K iSelect BeadChips was phenotyped for canopy wilting, carbon isotope composition ($\delta^{13}C$), nitrogen concentration, nitrogen isotope composition ($\delta^{15}N$), and transpiration response to $AgNO_3$. Additionally, 130 recombinant inbred lines (RILs) derived from 'Hutcheson' × PI 471938 genotyped with the SoySNP6K iSelect BeadChips were evaluated for canopy wilting. Genome-wide association analyses and composite interval mapping revealed genomic regions controlling these drought related traits in soybean, and new soybean accessions were identified with high numbers of

beneficial alleles and favorable breeding values for the traits. The germplasm and genomic regions identified through this research can be used to better understand the genetic architecture for these traits and be incorporated into elite germplasm to improve drought tolerance in soybean.

INDEX WORDS: Soybean, *Glycine max*, drought, canopy wilting, carbon isotope composition, water use efficiency, nitrogen concentration, nitrogen isotope composition, aquaporin, genome-wide association study (GWAS), quantitative trait loci (QTL) mapping

DISCOVERY OF GERMPLASM AND GENOMIC REGIONS TO IMPROVE SOYBEAN DROUGHT TOLERANCE

by

CLINTON JAMES STEKETEE

B.S., Michigan State University, 2012

M.S., University of Georgia, 2014

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

2018

DISCOVERY OF GERMPLASM AND GENOMIC REGIONS TO IMPROVE SOYBEAN DROUGHT TOLERANCE

by

CLINTON JAMES STEKETEE

Major Professor:	Zenglu Li
Committee:	H. Roger Boerma
	Chung-Jui Tsai
	Marc W. van Iersel
	Jason G. Wallace

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2018

DEDICATION

For my parents Jim and Karen, and my sister Kelsey – thank you for all of your love and support over the years.

ACKNOWLEDGEMENTS

I am grateful to many different people that have played a role in my professional development and in completing this dissertation research. My advisor, Dr. Zenglu Li, has been a great mentor and showed continuous positive support and encouragement as I progressed through this project. My committee members, Drs. H. Roger Boerma, CJ Tsai, Marc van Iersel, and Jason Wallace also provided valuable input and feedback on my dissertation. I am thankful to collaborators Drs. Bill Schapaugh, Tommy Carter, and Tom Sinclair for their help in making this project a success. I would also like to recognize fellow members of the soybean breeding and genetics lab and other IPBGG students/staff for their comradery, insights, and technical support: Dale Wood, Brice Wilson, Jeremy Nation, Earl Baxter, Ben Stewart-Brown, Dr. Josh Clevenger, Dr. Zach King, Tatyana Nienow, Kurk Lance, Ricky Zoller, Heather Kelley, Janette diMonda, Silas Childs, Ivy Tran, Liz Prenger, Nicole Bachleda, Mary Campbell, Dr. Justin Vaughn, Dr. Jeff Boehm, Dr. Miles Ingwers, Dr. Becky Tashiro, Gina Bishop, Congling Wu, Tom Maddox, Ethan Menke, and Brooks Arnold. This research and my fellowship were supported with funding from the United Soybean Board. I sincerely apologize to anyone who I may have omitted - I could not possibly provide the recognition that everyone deserves.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Soybean production and seed composition

Soybean [*Glycine max* (L.) Merrill] is among the most cultivated crops in the USA and throughout the world. In the USA, 31% of crop area planted was soybean in 2016, with only corn (35%) being planted on more hectares (ha). In 2016, the total value of the U.S. soybean crop was approximately $41 billion, and the USA was first in world soybean production with 117.2 million metric tons, followed closely by Brazil with 108 million metric tons. A total of 33.8 million ha of soybean were planted in the USA in 2016 with the national average yield being 3500 kg ha^{-1} (SoyStats, 2017).

The seed composition of soybean consists of approximately 20% oil and 40% protein, making it a valuable crop for oil production as well as a protein source for livestock. Soybean oil is comprised of different fatty acids with approximately 10% palmitic, 4% stearic, 22% oleic, 54% linoleic, and 10% linolenic acids present in most current commercially produced soybean (Wilson, 2004). Increased levels of oleic acid (greater than 75%) and lower levels of linolenic acid (less than 3%) have been targets for fatty acid composition improvement in soybean, because this reduces the need for hydrogenation of its oil and improves shelf life of these products (Pham et al., 2012). Compared to other vegetable sources of protein, soybean has the highest levels of crude protein, and its amino acid profile provides many of the nutritional requirements for swine and poultry production (Wilson, 2004). However, sulfur-containing amino acids methionine and cysteine are often deficient in soybean meal, and have to be

supplemented in soy-based animal feeds. Therefore, increasing the levels of these sulfur-containing amino acids is a breeding objective for soybean (Panthee et al., 2006). Simultaneous improvements in protein and oil content in soybean are challenging, because there is a negative correlation between protein and oil, and also a negative correlation between protein and grain yield (Patil et al., 2017). Approximately 86% of the residual dry mass of soybean seed is made up of carbohydrates including sucrose, starch, oligosaccharides (stachyose and raffinose), and other soluble sugars. Reducing the levels of stachyose and raffinose could improve the digestibility of soybean meal used in livestock diets (Wilson, 2004).

Importance of improving soybean drought tolerance

Drought stress is a significant issue threatening the agricultural productivity of soybean, and can reduce yields by 40% when comparing irrigated versus non-irrigated production in especially dry years (Specht et al., 1999). For breeding programs, drought is often defined as an insufficient supply of available soil water causing a reduction in crop yield. This decline in production is a result of the gap between demand for water and the supply of water available in its growing environment; the transpiration needs of the crop are not met (Monneveux et al., 2014). According to Blum (2011), there are three different types of drought that affect crop production. A prolonged period with less than average precipitation in a given location is referred to as meteorological drought, and typically comes before the other types of drought. Hydrological drought is brought about when water supplies from sources such as lakes, reservoirs, and aquifers fall below normal levels due to reduced precipitation or increased usage of water. The last type of drought is agricultural drought, which refers to insufficient moisture available for maximum or potential growth of crop plants. This type of drought can occur due to specific soil

conditions, topography, or other factors, even when precipitation is at average levels and can result in a range of yield losses from relatively minor to complete crop failure (Blum, 2011).

Irrigation is expensive and not a practical option for many soybean growing regions in the USA to alleviate stress caused by drought, thus cultivars with improved drought tolerance are needed to sustain and increase soybean production to feed a continuously growing world human population (Pathan et al., 2010). There is an expectation with climate change that events such as drought will be more frequent and extreme in the future (Dai, 2013). Producers of soybean must be prepared to deal with this important abiotic stress, but mitigating yield loss is currently difficult with few stress tolerant cultivars available. New soybean cultivars with improved drought tolerance will allow producers to reduce the impact of drought stress on soybean yields, and allow this important crop to utilize water more efficiently.

Effects of drought stress on soybean production

In 2012, approximately 9% of U.S. soybean hectares (ha) were irrigated (USDA-NASS, 2012). Given that the majority of soybean planted area is not irrigated throughout the USA, drought stress during the growing season can have a large effect on soybean yields. The impact of drought stress on soybean yields depends upon the region and time span of production. One study found that over a 50 year period (1958-2007), drought stress was associated with 13% of the variability of U.S. soybean and corn yields (Zipper et al., 2016). An estimate from another study found that U.S. average yield for soybean was suppressed by 30% from 1994-2013 due to year-to-year changes in precipitation and temperature, resulting in a loss of $11 billion (Mourtzinis et al., 2015). In 2012, a year with especially variable precipitation and water deficits in U.S. soybean producing areas, mean soybean yields in Arkansas for irrigated fields were 3248

kg ha^{-1} and were reduced to 1634 kg ha^{-1} under non-irrigated conditions. A similar trend was observed in Nebraska in 2012, with irrigated soybean yields of 4082 kg ha^{-1} and non-irrigated yields of 1675 kg ha^{-1} (USDA-NASS, 2018).

Reduction in seed quantity, quality, or weight and morphological changes to the plant are common effects of drought stress on soybean (Dornbos Jr. and Mullen, 1991; Frederick et al., 2001). The timing and duration of the stress also affect to what extent productivity is decreased, and several quantitative and qualitative parameters can be measured to determine the effect drought has on soybean production (Ku et al., 2013). Soybean growth is classified into separate vegetative and reproductive development stages. Vegetative stages start at V1 with one fully expanded trifoliolate, and continue to go up in number as new trifoliolates emerge. Reproductive stages start at R1 with one flower at any node and go through R8, which is full maturity (Fehr et al., 1971). Yield loss is typically greatest if water stress occurs during the reproductive growth stages (Korte et al., 1983; Brown et al., 1985; Kadhem et al., 1985; Eck et al., 1987; Desclaux et al., 2000; Frederick et al., 2001). Therefore, imposing water limiting conditions for screening purposes in the field and greenhouse takes place during the reproductive stages when possible. Drought stress has also been reported to reduce germination of harvested soybean seed that was exposed to drought conditions (Heatherly, 1993), and increase protein content while decreasing oil content under water deficit conditions (Vollmann et al., 2000).

Drought tolerance mechanisms in soybean

Most modern plant breeding studies utilize mild to intermediate drought stress to discover traits that mitigate yield losses in these growing conditions. This should be differentiated from finding traits and genotypes that perform well under severe drought (Purcell and Specht, 2004; Tuberosa,

2012). Some of these traits can result in yield reduction under favorable conditions, which would not be advantageous for soybean producers. For example, plants that use less water during times of drought stress may also use less water during times when water availability is high, which reduces the yield potential of the soybean plant (Blum, 2005).

There are several mechanisms that soybean plants can use to overcome drought stress, and many different definitions of these mechanisms and traits have been provided in literature. Drought escape and drought resistance are the most referenced mechanisms to reduce the effects of water stress. Drought escape is a management strategy that allows a plant to complete its life cycle when water supply is optimal prior to the onset of drought conditions (Pathan et al., 2010). An example of this is planting earlier maturity group (MG) soybean cultivars in southern U.S. regions than those commonly grown in the region, allowing soybean plants to avoid drought stress that typically occurs in dryer late summer months (Weaver-Missick, 1999). Drought resistance can be separated into avoidance and tolerance (Clark and Levitt, 1956). Drought or dehydration avoidance is the ability of a plant to avoid being stressed by allowing plant functions to continue via high plant water status or sustained cellular hydration. Enhanced capture of soil moisture, reduced water use, and osmotic adjustment to conserve cellular water content are examples of how plants can avoid dehydration (Blum, 2005). In soybean, traits that can lead to improved drought avoidance include high transpiration efficiency, fibrous and moderate lateral roots, deep taproot, low epidermal conductance, and high relative water content (Manavalan et al., 2009). Drought or dehydration tolerance is defined as the ability of the plant to continue metabolic activity and plant functions in a dehydrated state (Blum, 2005). Traits that are associated with dehydration tolerance in soybean include partial closure of stomata, leaf osmotic adjustment, high harvest index, and less ureide accumulation in petioles (Manavalan et al.,

2009). For purposes of this dissertation, all references to 'drought tolerance' encompass both forms of drought resistance for simplicity, and because in reality, soybean plants can never truly be completely 'resistant' to drought stress, but rather 'tolerate' it for periods of time.

Drought tolerance related traits and rationale for evaluation

A major obstacle for translation of the results of drought tolerance studies to improved soybean cultivars is the difficulty in rapidly and accurately phenotyping drought related traits. One main reason for this is that studying drought tolerance in the field can be difficult, as drought is unpredictable both spatially and temporally (Passioura, 2007). Additionally, directly selecting for grain yield under drought conditions is challenging as this trait has relatively low heritability, is controlled by many different genomic regions, and can have genotype by environment interactions (Rebetzke et al., 2013). Therefore, selection for drought tolerance using secondary or indirect traits is more common in breeding programs (Blum, 2011).

The traits to be phenotyped that ultimately affect soybean yield can either be classified as drought responsive or constitutive. The latter refers to traits that can also be expressed in well-watered conditions, while the former is expression of traits that are only visible under water deficit conditions (Tuberosa, 2012). Phenotyping for drought tolerance has tended to focus more on constitutively expressed traits that help to increase yield both with and without drought conditions (Blum, 2011). Ideally these traits to phenotype also have high heritability (greater than yield under stressed conditions) and enough genetic variability to detect differences among soybean genotypes that are evaluated. Also, measurement of the target traits should be inexpensive, reliable, and accurate (Tuberosa, 2012). Canopy wilting, water use efficiency (WUE), nitrogen fixation sensitivity to drought stress, and leaf hydraulic conductance are traits

that have been reported in soybean to understand and improve drought tolerance. In addition, root architecture, flowering time, canopy temperature, and use of remote sensing and imaging technologies are among the other traits that have also been studied in crop species.

Canopy wilting

Slow canopy wilting is observed when leaf wilting and loss of petiole turgidity are delayed during drought stress, and could lead to less yield reduction during water deficits for soybean. Canopy wilting is commonly phenotyped with a visual observation using a quantitative scale, and both scales of 0 (no wilting) to 100 (plant death), or 1 (no wilting) to 5 (plant death) have been used to measure the canopy wilting trait in soybean (King et al., 2009; Abdel-Haleem et al., 2012). Evaluation typically occurs in the earlier reproductive growth stages (R1-4) after the soybean plants have experienced drought stress for a period of time and is scored around solar noon (Abdel-Haleem et al., 2012).

A few plant introductions (PIs) have been previously identified as potential sources for drought tolerance improvement based on their slow wilting phenotype during stress conditions. One of these introductions, PI 416937, is a Japanese MG VI introduction identified in the 1980's as a line that wilted slower under water deficit conditions than other genotypes (Sloane et al., 1990). PI 416937 also has an extensive lateral root system and high root surface area according to previous research (Goldman et al., 1989; Sloane et al., 1990; Hudak and Patterson, 1996; Pantalone and Rebetzke, 1996). This PI has also been reported to have low stomatal conductance, which allows it to conserve water during drought stress conditions (Tanaka et al., 2010). Water conservation is the basis for the slow wilting trait found in PI 416937, based on research that showed it reaches a maximum transpiration rate near 2.0 kPa, while other

genotypes continue to increase their transpiration rates at higher vapor pressure deficit (VPD) conditions (Fletcher et al., 2007). PI 471938, a MG V introduction from Nepal, has also been reported to have the slow wilting trait, but the basis for this drought tolerance is still not known (Sadok et al., 2012). PI 416937 and PI 471938 are commonly used as checks in evaluations of soybean canopy wilting after water deficit stress has occurred. In addition, two other PIs (PI 567690 and PI 567731) were previously reported to have the slow wilting phenotype and reduced yield loss under water deficit conditions (Pathan et al., 2014).

Three possible combinations of physiological mechanisms have previously been proposed that could result in delayed canopy wilting phenotypes (Ries et al., 2012). One combination is high water use efficiency (WUE), high radiation use efficiency (RUE), and conservation of soil moisture which would allow soybean plants to produce biomass through more efficient use of transpired water in photosynthesis. Another combination proposed is low stomatal conductance, low RUE, low WUE, and conservation of soil moisture which would be better at conserving water during deficits, but could have reduced photosynthetic capacity due to low transpiration rates (Ries et al., 2012). In addition, the researchers also proposed deeper rooting as a third mechanism that could lead to slower soybean canopy wilting (Ries et al., 2012).

Carbon isotope composition

Water use efficiency can be defined in several different ways. At the crop production level, it is the ratio of grain yield to water used during growth. At the plant level, it can be defined as the amount of biomass produced per unit of water transpired. At the leaf level, it is the amount of photosynthetic carbon gained per unit of water transpired (Angus and van Herwaarden, 2001;

Gilbert et al., 2011). Soybean yields could potentially be improved under drought stress if genotypes with higher WUE are selected and used in breeding programs (Condon et al., 2004; Sinclair, 2012). However, due to the time- and labor-consuming nature of the measurements, directly selecting for actual WUE in a large number of genotypes is difficult and therefore correlated traits are often measured instead.

One such trait, carbon isotope composition, has been reported to closely correlate with WUE under certain conditions, especially in C3 plant species (O'Leary, 1981; Farquhar et al., 1982; Hubick and Farquhar, 1989; Condon et al., 1990, 1993). There are two common stable carbon isotopes, ^{12}C and ^{13}C, which constitute around 98.9% and 1.1% of the Earth's atmospheric carbon dioxide, respectively (O'Leary, 1981). C3 plants like soybean prefer the ^{12}C isotope, and discriminate against the heavier ^{13}C isotope during the process of carboxylation of CO_2 in the first products of photosynthesis by Rubisco (Condon et al., 2002). Carbon isotope composition of tissue samples can be expressed as either carbon isotope discrimination ($\Delta^{13}C$, CID) or carbon isotope ratio ($\delta^{13}C$), and is a measure of how much plants discriminate against the ^{13}C isotope during photosynthesis due to both enzymatic processes and stomatal limitations. Carbon isotope ratio is commonly measured with mass spectrometers, and compared to a PDB (belemnite from the Pee Dee Formation in South Carolina) standard. Relative abundance of $^{13}C/^{12}C$ for a sample is generally expressed in units per mil (‰) for $\delta^{13}C$ using the following equations (O'Leary, 1981):

$$R = {}^{13}CO_2/{}^{12}CO_2$$

$$\delta^{13}C\ (‰) = 1000\ (R_{sample} - R_{standard})/R_{standard}$$

Carbon isotope composition can be used as an indirect method for selection of genotypes with improved WUE and productivity in water stressed environments due to its correlation with

WUE in some environments. Carbon isotope discrimination ($\Delta^{13}C$, CID) has a negative relationship with WUE, and carbon isotope ratio ($\delta^{13}C$) has a positive relationship with WUE (Condon et al., 1990; Ehleringer et al., 1991; Ismail and Hall, 1992; Rebetzke et al., 2002). The differences in the relationship with WUE are a product of how $\Delta^{13}C$ and $\delta^{13}C$ are mathematically expressed as part of their calculations. The association between carbon isotope composition and WUE is due to the common relationship of the ratio of CO_2 inside and outside of the leaf. This is because as the ratio of internal to external CO_2 decreases, both carbon isotope composition and WUE increase.

Nitrogen fixation during drought stress

Soybean uses a symbiotic association with *Bradyrhizobia japonicum* to fix N_2 from the atmosphere. This fixation provides the majority of its total nitrogen (N) needs, but when water resources are limited, symbiotic N_2 fixation can be affected (Sprent, 1971). A previous simulation study to assess the benefits of altered soybean drought tolerance traits using 50 years of weather data across U.S. soybean growing areas found that sustained N_2 fixation had the greatest and most consistent yield advantage compared to four other traits (Sinclair et al., 2010). Certain genotypes of soybean are more sensitive than others to reductions in N_2 fixation during drought stress (Sall and Sinclair, 1991; Serraj and Sinclair, 1997; King and Purcell, 2006; Sinclair et al., 2007; King et al., 2014). One previous method used to determine sensitivity of genotypes to N_2 fixation during water deficit conditions was measuring differences in foliar N concentrations (Sall and Sinclair, 1991; King and Purcell, 2006; King et al., 2014). Sustained nitrogen fixation during drought stress was associated with low shoot N concentration under well-watered conditions in one study (King et al., 2014). In addition, nitrogen isotope

composition ($\delta^{15}N$, the ratio of ^{15}N to ^{14}N compared to air) is thought to be correlated with nitrogen fixation, and could be another useful nitrogen related trait to evaluate. The ^{15}N isotope is more prevalent in the soil compared to the atmosphere (Amarger et al., 1979; Shearer et al., 1980; Houngnandan et al., 2008), so plants that derive N only from the soil will have more ^{15}N compared to plants actively fixing N_2 from the atmosphere. This is because the amount of the ^{15}N isotope is not being diluted in the plant tissue by the ^{14}N isotope more commonly found in the air. Therefore, $\delta^{15}N$ could be an additional indicator of how much nitrogen fixation is affected by drought stress (Dhanapal et al., 2015b). The quantity of ^{15}N in leaf samples is expressed in units per mil (‰) according to the following equations:

$$R = {}^{15}N/{}^{14}N$$

$$\delta^{15}N\ (‰) = 1000\ (R_{sample} - R_{air\ N2})/R_{air\ N2}$$

Limited leaf hydraulic conductance

Vapor pressure deficit (VPD) in respect to plants is the difference between the vapor pressure inside the leaf compared to the vapor pressure of the air. Transpiration rates of plants vary depending on this difference in vapor pressure. Limiting transpiration could be desirable in certain conditions, because it allows plants to conserve soil water early in the season that can be used during drought stress periods later in the growing season; however, this may not be advantageous if sufficient soil water is available throughout all stages of development. In addition, it could result in improved crop WUE at high VPD due to limited gas exchange (Sinclair et al., 2017). Sinclair et al. (2008) showed that the soybean accession PI 416937 likely has the limited transpiration rate trait as a result of low hydraulic conductivity between the xylem and into the guard cells at high VPD. It was further hypothesized that aquaporins (AQP), water-

transporting proteins, play a major role in hydraulic conductivity (Heinen et al., 2009), and less sensitivity of aquaporin proteins to certain chemical inhibitors can be correlated with limited hydraulic conductance (Sadok and Sinclair, 2010b). Sensitivity (or insensitivity) to silver nitrate ($AgNO_3$) was previously tested by subjecting de-rooted soybean shoots to a $AgNO_3$ solution, and comparing transpiration rates in the solution to water. Genotypes that were insensitive to the $AgNO_3$ treatment were hypothesized to lack a protein-mediated water pathway that is sensitive to this inhibitor, while a hypothesis that silver nitrate blocks certain aquaporins could explain why transpiration rate is decreased in other genotypes. Also, genotypes that expressed the limited transpiration trait and low hydraulic conductance, including PI 416937, tended to show relative insensitivity to silver nitrate compared to other soybean genotypes (Sadok and Sinclair, 2010a). Therefore, it was further hypothesized that genotypes expressing these traits likely do not have silver nitrate sensitive AQPs, or have lower levels of these sensitive AQPs compared to other genotypes. Given this relationship of limited transpiration rate and low hydraulic conductance with sensitivity to chemical inhibitors, soybean genotypes could be characterized by this procedure as a way to identify drought tolerant germplasm and determine differences in aquaporin populations among genotypes for improvement of drought tolerance.

Root architecture

Roots play an important role in water uptake, and the size and architecture of a root system can vary depending on the physical and chemical properties of the soil (Passioura, 1983). The root ideotype best suited for a given drought stress condition will depend on availability of soil moisture and the distribution of water (Yu et al., 2007); however, most research in crops such as corn, soybean, rice, and chickpea has shown that deeper and more vigorous root systems led to

higher yields under drought (Nguyen et al., 1997; Tuberosa et al., 2003, 2011; Sadok and Sinclair, 2011; Varshney et al., 2011). Phenotyping for root architecture is often destructive and time consuming, but methods such as “shovelomics”, rhizotrons, and capacitance meters have been used to accurately measure root traits in the field (Johnson et al., 2001; McBride et al., 2008; Trachsel et al., 2011). Controlled condition experiments in greenhouses and growth chambers using hydroponics, gel- or soil-filled chambers, or other methods can provide higher resolution of traits using imaging or x-ray microtomography (Gregory et al., 2009; Metzner et al., 2015), but the results may not always translate to field conditions. Larger root systems or certain root features may not always translate to improved drought tolerance (Comas et al., 2013), but the ideal architecture for a given drought stress condition could lead to improved yields under water deficit conditions.

Flowering time

Flowering time is a key component of a plant being able to grow well in environments that vary in water availability over the course of the growing season (Richards, 2006). In maize, the anthesis-silking interval (ASI) can be targeted by breeders to improve drought tolerance. ASI has moderate heritability, is relatively easy to measure, and tends to be negatively correlated with yield under drought stressed conditions (Chapman and Edmeades, 1999). Differences in phenology (including flowering time) can bias results from drought tolerance experiments if it is not properly accounted for in the experimental design or statistical analyses. A plant with earlier flowering time may appear to have favorable drought tolerance traits, but this could be a result of drought escape, rather than inherent advantageous attributes (Soltani and Sinclair, 2012).

Canopy temperature

Canopy or leaf temperature can be measured via thermal imaging or infrared thermometers. These measurements reflect the amount of transpiration occurring in the plants with warmer canopies or leaves indicating less transpiration occurring and cooler canopies indicating more transpiration (Jones et al., 2009). Canopy temperature measurements integrate what is happening at the root, leaf, stoma, and canopy levels, and provide an additional way to measure stomatal conductance (Yousfi et al., 2016). Environmental conditions such as air temperature and wind speed can affect these measurements, as well as the leaf orientation towards the sun or light source at the time of image collection (Jones et al., 2009; Berger et al., 2010). A recent study in soybean found that differences in canopy temperature were generally associated with soil water availability, differences in the slow canopy wilting trait, and grain yield (Bai and Purcell, 2018). Canopy temperature depression (CTD) uses the temperature measured at the canopy surface and compares it to the temperature of the atmosphere at the time of the measurement (Reynolds et al., 2009). Under drought stress, plants with higher CTD (cooler canopy temperature) avoid excessive plant dehydration by using more soil available water (Blum, 2011).

Remote sensing and imaging for evaluation of drought stress

Remote sensing with vegetative index imaging in the visible band, near infrared (NIR), or longwave infrared (LWIR) of the electromagnetic spectrum can be useful for drought tolerance phenotyping as well. As field and greenhouse based high-throughput phenotyping platforms advance, this non-destructive phenotyping approach could allow for rapid screening of a large number of genotypes (Araus and Cairns, 2014). One physiological change that occurs under drought conditions is decreasing leaf water content. NIR measurements of drought stressed

plants can be used to determine the amount of water in plant leaves (Seelig et al., 2008, 2009). Red, green, blue (RGB) and visible spectrum imaging have been successful in identification of early onset of senescence, decreased growth rate, and biomass accumulation (Rajendran et al., 2009). Overall, the trait(s) in which a plant breeder should focus their research and improvement efforts depend on the target environment for production of the crop along with resource constraints of the program.

Genetic mapping for soybean drought tolerance related traits

The USDA maintains a collection of around 20,000 soybean accessions, with approximately 170,000 accessions in worldwide germplasm collections. However, only a limited number of these lines have been screened for numerous stress tolerances over the years and been used for drought tolerance improvement (Carter et al., 2004). Also, there is potential to make *Glycine max* (cultivated soybean) by *Glycine soja* (wild soybean) crosses that could generate new lines with improved tolerance. Wild soybean has greater allelic diversity than cultivated soybean, and could be exploited to create new drought tolerant cultivars (Lam et al., 2010). Identification of new accessions with beneficial alleles for traits such as slow canopy wilting, improved WUE, sustained nitrogen fixation under drought, and limited hydraulic conductance could lead to the development of soybean cultivars with improved drought tolerance. Soybean has a reference genome (Schmutz et al., 2010), and also genome-wide SoySNP50K (Song et al., 2013) and SoySNP6K SNP (Song et al., 2014) marker chips to help enable mapping and breeding efforts. The entire USDA soybean germplasm has been genotyped with the SoySNP50K chip (Song et al., 2015).

Linkage mapping

Given that drought tolerance is mostly considered to be a complex, quantitative trait, discovery of QTL that control a large percentage of the variation for drought related traits could help develop improved cultivars (Cattivelli et al., 2008). Linkage mapping can be a useful tool for identifying QTLs for traits of interest. The general steps for linkage-based QTL mapping are: 1) develop a mapping population by crossing lines that differ for a trait, 2) genotype the mapping population with a set of DNA markers, 3) phenotype for the trait of interest, and 4) associate the genotypic and phenotypic data. This form of QTL mapping has its disadvantages including potentially low mapping resolution (Myles et al., 2009), but can provide regions of interest that breeders can select for when attempting to produce new, improved cultivars. The most commonly used methods of linkage mapping are single-marker analysis, interval mapping, and composite interval mapping (Sehgal et al., 2016). Two popular methods of conducting QTL mapping are with the software package Windows QTL Cartographer (Wang et al., 2012), and with the R package r/QTL (Broman et al., 2003).

Previous linkage mapping studies have identified QTL that control canopy wilting, carbon isotope discrimination, WUE, nitrogen concentration, or leaf hydraulic conductance in soybean. Four QTLs for canopy wilting were mapped by Charlson et al. (2009) using a KS4895 × 'Jackson' recombinant inbred line (RIL) population. Another study with RILs derived from Kefeng No. 1 × Nannong1138-2 identified a total of eight QTLs for canopy wilting under greenhouse and field conditions (Du et al., 2009). Abdel-Haleem et al. (2012) found seven QTLs using 150 RILs from 'Benning' × PI 416937 that explained 75% of the canopy wilting variation. Using five different populations, Hwang et al. (2015) identified eight clusters that had QTLs from at least two of the five populations. These regions were further refined using a meta-

analysis to reduce the size of the confidence intervals and refine the map positions (Hwang et al., 2016). For carbon isotope discrimination and WUE, 14 loci have been previously identified using linkage mapping (Mian et al., 1996, 1998; Specht et al., 2001). Four QTLs for foliar N concentration were previously identified on Chr 13, 16, and 17 using a KS4895 × Jackson RIL population (Hwang et al., 2013). A linkage mapping study for leaf hydraulic conductance with transpiration response to silver nitrate aquaporin inhibitor as a surrogate measurement for the trait identified four QTLs (Carpentieri-Pipolo et al., 2011).

Genome-wide association mapping

Genome-wide association studies (GWAS) are a useful way to identify genomic regions associated with traits of interest, and often have higher resolution compared to traditional QTL mapping. These types of studies have become more commonplace in soybean as genomic resources have expanded, including a reference genome (Schmutz et al., 2010) and availability of genome-wide SNP markers (Song et al., 2013). The number of markers needed to provide adequate coverage of the genome depends on the level of linkage disequilibrium (LD) between markers and the rate at which LD decay occurs (Ersoz et al., 2007). Outcrossing plant species like maize require more genetic markers due to faster LD decay, whereas inbred crops like soybean require lower marker density due to slower LD decay (Flint-Garcia et al., 2003; Hyten et al., 2007).

The basic premise of GWAS is to associate a particular phenotype with a given genotype using linear regression. However, population stratification in the form of population structure or relatedness needs to be accounted for, otherwise false positives are likely to be identified using naïve linear regression models (Flint-Garcia et al., 2003). Most models account for population

structure by including a Q matrix (proportion individual belongs to subpopulation or principal components) or K matrix (kinship, relationship among individuals), or both (Myles et al., 2009). Several different statistical models have been developed to conduct GWAS and to improve their statistical power. Methods have evolved from simple naïve methods (t-test) to mixed linear model (MLM) which control for population structure and relatedness, to compressed MLM (CMLM) which decrease the effective sample size by clustering individuals into groups, and recently to multiple loci mixed models (MLMM) which incorporates multiple markers simultaneously as covariates in a MLM (Zhang et al., 2010; Lipka et al., 2012; Liu et al., 2016). These advances have helped to increase statistical power by fitting covariates such as Q, K, and pseudo quantitative trait nucleotides (QTNs), and reducing confounding issues between testing markers and these covariates (Liu et al., 2016). Recently, Fixed and random model Circulating Probability Unification (FarmCPU) was developed to completely remove the cofounding between kinship and testing markers by dividing MLMM into fixed and random effect models (Liu et al., 2016). This enables FarmCPU to correct for potential false positives due to population stratification without compromising markers truly associated with the trait of interest.

Several studies in soybean have been reported using the GWAS approach for different traits including canopy wilting (Kaler et al., 2017b), carbon isotope composition (Dhanapal et al., 2015a; Kaler et al., 2017a), and nitrogen related traits (Dhanapal et al., 2015b). These studies utilized the same panel of 373 MG IV soybean genotypes, but differed in their number of markers and GWAS models. In addition, genomic regions for seed composition (Hwang et al., 2014; Vaughn et al., 2014; Bandillo et al., 2015; Cao et al., 2017), ureide concentration (Ray et al., 2015), chlorophyll traits (Dhanapal et al., 2016), insect resistance (Chang and Hartman, 2017), salt tolerance (Patil et al., 2016; Zeng et al., 2017), agronomic and phenology traits

(Zhang et al., 2015; Contreras-Soto et al., 2017; Li et al., 2017), and local adaptation (Bandillo et al., 2017) have been identified using association mapping approaches. The combination of conventional breeding, QTL mapping and discovery, GWAS, marker-assisted breeding, and other 'omic' technologies will be useful in helping to generate drought tolerant soybean cultivars.

Development of germplasm and cultivars with drought tolerance

Conventional breeding by integrating or pyramiding desirable favorable alleles for drought related traits into elite germplasm via crosses is one strategy for improving soybean drought tolerance. Trait introgression with conventional breeding can be time consuming, but has been successful in developing cultivars and germplasm lines with improved drought tolerance. 'N7001' was developed by the USDA soybean breeding program in North Carolina by hybridizing PI 416937, an exotic slow wilting genotype, with an adapted USDA breeding line, N77-144, and has been used to create slow wilting germplasm lines (Carter et al., 2003; Hufstetler et al., 2007; Devi et al., 2014). 'USDA-N8002' is a MG VIII germplasm line with high yield and drought resistance, and was derived from slow canopy wilting accessions PI 471938 (25% by pedigree) and PI 416937 (12.5% by pedigree) (Carter et al., 2016). USDA-N8002 exhibits delayed wilting, limited transpiration at high VPD, and sustained nitrogen fixation during water deficits (Carter et al., 2016). Jindou 21 is a Chinese drought tolerant cultivar that was developed by crossing a low yielding, slightly drought tolerant cultivar with a high yielding line with good drought tolerance, which was then crossed with Jindou 14. This improved genotype has become a popular cultivar in dryer regions of China (Ku et al., 2013). Two germplasm lines, R01-416F and R01-581F, were developed by the Arkansas Agricultural

Experiment Station which possess improved yield and nitrogen fixation under drought stress (Chen et al., 2007). These lines were derived from a Jackson × KS4895 cross, with Jackson likely being the source of the sustained N_2 fixation under drought. They were identified as having improved nitrogen fixation by screening for both nitrogen content in the field, as well as through the use of a flow-through acetylene reduction assay in a greenhouse experiment (Chen et al., 2007). In addition, a study found that breeding for soybean grain yield under water-deficit field environments led to the development of 10 elite lines, with most of them possessing the limited transpiration rate under elevated VPD conditions and sustained N_2 fixation during drought stress (Devi et al., 2014). This demonstrated that empirical selection for improved yield and agronomic performance under drought stressed conditions can lead to the development of elite lines expressing improved drought tolerance related traits.

In corn, drought tolerant commercial hybrids have been developed with similar yields to non-drought tolerant hybrids under favorable conditions, and have better yields under drought stress (Adee et al., 2016). AQUAmax® (DuPont Pioneer) and Agrisure Artesian® (Syngenta) maize hybrids were conventionally developed using molecular marker-assisted breeding for increased yield under both unfavorable (drought) and favorable (well-watered) conditions (Gaffney et al., 2015; Adee et al., 2016). DroughtGard® (Monsanto) was developed using both conventional plant breeding and the introduction of a bacterial cold shock protein B, and has improved drought tolerance in certain conditions (Nemali et al., 2015). Genetic engineering has also been used in the development of a drought tolerant sugarcane expressing choline dehydrogenase (betA), which is approved to be grown in Indonesia (Marshall, 2014). In soybean, genetic engineering and overexpression of many different types of genes has led to improved drought tolerance (De Ronde et al., 2004; Xue et al., 2007; Seo et al., 2012; Fuganti-Pagliarini et

al., 2017; Ning et al., 2017). Conventional breeding methods, genetic engineering, along with more recent gene editing technologies such as CRISPR, will continue to play an important role in developing new drought tolerant germplasm lines and cultivars.

Objectives

The primary objective of this research was to identify genomic regions and germplasm useful for improving soybean drought tolerance. A panel of over 200 genetically diverse soybean accessions genotyped with the SoySNP50K iSelect BeadChips was evaluated for canopy wilting, carbon and nitrogen related traits, and transpiration response to $AgNO_3$. Additionally, 130 RILs derived from Hutcheson × PI 471938 genotyped with the SoySNP6K iSelect BeadChips were evaluated for canopy wilting. Genome-wide association analyses and composite interval mapping were used to identify genomic regions controlling these traits, and we also examined the diverse panel for new sources of drought tolerance. This dissertation is split into two primary chapters that focus on:

1) Discovery of genomic regions associated with canopy wilting using both association and linkage mapping approaches, comparing and confirming QTL from previous studies, and identification of new germplasm with favorable breeding values and alleles for canopy wilting.
2) Unraveling the genetic architecture for $\delta^{13}C$, $\delta^{15}N$, nitrogen concentration, and transpiration response to silver nitrate using association mapping, discovery of germplasm with favorable breeding values for these traits, and understanding the relationships among all traits evaluated.

References

Abdel-Haleem, H., T.E. Carter, L.C. Purcell, C.A. King, L.L. Ries, P. Chen, W. Schapaugh, T.R. Sinclair, and H.R. Boerma. 2012. Mapping of quantitative trait loci for canopy-wilting trait in soybean (Glycine max L. Merr). Theor. Appl. Genet. 125(5): 837–46.

Adee, E., K. Roozeboom, G.R. Balboa, A. Schlegel, and I.A. Ciampitti. 2016. Drought-tolerant corn hybrids yield more in drought-stressed environments with no penalty in non-stressed environments. Front. Plant Sci. 7: 1534.

Amarger, N., A. Mariotti, F. Mariotti, J.C. Durr, C. Bourguignon, and B. Lagacherie. 1979. Estimate of symbiotically fixed nitrogen in field grown soybeans using variations in 15N natural abundance. Plant Soil 52(2): 269–280.

Angus, J.F., and A.F. van Herwaarden. 2001. Increasing water use and water use efficiency in dryland wheat. Agron. J. 93: 290–298.

Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. 19(1): 52–61.

Bai, H., and L.C. Purcell. 2018. Aerial canopy temperature differences between fast- and slow-wilting soya bean genotypes. J. Agron. Crop Sci.: 1–9.

Bandillo, N.B., J.E. Anderson, M.B. Kantar, R.M. Stupar, J.E. Specht, G.L. Graef, and A.J. Lorenz. 2017. Dissecting the genetic basis of local adaptation in soybean. Sci. Rep. 7: 17195.

Bandillo, N., D. Jarquin, Q. Song, R. Nelson, P. Cregan, J. Specht, and A. Lorenz. 2015. A population structure and genome-wide association analysis on the USDA soybean germplasm collection. Plant Genome 8(3): 1–13.

Berger, B., B. Parent, and M. Tester. 2010. High-throughput shoot imaging to study drought

responses. J. Exp. Bot. 61(13): 3519–28.

Blum, A. 2005. Drought resistance, water-use efficiency, and yield potential — are they compatible, dissonant, or mutually exclusive? Aust. J. Agric. Res. 56: 1159–1168.

Blum, A. 2011. Plant breeding for water-limited environments. Springer, New York.

Broman, K.W., H. Wu, Ś. Sen, and G.A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19(7): 889–890.

Brown, E.A., C.E. Caviness, and D.A. Brown. 1985. Response of selected soybean cultivars to soil moisture deficit. Agron. J. 77: 274–278.

Cao, Y., S. Li, Z. Wang, F. Chang, J. Kong, J. Gai, and T. Zhao. 2017. Identification of major quantitative trait loci for seed oil content in soybeans by combining linkage and genome-wide association mapping. Front. Plant Sci. 8: 1222.

Carpentieri-Pipolo, V., A.E. Pipolo, H. Abdel-Haleem, H.R. Boerma, and T.R. Sinclair. 2011. Identification of QTLs associated with limited leaf hydraulic conductance in soybean. Euphytica 186: 679–686.

Carter, T.E., J.W. Burton, D.T. Bowman, Z. Cui, X. Zhou, M.R. Villagarcia, A.S. Niewoehner, and M.O. Fountain. 2003. Registration of "N7001" Soybean. Crop Sci. 43: 1126–1127.

Carter, T.E., R.L. Nelson, C.H. Sneller, and Z. Cui. 2004. Genetic diversity in soybean. p. 303–416. *In* Soybeans: Improvement, Production, and Uses. 3rd ed. ASA, CSSA, and SSSA.

Carter, T.E., S.M. Todd, and A.M. Gillen. 2016. Registration of "USDA-N8002" soybean cultivar with high yield and abiotic stress resistance traits. J. Plant Regist. 10: 238–245.

Cattivelli, L., F. Rizza, F.-W. Badeck, E. Mazzucotelli, A.M. Mastrangelo, E. Francia, C. Marè, A. Tondelli, and A.M. Stanca. 2008. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. F. Crop. Res. 105: 1–14.

Chang, H.-X., and G.L. Hartman. 2017. Characterization of insect resistance loci in the USDA soybean germplasm collection using genome-wide association studies. Front. Plant Sci. 8: 670.

Chapman, S.C., and G.O. Edmeades. 1999. Selection improves drought tolerance in tropical maize populations. Crop Sci. 39: 1315–1324.

Charlson, D. V, S. Bhatnagar, C.A. King, J.D. Ray, C.H. Sneller, T.E. Carter, and L.C. Purcell. 2009. Polygenic inheritance of canopy wilting in soybean [Glycine max (L.) Merr.]. Theor. Appl. Genet. 119: 587–594.

Chen, P., C.H. Sneller, L.C. Purcell, T.R. Sinclair, C.A. King, and T. Ishibashi. 2007. Registration of soybean germplasm lines R01-416F and R01-581F for improved yield and nitrogen fixation under drought stress. J. Plant Regist. 1(2): 166–167.

Clark, J., and J. Levitt. 1956. The basis of drought resistance in the soybean plant. Physiol. Plant. 9: 598–607.

Comas, L.H., S.R. Becker, V.M. V. Cruz, P.F. Byrne, and D.A. Dierig. 2013. Root traits contributing to plant productivity under drought. Front. Plant Sci. 4: 442.

Condon, A.G., G.D. Farquhar, and R.A. Richards. 1990. Genotypic variation in carbon isotope discrimination and transpiration efficiency in wheat. Leaf gas-exchange and whole plant studies. Aust. J. Plant Physiol. 17: 9–22.

Condon, A.G., R.A. Richards, and G.D. Farquhar. 1993. Relationships between carbon isotope discrimination, water use efficiency and transpiration efficiency for dryland wheat. Aust. J. Agric. Res. 44(8): 1693–1711.

Condon, A.G., R. Richards, G. Rebetzke, and G. Farquhar. 2002. Improving intrinsic water use efficiency and crop yield. Crop Sci. 42: 122–131.

Condon, A.G., R.A. Richards, G.J. Rebetzke, and G.D. Farquhar. 2004. Breeding for high water-use efficiency. J. Exp. Bot. 55(407): 2447–2460.

Contreras-Soto, R.I., F. Mora, M.A.R. de Oliveira, W. Higashi, C.A. Scapim, and I. Schuster. 2017. A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-based haplotype analysis. PLoS One 12(2): e0171105.

Dai, A.G. 2013. Increasing drought under global warming in observations and models. Nat. Clim. Chang. 3(1): 52–58.

Desclaux, D., T. Huynh, and P. Roumet. 2000. Identification of soybean plant characteristics that indicate the timing of drought stress. Crop Sci. 40: 716–722.

Devi, J.M., T.R. Sinclair, P. Chen, and T.E. Carter. 2014. Evaluation of elite southern maturity soybean breeding lines for drought-tolerant traits. Agron. J. 106(6): 1947–1954.

Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C. Andy King, P.B. Cregan, Q. Song, and F.B. Fritschi. 2015a. Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [Glycine max (L.) Merr.] genotypes. Theor. Appl. Genet. 128(1): 73–91.

Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, and F.B. Fritschi. 2016. Genome-wide association mapping of soybean chlorophyll traits based on canopy spectral reflectance and leaf extracts. BMC Plant Biol. 16: 174.

Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C.A. King, and F.B. Fritschi. 2015b. Genome-wide association analysis of diverse soybean genotypes reveals novel markers for nitrogen traits. Plant Genome 8(3): 1–15.

Dornbos Jr., D.L., and R.E. Mullen. 1991. Influence of stress during soybean seed fill on seed weight, germination, and seedling growth rate. Can. J. Plant Sci. 71: 373–383.

Du, W., D. Yu, and S. Fu. 2009. Detection of quantitative trait loci for yield and drought tolerance traits in soybean using a recombinant inbred line population. J. Integr. Plant Biol. 51(9): 868–878.

Eck, H.V., A.C. Mathers, and J.T. Musick. 1987. Plant water stress at various growth stages and growth and yield of soybeans. F. Crop. Res. 17(1): 1–16.

Ehleringer, J.R., S. Klassen, C. Clayton, D. Sherrill, M. Fullerholbrook, Q.N. Fu, and T.A. Cooper. 1991. Carbon isotope discrimination and transpiration efficiency in common bean. Crop Sci. 31(6): 1611–1615.

Ersoz, E.S., J. Yu, and E.S. Buckler. 2007. Applications of linkage disequilibrium and association mapping in crop plants. p. 97–119. *In* Genomics-Assisted Crop Improvement.

Farquhar, G., M. O'Leary, and J. Berry. 1982. On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves. Aust. J. Plant Physiol. 9(2): 121–137.

Fehr, W.R., C.E. Caviness, D.T. Burmood, and J.S. Pennington. 1971. Stage of development descriptions for soybeans, Glycine max (L.) Merrill. Crop Sci. 11: 929–931.

Fletcher, A.L., T.R. Sinclair, and L.H. Allen. 2007. Transpiration responses to vapor pressure deficit in well watered "slow-wilting" and commercial soybean. Environ. Exp. Bot. 61(2): 145–151.

Flint-Garcia, S.A., J.M. Thornsberry, and E.S. Buckler. 2003. Structure of linkage disequilibrium in plants. Annu. Rev. Plant Biol. 54: 357–374.

Frederick, J.R., C.R. Camp, and P.J. Bauer. 2001. Drought-stress effects on branch and main stem seed yield and yield components of determinate soybean. Crop Sci. 41: 759–763.

Fuganti-Pagliarini, R., L.C. Ferreira, F.A. Rodrigues, H.B.C. Molinari, S.R.R. Marin, M.D.C.

Molinari, J. Marcolino-Gomes, L.M. Mertz-Henning, J.R.B. Farias, M.C.N. de Oliveira, N. Neumaier, N. Kanamori, Y. Fujita, J. Mizoi, K. Nakashima, K. Yamaguchi-Shinozaki, and A.L. Nepomuceno. 2017. Characterization of soybean genetically modified for drought tolerance in field conditions. Front. Plant Sci. 8: 448.

Gaffney, J., J. Schussler, C. Löffler, W. Cai, S. Paszkiewicz, C. Messina, J. Groeteke, J. Keaschall, and M. Cooper. 2015. Industry-scale evaluation of maize hybrids selected for increased yield in drought-stress conditions of the US corn belt. Crop Sci. 55: 1608–1618.

Gilbert, M.E., M.A. Zwieniecki, and N.M. Holbrook. 2011. Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. J. Exp. Bot. 62: 2875–2887.

Goldman, I.L., T.E. Carter, and R.P. Patterson. 1989. Differential genotypic response to drought stress and subsoil aluminum in soybean. Crop Sci. 29(2): 330–334.

Gregory, P.J., a. G. Bengough, D. Grinev, S. Schmidt, W.T.B. Thomas, T. Wojciechowski, and I.M. Young. 2009. Root phenomics of crops: Opportunities and challenges. Funct. Plant Biol. 36(11): 922–929.

Heatherly, L.G. 1993. Drought stress and irrigation effects on germination of harvested soybean seed. Crop Sci. 33(4): 777–781.

Heinen, R.B., Q. Ye, and F. Chaumont. 2009. Role of aquaporins in leaf physiology. J. Exp. Bot. 60(11): 2971–2985.

Houngnandan, P., R.G.H. Yemadje, S.O. Oikeh, C.F. Djidohokpin, P. Boeckx, and O. Van Cleemput. 2008. Improved estimation of biological nitrogen fixation of soybean cultivars (Glycine max L. Merril) using 15N natural abundance technique. Biol. Fertil. Soils 45(2):

175–183.

Hubick, K., and G. Farquhar. 1989. Carbon isotope discrimination and the ratio of carbon gained to water lost in barley cultivars. Plant. Cell Environ. 12: 795–804.

Hudak, C.M., and R.P. Patterson. 1996. Root distribution and soil moisture depletion pattern of a drought-resistant soybean plant introduction. Agron. J. 88(3): 478–485.

Hufstetler, E.V., H.R. Boerma, T.E. Carter, and H.J. Earl. 2007. Genotypic variation for three physiological traits affecting drought tolerance in soybean. Crop Sci. 47: 25–35.

Hwang, S., C.A. King, P. Chen, J.D. Ray, P.B. Cregan, T.E. Carter, Z. Li, H. Abdel-Haleem, K.W. Matson, W. Schapaugh, and L.C. Purcell. 2016. Meta-analysis to refine map position and reduce confidence intervals for delayed-canopy-wilting QTLs in soybean. Mol. Breed. 36(7): 91.

Hwang, S., C.A. King, M.K. Davies, J.D. Ray, P.B. Cregan, and L.C. Purcell. 2013. QTL analysis of shoot ureide and nitrogen concentrations in soybean [Glycine max (L.) Merr.]. Crop Sci. 53(6): 2421–2433.

Hwang, S., C.A. King, J.D. Ray, P.B. Cregan, P. Chen, T.E. Carter, Z. Li, H. Abdel-Haleem, K.W. Matson, W. Schapaugh, and L.C. Purcell. 2015. Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. Theor. Appl. Genet. 128(10): 2047–2065.

Hwang, E.-Y., Q. Song, G. Jia, J.E. Specht, D.L. Hyten, J. Costa, and P.B. Cregan. 2014. A genome-wide association study of seed protein and oil content in soybean. BMC Genomics 15: 1.

Hyten, D.L., I.Y. Choi, Q. Song, R.C. Shoemaker, R.L. Nelson, J.M. Costa, J.E. Specht, and P.B. Cregan. 2007. Highly variable patterns of linkage disequilibrium in multiple soybean

populations. Genetics 175(4): 1937–1944.

Ismail, A.M., and A.E. Hall. 1992. Correlation between water-use efficiency and carbon isotope discrimination in diverse cowpea genotypes and isogenic lines. Crop Sci. 32(3): 7–12.

Johnson, M.G., D.T. Tingey, D.L. Phillips, and M.J. Storm. 2001. Advancing fine root research with minirhizotrons. Environ. Exp. Bot. 45(3): 263–289.

Jones, H., R. Serraj, and B. Loveys. 2009. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. Funct. Plant Biol. 36: 978–989.

Kadhem, F.A., J.E. Specht, and J.H. Williams. 1985. Soybean irrigation serially timed during stages R1 to R6. I. Agronomic responses. Agron. J. 77(2): 291.

Kaler, A.S., A.P. Dhanapal, J.D. Ray, C.A. King, F.B. Fritschi, and L.C. Purcell. 2017a. Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. Crop Sci. 57: 1–16.

Kaler, A.S., J.D. Ray, W.T. Schapaugh, C.A. King, and L.C. Purcell. 2017b. Genome-wide association mapping of canopy wilting in diverse soybean genotypes. Theor. Appl. Genet. 130(10): 2203–2217.

King, C.A., and L.C. Purcell. 2006. Genotypic variation for shoot N concentration and response to water deficits in soybean. Crop Sci. 46(6): 2396–2402.

King, C.A., L.C. Purcell, A. Bolton, and J.E. Specht. 2014. A possible relationship between shoot N concentration and the sensitivity of N2 fixation to drought in soybean. Crop Sci. 54(2): 746–756.

King, C.A., L.C. Purcell, and K.R. Brye. 2009. Differential wilting among soybean genotypes in response to water deficit. Crop Sci. 49(1): 290–298.

Korte, L.L., J.H. Williams, J.E. Specht, and R.C. Sorensen. 1983. Irrigation of soybean genotypes during reproductive ontogeny. I. Agronomic responses. Crop Sci. 23(3): 521–527.

Ku, Y., W. Au-Yeung, and Y. Yung. 2013. Drought stress and tolerance in soybean. *In* Board, J. (ed.), A Comprehensive Survey of International Soybean Research - Genetics, Physiology, Agronomy and Nitrogen Relationships. InTech.

Lam, H.-M., X. Xu, X. Liu, W. Chen, G. Yang, F.-L. Wong, M.-W. Li, W. He, N. Qin, B. Wang, J. Li, M. Jian, J. Wang, G. Shao, J. Wang, S.S.-M. Sun, and G. Zhang. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat. Genet. 42(12): 1053–1059.

Li, S., Y. Cao, J. He, T. Zhao, and J. Gai. 2017. Detecting the QTL-allele system conferring flowering date in a nested association mapping population of soybean using a novel procedure. Theor. Appl. Genet. 130(11): 2297–2314.

Lipka, A.E., F. Tian, Q. Wang, J. Peiffer, M. Li, P.J. Bradbury, M.A. Gore, E.S. Buckler, and Z. Zhang. 2012. GAPIT: Genome association and prediction integrated tool. Bioinformatics 28(18): 2397–2399.

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

Manavalan, L.P., S.K. Guttikonda, L.-S. Tran, and H.T. Nguyen. 2009. Physiological and molecular approaches to improve drought resistance in soybean. Plant Cell Physiol. 50(7): 1260–1276.

Marshall, A. 2014. Drought-tolerant varieties begin global march. Nat. Biotechnol. 32(4): 308.

McBride, R., M. Candido, and J. Ferguson. 2008. Estimating root mass in maize genotypes using the electrical capacitance method. Arch. Agron. Soil Sci. 54(2): 215–226.

Metzner, R., A. Eggert, D. van Dusschoten, D. Pflugfelder, S. Gerth, U. Schurr, N. Uhlmann, and S. Jahnke. 2015. Direct comparison of MRI and X-ray CT technologies for 3D imaging of root systems in soil: Potential and challenges for root trait quantification. Plant Methods 11: 17.

Mian, M.A.R., D.A. Ashley, and H.R. Boerma. 1998. An additional QTL for water use efficiency in soybean. Crop Sci. 38(2): 390–393.

Mian, M.A.R., T.E. Carter, W.A. Parrott, R. Wells, M.A. Bailey, D.A. Ashley, and H.R. Boerma. 1996. Molecular markers associated with water use efficiency and leaf ash in soybean. Crop Sci. 36: 1252–1257.

Monneveux, P., J. Ribaut, and A. Okono. 2014. Drought phenotyping in crops: From theory to practice (P Monneveux, J-M Ribaut, and A Okono, Eds.). Frontiers Media SA.

Mourtzinis, S., J.E. Specht, L.E. Lindsey, W.J. Wiebold, J. Ross, E.D. Nafziger, H.J. Kandel, N. Mueller, P.L. Devillez, F.J. Arriaga, and S.P. Conley. 2015. Climate-induced reduction in US-wide soybean yields underpinned by region-and in-season-specific responses. Nat. Plants 1(February): 8–11.

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z. Zhang, D.E. Costich, and E.S. Buckler. 2009. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21(8): 2194–2202.

Nemali, K.S., C. Bonin, F.G. Dohleman, M. Stephens, W.R. Reeves, D.E. Nelson, P. Castiglioni, J.E. Whitsel, B. Sammons, R.A. Silady, D. Anstrom, R.E. Sharp, O.R. Patharkar, D. Clay, M. Coffin, M.A. Nemeth, M.E. Leibman, M. Luethy, and M. Lawson. 2015. Physiological

responses related to increased grain yield under drought in the first biotechnology-derived drought-tolerant maize. Plant, Cell Environ. 38(9): 1866–1880.

Nguyen, H.T., R.C. Babu, and A. Blum. 1997. Breeding for drought resistance in rice: physiology and molecular genetics considerations. Crop Sci. 37: 1426–1434.

Ning, W., H. Zhai, J. Yu, S. Liang, X. Yang, X. Xing, J. Huo, T. Pang, Y. Yang, and X. Bai. 2017. Overexpression of Glycine soja WRKY20 enhances drought tolerance and improves plant yields under drought stress in transgenic soybean. Mol. Breed. 37: 19.

O'Leary, M.H. 1981. Carbon isotope fractionation in plants. Phytochemistry 20(4): 553–567.

Pantalone, V., and G. Rebetzke. 1996. Phenotypic evaluation of root traits in soybean and applicability to plant breeding. Crop Sci. 36: 456–459.

Panthee, D.R., V.R. Pantalone, C.E. Sams, A.M. Saxton, D.R. West, J.H. Orf, and A.S. Killam. 2006. Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds. Theor. Appl. Genet. 112(3): 546–553.

Passioura, J.B. 1983. Roots and drought resistance. Agric. Water Manag. 7: 265–280.

Passioura, J. 2007. The drought environment: Physical, biological and agricultural perspectives. J. Exp. Bot. 58(2): 113–117.

Pathan, S.M., J.-D. Lee, D.A. Sleper, F.B. Fritschi, R.E. Sharp, T.E. Carter, R.L. Nelson, C.A. King, W.T. Schapaugh, M.R. Ellersieck, H.T. Nguyen, and J.G. Shannon. 2014. Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought. J. Agron. Crop Sci. 200(3): 231–236.

Pathan, S., H.T. Nguyen, R.E. Sharp, and J.G. Shannon. 2010. Soybean improvement for drought, salt and flooding tolerance. Korean J. Breed. Sci. 42(4): 329–338.

Patil, G., T. Do, T.D. Vuong, B. Valliyodan, J.D. Lee, J. Chaudhary, J.G. Shannon, and H.T.

Nguyen. 2016. Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. Sci. Rep. 6: 19199.

Patil, G., R. Mian, T. Vuong, V. Pantalone, Q. Song, P. Chen, G.J. Shannon, T.C. Carter, and H.T. Nguyen. 2017. Molecular mapping and genomics of soybean seed protein: a review and perspective for the future. Theor. Appl. Genet. 130(10): 1975–1991.

Pham, A.T., J.G. Shannon, and K.D. Bilyeu. 2012. Combinations of mutant FAD2 and FAD3 genes to produce high oleic acid and low linolenic acid soybean oil. Theor. Appl. Genet. 125(3): 503–515.

Purcell, L.C., and J.E. Specht. 2004. Physiological traits for ameliorating drought stress. p. 569–620. *In* Soybeans: Improvement, Production, and Uses. ASA, CSSA, and SSSA.

Rajendran, K., M. Tester, and S.J. Roy. 2009. Quantifying the three main components of salinity tolerance in cereals. Plant, Cell Environ. 32(3): 237–249.

Ray, J.D., A.P. Dhanapal, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C.A. King, D. Boykin, P.B. Cregan, Q. Song, and F.B. Fritschi. 2015. Genome-wide association study of ureide concentration in diverse maturity group IV soybean [Glycine max (L.) Merr.] accessions. G3 5(11): 2391–2403.

Rebetzke, G.J., K. Chenu, B. Biddulph, C. Moeller, D.M. Deery, A.R. Rattey, D. Bennett, E.G. Barrett-Lennard, and J.E. Mayer. 2013. A multisite managed environment facility for targeted trait and germplasm phenotyping. Funct. Plant Biol. 40: 1–13.

Rebetzke, G.J., A.G. Condon, R.A. Richards, and G.D. Farquhar. 2002. Selection for reduced carbon isotope discrimination increases aerial biomass and grain yield of rainfed bread wheat. Crop Sci. 42(3): 739–745.

Reynolds, M., Y. Manes, A. Izanloo, and P. Langridge. 2009. Phenotyping approaches for

physiological breeding and gene discovery in wheat. Ann. Appl. Biol. 155(3): 309–320.

Richards, R.A. 2006. Physiological traits used in the breeding of new cultivars for water-scarce environments. Agric. Water Manag. 80: 197–211.

Ries, L.L., L.C. Purcell, T.E. Carter, J.T. Edwards, and C.A. King. 2012. Physiological traits contributing to differential canopy wilting in soybean under drought. Crop Sci. 52: 272–281.

De Ronde, J.A., R.N. Laurie, T. Caetano, M.M. Greyling, and I. Kerepesi. 2004. Comparative study between transgenic and non-transgenic soybean lines proved transgenic lines to be more drought tolerant. Euphytica 138(2): 123–132.

Sadok, W., M.E. Gilbert, M.A.S. Raza, and T.R. Sinclair. 2012. Basis of slow-wilting phenotype in soybean PI 471938. Crop Sci. 52(3): 1261–1269.

Sadok, W., and T.R. Sinclair. 2010a. Genetic variability of transpiration response of soybean [Glycine max (L.) Merr.] shoots to leaf hydraulic conductance inhibitor AgNO3. Crop Sci. 50(4): 1423–1430.

Sadok, W., and T.R. Sinclair. 2010b. Transpiration response of "slow-wilting" and commercial soybean (Glycine max (L.) Merr.) genotypes to three aquaporin inhibitors. J. Exp. Bot. 61(3): 821–829.

Sadok, W., and T.R. Sinclair. 2011. Crops yield increase under water-limited conditions: Review of recent physiological advances for soybean genetic improvement. p. 325–349. *In* Advances in Agronomy. Elsevier Inc.

Sall, K., and T.R. Sinclair. 1991. Soybean genotypic differences in sensitivity of symbiotic nitrogen fixation to soil dehydration. Plant Soil 133(1): 31–37.

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J.

Thelen, J. Cheng, D. Xu, U. Hellsten, G.D. May, Y. Yu, T. Sakurai, T. Umezawa, M.K. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, B. Abernathy, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.-C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, and S.A. Jackson. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463(7278): 178–183.

Seelig, H.D., A. Hoehn, L.S. Stodieck, D.M. Klaus, W.W. Adams, and W.J. Emery. 2009. Plant water parameters and the remote sensing R 1300/R 1450 leaf water index: Controlled condition dynamics during the development of water deficit stress. Irrig. Sci. 27(5): 357–365.

Seelig, H. -D., A. Hoehn, L.S. Stodieck, D.M. Klaus, W.W. Adams III, and W.J. Emery. 2008. The assessment of leaf water content using leaf reflectance ratios in the visible, near, and short wave infrared. Int. J. Remote Sens. 29(13): 3701–3713.

Sehgal, D., R. Singh, and V.R. Rajpal. 2016. Quantitative trait loci mapping in plants: Concepts and approaches. p. 31–60. *In* Molecular Breeding for Sustainable Crop Improvement. Springer International Publishing.

Seo, J.S., H.B. Sohn, K. Noh, C. Jung, J.H. An, C.M. Donovan, D.A. Somers, D.I. Kim, S.C. Jeong, C.G. Kim, H.M. Kim, S.H. Lee, Y. Do Choi, T.W. Moon, C.H. Kim, and J.J. Cheong. 2012. Expression of the Arabidopsis AtMYB44 gene confers drought/salt-stress tolerance in transgenic soybean. Mol. Breed. 29(3): 601–608.

Serraj, R., and T.R. Sinclair. 1997. Variation among soybean cultivars in dinitrogen fixation response to drought. Agron. J. 89: 963–969.

Shearer, G., D.H. Kohl, and J.E. Harper. 1980. Distribution of N among plant parts of nodulating

and nonnodulating isolines of soybeans. Plant Physiol. 66: 57–60.

Sinclair, T. 2012. Is transpiration efficiency a viable plant trait in breeding for crop improvement? Funct. Plant Biol. 39: 359–365.

Sinclair, T.R., J. Devi, A. Shekoofa, S. Choudhary, W. Sadok, V. Vadez, M. Riar, and T. Rufty. 2017. Limited-transpiration response to high vapor pressure deficit in crop species. Plant Sci. 260: 109–118.

Sinclair, T.R., C.D. Messina, A. Beatty, and M. Samples. 2010. Assessment across the United States of the benefits of altered soybean drought traits. Agron. J. 102(2): 475–482.

Sinclair, T.R., L.C. Purcell, C.A. King, C.H. Sneller, P. Chen, and V. Vadez. 2007. Drought tolerance and yield increase of soybean resulting from improved symbiotic N2 fixation. F. Crop. Res. 101: 68–71.

Sinclair, T.R., M.A. Zwieniecki, and N.M. Holbrook. 2008. Low leaf hydraulic conductance associated with drought tolerance in soybean. Physiol. Plant. 132(4): 446–451.

Sloane, R.J., R.P. Patterson, and T.E. Carter. 1990. Field drought tolerance of a soybean plant introduction. Crop Sci. 30: 118–123.

Soltani, A., and T.R. Sinclair. 2012. Optimizing chickpea phenology to available water under current and future climates. Eur. J. Agron. 38: 22–31.

Song, Q., D.L. Hyten, G. Jia, C. V Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS One 8(1): e54985.

Song, Q., D.L. Hyten, G. Jia, C. V Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2015. Fingerprinting soybean germplasm and its utility in genomic research. G3 (Bethesda). 5(10): 1999–2006.

Song, Q., G. Jia, C. Quigley, E. Fickus, D. Hyten, R. Nelson, and P. Cregan. 2014. Soybean BARC- SoySNP6K Beadchip—a tool for soybean genetics research. p. Abstract no.: P306. *In* Poster presented at: Plant and Animal Genome XXII Conference.

SoyStats. 2017. A reference guide to important soybean facts and figures.

Specht, J.E., K. Chase, M. Macrander, G.L. Graef, J. Chung, J.P. Markwell, M. Germann, J.H. Orf, and K.G. Lark. 2001. Soybean response to water: A QTL analysis of drought tolerance. Crop Sci. 41: 493–509.

Specht, J., D. Hume, and S. Kumudini. 1999. Soybean yield potential—A genetic and physiological perspective. Crop Sci. 39: 1560–1570.

Sprent, J.I. 1971. The effects of water stress on nitrogen-fixing root nodules: I. Effects on the physiology of detached soybean nodules. New Phytol. 70: 9–17.

Tanaka, Y., K. Fujii, and T. Shiraiwa. 2010. Variability of leaf morphology and stomatal conductance in soybean [Glycine max (L.) Merr.] cultivars. Crop Sci. 50(6): 2525–2532.

Trachsel, S., S.M. Kaeppler, K.M. Brown, and J.P. Lynch. 2011. Shovelomics: High throughput phenotyping of maize (Zea mays L.) root architecture in the field. Plant Soil 341: 75–87.

Tuberosa, R. 2012. Phenotyping for drought tolerance of crops in the genomics era. Front. Physiol. 3: 347.

Tuberosa, R., S. Salvi, S. Giuliani, M.C. Sanguineti, E. Frascaroli, S. Conti, and P. Landi. 2011. Genomics of root architecture and functions in maize. p. 179–204. *In* de Oliveira, A., Varshney, R.K. (eds.), Root Genomics. Springer Berlin Heidelberg, Berlin, Heidelberg.

Tuberosa, R., S. Salvi, M.C. Sanguineti, M. Maccaferri, S. Giuliani, and P. Landi. 2003. Searching for quantitative trait loci controlling root traits in maize: A critical appraisal. Plant Soil 255: 35–54.

USDA-NASS. 2012. USDA-National Agricultural Statistics Service (USDA-NASS). Census Agric.: Table 37.

USDA-NASS. 2018. Quick StatsAvailable at https://quickstats.nass.usda.gov/results/C42D1446-BB28-3283-AA26-528469130A62 (verified 25 March 2018).

Varshney, R.K., L. Pazhamala, J. Kashiwagi, P.M. Gaur, L. Krishnamurthy, and D. Hoisington. 2011. Genomics and physiological approaches for root trait breeding to improve drought tolerance in chickpea (Cicer arietinum L.). p. 233–250. *In* de Oliveira, A., Varshney, R.K. (eds.), Root Genomics. Springer Berlin Heidelberg, Berlin, Heidelberg.

Vaughn, J.N., R.L. Nelson, Q. Song, P.B. Cregan, and Z. Li. 2014. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. G3 (Bethesda). 4(11): 2283–2294.

Vollmann, J., C.N. Fritz, H. Wagentristl, and P. Ruckenbauer. 2000. Environmental and genetic variation of soybean seed protein content under Central European growing conditions. J. Sci. Food Agric. 80(9): 1300–1306.

Wang, S.C., J. Basten, and Z.B. Zeng. 2012. Windows QTL Cartographer 2.5.

Weaver-Missick, T. 1999. Early-season soybeans. Agric. Res. 47: 8–9.

Wilson, R.F. 2004. Seed composition. p. 621–677. *In* Soybeans: improvement, production, and uses.

Xue, R.G., B. Zhang, and H.F. Xie. 2007. Overexpression of a NTR1 in transgenic soybean confers tolerance to water stress. Plant Cell. Tissue Organ Cult. 89: 177–183.

Yousfi, S., N. Kellas, L. Saidi, Z. Benlakehal, L. Chaou, D. Siad, F. Herda, M. Karrou, O. Vergara, A. Gracia, J.L. Araus, and M.D. Serret. 2016. Comparative performance of remote sensing methods in assessing wheat performance under Mediterranean conditions. Agric.

Water Manag. 164: 137–147.

Yu, G.R., J. Zhuang, K. Nakayama, and Y. Jin. 2007. Root water uptake and profile soil water as affected by vertical root distribution. Plant Ecol. 189: 15–30.

Zeng, A., P. Chen, K. Korth, F. Hancock, A. Pereira, K. Brye, C. Wu, and A. Shi. 2017. Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. Mol. Breed. 37: 30.

Zhang, Z., E. Ersoz, C.-Q. Lai, R.J. Todhunter, H.K. Tiwari, M. a Gore, P.J. Bradbury, J. Yu, D.K. Arnett, J.M. Ordovas, and E.S. Buckler. 2010. Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. 42(4): 355–360.

Zhang, J., Q. Song, P.B. Cregan, and G.-L. Jiang. 2015. Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (Glycine max). Theor. Appl. Genet. 129: 117–130.

Zipper, S.C., J. Qiu, and C.J. Kucharik. 2016. Drought effects on US maize and soybean production: spatiotemporal patterns and historical changes. Environ. Res. Lett. 11(9): 94021.

CHAPTER 2

GENOME-WIDE ASSOCIATION ANALYSIS AND LINKAGE MAPPING REVEAL GENOMIC REGIONS CONTROLLING CANOPY WILTING IN SOYBEAN[1]

[1] C.J. Steketee, W.T. Schapaugh, T.E. Carter Jr., and Z. Li. To be submitted to *Frontiers in Plant Science*.

Abstract

Drought stress causes the greatest soybean [*Glycine max* (L.) Merr.] yield losses among the abiotic stresses in rain-fed U.S. growing areas. Because less than 10% of U.S. soybean hectares are irrigated, combating this stress requires soybean plants which possess physiological mechanisms to tolerate drought for a period of time. Phenotyping for these mechanisms is challenging, and the genetic architecture for these traits is poorly understood. A morphological trait, slow or delayed canopy wilting, has been observed in a few exotic plant introductions (PIs), and may lead to yield improvement in drought stressed fields. In this study, we visually scored wilting during stress for a panel of 209 genetically diverse soybean lines genotyped with the SoySNP50K iSelect BeadChip, and 130 F_4 RILs derived from 'Hutcheson' and PI 471938 genotyped with the SoySNP6K iSelect BeadChip. Field evaluation of canopy wilting was conducted under rain-fed conditions at two locations (Athens, GA and Salina, KS). The panel was evaluated in 2015 and 2016, and the RIL population in 2016 only. Substantial variation in canopy wilting was observed among the genotypes in both the panel and RIL population. Using a genome-wide association mapping approach, 47 unique SNPs that tagged 45 loci were associated with canopy wilting in at least one environment, with five regions identified both in a single environment and using across all environments data. Several new soybean accessions were identified with canopy wilting superior to those of commonly-grown check genotypes. Additionally, two QTL were identified for canopy wilting from the RIL population using composite interval mapping. The germplasm and genomic regions identified through this research can be used to better understand the slow canopy wilting trait and be incorporated into elite germplasm to improve drought tolerance in soybean.

Keywords: soybean, *Glycine max*, drought tolerance, canopy wilting, genome-wide association study (GWAS), linkage mapping

Introduction

Soybean [*Glycine max* (L.) Merrill] is the world's leading oilseed crop, and is used to produce vegetable oil, protein feed for livestock, biodiesel, and many soyfoods. Drought stress is the most significant abiotic threat to the agricultural productivity of soybean around the world, and can reduce yield by more than 40 percent (Specht et al., 1999; Purcell and Specht, 2004).

Loss of turgor and leaf droop, known commonly as canopy wilting, is an often observed response to drought stress in soybean. Some exotic soybean types exhibit a slow or delayed canopy wilting response to drought, which may reflect favorable underlying plant mechanisms to access soil moisture, conserve soil moisture prior to stress, or use water more efficiently. One mechanism related to water conservation is to restrict transpiration early in the growing season whenever vapor pressure deficit (VPD) is high, so that plants can utilize saved soil water during pod filling when drought stress in soybean is usually more detrimental to yield. Plant introduction (PI) 416937 is a Japanese maturity group (MG) VI introduction identified in the 1980's as exhibiting slower wilting under water deficit conditions than existing cultivars (Sloane et al., 1990; "USDA-ARS National Genetic Resources Program," 2018a). This PI has an extensive lateral root system and high root surface area (Goldman et al., 1989; Sloane et al., 1990; Hudak and Patterson, 1996; Pantalone and Rebetzke, 1996) and low stomatal conductance (Tanaka et al., 2010). In a study that evaluated PI 416937 in high VPD conditions, it reached a maximum transpiration rate near 2.0 kPa, whereas other genotypes continued to increase transpiration rates at much greater than 2.0 kPa (Fletcher et al., 2007). This result indicated that

water conservation during vegetative growth may be the basis for the slow wilting trait found in PI 416937. PI 471938, a MG V introduction from Nepal, also exhibits the slow wilting trait, but the basis for this trait is unknown (Sadok et al., 2012; Bagherzadi et al., 2017; "USDA-ARS National Genetic Resources Program," 2018b). PI 471938 was also previously identified as expressing N_2 fixation tolerance to soil drying (Sinclair et al., 2000; Devi and Sinclair, 2013; Riar et al., 2018) These two plant introductions are being used as sources of slow wilting in applied breeding programs (Devi et al., 2014; Carter et al., 2016). Two additional MG III PIs were also previously identified that have reduced yield loss under drought stress and delayed leaf wilting (Pathan et al., 2014).

The canopy wilting trait has been mapped using linkage and genome-wide association mapping approaches. Charlson et al. (2009) mapped four QTLs on chromosomes (Chr) 8, 13, 14, and 17 that collectively explained 47% of phenotypic variation in a KS4895 × 'Jackson' RIL population (Johnson, 1958; Hwang et al., 2015a). A total of eight QTLs were identified for canopy wilting under field and greenhouse conditions in Du et al. (2009) using 184 RILs derived from Kefeng No. 1 × Nannong1138-2. Using 150 RILs derived from the hybridization of 'Benning' (Boerma et al., 1997) and PI 416937, seven QTLs were identified by Abdel-Haleem et al. (2012) that explained 75% of the variation in canopy wilting. Hwang et al. (2015) identified eight QTL clusters that had QTLs from at least two of five different RIL populations (93705 KS4895 × Jackson, 08705 KS4895 × Jackson, KS4895 × PI 424140, 'A5959' × PI 416937, and Benning × PI 416937) that are responsible for canopy wilting. With these same populations, Hwang et al. (2016) performed a meta-analysis on nine QTLs with refined map positions and reduced confidence intervals for the eight QTL clusters reported by Hwang et al. (2015). Most recently, Kaler et al. (2017b) used a genome-wide association analysis of 373 MG IV genotypes

to identify 61 single nucleotide polymorphism (SNP) markers associated with canopy wilting, which tagged 51 different genetic loci. Of the 373 genotypes they tested, 185 genotypes had lower canopy wilting scores across environments than PI 416937 (Kaler et al., 2017b).

There are approximately 170,000 soybean accessions maintained in germplasm collections worldwide, and the USDA maintains a collection of around 20,000 accessions. However, only a limited number of these genotypes have been screened for stress tolerance, and few have been identified as tolerant and used in soybean breeding programs to improve drought tolerance (Carter et al., 1999, 2004, 2016; Devi et al., 2014; Sinclair et al., 2016). Identification of new accessions with beneficial alleles for drought tolerance related traits including slow canopy wilting could help in the development of drought tolerant soybean cultivars. To aid in the search for beneficial alleles, the SoySNP50K and SoySNP6K iSelect BeadChips are available for high-throughput genotyping that supports QTL mapping efforts. In addition, the entire USDA soybean germplasm collection has been genotyped with the SoySNP50K iSelect BeadChips (Song et al., 2013, 2014, 2015).

Genome-wide association studies (GWAS) allow for the opportunity to map the genomic regions for traits of interest by utilizing diverse soybean germplasm and populations. Use of association panels can increase the mapping resolution compared to traditional QTL mapping (Deshmukh et al., 2014). Population structure in these panels can occur if some of the genotypes are more related to each other compared to the rest of the population. Failure to correct for population stratification in GWAS models can lead to false positives, especially if the trait of interest is correlated with the structure of the panel (Wang et al., 2005). Several studies in soybean have been reported using the GWAS approach with SNP markers for many different traits, such as seed composition (Hwang et al., 2014; Vaughn et al., 2014; Bandillo et al., 2015;

Cao et al., 2017), salt tolerance (Patil et al., 2016; Zeng et al., 2017), carbon isotope composition (Dhanapal et al., 2015; Kaler et al., 2017a), ureide concentration (Ray et al., 2015), agronomic traits (Zhang et al., 2015; Contreras-Soto et al., 2017; Li et al., 2017), chlorophyll traits (Dhanapal et al., 2016), local adaptation (Bandillo et al., 2017), insect resistance (Chang and Hartman, 2017), and canopy wilting (Kaler et al., 2017b). These types of studies have provided a useful way to identify potential genomic regions with high resolution and candidate genes or QTLs for traits of interest.

The objectives of this study were to: i) evaluate a genetically diverse panel of soybean genotypes and RIL population in repeated field experiments for canopy wilting, ii) identify new germplasm that possesses the slow canopy wilting trait, iii) perform a population structure analysis to characterize the panel and determine extent subpopulation structure will affect mapping, and iv) elucidate genomic regions responsible for canopy wilting using both association and linkage mapping approaches.

Materials and methods

Plant materials and panel selection

When selecting the panel for the present study, approximately 600 soybean accessions were chosen initially from the USDA collection based on geographic origin and low annual precipitation. The 600 were truncated to ~170 accessions by examining the diversity among the accessions based on SNP genotype profiles. Only PIs with less than 85% similarity to each other based on these SNP genotypes were included in the panel. An additional 40 newly developed breeding lines with enhanced drought-related traits, as well as drought tolerant and susceptible checks, were added to bring the total number of genotypes in the panel to 209. These 209

genotypes were derived from 30 countries and range from MG III-IX, with 75% of the lines in MG VI-VIII (groups commonly grown in Georgia).

A cross between 'Hutcheson' (PI 518664) and PI 471938 was made in 1998 at Raleigh, NC, USA. Hutcheson is a MG V cultivar developed by Virginia Tech (Buss et al., 1988). PI 471938 is a MG V plant introduction characterized previously as a slow wilting soybean genotype (Carter et al., 1999; Hufstetler et al., 2007; Sadok et al., 2012). The F_1 seed from this cross were grown at the USDA Tropical Agricultural Research Station in Isabela, Puerto Rico. The F_2 to F_4 generations were advanced by single seed decent (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. Three replicates of each parent were also evaluated with this RIL population as check genotypes.

Genotype data and quality control

All but nine of the genotypes in the panel were previously genotyped with the SoySNP50K iSelect BeadChips (Song et al., 2013). These nine accessions were genotyped using the same procedure at the USDA Soybean Genomics and Improvement Lab in Beltsville, MD (Song et al., 2013). Briefly, 15 seed from each of these nine accessions were grown in a single 32 oz. styrofoam cup in a greenhouse at the University of Georgia, Athens, GA, USA. After approximately 2 weeks, leaf tissue was harvested and bulked in a 50 mL tube. The tissue was then placed in a lyophilizer for 2 days, and ground into a fine powder using a Geno/Grinder (SPEX SamplePrep, Metuchen, New Jersey, USA). DNA was extracted and then genotyped with the 50K chip. All SNP marker data was downloaded from SoyBase (Grant et al., 2010) for the remaining 200 accessions. A total of 42,079 SNP markers were available from these genotyping

efforts. The final number employed for mapping was from 34,808-35,233 markers (depending on environment) after removing markers with minor allele frequencies below 0.05. The number of markers used varied slightly due to the different number of accessions evaluated (185-209 accessions) in each environment. The Glyma.Wm82.a2 reference genome physical positions were used to determine the locations of the SNPs used in the analysis.

For the Hutcheson × PI 471938 RIL population, 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 for the association panel. 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 for creation of 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 1,212 polymorphic SNP markers to be used in QTL mapping.

Evaluation of canopy wilting

Canopy wilting was evaluated for the association panel in Athens, GA, USA in 2015 (GA-15-PANEL) and 2016 (GA-16-PANEL) and Salina, KS, USA in 2015 (KS-15-PANEL) and 2016 (KS-16-PANEL) in rain-fed field plots after extended periods of little or no rainfall. Genotypes were planted as two-row plots in each environment, employing a randomized complete block design with three replications. Experiments were sown on 16 June 2015 (GA-15-PANEL), 8 June 2016 (GA-16-PANEL), 10 June 2015 (KS-15-PANEL), and 8 June 2016 (KS-16-PANEL).

All plots for the four environments were planted with 0.76 m row spacing at a seeding density of 32 seed m^{-2}. For GA-15-PANEL and GA-16-PANEL, the plots were 2.43 m in length, and for KS-15-PANEL and KS-16-PANEL the plots were 3.65 m long.

The Hutcheson × PI 471938 RIL population was evaluated in Athens, GA in 2016 (GA-16-RIL) and Salina, KS in 2016 (KS-16-RIL). One-row plots were planted at both locations with three replications in GA-16-RIL and two replications in KS-16-RIL using a randomized complete block design. The experiments were sown at each location on the same day, 8 June 2016. Row spacing, seeding density, and plot length were the same as described for each location with the association panel.

Wilting was rated for both populations 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. Volumetric water content (VWC) was measured with a single Decagon GS1 soil moisture probe placed approximately 15 cm below the soil surface in one corner of the field plots to measure available soil water at the time the scores were recorded.

For the GWAS study, wilting scores in Athens, GA were taken by two raters in 2015 and three raters in 2016. In 2015, a single canopy wilting rating was taken on 29 July (most genotypes in vegetative stages, 17% VWC) for the Athens, GA plots and four ratings were taken in 2016 between 25 August and 16 September (most genotypes in pod filling stage, ~5-8% VWC). One rater recorded canopy wilting scores for Salina, KS in both years. In 2015, four

ratings were taken between 12 August and 28 September (most genotypes in vegetative stage in first rating and pod filling stage at last rating, ~18-20% VWC), and in 2016 three ratings were taken between 26 July and 4 August (during flowering, VWC data not available). Mean ratings for an individual plot over dates and raters were employed as the phenotypic wilting score given that correlations between raters and rating dates were generally high (data not shown). All 209 entries were evaluated in GA-15-PANEL. However, because of seed availability and quality, 206 entries were evaluated in GA-16-PANEL and KS-16-PANEL, and 185 entries in KS-15-PANEL.

For the RIL study, the mean of three ratings taken between 26 August and 16 September (during pod filling) by three raters was used as the phenotypic score for the GA-16-RIL environment. A single rating taken on 27 July 2016 (during flowering) by one rater was used as the phenotypic score for the KS-16-RIL environment. In 2017, we attempted to evaluate this RIL population at three additional environments (Athens, GA, Salina, KS, and Sandhills, NC), but no canopy wilting scores were recorded because water stress conditions were minimal.

Population structure analyses

Population structure was determined using fastSTRUCTURE (Raj et al., 2014), principal coordinate analysis, and by constructing a dendrogram based on the 50K SNP data for the association panel. The fastSTRUCTURE program was run in default settings with the simple option on the 209 soybean genotypes testing for subpopulations (K) ranging from K = 2 - 10. As part of the fastSTRUCTURE package, the python script ChooseK was used to choose the number of subpopulations that maximize the marginal likelihood. Principal coordinate analysis was performed using the GAPIT R package (Lipka et al., 2012) and visualized with TIBCO Spotfire (TIBCO Software Inc., Palo Alto, CA, USA). The neighbor-joining clustering algorithm

in TASSEL version 5.0 (Bradbury et al., 2007) was used to build a dendrogram, which was visualized with FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

Statistical analyses

Analyses of variance (ANOVA) were conducted using PROC GLM in SAS version 9.4 (SAS Institute, 2014). For both populations, a model for canopy wilting scores was created with genotype treated as a fixed effect, and environment, replication within environment, and genotype by environment interaction treated as random effects. 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 after Holland et al. (2002) with the variance components being calculated with PROC MIXED of SAS version 9.4 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 for the association panel using JMP Pro (JMP®, Version 13, SAS Institute Inc., Cary, NC, USA). The model was built by treating genotype, environment, genotype by environment interaction, and replication within environment as random variables using the Standard Least Squares personality and REML method. For individual environments for both the association panel and RIL population, only genotype and replication were used and treated as random to calculate BLUPs. Use of BLUP values for each genotype across and within environments helped to account for variation caused by environmental factors and missing data, and these BLUPs were used as the phenotype values for subsequent GWAS and QTL analyses.

Genome-wide association analyses

Genome-wide association analyses were performed using Fixed and random model Circulating Probability Unification (FarmCPU) (Liu et al., 2016). This R package uses a multiple loci linear mixed model (MLMM) which incorporates the most significant SNP markers as covariates in a modified mixed linear model (MLM), and uses fixed and random effect models iteratively to eliminate confounding between kinship and the markers being tested. This method helps to improve statistical power to detect significant markers associated with a particular phenotype and is computationally efficient. A total of 34,808-35,233 genome-wide SNP markers were used for the analysis after removing markers with minor allele frequencies (MAF) below 0.05. The number of markers used varied slightly due to the different number of accessions evaluated (185-209 accessions) in each environment. The differences in accession number affected which SNP markers were included for each genotype file when the MAF of the marker was close to 0.05. Manhattan plots were visualized with the 'qqman' R package (Turner, 2014) using p-values generated from the FarmCPU output.

A Bonferroni threshold ($p < 2.83E\text{-}07$, $-\log_{10}(P) > 6.55$) is overly strict when the linkage disequilibrium among genetic markers is large, which is generally the case with soybean (Hyten et al., 2007). Therefore, a p-value threshold of ($p < 0.0001$; $-\log_{10}(P) > 4$) was used, which is less stringent than the Bonferroni-corrected threshold, but more stringent than the threshold used in Kaler et al. (2017b) and many other soybean GWAS studies using SoySNP50K genotype data. This threshold was used to identify SNPs that were significantly associated with the canopy wilting trait.

Pairwise estimates of D′ and r^2 were calculated by chromosome using Haploview version 4.2 software (Barrett et al., 2005). Linkage disequilibrium (LD) blocks were estimated using the

Solid Spine of LD option using $D' > 0.8$ to extend the spine. Significant SNPs associated with canopy wilting were deemed part of the same locus (genomic region) controlling the trait if they were in the same LD block. Allelic effects were calculated by taking the difference in mean canopy wilting score between the two alleles at a particular SNP, and the direction, negative or positive, of the allelic effect estimates are relative to the alphabetical order of the nucleotides at each particular marker. For example, if the nucleotides at a particular SNP are "A" and "C", then a positive allelic effect (e.g. 3) indicates that possessing the "C" allele will increase the phenotype by three units. Therefore, a negative effect value indicates that an individual possessing the second nucleotide alphabetically for this SNP would have a lower canopy wilting score, while a positive effect value would indicate that having the second nucleotide alphabetically would have increased the score. For this study, the allelic effects are based on BLUP values, not actual canopy wilting scores, so while the overall effect is relevant, it does not directly apply to raw canopy wilting scores. The amount of phenotypic variation explained (R^2) by SNPs was calculated using a simple linear regression in R with lm(BLUP ~ SNP) for an individual SNP and lm(BLUP ~ SNP_1 + SNP_2 + …) for all significant SNPs in a given environment used as the models.

Breeding values for accessions evaluated in the association panel were calculated by adding up the allelic effects for all SNPs significantly associated ($p < 0.0001$; $-\log_{10}(P) > 4$) with canopy wilting in each individual environment and with the across all environments BLUPs. Breeding values from across the individual environments (GA-15-PANEL, GA-16-PANEL, KS-15-PANEL, and KS-16-PANEL) were also summed. The allelic effect for a given accession was considered negative if the allele contributed to lower canopy wilting scores. In contrast, if the allele increased canopy wilting score, it was considered a positive value. Therefore, a more

negative breeding value indicated an accession had a sum of allelic effects across all significant SNPs that was more favorable towards reduced canopy wilting. If the allele at a particular SNP was heterozygous or missing for a genotype, it was omitted in the breeding value calculation.

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). Logarithm of odds (LOD) criterion of greater than six was used to establish linkage groups. As necessary, some of the 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 walk speed of 1 cM and 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 thresholds were LOD = 3.4 and LOD = 3.5 for the GA-16-RIL and KS-16-RIL environments, respectively. MapChart 2.30 (Voorrips, 2002) was used to visualize the genetic maps and QTL mapping results.

Candidate gene identification

SNPs from the GWAS that met the threshold of $-\log_{10}(P) > 4$ and QTLs with logarithm of odds (LOD) scores greater than 3.4 (GA-16-RIL) or 3.5 (KS-16-RIL) were used to identify nearby candidate genes. Candidate genes and their functional annotation were identified using the Glyma2.1 gene models in SoyBase for models within plus or minus 10 kb of the SNP physical position.

Results

Canopy wilting for GWAS panel

Substantial variation for canopy wilting was observed among the genotypes within the panel across the four environments tested. In general, canopy wilting scores were higher (more severe wilting) in Athens, GA compared to Salina, KS in both 2015 and 2016 (Figure 2.1). Genotypes, environments, and their interactions were statistically significant ($p < 0.05$) for canopy wilting scores (Table 2.1). Correlations of canopy wilting scores based on genotype means among the environments ranged from r = 0.29 (GA-15-PANEL/KS-15-PANEL) to r = 0.61 (GA-15-PANEL/KS-16-PANEL). Broad sense heritability of canopy wilting on an entry-mean basis for each environment was 62% (GA-15-PANEL), 75% (GA-16-PANEL), 38% (KS-15-PANEL), 74% (KS-16-PANEL), and 34% across all environments (ALL-PANEL).

The 209 genotypes were ranked from lowest to highest canopy wilting score within each environment, and then over the four environments (Tables 2.2 and 2.S1). Numerically, 106 genotypes exhibited less wilting than the slow canopy wilting check, PI 416937. Additionally, 68 lines exhibited less wilting, numerically, than slow canopy wilting check, PI 471938. Thirty

eight and 44 lines exhibited numerically greater wilting than fast wilting cultivars Benning and Hutcheson, respectively (Table 2.S1).

Canopy wilting for RIL population

Much like the association panel, the RIL population exhibited a wide range of canopy wilting among the RILs and the wilting was more severe in the Georgia environment (Figure 2.1 and Table 2.S2). Genotypes, environments, and their interactions were statistically significant ($p < 0.05$) for canopy wilting scores (Table 2.1). Correlation between the GA-16-RIL and KS-16-RIL environment canopy wilting scores was low ($r = 0.02$). Given this low correlation between two environments and the significant genotype by environment interaction, QTL mapping was conducted separately by each environment. Broad sense heritability of canopy wilting on an entry-mean basis for each environment was 32% (GA-16-RIL) and 36% (KS-16-RIL). Only N98-7291 ranked lower than the slow wilting parent, PI 471938. Forty nine of the RILs had faster wilting based on mean performance over environments compared to the fast wilting parent, Hutcheson (Table 2.S2).

Population structure for GWAS panel

The first two principal coordinates were visualized and colored by continent of origin (Figure 2.2A). All of the North American lines represent genotypes from the USA, and the majority (88/118) of Asian genotypes are comprised of genotypes from China. The genotypes of U.S. origin were more tightly clustered than were genotypes from China (Figure 2.2A). The first four principal coordinates explained approximately 19% of the variation in the data set, and were used as covariates in the GWAS model to help correct for potential population stratification

(Figure 2.S1). The North American genotypes had shorter distances between accessions compared to Asian and African genotypes based on the neighbor joining dendrogram analysis, indicating they are more closely related, which concurs with the principal coordinates analysis. Genotypes from Asia tended to group close to one another based on this analysis, but also intermixed with lines of African origin (Figure 2.2B). A continuous increase in marginal likelihood with increasing K was observed for the fastSTRUCTURE analysis, meaning as each sequential K value (increasing from 2 to 10) was tested it was deemed the K value that best characterized the population structure of the panel. This indicated that little structure is apparent for this population, because fastSTRUCTURE was not able to settle on an optimal K value to describe this panel within K = 2 - 10 (Westbrook et al., 2015) (Figure 2.S2).

GWAS of canopy wilting trait

Across and within environments, a total of 47 unique SNPs were identified that tagged 45 loci that are associated with canopy wilting from the GWAS analysis (Figure 2.3 and Table 2.3). Five of these SNPs (ss715580569, ss715593787, ss715596504, ss715610566, and ss715638586) on Chr 1, 6, 7, 11, and 20, respectively, were found to be significant ($p < 0.0001$; $-\log_{10}(P) > 4$) both in individual environments and with the BLUP value calculated using the canopy wilting scores across all environments (ALL-PANEL). Two additional regions, loci 9 and 30 on Chr 5 and 12, respectively, were identified in two of the four environments.

The quantile-quantile (QQ) plots follow the expected diagonal, with a sudden uptick for statistically significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNP markers in each different environment (Figure 2.3). This linear pattern of expected versus observed p-values also does not have a slope greater than one, which indicates the first four principal coordinates included in the GWAS

model adequately accounted for population stratification in this panel of genotypes. The deviation of the markers from this diagonal occurs at or greater than $-\log_{10}(P) = 4$ in each environment, which was the threshold we used to determine if a SNP marker was significantly associated with the canopy wilting trait.

Allelic effects across all significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs ranged from -4.74 to 3.79 (Table 2.3). The phenotypic variation for canopy wilting explained by an individual significant SNP (R^2) ranged from 0.01-30.0%, with a mean of 7.1% (Table 2.3). The R^2 for all significant SNPs in a given environment was 51% (GA-15-PANEL), 56% (GA-16-PANEL), 60% (KS-15-PANEL), 42% (KS-16-PANEL), and 68% (ALL-PANEL). The number of beneficial alleles each genotype possessed was determined by counting the number of alleles with effects that reduced canopy wilting score from all the significant SNPs. Overall, the number of beneficial alleles ranged from 14 to 34. The 10 slowest wilting genotypes had 23 to 34 beneficial alleles, while the 10 fastest wilting genotypes had 14 to 21 beneficial alleles (Table 2.2 and Table 2.S1). Summed breeding values across the individual environments ranged from -31.79 to 31.75 overall, and the 10 slowest wilting genotypes had breeding values ranging from -31.79 to 0.24 while the 10 fastest wilting genotypes had a range from 5.45 to 31.75 (Table 2.2 and Table 2.S1). Negative breeding values indicate that the genotype had a sum of allelic effects across the significant SNPs that was more favorable towards reduced canopy wilting scores. Positive breeding values indicate that the genotype had a sum of allelic effects across the significant SNPs that was less favorable towards reduced canopy wilting scores.

QTL mapping of canopy wilting trait for RIL population

Using CIM, QTLs were identified for canopy wilting in each of the two environments (GA-16-RIL and KS-16-RIL) tested (Table 2.4 and Figure 2.4). For GA-16-RIL, a QTL (Wilt-GA-16-RIL) was identified on Chr 13 which explained 12% of the phenotypic variation for this trait. The peak marker for the Wilt-GA-16-RIL QTL was only 31 kb from locus 32 identified with the association mapping approach. A confidence interval (CI) which included all markers that met the genome-wide logarithm of odds (LOD) threshold within the QTL spanned 2.3 Mb [27,485,765 to 29,808,335 base pairs (bp)]; the peak for this QTL was at 29,481,243 bp. For the KS-16-RIL environment, one QTL on Chr 16 (Wilt-KS-16-RIL) explained 11% of the phenotypic variation for canopy wilting score. This QTL has a CI that spanned 6.3 Mb (7,364,708 to 13,657,908 bp), with the peak at 7,851,145 bp. No significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs were found within this CI with association mapping in the current study. Both of these QTLs had negative additive effects calculated from the CIM. This negative effect indicated the mean canopy wilting score for RILs possessing the allele from PI 471938 was lower than RILs possessing the allele from Hutcheson.

Identification of candidate genes for the canopy wilting trait

The median distance between SNP markers used in the GWAS was 9 kb, and the mean distance was 26 kb. Although identifying all gene models in LD with significant SNPs would be ideal, we focused our efforts on models in close proximity (within plus or minus 10 kb), which approximately spans this distance between markers. One hundred sixteen candidate genes were found within plus or minus 10 kb of the 47 significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs for the GWAS (Table 2.S3). Using the same distance (plus or minus 10 kb), six gene models were

identified near the SNP markers with the highest LOD score for each QTL identified with linkage mapping (Table 2.S4).

Discussion

Canopy wilting for GWAS panel

In this study, we evaluated a panel of 209 soybean genotypes in four environments and 130 RILs in two environments for canopy wilting score after extended periods of drought stress. Plant introductions in the panel were primarily MG VI-VIII, and many were never evaluated previously for drought tolerance related traits. There was substantial genetic variation for canopy wilting within each environment. Canopy wilting is a complex, quantitative trait (Charlson et al., 2009; Hwang et al., 2016), and our phenotypic data further confirm this notion (Figure 2.1).

Slow wilting PI 603535, a MG VIII accession from China, had a mean score of 6 across all environments. Nine of 10 genotypes with the slowest wilting ranking originated from China. Among the fast wilting genotypes in the panel, PI 330635, a MG VII accession from South Africa, had a mean canopy wilting score of 39 across environments (Table 2.2). Soybean genotypes PI 416937 and PI 471938 were included as slow wilting checks in the association panel studies, and many accessions had lower mean wilting scores (less wilting) than these check genotypes. One hundred and six genotypes had lower canopy wilting scores than PI 416937 and 68 genotypes had lower canopy wilting scores than PI 471938 (Table 2.S1). However, the mean canopy wilting scores across all environments evaluated with the association panel for PI 416937 and PI 471938 were 18 and 14, respectively, which are also relatively low scores (Table 2.S1). These newly identified PIs could be sources of parental stock that could be exploited by soybean breeders to improve the canopy wilting trait, especially the accessions with favorable alleles

different than PI 416937 and PI 471938, and with negative breeding values. In Kaler et al. (2017b), 185 of the 373 genotypes they tested had scores lower than PI 416937. The current study and Kaler et al. (2017b) demonstrate there is likely more variation and potential for improvement of the canopy wilting trait than previously reported, but testing these new slow wilting genotypes in more environments is necessary to further confirm they will consistently exhibit this trait in different locations and stress severities.

Physiological mechanisms for canopy wilting and relationship to other traits

Slow canopy wilting could lead to less yield reduction during drought stress for soybeans. 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 high water use efficiency (WUE), high radiation use efficiency (RUE), and conservation of soil moisture. Genotypes in this group would utilize transpired water for biomass production more efficiently, and higher RUE would be expected in both drought stressed and optimal growing conditions. The second combination is low stomatal conductance, low RUE, low WUE, and conservation of soil moisture. The genotypes in this group would have low transpiration which 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 selection which can show relationships with yield under stressed conditions. Leaf rolling reduces exposed leaf area, and thereby decreases transpiration and reduces light interception, and can be observed in crops such as maize, rice, and wheat. The earlier leaf rolling occurs in a given day or 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. 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 stay green, root architecture, and canopy temperature depression can impact a plant's ability to tolerate drought stress and have been evaluated in a number of crop species (Tuberosa, 2012). Slow canopy wilting in soybean is potentially related to many possible secondary physiological mechanisms that can be evaluated to improve our understanding of soybean drought physiology and potentially improve yield under stressed conditions.

Relationship of canopy wilting, days to flowering, and maturity group

Canopy wilting scores were overall higher in Georgia compared to Kansas in both 2015 and 2016 for the association panel (Figure 2.1). Given that the panel we evaluated consisted of genetically diverse genotypes (most of which were plant introductions) with some variation in phenology (flowering time, height, root mass, etc.) these scores could potentially be affected by varying degrees of competition for water resources from neighboring plots due to these factors.

However, 75% of the lines we tested were from a relatively narrow range of maturities (MG VI-VIII), so the majority of the accessions experienced drought stress at the same growth stage.

Days to flowering (DTF) was recorded in the GA-15-PANEL and GA-16-PANEL environments as the number of days from planting until 50% of the plants in a plot reached the R1 (first bloom) stage of development. Maturity groups (MG) for all accessions were obtained from the USDA GRIN website or were provided by the breeder who developed the line. Correlations between wilting score and MG ($r = -0.21 – 0.13$), and wilting score and DTF ($r = -0.05 – 0.15$) in single environments were relatively low and not consistently positive or negative in the current study. Across all environments, the correlation ($r = 0.01$) was also low for wilting score and MG. There did not appear to be a relationship between mean rank across environments and MG or DTF in the current study (Figure 2.S3). Additionally, previous work showed that canopy wilting ranking among genotypes was consistent across multiple MGs (IV to VII), ratings in different years/growth stages, and in different row spacing (King et al., 2009). In the current study, canopy wilting was primarily evaluated during the reproductive growth stages, with the exception of GA-15-PANEL and first rating in KS-15-PANEL, which were rated during the late vegetative growth stages for most genotypes. Water stress during the early reproductive growth stages has the greatest impact on reducing soybean yield as a result of the plants producing fewer pods, and in turn, less seed (Manavalan et al., 2009). Therefore, slow canopy wilting during reproductive growth stages could be a good indicator of drought tolerance and ability to maintain yield potential during water stress.

Genotype by environment interaction and heritability

Although genotype by environment interactions were significant ($p < 0.05$) with both populations (Table 2.1) and the severity of wilting experienced in the four environments varied (Figure 2.1), the correlations between wilting scores across environments were relatively high ($r = 0.29$-0.61) for the association panel, indicating that the genotypes tested from this panel wilted similarly across environments. 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., 2015b; Kaler et al., 2017b).

Population structure analyses with panel

Soybean was first domesticated in China (Hyten et al., 2006), and the accessions of Chinese origin from the association panel had the least tight cluster in the principal coordinates plot compared to other countries, and exhibited the greatest distance between accessions in the dendrogram, indicating they had the most diversity of the accessions tested (Figure 2.2A and Figure 2.2B). Lines from the USA were tightly clustered in the principal coordinates plot, and had short distances apart from one another in the dendrogram (Figures 2.2A and 2.2B). The genetic base used in North American soybean breeding has been characterized as being narrow, with only a few common ancestors explaining the majority of diversity for these breeding materials (Carter et al., 2004). The tight clustering based on principal coordinates and short distance between accessions in the dendrogram is a reflection of this narrow genetic diversity of elite U.S. soybean breeding lines. Given that the panel of soybean genotypes used for this study was explicitly chosen to be genetically diverse based on genome-wide 50K SNP data, the lack of

ability for fastSTRUCTURE to select an optimal number of K groups was expected. Based on the combination of the results of these population structure analyses, we determined that this panel had little or moderate population structure present that would affect the GWAS and cause false positives. As is commonly done with association mapping, we did include the first four principal coordinates in our GWAS model as a way to help reduce the possibility of potential population structure affecting the mapping results (Price et al., 2006).

Comparisons of genetic mapping for canopy wilting to previous studies

We identified 47 unique SNPs that tagged 45 loci that are associated with canopy wilting using a genome-wide association mapping approach. Five of these SNPs were found both in individual environments and when using the BLUP value calculated using the canopy wilting scores across all environments (ALL-PANEL). On Chr 5 and 12, two physically close SNPs found in two different environments within the same LD block were significantly associated with canopy wilting. Overall, significant SNPs were identified on 18 of the 20 soybean chromosomes, with only Chr 4 and 16 not having any marker-trait associations. The R^2 for all significant SNPs in a given environment was 51% (GA-15-PANEL), 56% (GA-16-PANEL), 60% (KS-15-PANEL), 42% (KS-16-PANEL), and 68% (ALL-PANEL). Using a linkage mapping approach, QTLs for canopy wilting on Chr 13 (Wilt-GA-16-RIL) and 16 (Wilt-KS-16-RIL) were identified. These QTLs accounted for 12 and 11% of the phenotypic variation for this trait, respectively (Figure 2.4 and Table 2.4). The Wilt-GA-16-RIL and Wilt-KS-16-RIL QTLs both had negative additive effects, with the allele coming from the PI 471938 parent reducing the canopy wilting score.

Several reports of QTL or genomic regions that control canopy wilting in soybean have been previously reported, and many are numbered with their approximate physical locations on

the SoyBase website. For the GWAS results, Loci 4 (ss715583580) and 5 (ss715582774) found on Chr 2 are in the same location or near three meta-QTLs identified in Hwang et al. (2016). Loci 32 (ss715614812) and 33 (ss715616411) flank the Canopy wilt 1-4 QTL found in Charlson et al. (2009) on Chr 13, and locus 34 (ss715618174) is co-located with the Canopy wilt 2-5 QTL identified in Abdel-Haleem et al. (2012) on Chr 14. Additionally, locus 19 on Chr 8 (Canopy wilt 3-2), 31 on Chr 12 (Canopy wilt 4-5), 40 on Chr 17 (Canopy wilt 3-10), 41 on Chr 17 (Canopy wilt 4-2 and 5-2), and 44 on Chr 19 (Canopy wilt 5-4), are within the confidence intervals of QTLs previously mapped across multiple populations (Hwang et al., 2015b). Of the 47 SNPs identified from GWAS associated with canopy wilting in the current study, a total of 10 were found in or near the same location as canopy wilting QTLs identified from the linkage mapping studies described above. Additionally, the peak marker of Wilt-GA-16-RIL identified on Chr 13 in the current study is only 31 kb from the locus 32 identified from the association mapping results and is also near the Canopy wilt 1-4 QTL (Charlson et al., 2009) (Figure 2.5). The overlapping regions and consistent QTLs across both GWAS and linkage mapping help provide validation that these loci are associated with the canopy wilting trait and could be the targets of improvement efforts.

Kaler et al. (2017b) recently used an association mapping approach to identify 61 SNP markers tagging 51 different loci for canopy wilting. A comparison of the significant SNPs found in the current study and Kaler et al. (2017b) was conducted (Figure 2.5). In total, eight SNPs tagging seven genomic regions were found on Chr 1, 7, 8, 12, 15, and 19 in the current study that are near SNPs identified in Kaler et al. (2017b). The main difference in our study compared to Kaler et al. (2017b) is that later maturity group soybeans (VI-VIII vs. IV) were used in the current study. The genomic regions that are consistent across maturity groups and many

different environments show promise as selection targets for improving this trait. Directing research efforts towards the genomic regions found in common between the current and previous GWAS studies could yield favorable alleles for the improvement of the canopy wilting trait in soybean.

Candidate genes at canopy wilting significant genomic regions

The SNP with the greatest absolute allelic effect (4.74) was found on Chr 1 (ss715578894) and had a MAF of 0.05 (Table 2.3). This SNP is located in the coding region of Glyma.01g001900, which has response to stress as one of its gene ontology terms (Table 2.S3). The SNP with the second greatest absolute allelic effect (3.79) was found on Chr 6 (ss715593748) and had a MAF of 0.13 (Table 2.3). The closest gene model to this SNP is Glyma.06g195100 (located ~4 kb away), which encodes a protein for atypical CYS HIS rich thioredoxin 4, and has a biological function of cell redox homeostasis (Table 2.S3). Modulators of cell redox homeostasis such as ascorbate, peroxiredoxins, and glutaredoxin, regulate the accumulation of reactive oxygen species (ROS) and can play a role in abiotic stress tolerance (Kapoor et al., 2015). Therefore, Glyma.06g195100 could be a potential target for improvement of the slow canopy wilting trait.

Three SNPs (ss715593787, ss715606644, and ss715613671) explained more than 20% of the phenotypic variation in a given environment (Table 2.3). A gene model (Glyma.06g200100) located 7 kb away from ss715593787 (Locus 14) on Chr 6 is part of the N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) superfamily (Table 2.S3), which has been hypothesized to regulate plant resistance mechanisms to abiotic stresses. In a recent study, overexpression of *GsSNAP33*, a SNAP25-type protein of this superfamily, from *G. soja* improved salt and drought tolerances in *Arabidopsis thaliana* (Nisa et al., 2017). Locus 27 has a

gene model (Glyma.10g151500) 958 bp away from ss715606644 on Chr 10 which encodes a RING/U-box superfamily protein (Table 2.S3). RING-type E3 ubiquitin ligases including DREB2A-interacting proteins DRIP1 and DRIP2 have been previously shown to play a role in drought response (Qin et al., 2008). GmRFP1 functions as a RING-type E3 ubiquitin ligase and is down-regulated by drought and cold stress, but is induced by ABA and salt stress suggesting in may be involved in abiotic stress response (Du et al., 2010). Given their relationship with drought stress response and improvement, these two gene models could be targeted for understanding and improving canopy wilting in soybean.

Four gene models are located within 10 kb on either side of the peak of Wilt-GA-16-RIL – these models have gene ontology terms including protein binding and kinase activity, ATP binding, and intracellular protein and vesicle-mediated transport. Two gene models are located near the peak of Wilt-KS-16-RIL including Glyma.16g076700 which encodes a protein for Xyloglucanase 113 (Table 2.S4). Xyloglucans are hemicelluloses that make up a large portion of the primary cell walls of dicots. In one study, constitutive expression of a xyloglucan endotransglucosylase/hydrolase homolog from hot pepper increased Arabidopsis plants' ability to tolerate drought and salt tolerance (Cho et al., 2006). Therefore, Glyma.16g076700 could play a role in improving the slow canopy wilting trait.

Conclusions

Using 209 genetically diverse soybean genotypes, a genome-wide association mapping approach identified 47 unique SNPs tagging 45 loci that are significantly associated with the canopy wilting trait. Five of these SNPs were identified in an individual environment, as well as across environments. Of these 47 SNPs, 10 were found in or near the same location as previous canopy

wilting QTLs identified from linkage mapping studies. In addition, eight SNPs mapped to seven genomic regions were found near regions identified in a previous association mapping study for canopy wilting. Two QTLs for canopy wilting were identified using CIM, of which one was also found in the GWAS from the current study, which is near a previously published QTL. Candidate genes located at these genomic regions were identified that could help to understand the functions of these genes and improve canopy wilting in soybean. The genomic regions discovered across environments and studies, in addition to the new slow wilting germplasm identified with favorable alleles, can be exploited by breeders to improve soybean drought tolerance.

References

Abdel-Haleem, H., T.E. Carter, L.C. Purcell, C.A. King, L.L. Ries, P. Chen, W. Schapaugh, T.R. Sinclair, and H.R. Boerma. 2012. Mapping of quantitative trait loci for canopy-wilting trait in soybean (Glycine max L. Merr). Theor. Appl. Genet. 125(5): 837–46.

Bagherzadi, L., T.R. Sinclair, M. Zwieniecki, F. Secchi, W. Hoffmann, T.E. Carter, and T.W. Rufty. 2017. Assessing water-related plant traits to explain slow-wilting in soybean PI 471938. J. Crop Improv. 31(3): 400–417.

Bandillo, N.B., J.E. Anderson, M.B. Kantar, R.M. Stupar, J.E. Specht, G.L. Graef, and A.J. Lorenz. 2017. Dissecting the genetic basis of local adaptation in soybean. Sci. Rep. 7: 17195.

Bandillo, N., D. Jarquin, Q. Song, R. Nelson, P. Cregan, J. Specht, and A. Lorenz. 2015. A population structure and genome-wide association analysis on the USDA soybean germplasm collection. Plant Genome 8(3): 1–13.

Barrett, J.C., B. Fry, J. Maller, and M.J. Daly. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21(2): 263–265.

Boerma, H.R., R.S. Hussey, D.V. Phillips, E.D. Wood, G.B. Rowan, and S.L. Finnerty. 1997. Registration of "Benning" soybean. Crop Sci. 37: 1982.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633–2635.

Brim, C.A. 1966. A modified pedigree method of selection in soybeans. Crop Sci. 6: 220.

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

Cao, Y., S. Li, Z. Wang, F. Chang, J. Kong, J. Gai, and T. Zhao. 2017. Identification of major quantitative trait loci for seed oil content in soybeans by combining linkage and genome-wide association mapping. Front. Plant Sci. 8: 1222.

Carter, T.E., R.L. Nelson, C.H. Sneller, and Z. Cui. 2004. Genetic diversity in soybean. p. 303–416. *In* Soybeans: Improvement, Production, and Uses. 3rd ed. ASA, CSSA, and SSSA.

Carter, T.E., P.I. De Souza, and L.C. Purcell. 1999. Recent advances in breeding for drought and aluminum resistance in soybean. p. 106–125. *In* Kauffman, H. (ed.), World Soybean Conference VI. Champaign, IL.

Carter, T.E., S.M. Todd, and A.M. Gillen. 2016. Registration of "USDA-N8002" soybean cultivar with high yield and abiotic stress resistance traits. J. Plant Regist. 10: 238–245.

Chang, H.-X., and G.L. Hartman. 2017. Characterization of insect resistance loci in the USDA soybean germplasm collection using genome-wide association studies. Front. Plant Sci. 8: 670.

Charlson, D. V, S. Bhatnagar, C.A. King, J.D. Ray, C.H. Sneller, T.E. Carter, and L.C. Purcell. 2009. Polygenic inheritance of canopy wilting in soybean [Glycine max (L.) Merr.]. Theor. Appl. Genet. 119: 587–594.

Cho, S.K., J.E. Kim, J.A. Park, T.J. Eom, and W.T. Kim. 2006. Constitutive expression of abiotic stress-inducible hot pepper CaXTH3, which encodes a xyloglucan endotransglucosylase/hydrolase homolog, improves drought and salt tolerance in transgenic Arabidopsis plants. FEBS Lett. 580(13): 3136–3144.

Contreras-Soto, R.I., F. Mora, M.A.R. de Oliveira, W. Higashi, C.A. Scapim, and I. Schuster. 2017. A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-based haplotype analysis. PLoS One 12(2): e0171105.

Deshmukh, R., H. Sonah, G. Patil, W. Chen, S. Prince, R. Mutava, T. Vuong, B. Valliyodan, and H.T. Nguyen. 2014. Integrating omic approaches for abiotic stress tolerance in soybean. Front. Plant Sci. 5: 244.

Devi, M.J., and T.R. Sinclair. 2013. Nitrogen fixation drought tolerance of the slow-wilting soybean PI 471938. Crop Sci. 53(5): 2072–2078.

Devi, J.M., T.R. Sinclair, P. Chen, and T.E. Carter. 2014. Evaluation of elite southern maturity soybean breeding lines for drought-tolerant traits. Agron. J. 106(6): 1947–1954.

Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C. Andy King, P.B. Cregan, Q. Song, and F.B. Fritschi. 2015. Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [Glycine max (L.) Merr.] genotypes. Theor. Appl. Genet. 128(1): 73–91.

Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, and F.B. Fritschi. 2016. Genome-wide association mapping of soybean chlorophyll traits based on canopy spectral reflectance and leaf extracts. BMC Plant Biol. 16: 174.

Du, Q.L., W.Z. Cui, C.H. Zhang, and D.Y. Yu. 2010. GmRFP1 encodes a previously unknown RING-type E3 ubiquitin ligase in soybean (Glycine max). Mol. Biol. Rep. 37(2): 685–693.

Du, W., D. Yu, and S. Fu. 2009. Detection of quantitative trait loci for yield and drought tolerance traits in soybean using a recombinant inbred line population. J. Integr. Plant Biol. 51(9): 868–878.

Fletcher, A.L., T.R. Sinclair, and L.H. Allen. 2007. Transpiration responses to vapor pressure deficit in well watered “slow-wilting” and commercial soybean. Environ. Exp. Bot. 61(2): 145–151.

Goldman, I.L., T.E. Carter, and R.P. Patterson. 1989. Differential genotypic response to drought

stress and subsoil aluminum in soybean. Crop Sci. 29(2): 330–334.

Grant, D., R.T. Nelson, S.B. Cannon, and R.C. Shoemaker. 2010. SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res. 38: 843–846.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2010. Estimating and interpreting heritability for plant breeding: An update. *In* Janick, J. (ed.), Plant Breeding Reviews. John Wiley and Sons, Inc., Oxford, UK.

Hudak, C.M., and R.P. Patterson. 1996. Root distribution and soil moisture depletion pattern of a drought-resistant soybean plant introduction. Agron. J. 88(3): 478–485.

Hufstetler, E.V., H.R. Boerma, T.E. Carter, and H.J. Earl. 2007. Genotypic variation for three physiological traits affecting drought tolerance in soybean. Crop Sci. 47: 25–35.

Hwang, S., C.A. King, P. Chen, J.D. Ray, P.B. Cregan, T.E. Carter, Z. Li, H. Abdel-Haleem, K.W. Matson, W. Schapaugh, and L.C. Purcell. 2016. Meta-analysis to refine map position and reduce confidence intervals for delayed-canopy-wilting QTLs in soybean. Mol. Breed. 36(7): 91.

Hwang, S., C.A. King, M.K. Davies, D. V. Charlson, J.D. Ray, P.B. Cregan, C.H. Sneller, P. Chen, T.E. Carter, and L.C. Purcell. 2015a. Registration of the KS4895 × Jackson soybean mapping population, AR93705. J. Plant Regist. 9(2): 266.

Hwang, S., C.A. King, J.D. Ray, P.B. Cregan, P. Chen, T.E. Carter, Z. Li, H. Abdel-Haleem, K.W. Matson, W. Schapaugh, and L.C. Purcell. 2015b. Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. Theor. Appl. Genet. 128(10): 2047–2065.

Hwang, E.-Y., Q. Song, G. Jia, J.E. Specht, D.L. Hyten, J. Costa, and P.B. Cregan. 2014. A genome-wide association study of seed protein and oil content in soybean. BMC Genomics

15: 1.

Hyten, D.L., I.Y. Choi, Q. Song, R.C. Shoemaker, R.L. Nelson, J.M. Costa, J.E. Specht, and P.B. Cregan. 2007. Highly variable patterns of linkage disequilibrium in multiple soybean populations. Genetics 175(4): 1937–1944.

Hyten, D.L., Q. Song, Y. Zhu, I.-Y. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker, and P.B. Cregan. 2006. Impacts of genetic bottlenecks on soybean genome diversity. PNAS 103(45): 16666–16671.

Johnson, H.W. 1958. Registration of soybean varieties, VI. Agron. J 50(22): 690–691.

Kaler, A.S., A.P. Dhanapal, J.D. Ray, C.A. King, F.B. Fritschi, and L.C. Purcell. 2017a. Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. Crop Sci. 57: 1–16.

Kaler, A.S., J.D. Ray, W.T. Schapaugh, C.A. King, and L.C. Purcell. 2017b. Genome-wide association mapping of canopy wilting in diverse soybean genotypes. Theor. Appl. Genet. 130(10): 2203–2217.

Kapoor, D., R. Sharma, N. Handa, H. Kaur, A. Rattan, P. Yadav, V. Gautam, R. Kaur, and R. Bhardwaj. 2015. Redox homeostasis in plants under abiotic stress: role of electron carriers, energy metabolism mediators and proteinaceous thiols. Front. Environ. Sci. 3: 13.

King, C.A., L.C. Purcell, and K.R. Brye. 2009. Differential wilting among soybean genotypes in response to water deficit. Crop Sci. 49(1): 290–298.

Li, S., Y. Cao, J. He, T. Zhao, and J. Gai. 2017. Detecting the QTL-allele system conferring flowering date in a nested association mapping population of soybean using a novel procedure. Theor. Appl. Genet. 130(11): 2297–2314.

Lipka, A.E., F. Tian, Q. Wang, J. Peiffer, M. Li, P.J. Bradbury, M.A. Gore, E.S. Buckler, and Z.

Zhang. 2012. GAPIT: Genome association and prediction integrated tool. Bioinformatics 28(18): 2397–2399.

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

Manavalan, L.P., S.K. Guttikonda, L.-S. Tran, and H.T. Nguyen. 2009. Physiological and molecular approaches to improve drought resistance in soybean. Plant Cell Physiol. 50(7): 1260–1276.

Nisa, Z. un, A.I. Mallano, Y. Yu, C. Chen, X. Duan, S. Amanullah, A. Kousar, A.W. Baloch, X. Sun, D. Tabys, and Y. Zhu. 2017. GsSNAP33, a novel Glycine soja SNAP25-type protein gene: Improvement of plant salt and drought tolerances in transgenic Arabidopsis thaliana. Plant Physiol. Biochem. 119: 9–20.

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

Pantalone, V., and G. Rebetzke. 1996. Phenotypic evaluation of root traits in soybean and applicability to plant breeding. Crop Sci. 36: 456–459.

Pathan, S.M., J.-D. Lee, D.A. Sleper, F.B. Fritschi, R.E. Sharp, T.E. Carter, R.L. Nelson, C.A. King, W.T. Schapaugh, M.R. Ellersieck, H.T. Nguyen, and J.G. Shannon. 2014. Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought. J. Agron. Crop Sci. 200(3): 231–236.

Patil, G., T. Do, T.D. Vuong, B. Valliyodan, J.D. Lee, J. Chaudhary, J.G. Shannon, and H.T. Nguyen. 2016. Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. Sci. Rep. 6: 19199.

Price, A.L., N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38(8): 904–909.

Purcell, L.C., and J.E. Specht. 2004. Physiological traits for ameliorating drought stress. p. 569–620. *In* Soybeans: Improvement, Production, and Uses. ASA, CSSA, and SSSA.

Qin, F., Y. Sakuma, L.-S.P. Tran, K. Maruyama, S. Kidokoro, Y. Fujita, M. Fujita, T. Umezawa, Y. Sawano, K.-I. Miyazono, M. Tanokura, K. Shinozaki, and K. Yamaguchi-Shinozaki. 2008. Arabidopsis DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene expression. Plant Cell 20: 1693–1707.

Raj, A., M. Stephens, and J.K. Pritchard. 2014. FastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197(2): 573–589.

Rauf, S., J.M. Al-khayri, M. Zaharieva, P. Monneveux, and F. Khalil. 2016. Breeding strategies to enhance drought tolerance in crops. p. 1–70. *In* Advances in Plant Breeding Strategies: Agronomic, Abiotic and Biotic Stress Traits.

Ray, J.D., A.P. Dhanapal, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C.A. King, D. Boykin, P.B. Cregan, Q. Song, and F.B. Fritschi. 2015. Genome-wide association study of ureide concentration in diverse maturity group IV soybean [Glycine max (L.) Merr.] accessions. G3 5: 2391–2403.

Riar, M.K., P. Cerezini, A. Manandhar, T.R. Sinclair, Z. Li, and T.E. Carter. 2018. Expression of drought-tolerant N fixation in heterogeneous inbred families derived from PI471938 and Hutcheson soybean. Crop Sci. 58: 364–369.

Ries, L.L., L.C. Purcell, T.E. Carter, J.T. Edwards, and C.A. King. 2012. Physiological traits

contributing to differential canopy wilting in soybean under drought. Crop Sci. 52: 272–281.

Sadok, W., M.E. Gilbert, M.A.S. Raza, and T.R. Sinclair. 2012. Basis of slow-wilting phenotype in soybean PI 471938. Crop Sci. 52(3): 1261–1269.

SAS Institute. 2014. The SAS system for Windows. Release 9.4.

Sinclair, T.R., J.M. Devi, and T.E. Carter. 2016. Limited-transpiration trait for increased yield for water-limited soybean: From model to phenotype to genotype to cultivars. p. 129–146. *In* Yin, X., Struik, P.C. (eds.), Crop Systems Biology: Narrowing the gaps between crop modelling and genetics. Springer International Publishing, Cham.

Sinclair, T.R., L.C. Purcell, V. Vadez, R. Serraj, C.A. King, and R. Nelson. 2000. Identification of soybean genotypes with N fixation tolerance to water deficits. Crop Sci. 40: 1803–1809.

Sloane, R.J., R.P. Patterson, and T.E. Carter. 1990. Field drought tolerance of a soybean plant introduction. Crop Sci. 30: 118–123.

Song, Q., D.L. Hyten, G. Jia, C. V Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS One 8(1): e54985.

Song, Q., D.L. Hyten, G. Jia, C. V Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2015. Fingerprinting soybean germplasm and its utility in genomic research. G3 (Bethesda). 5(10): 1999–2006.

Song, Q., G. Jia, C. Quigley, E. Fickus, D. Hyten, R. Nelson, and P. Cregan. 2014. Soybean BARC- SoySNP6K Beadchip—a tool for soybean genetics research. p. Abstract no.: P306. *In* Poster presented at: Plant and Animal Genome XXII Conference.

Specht, J., D. Hume, and S. Kumudini. 1999. Soybean yield potential—A genetic and

physiological perspective. Crop Sci. 39: 1560–1570.

Tanaka, Y., K. Fujii, and T. Shiraiwa. 2010. Variability of leaf morphology and stomatal conductance in soybean [Glycine max (L.) Merr.] cultivars. Crop Sci. 50(6): 2525–2532.

Tuberosa, R. 2012. Phenotyping for drought tolerance of crops in the genomics era. Front. Physiol. 3: 347.

Turner, S.D. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv Prepr.

USDA-ARS National Genetic Resources Program. 2018a. Glycine max Merr. PI 416937 Germplasm Resour. Inf. Netw.

USDA-ARS National Genetic Resources Program. 2018b. Glycine max Merr. PI 471938 Germplasm Resour. Inf. Netw.

Vaughn, J.N., R.L. Nelson, Q. Song, P.B. Cregan, and Z. Li. 2014. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. G3 (Bethesda). 4(11): 2283–2294.

Voorrips, R.E. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. J. Hered. 93(1): 77–78.

Wang, W.Y.S., B.J. Barratt, D.G. Clayton, and J.A. Todd. 2005. Genome-wide association studies: theoretical and practical concerns. Nat. Rev. Genet. 6(2): 109–118.

Wang, S.C., J. Basten, and Z.B. Zeng. 2012. Windows QTL Cartographer 2.5.

Westbrook, J., V. Chhatre, L. Wu, S. Chamala, L. Neves, P. Muñoz, P. Martínez-García, D. Neale, M. Kirst, K. Mockaitis, C. Nelson, G. Peter, J. Davis, and C. Echt. 2015. A consensus genetic map for Pinus taeda and Pinus elliottii and extent of linkage disequilibrium in two genotype-phenotype discovery populations of Pinus taeda. G3 Genes

Genomes Genet. 5: 1685–1694.

Zeng, A., P. Chen, K. Korth, F. Hancock, A. Pereira, K. Brye, C. Wu, and A. Shi. 2017. Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. Mol. Breed. 37: 30.

Zhang, J., Q. Song, P.B. Cregan, and G.-L. Jiang. 2015. Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (Glycine max). Theor. Appl. Genet. 129: 117–130.

Figures and tables

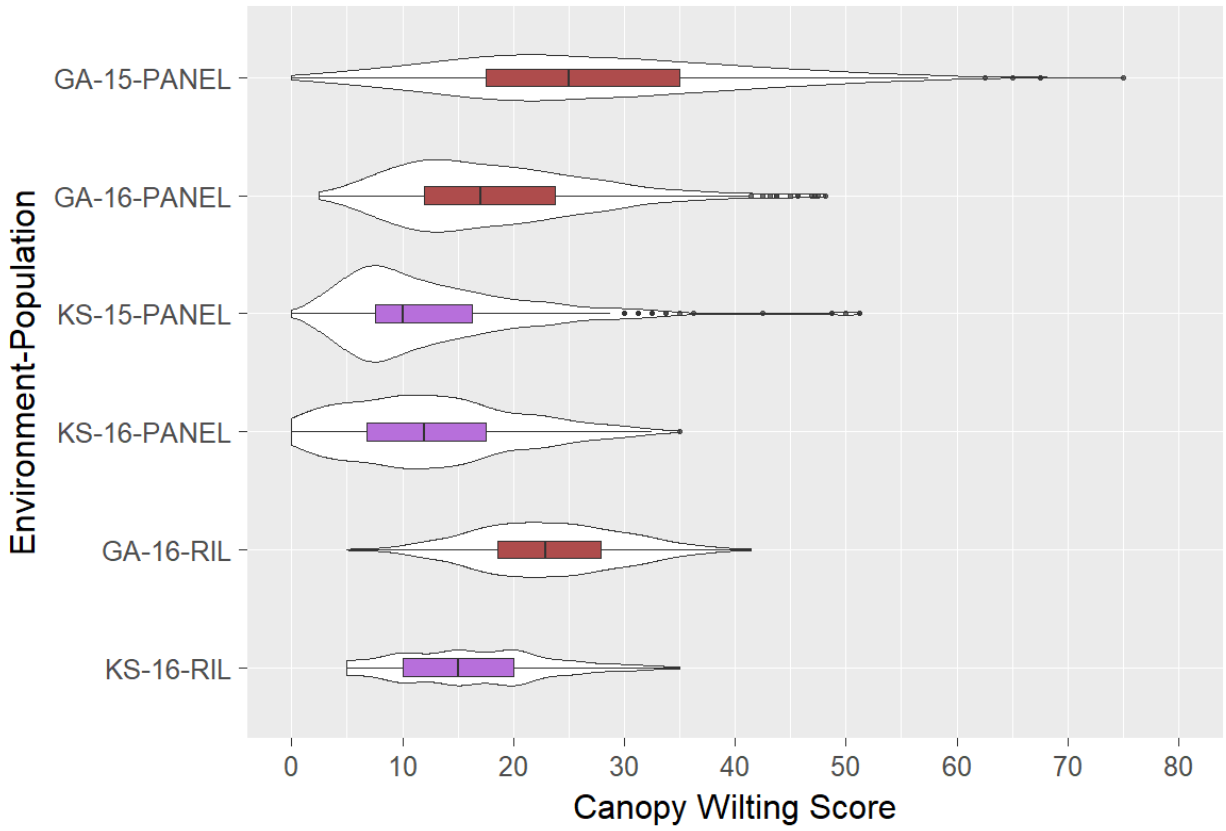


Figure 2.1. Violin plots with boxplots inside showing the distribution of canopy wilting scores for the panel and RIL population across environments. Environments are named as Location-Year-Population, with Georgia (GA) and Kansas (KS) as locations, 2015 (15) and 2016 (16) as years, and an association panel (PANEL) and Hutcheson × PI 471938 RILs (RIL) as the populations.

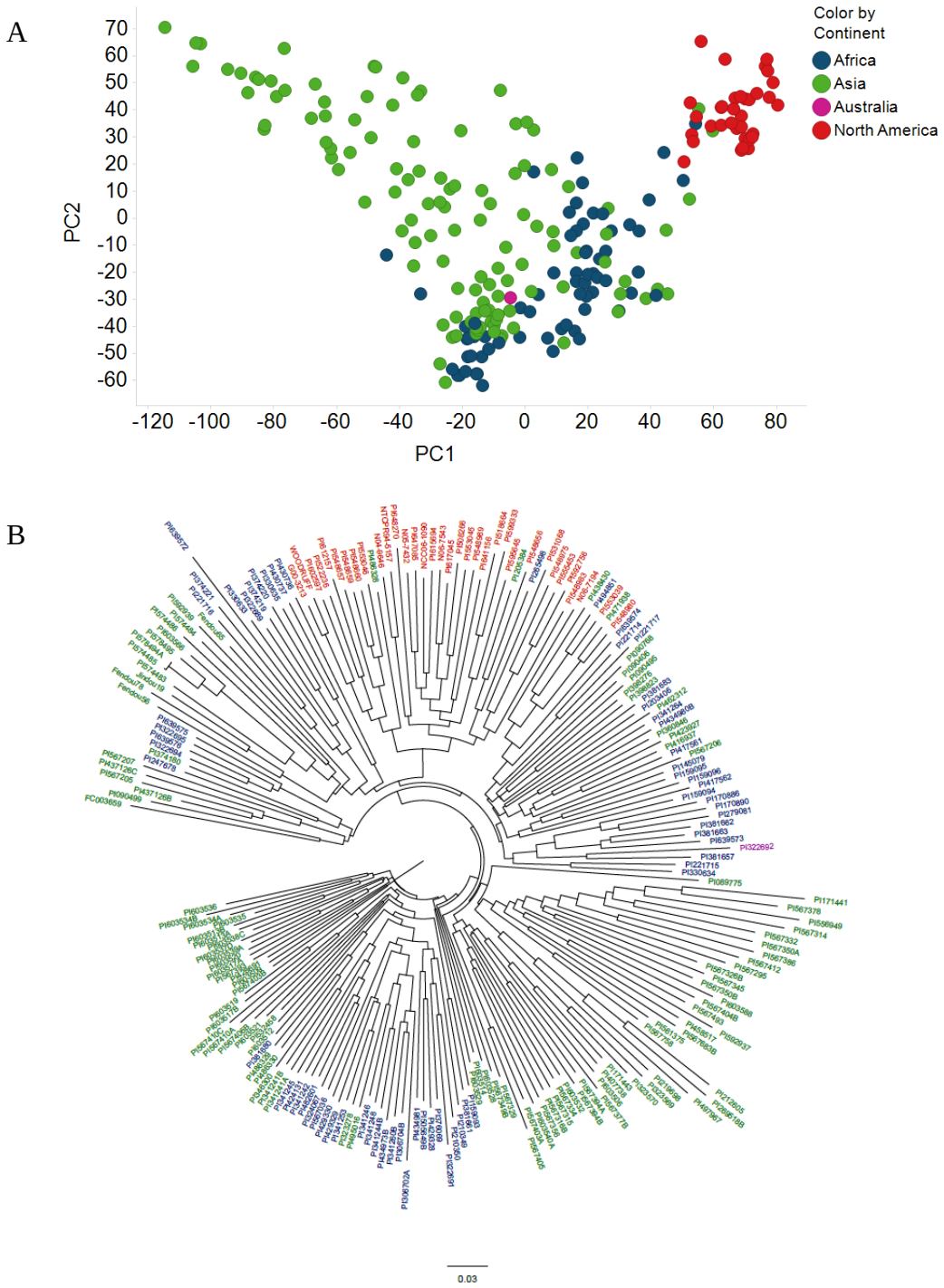


Figure 2.2. A) Plot of first and second principal coordinates for the diverse panel of soybean accessions evaluated. Each individual soybean genotype is colored by their continent of origin. B) Dendrogram using neighbor joining clustering algorithm in TASSEL visualized in FigTree for the association panel. Genotypes are colored by their continent of origin: red = North America, blue = Africa, green = Asia, and purple = Australia.

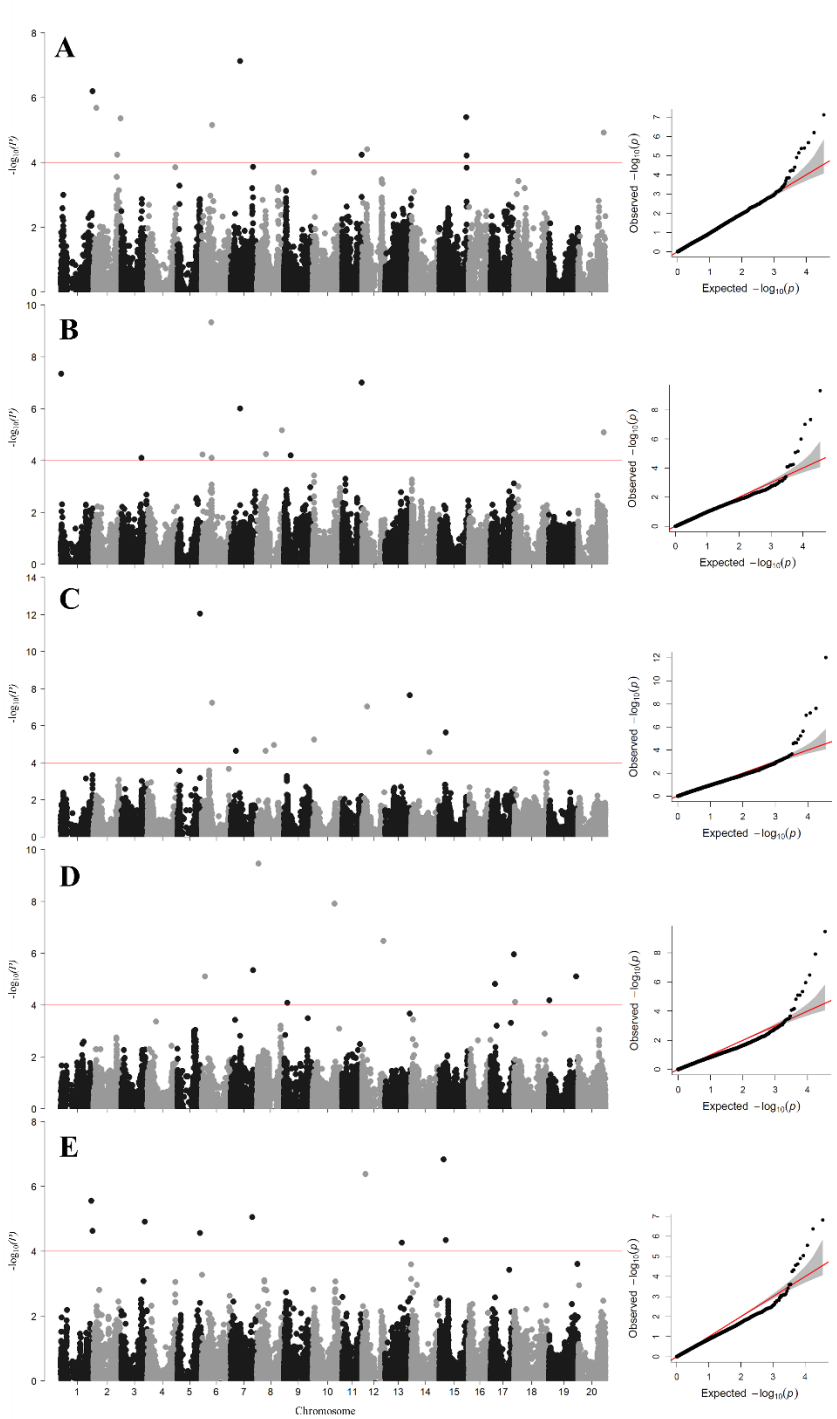


Figure 2.3. Genome-wide Manhattan plots for A) ALL-PANEL, B) GA-15-PANEL, C) GA-16-PANEL, D) KS-15-PANEL, and E) KS-16-PANEL. The X-axis is the genomic position of SNPs by chromosome across the soybean genome, and the Y-axis is the $-\log_{10}$ of the p-values obtained from the GWAS model. Significance threshold $-\log_{10}(P) > 4$ (red line). The quantile-quantile (QQ) plots to the right of each Manhattan plot show the expected versus observed p-values of each SNP tested in the GWAS models.

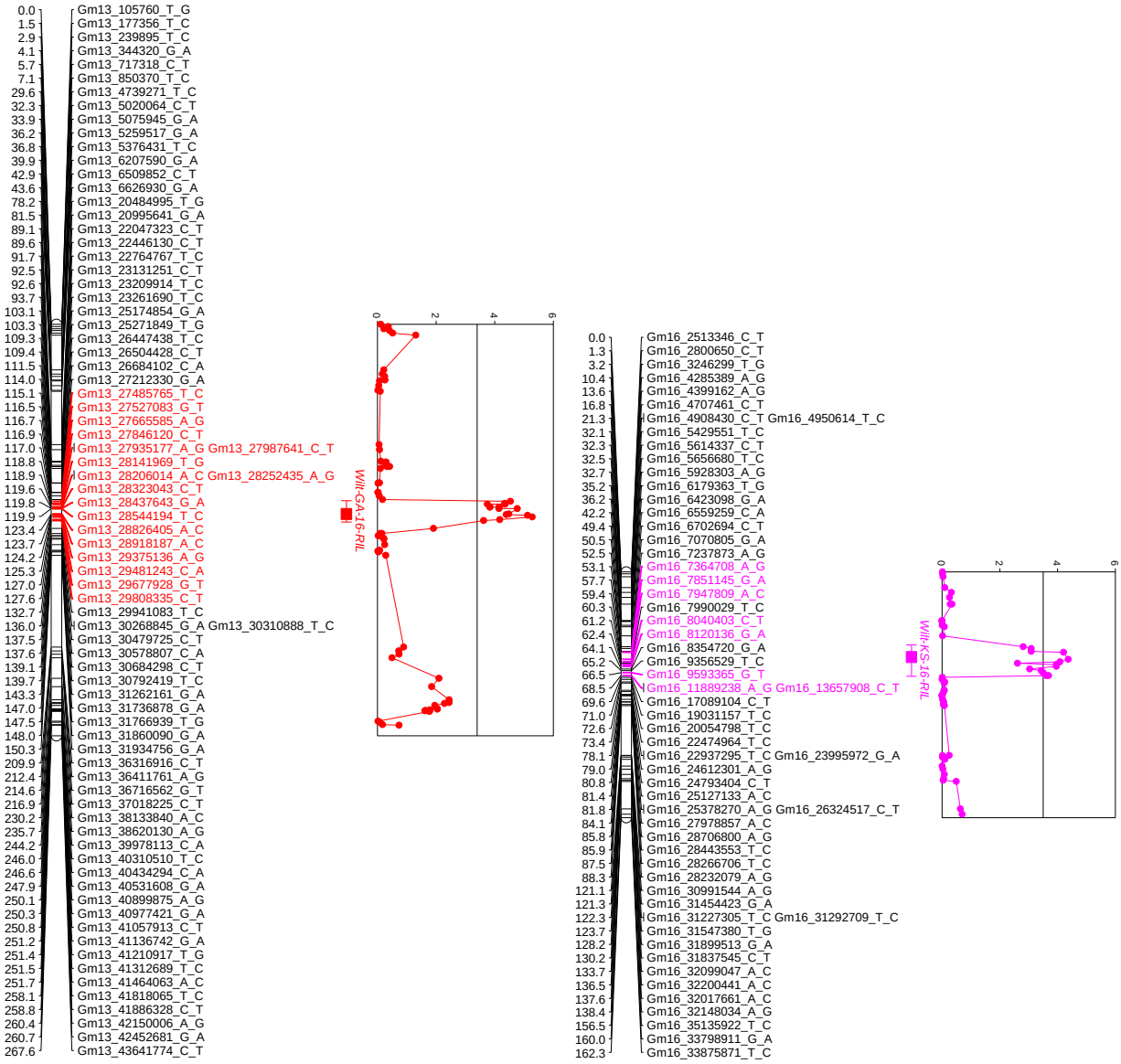


Figure 2.4. Composite interval mapping for canopy wilting in the RIL population derived from Hutcheson × PI 471938. Genetic maps with cM positions and graphs of mapping results for chromosomes with QTL meeting logarithm of odds (LOD) significance thresholds are shown for Chr 13 (GA-16-RIL) and Chr 16 (KS-16-RIL). Locus names colored red (GA-16-RIL; LOD = 3.4) and purple (KS-16-RIL; LOD = 3.5) met their respective genome-wide LOD thresholds.

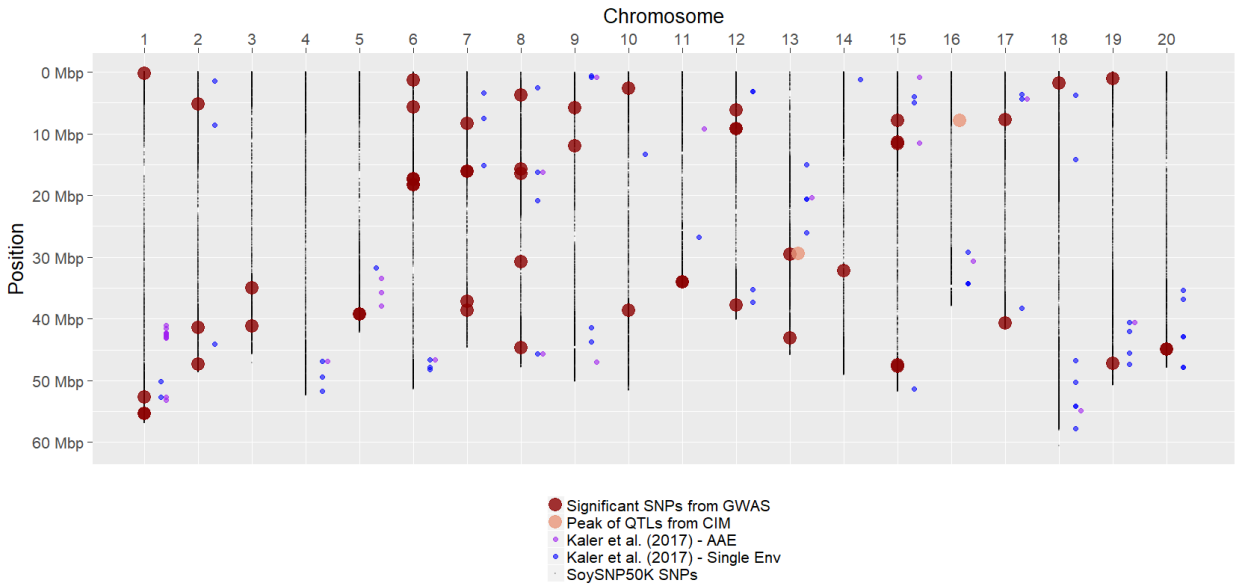


Figure 2.5. Location and comparison of SNPs significantly associated with canopy wilting based on association and linkage mapping results. Physical positions are based on the Glyma.Wm82.a2 version of the soybean genome. SNPs identified in GWAS that met $-\log_{10}(P) > 4$ significance threshold and SNPs with the highest LOD score from composite interval mapping (CIM) are shown as large red and salmon colored circles, respectively. Average of all environments (AAE) and single environment (Env) significant SNPs from Kaler et al. (2017b) are shown as purple and blue circles, respectively. Position locations were converted from version 1 to 2 of the soybean genome assembly for the Kaler et al. (2017b) SNPs, so that comparisons were made using the same physical positions. ss715637687 found in AAE for Kaler et al. (2017b) is not in version 2 of soybean genome assembly, and therefore not included in this comparison.

Table 2.1. Summary of analyses of variance (ANOVA) for effects of genotype (G), environment (E), and their interaction based on association panel and RIL population canopy wilting scores. The G × E MS was used as the denominator of the F Value for significance testing.

Association Panel				Hutcheson × PI 471938 RILs			
Source	DF	F Value	P > F	Source	DF	F Value	P > F
Genotype (G)	208	13.6	<0.0001	Genotype (G)	129	2.3	<0.0001
Environment (E)	3	827.1	<0.0001	Environment (E)	1	275.5	<0.0001
G × E	594	3.2	<0.0001	G × E	129	2.3	<0.0001

Table 2.2. Canopy wilting scores for the 10 genotypes with the lowest and highest scores based on mean ranking across environments for the association panel along with four check genotypes. Each environment was ranked individually, and the mean of those rankings was used to rank all of the 209 genotypes tested. Canopy wilting scores shown are the mean of all reps within each respective environment. A full table of all accessions tested and their canopy wilting scores is provided in the supplementary materials (Table 2.S1).

Accession	Name	Country[a]	MG[b]	Canopy Wilting Score					Rank	Beneficial Alleles[c]	Breeding Value[d]
				ALL-PANEL	GA-15-PANEL	GA-16-PANEL	KS-15-PANEL	KS-16-PANEL			
				Slow wilting							
PI603535	Hei zong huang dou	China	VIII	6	10	6	6	1	1	28	-13.96
PI603513A	Xiao niu mao huang	China	VIII	7	9	7	6	5	2	29	-12.20
PI603529	Hei huang dou	China	VIII	8	12	10	-	2	3	25	0.24
PI603513B	-	China	VIII	8	13	10	-	3	4	28	-4.23
PI458517	Xiao Wuyie	China	III	9	13	9	-	6	5	27	-1.47
Fendou65	-	China	IV	8	12	8	3	9	6	34	-31.79
PI603588	Jing si dou	China	V	8	8	10	7	6	6	25	-5.51
PI603534A	Da niu mao huang	China	VII	8	7	12	8	5	8	28	-10.29
PI219698	Kulat	Pakistan	VI	10	18	11	-	3	9	23	-4.16
PI532458	Ba yue bao	China	VIII	10	23	10	3	3	10	27	-13.00
				Checks							
PI471938	-	Nepal	V	14	16	15	6	18	68	29	-10.22
PI416937	Houjaku Kuwazu	Japan	VI	18	39	16	8	11	107	26	-4.66
PI518664	Hutcheson	United States	V	22	25	26	27	12	165	24	-2.10
PI595645	Benning	United States	VII	23	38	24	11	20	171	23	-2.35
				Fast wilting							
PI567377B	(Ba yue zha)	China	VI	34	63	32	13	26	200	17	31.75
PI159096	41S77	South Africa	VII	31	52	29	25	17	201	18	16.63
PI648270	Osage	United States	V	29	43	27	26	22	202	20	5.45
PI381663	Kakira 1	Uganda	VI	35	55	45	20	20	203	18	20.95
NCC06-1090	-	United States	VI	32	39	39	24	28	204	18	11.32
PI639573	-	Burundi	VIII	33	39	37	30	25	205	21	17.97
PI574486	Jin dou 13	China	III	36	58	26	-	24	206	21	8.20
PI599333	Musen	United States	VI	33	53	32	20	27	207	17	13.57
PI417562	54.S.30 DL/64/185	South Africa	VI	36	46	36	36	25	208	14	28.14
PI330635	-	South Africa	VII	39	53	40	-	25	209	14	24.02
Mean				17.8	26.9	18.4	12.7	12.6			
LSD ($\alpha = 0.05$)				5.3	12.9	6.8	10.5	6.1			

[a] Country of origin of the accession based on GRIN data.

[b] Maturity group.

[c] Number of alleles from all significant SNPs with an effect that reduces canopy wilting score.

[d] Breeding value determined by adding the allelic effects for all significant SNPs individually by environment, and then summing the breeding values across individual environments.

Table 2.3. SNPs that met significance level of $-\log_{10}(P) > 4$ for the GWAS of canopy wilting.

Locus[a]	Chr.[b]	Pos.[c]	SNP	$-\log_{10}(P)$	MAF[d]	Effect[e]	R^2 (%)	Env[f]
1	1	283533	ss715578894	7.33	0.05	-4.74	6.39	GA-15-PANEL
2	1	52705825	ss715580284	5.55	0.28	1.20	1.29	KS-16-PANEL
3	1	55375763	ss715580569	6.19	0.28	-1.09	12.25	ALL-PANEL*
	1	55375763	ss715580569	4.62	0.29	-1.20	14.50	KS-16-PANEL*
4	2	5262408	ss715583580	5.68	0.37	0.89	3.43	ALL-PANEL
5	2	41482648	ss715582774	4.23	0.06	-1.56	5.54	ALL-PANEL
6	2	47303895	ss715583400	5.36	0.39	0.82	0.45	ALL-PANEL
7	3	34961318	ss715585532	4.09	0.33	-1.49	0.37	GA-15-PANEL
8	3	41162031	ss715586200	4.90	0.10	-1.75	0.44	KS-16-PANEL
9	5	39242165	ss715592087	4.55	0.21	-1.19	2.13	KS-16-PANEL
	5	39247507	ss715592085	12.04	0.20	2.28	10.32	GA-16-PANEL
10	6	1270895	ss715592929	4.22	0.38	1.56	0.05	GA-15-PANEL
11	6	5698957	ss715595359	5.10	0.15	0.92	7.54	KS-15-PANEL
12	6	17266849	ss715593709	9.33	0.15	-3.56	10.79	GA-15-PANEL
13	6	17439802	ss715593748	4.09	0.13	3.79	3.18	GA-15-PANEL
14	6	18315510	ss715593787	5.15	0.36	-0.95	27.51	ALL-PANEL*
	6	18315510	ss715593787	7.21	0.37	-1.66	29.96	GA-16-PANEL*
15	7	8335280	ss715598799	4.62	0.11	-1.64	3.99	GA-16-PANEL
16	7	16111431	ss715596504	7.13	0.48	-1.01	15.91	ALL-PANEL*
	7	16111431	ss715596504	6.00	0.48	-1.78	15.18	GA-15-PANEL*
17	7	37177741	ss715597558	5.05	0.37	-1.06	5.77	KS-16-PANEL
18	7	38622942	ss715597793	5.34	0.19	0.92	4.96	KS-15-PANEL
19	8	3786749	ss715601594	9.45	0.09	2.06	10.76	KS-15-PANEL
20	8	15731880	ss715599729	4.63	0.13	-1.60	0.02	GA-16-PANEL
21	8	16529057	ss715599836	4.24	0.12	2.35	0.26	GA-15-PANEL
22	8	30788472	ss715601359	4.94	0.05	-2.33	0.18	GA-16-PANEL
23	8	44698680	ss715602239	5.16	0.20	-1.98	2.60	GA-15-PANEL
24	9	5849654	ss715605197	4.07	0.37	0.65	0.01	KS-15-PANEL
25	9	11970660	ss715603000	4.20	0.44	1.54	1.92	GA-15-PANEL
26	10	2691984	ss715606025	5.23	0.08	2.37	1.28	GA-16-PANEL
27	10	38659084	ss715606644	7.91	0.18	1.31	20.24	KS-15-PANEL
28	11	34070012	ss715610566	4.23	0.30	0.78	14.08	ALL-PANEL*
	11	34070012	ss715610566	7.01	0.30	2.36	16.51	GA-15-PANEL*
29	12	6218392	ss715613249	6.37	0.24	1.56	13.43	KS-16-PANEL
30	12	9209526	ss715613671	7.03	0.13	2.21	20.06	GA-16-PANEL
	12	9229120	ss715613672	4.40	0.22	0.92	18.44	ALL-PANEL
31	12	37774430	ss715612794	6.47	0.29	1.22	17.25	KS-15-PANEL
32	13	29512635	ss715614812	4.25	0.12	-1.37	4.31	KS-16-PANEL
33	13	43102597	ss715616411	7.63	0.10	-2.50	5.96	GA-16-PANEL

34	14	32173175	ss715618174	4.56	0.49	-1.03	0.15	GA-16-PANEL
35	15	7892897	ss715623086	6.82	0.18	1.71	4.27	KS-16-PANEL
36	15	11342488	ss715620317	5.63	0.13	-1.60	0.01	GA-16-PANEL
37	15	11599443	ss715620339	4.33	0.32	-1.09	0.26	KS-16-PANEL
38	15	47509927	ss715622189	5.39	0.20	-1.02	1.15	ALL-PANEL
39	15	47732356	ss715622244	4.21	0.31	-0.88	4.59	ALL-PANEL
40	17	7803078	ss715628209	4.81	0.33	-0.84	0.86	KS-15-PANEL
41	17	40688330	ss715627773	5.96	0.38	-0.75	0.62	KS-15-PANEL
42	18	1817213	ss715629399	4.10	0.14	-0.95	17.58	KS-15-PANEL
43	19	1055898	ss715633017	4.17	0.22	-0.71	1.77	KS-15-PANEL
44	19	47203967	ss715635643	5.09	0.28	0.75	1.25	KS-15-PANEL
45	20	44939180	ss715638586	4.91	0.33	-0.94	4.41	ALL-PANEL*
	20	44939180	ss715638586	5.08	0.33	-2.03	3.89	GA-15-PANEL*

[a] If multiple SNPs were identified in the same linkage disequilibrium (LD) block they were deemed part of the same locus (genomic region).

[b] Chromosome.

[c] Glyma.Wm82.a2 physical position.

[d] Minor allele frequency.

[e] Allelic effects were calculated by taking the difference in mean canopy wilting score between the two alleles at a particular SNP, and the direction, negative or positive, of the allelic effect estimates are relative to the alphabetical order of the nucleotides at each particular marker.

[f] Environment written as location-year-population.

* Same SNP identified in both a single environment and across all environments.

Table 2.4. QTLs for canopy wilting identified with composite interval mapping (CIM) for the Hutcheson × PI 471938 RIL population.

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]
Wilt-GA-16-RIL	13	Gm13_29481243_C_A	125.29	115.06-127.63	29,481,243	27,485,765-29,808,335	5.28	-0.98	12	GA-16
Wilt-KS-16-RIL	16	Gm16_7851145_G_A	57.67	53.1-68.47	7,851,145	7,364,708-13,657,908	4.37	-1.00	11	KS-16

[a] Chromosome

[b] Position in centiMorgans based on 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 written as location-year

Appendix A: Supplemental figures

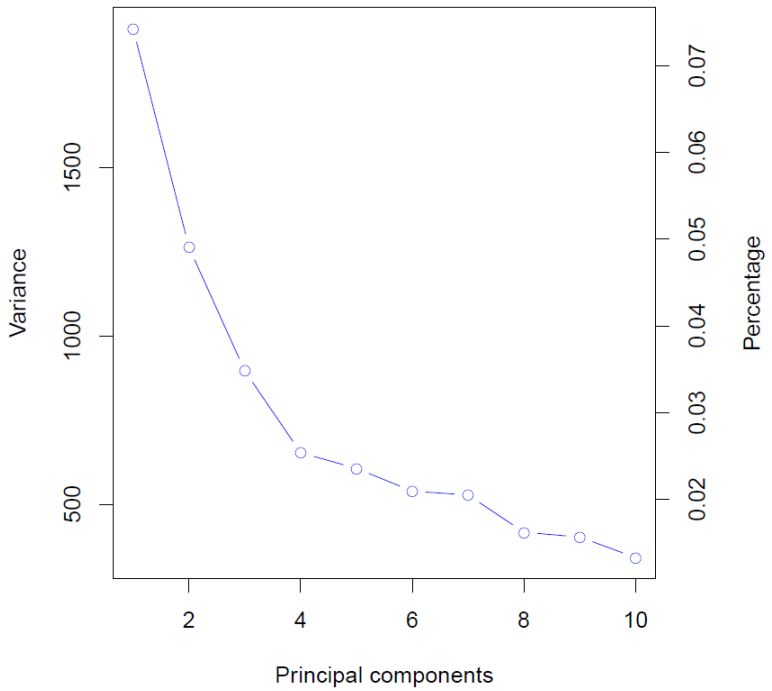


Figure 2.S1. Scree plot showing amount of genetic variation explained by the first 10 principal coordinates for the association panel. The amount of variation explained starts to level off at four coordinates, and therefore the first four coordinates were used as covariates in the GWAS models to help correct for potential population structure.

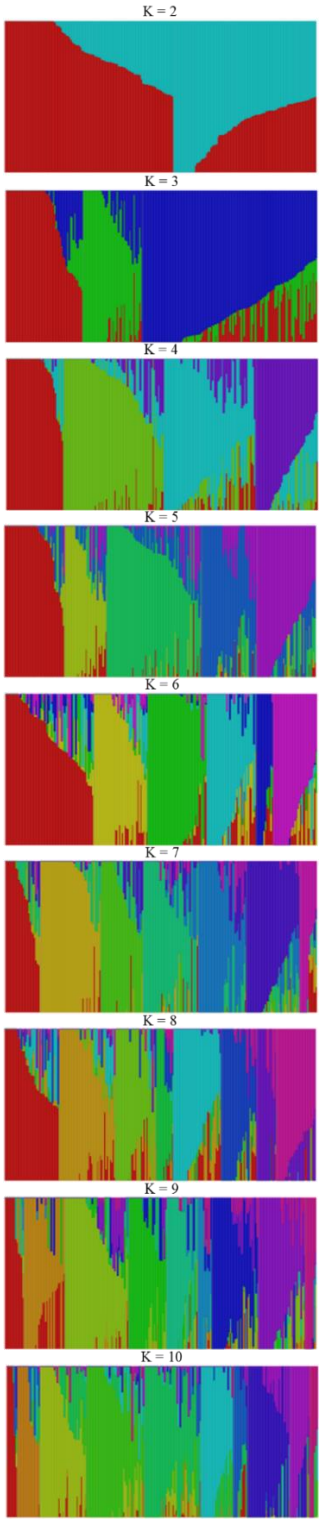
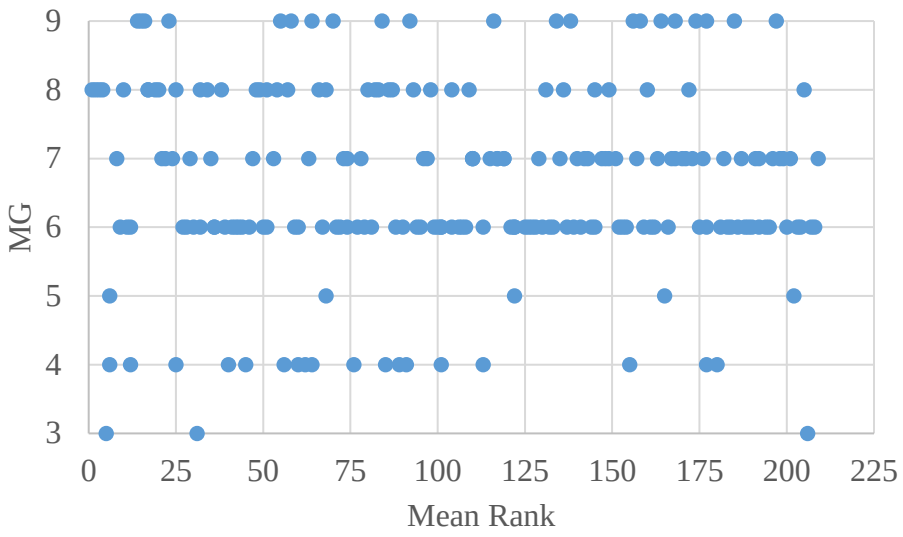


Figure 2.S2. Bayesian clustering (fastSTRUCTURE) results for K = 2 - 10 using simple option for the association panel. Color in vertical bars represents proportion individual belongs to each K group.

A



B

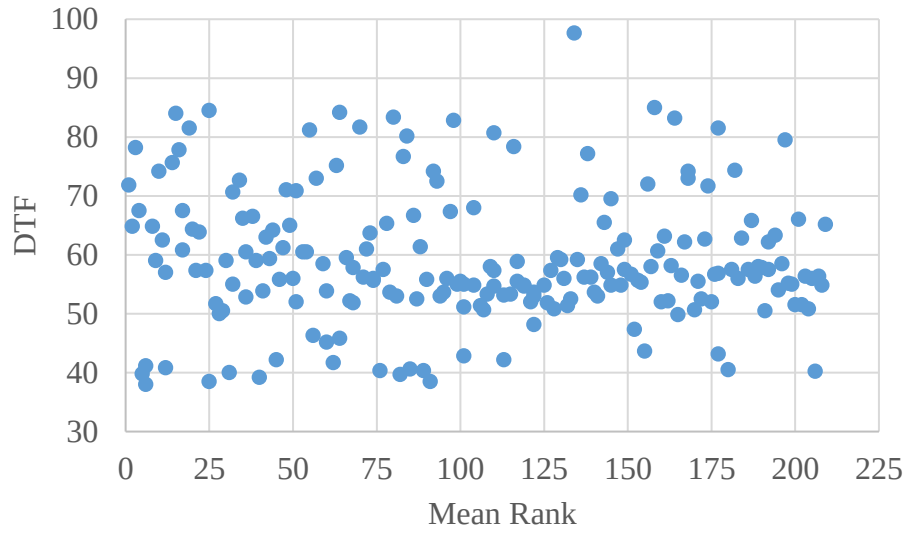


Figure 2.S3. Plots showing relationship of canopy wilting scores of the lines ranked across environments with A) maturity group (MG), and B) days to flowering (DTF) for the association panel.

Appendix B: Supplemental tables

Table 2.S1. Association panel canopy wilting scores, ranking, number of beneficial alleles, and breeding values. Each environment was ranked individually, and the mean of those rankings was used to rank all of the 209 genotypes tested. Canopy wilting scores shown are the mean of all reps within each respective environment.

				Canopy Wilting Score							
Accession	Name	Country[a]	MG[b]	ALL-PANEL	GA-15-PANEL	GA-16-PANEL	KS-15-PANEL	KS-16-PANEL	Rank	Beneficial Alleles[c]	Breeding Value[d]
PI603535	Hei zong huang dou	China	VIII	6	10	6	6	1	1	28	-13.96
PI603513A	Xiao niu mao huang	China	VIII	7	9	7	6	5	2	29	-12.20
PI603529	Hei huang dou	China	VIII	8	12	10	-	2	3	25	0.24
PI603513B	-	China	VIII	8	13	10	-	3	4	28	-4.23
PI458517	Xiao Wuyie	China	III	9	13	9	-	6	5	27	-1.47
Fendou65	-	China	IV	8	12	8	3	9	6	34	-31.79
PI603588	Jing si dou	China	V	8	8	10	7	6	6	25	-5.51
PI603534A	Da niu mao huang	China	VII	8	7	12	8	5	8	28	-10.29
PI219698	Kulat	Pakistan	VI	10	18	11	-	3	9	23	-4.16
PI532458	Ba yue bao	China	VIII	10	23	10	3	3	10	27	-13.00
PI269518B	(Koolat)	Pakistan	VI	11	21	9	-	3	11	27	-3.19
PI567405	Wei zi dou	China	VI	9	10	12	7	6	12	28	-6.02
PI574484	Jin dou No. 6	China	IV	9	12	7	10	5	12	30	-24.95
PI567036	-	Morocco	IX	11	23	10	4	6	14	31	-11.65
PI341241B	(Seminole)	Israel	IX	10	16	9	6	8	15	27	-0.41
PI341248	Sangalo	Tanzania	IX	10	23	12	4	2	16	30	-11.11
PI603521	Huang dou	China	VIII	10	18	8	9	5	17	29	-5.45
PI603534B	(Da niu mao huang)	China	VIII	9	13	13	8	3	17	27	-16.22
PI429328	-	Nigeria	VIII	10	24	12	4	1	19	29	-12.32
PI603536	Hui huang dou	China	VIII	10	18	10	8	4	20	30	-12.72
PI603537D	(Niu yan jing quan zi)	China	VII	10	16	16	3	5	21	27	-2.53
PI603528	Hei ke zha	China	VII	9	12	10	11	5	22	32	-15.09
PI434973B	(Malayan)	Nigeria	IX	14	23	-	5	-	23	26	2.84
PI567315	Hong huang dou	China	VII	9	9	16	6	6	24	29	-11.29
Jindou19	-	China	IV	9	9	6	13	6	25	23	-23.32
PI210349	Jubiltan 65	Mozambique	VIII	12	25	9	-	3	25	28	-5.02
PI567326B	(Huang dou)	China	VI	11	18	13	7	6	27	27	-8.36
PI567316B	(Hong huang dou)	China	VI	10	8	16	-	6	28	27	-9.34
PI497967	-	India	VII	12	19	14	-	4	29	26	4.61
PI212605	-	Afghanistan	VI	11	18	15	8	4	30	26	1.08
PI603566	Jin dou No. 4	China	III	11	18	6	-	11	31	29	-12.56
PI434981	Indo 226	Central African Republic	VIII	13	26	11	-	2	32	24	-3.98

PI567349B	(Shu pi huang dou)	China	VI	10	11	15	7	9	32	25	-4.62
PI486330	Macs-75	India	VIII	12	24	12	8	3	34	28	-8.83
PI323570	-	India	VII	11	19	11	8	8	35	26	-1.51
PI567356	Zao bai huang dou	China	VI	10	9	11	12	8	36	31	-20.58
PI567394B	-	China	VI	10	8	19	6	7	36	29	-3.73
PI567295	Bian huang dou	China	VIII	11	18	17	7	4	38	25	5.32
PI567345	Niu mao huang	China	VI	10	9	14	9	7	39	26	-3.77
PI592939	Jin dou 16	China	IV	9	2	7	25	4	40	29	-19.12
PI567314	Hei you huang dou	China	VI	12	16	15	-	6	41	24	3.61
PI603517B	-	China	VI	11	17	15	8	6	42	28	-10.01
PI567406B	(Wu se da dou)	China	VI	10	6	14	-	10	43	30	-9.95
PI603512	Jin man dou	China	VI	12	17	17	5	8	44	25	-5.05
PI578494A	Jin dou No. 1	China	IV	10	3	9	23	5	45	28	-8.04
NTCPR94-5157	-	United States	VI	12	22	10	7	11	46	33	-18.98
PI603540A	Hei huang dou	China	VII	12	16	15	8	8	47	28	-7.74
PI374180	-	India	VIII	13	24	11	5	11	48	29	-7.86
PI486329	-	India	VIII	13	24	10	9	8	49	27	-5.42
PI603517A	Lao shu pi	China	VI	13	20	19	7	5	50	28	-11.66
PI407738	-	China	VI	12	15	19	7	7	51	25	1.68
PI486328	Birsa Soybean-1	India	VIII	13	25	9	6	12	51	27	-9.09
PI330634	-	South Africa	VII	13	22	13	9	7	53	26	-6.63
PI603509	Huang dou	China	VIII	13	28	12	10	2	54	26	-3.60
PI495016	Nuwara Eliya Local	Sri Lanka	IX	14	30	9	7	9	55	24	4.22
PI567758	Pei xian tu shan da ping ding huang	China	IV	12	18	4	15	10	56	27	-8.03
PI381661	Bukalasa 6	Uganda	VIII	14	33	15	5	3	57	23	3.86
PI341260B	(HLS 219)	Tanzania	IX	16	23	-	8	-	58	24	3.87
PI567378	Ba yue zha	China	VI	12	10	17	5	15	59	23	2.65
PI171441	Mud-bean	China	VI	14	19	17	-	7	60	20	21.02
PI556949	Ke feng No. 1	China	IV	14	17	13	-	12	60	23	14.74
PI578495	Jin dou No. 4	China	IV	13	16	8	19	9	62	28	-9.84
PI381680	S7	Uganda	VII	14	25	11	12	8	63	25	-3.16
PI341253	CMS	Sudan	IX	14	24	18	9	6	64	26	0.04
PI592937	Jin dou 14	China	IV	14	24	17	3	12	64	22	8.40
N05-7432	-	United States	VIII	14	23	18	10	7	66	29	-18.28
PI567394A	Jiu yue han	China	VI	15	18	19	-	8	67	27	-0.62
N06-7194	-	United States	VIII	14	23	15	9	9	68	26	-10.67
PI471938	-	Nepal	V	14	16	15	6	18	68	29	-10.22
PI341244B	(Yellow Kedele)	Tanzania	IX	16	36	13	8	6	70	27	-9.30
PI603520	Huang dou	China	VI	15	26	12	15	6	71	24	0.88
PI567334	Jiang dou zi	China	VI	14	20	20	5	11	72	25	0.27
PI567393	Jiu yue han	China	VII	15	22	23	5	9	73	22	8.80
N04-9646	-	United States	VII	15	28	11	17	3	74	25	-3.35
PI603519	Lu da dou	China	VI	13	14	15	18	7	74	27	-10.15

Fendou78	-	China	IV	13	8	8	22	14	76	26	-16.79
PI603506	Xiao ke zao huang dou	China	VI	15	23	20	8	9	77	27	-2.38
PI346300	-	India	VII	15	30	10	15	6	78	27	-10.88
PI567350B	(Shu pi huang dou)	China	VI	15	20	19	6	13	79	24	2.75
PI429330	-	Nigeria	VIII	16	25	13	-	10	80	28	-2.01
PI548983	Tracy	United States	VI	15	18	19	12	10	81	30	-18.05
PI639572	-	Ghana	VIII	14	16	11	13	18	82	26	-8.39
PI462312	Ankur	India	VIII	15	26	13	7	17	83	27	-1.71
PI341246	CNS	Tanzania	IX	15	24	13	13	10	84	23	9.99
PI574485	Jin dou No. 9	China	IV	14	11	14	19	11	85	28	-8.04
PI428691	-	India	VIII	16	23	11	11	17	86	21	3.07
PI639576	-	Burundi	VIII	16	28	14	10	11	87	25	-3.75
PI567404B	(Wang shan hou)	China	VI	17	19	24	-	7	88	21	18.53
PI574483	Jin dou No. 5	China	IV	17	19	10	29	11	89	25	-6.55
PI567683B	(Zheng zhou niu yao qi)	China	VI	17	28	23	7	9	90	22	5.54
Fendou56	-	China	IV	15	18	8	18	18	91	25	-18.17
PI322692	Max C.P1159A8	Australia	IX	17	23	16	6	23	92	26	-7.24
PI247678	Herman	Zaire	VIII	16	31	13	9	12	93	22	3.83
PI089775	-	China	VI	16	15	19	-	15	94	24	7.92
PI567412	Yi wo feng	China	VI	17	20	24	-	8	95	23	17.64
PI548989	Ransom	United States	VII	16	26	13	13	13	96	29	-15.00
PI323569	-	India	VII	20	39	13	-	7	97	23	-0.62
PI203406	-	South Africa	VIII	18	26	14	5	27	98	24	1.46
PI171443	Tea-bean	China	VI	17	21	24	15	7	99	24	4.20
PI567329	Huang huang dou	China	VI	17	28	17	6	18	100	26	-2.71
PI423927	Tousan 93	Japan	IV	17	29	14	13	11	101	24	-6.88
PI548980	Hood	United States	VI	17	23	25	10	11	101	31	-22.91
PI603514	Ni ba dou	China	VI	17	26	20	15	7	101	24	-0.88
PI265498	-	Zaire	VIII	19	36	10	8	22	104	22	2.37
PI548656	Lee	United States	VI	17	23	13	16	15	104	26	-6.28
PI567205	GL2671/89	Georgia	VI	17	23	22	15	8	106	21	6.71
PI416937	Houjaku Kuwazu	Japan	VI	18	39	16	8	11	107	26	-4.66
PI090406	-	China	VI	18	23	26	12	11	108	24	2.47
PI553045	Cook	United States	VIII	18	24	12	13	22	109	28	-11.27
PI221715	-	South Africa	VII	17	23	17	16	12	110	25	-4.14
PI429329	-	Nigeria	VII	19	38	18	5	15	110	24	3.39
PI567410C	-	China	VII	18	33	24	7	10	110	25	0.18
PI398823	-	South Korea	IV	18	33	13	14	12	113	27	-11.68
PI567386	Huang da dou (1)	China	VI	19	31	27	8	9	113	23	12.41
PI615694	N7001	United States	VII	17	19	24	13	13	115	28	-12.12
PI322691	Jubiltan 109	Mozambique	IX	19	29	18	4	26	116	22	8.11
PI145079	Hernon No. 6	Zimbabwe	VII	18	23	19	19	10	117	22	7.89
WOODRUFF	-	United States	VII	18	22	21	13	16	117	28	-14.49

G00-3213	-	United States	VII	18	24	19	15	13	119	26	-9.02
PI567403A	Shuan huang dou	China	VII	18	32	20	16	6	119	23	5.46
PI090495	-	China	VI	18	28	25	9	12	121	24	2.24
PI561375	Qi huang No. 1	China	V	19	38	13	9	17	122	19	12.63
PI567332	Huo huang dou	China	VI	19	33	22	13	10	122	20	14.70
PI603532	Hong li huang dou	China	VI	19	29	22	13	11	122	24	5.94
PI322694	Hernnon	Zimbabwe	VI	19	34	20	14	9	125	29	-7.62
FC003659	Da Wu Don	China	VI	19	23	27	11	15	126	24	-1.19
PI437126B	-	Georgia	VI	19	28	21	11	15	127	23	2.38
PI090499	Black and white	China	VI	19	24	23	9	19	128	21	4.88
PI553046	Gasoy 17	United States	VII	20	38	13	13	16	129	24	-3.23
PI170890	-	South Africa	VI	19	28	23	13	12	130	25	1.17
PI603538C	(Wan dou zao)	China	VIII	19	22	23	13	18	131	25	-0.43
PI417561	48.S.103 DL/63/180	South Africa	VI	20	18	25	28	10	132	24	-4.02
PI205384	-	Pakistan	VI	19	32	15	15	15	133	24	-1.75
PI210350	Jubiltan 67	Mozambique	IX	22	43	14	7	23	134	20	11.61
PI522236	Thomas	United States	VII	22	47	10	13	16	135	24	-1.54
PI381657	3H55 F4/9/2	Uganda	VIII	19	31	21	12	13	136	24	5.63
PI437126C	-	Georgia	VI	20	29	19	21	10	137	22	0.82
PI341241A	Seminole	Israel	IX	19	28	15	16	18	138	24	0.57
PI322695	Bicolor do Cuima	Angola	VI	22	46	18	8	16	139	26	2.01
PI647085	N7002	United States	VII	20	30	20	11	17	140	30	-15.42
PI603539A	Huang dou	China	VI	19	28	20	17	12	141	23	-0.51
PI159093	34S51	South Africa	VII	19	30	17	16	14	142	20	9.74
PI159095	41S31	South Africa	VII	20	38	20	11	12	143	24	3.60
PI508266	Young	United States	VI	20	30	21	13	15	144	30	-9.20
PI376069	DRO 9	Cameroon	VIII	22	38	11	13	27	145	20	11.98
PI567207	-	Georgia	VI	20	28	25	12	15	145	21	1.96
PI221716	-	South Africa	VII	20	30	19	19	11	147	25	1.36
PI531068	Stonewall	United States	VII	22	44	16	7	22	148	25	-4.81
PI567410A	Yang huang dou	China	VII	20	20	30	17	13	149	22	10.06
PI612157	Prichard	United States	VIII	21	25	14	25	18	149	22	4.22
N06-7543	-	United States	VII	20	33	22	17	8	151	23	-6.11
PI567206	GL2674/90	Georgia	VI	20	27	24	15	14	152	25	1.06
PI602597	Boggs	United States	VI	21	38	15	15	17	153	22	3.50
PI548975	Centennial	United States	VI	21	35	22	9	17	154	26	-6.59
PI567493	Huang dou	China	IV	20	32	15	16	18	155	18	14.31
PI341242	Hernon 247	Tanzania	IX	21	28	28	8	23	156	27	-0.74
PI548657	Jackson	United States	VII	21	36	22	10	17	157	25	-3.81
PI306702A	3H/1	Kenya	IX	21	37	12	16	21	158	23	1.93
PI221714	-	South Africa	VI	22	33	31	14	10	159	21	8.91
PI639575	-	Burundi	VIII	22	37	17	22	13	160	23	2.65
PI374220	Geduld	South Africa	VI	21	34	22	20	10	161	23	9.60

PI381683	S36	Uganda	VI	21	28	26	14	17	162	22	5.04
PI548659	Braxton	United States	VII	21	31	19	15	19	163	23	-4.15
PI323278	K-30	Pakistan	IX	23	45	20	8	19	164	24	7.95
PI518664	Hutcheson	United States	V	22	25	26	27	12	165	24	-2.10
PI341264	-	Liberia	VI	21	26	22	17	21	166	26	-4.71
PI567403B	-	China	VII	23	37	25	12	17	167	22	8.10
PI322689	Improved	Angola	VII	22	36	24	15	14	168	21	10.73
PI482601	-	Zimbabwe	IX	22	29	23	15	20	168	23	9.37
PI438430	-	Israel	VII	22	27	29	17	16	170	25	-7.59
PI595645	Benning	United States	VII	23	38	24	11	20	171	23	-2.35
PI639574	-	Burundi	VIII	24	37	30	14	13	172	23	0.64
PI555453	Hagood	United States	VII	23	32	17	21	21	173	25	-1.30
PI434980B	(Indo 180)	Central African Republic	IX	25	35	23	11	30	174	23	1.51
PI592756	Dillon	United States	VI	22	30	26	16	18	175	25	-1.30
PI641156	NC-Raleigh	United States	VII	24	43	17	17	19	176	23	-2.89
PI090768	-	China	VI	25	41	30	14	14	177	21	8.93
PI341245	Avoyelles	Tanzania	IX	24	38	30	14	16	177	25	4.45
PI398276	Chirpan 90 (Bulgaria)	South Korea	IV	25	43	15	18	23	177	24	2.05
PI360846	Shiroge-9	Japan	IV	25	33	18	26	22	180	24	-4.89
PI430736	Kudu	Zimbabwe	VI	24	38	20	23	17	181	22	6.49
PI324067	Hernon 237	Zimbabwe	VII	26	43	30	13	17	182	21	14.05
PI494851	-	Zambia	VI	25	41	27	13	20	183	23	7.60
PI567350A	Shu pi huang dou	China	VI	28	52	24	10	26	184	19	24.26
PI306704B	(7H/101)	Kenya	IX	28	40	-	15	-	185	19	15.65
PI170886	-	South Africa	VI	28	37	35	-	13	186	22	9.50
PI159094	35S377	South Africa	VII	27	43	34	10	21	187	19	16.48
PI221717	-	South Africa	VI	25	41	29	18	14	188	23	4.10
PI374221	Welkom	South Africa	VI	27	34	29	29	15	189	19	16.13
PI553039	Davis	United States	VI	25	40	24	15	21	190	25	0.65
PI279081	Masterpiece	South Africa	VII	30	34	47	26	14	191	18	21.49
PI374219	Blyvoor	South Africa	VI	29	48	33	25	11	192	20	12.92
PI548660	Bragg	United States	VII	26	44	19	22	21	192	22	-1.02
PI381662	Hernon 49	Uganda	VI	27	38	32	14	26	194	22	10.16
PI617045	NC-Roy	United States	VI	26	37	31	19	19	195	22	-3.06
PI330633	-	South Africa	VII	26	38	28	18	20	196	20	9.57
PI505649B	-	Zambia	IX	28	31	26	24	30	197	15	22.04
PI424131	Buffalo	Zimbabwe	VII	27	36	26	23	25	198	22	14.36
PI430737	Oribi	Zimbabwe	VII	31	49	27	-	16	199	18	21.13
PI567377B	(Ba yue zha)	China	VI	34	63	32	13	26	200	17	31.75
PI159096	41S77	South Africa	VII	31	52	29	25	17	201	18	16.63
PI648270	Osage	United States	V	29	43	27	26	22	202	20	5.45
PI381663	Kakira 1	Uganda	VI	35	55	45	20	20	203	18	20.95
NCC06-1090	-	United States	VI	32	39	39	24	28	204	18	11.32
PI639573	-	Burundi	VIII	33	39	37	30	25	205	21	17.97
PI574486	Jin dou 13	China	III	36	58	26	-	24	206	21	8.20
PI599333	Musen	United States	VI	33	53	32	20	27	207	17	13.57

PI417562	54.S.30 DL/64/185	South Africa	VI	36	46	36	36	25	208	14	28.14
PI330635	-	South Africa	VII	39	53	40	-	25	209	14	24.02
Mean				17.8	26.9	18.4	12.7	12.6			
LSD ($\alpha = 0.05$)				5.3	12.9	6.8	10.5	6.1			

[a] Country of origin of the accession based on GRIN data.

[b] Maturity group.

[c] Number of alleles from all significant SNPs with an effect that reduces canopy wilting score.

[d] Breeding value determined by adding the allelic effects for all significant SNPs individually by environment, and then summing the breeding values across individual environments.

Table 2.S2. Hutcheson × PI 471938 RIL population canopy wilting scores and ranking. Each environment was ranked individually, and the mean of those rankings was used to rank all of the RILs tested. Canopy wilting scores shown are the mean of all reps within each respective environment.

	Canopy Wilting Score		
Name	GA-16	KS-16	Rank
N98-7291	14.8	7.5	1
PI471938	14.6	10.0	2
N98-7276	17.9	7.5	3
N98-7215	16.4	10.0	4
N98-7308	18.8	7.5	5
N98-7266	19.0	5.0	6
N98-7335	15.7	12.5	7
N98-7285	16.2	12.5	8
N98-7324	16.9	12.5	9
N98-7194	20.2	5.0	10
N98-7295	18.6	12.5	11
N98-7273	18.8	12.5	12
N98-7326	18.8	12.5	12
N98-7256	20.0	10.0	14
N98-7274	14.8	15.0	15
N98-7281	20.5	10.0	16
N98-7238	19.3	12.5	16
N98-7275	19.3	12.5	18
N98-7198	16.0	15.0	19
N98-7243	20.0	12.5	20
N98-7321	21.7	10.0	21
N98-7245	20.7	12.5	22
N98-7261	21.0	12.5	23
N98-7296	21.2	12.5	24
N98-7262	22.4	10.0	25
N98-7193	18.8	15.0	26
N98-7216	21.7	12.5	27
N98-7240	21.7	12.5	27
N98-7190	15.2	17.5	29
N98-7180	24.0	7.5	30
N98-7271	24.0	7.5	30
N98-7233	21.9	12.5	32
N98-7332	24.3	7.5	33
N98-7265	16.9	17.5	33
N98-7267	19.5	15.0	33

N98-7283	23.6	10.0	36
N98-7272	17.4	17.5	37
N98-7293	20.0	15.0	38
N98-7177	24.8	7.5	39
N98-7249	24.0	10.0	39
N98-7287	23.1	12.5	41
N98-7337	24.8	10.0	42
N98-7309	15.0	20.0	42
N98-7232	21.4	15.0	44
N98-7255	16.7	20.0	45
N98-7214	17.1	20.0	46
N98-7278	24.8	12.5	47
N98-7306	19.8	17.5	47
N98-7185	17.9	20.0	49
N98-7327	17.9	20.0	49
N98-7317	22.1	15.0	51
N98-7223	22.4	15.0	52
N98-7292	22.4	15.0	52
N98-7318	18.1	20.0	52
N98-7209	25.0	12.5	55
N98-7328	18.3	20.0	55
N98-7188	27.4	10.0	57
N98-7225	22.9	15.0	58
N98-7310	22.9	15.0	58
N98-7331	29.3	7.5	60
N98-7289	28.3	10.0	61
N98-7218	26.2	12.5	62
N98-7182	19.5	20.0	63
N98-7323	23.8	15.0	64
N98-7248	19.8	20.0	65
N98-7181	28.1	12.5	66
N98-7189	28.1	12.5	66
N98-7219	28.1	12.5	66
N98-7270	30.7	10.0	69
N98-7192	28.3	12.5	69
N98-7196	18.8	22.5	69
N98-7259	17.6	25.0	72
N98-7206	25.0	15.0	73
N98-7264	22.9	17.5	73
N98-7234	21.2	20.0	73
N98-7319	21.2	20.0	73
N98-7174	31.4	10.0	77
N98-7226	29.3	12.5	77

N98-7330	18.8	25.0	79
N98-7320	21.7	20.0	80
N98-7222	25.5	15.0	81
N98-7171	20.2	22.5	82
Hutcheson	24.5	17.5	83
N98-7280	24.8	17.5	84
N98-7247	20.5	22.5	84
N98-7229	31.4	12.5	84
N98-7211	31.9	12.5	87
N98-7311	19.3	27.5	88
N98-7322	22.6	20.0	88
N98-7298	20.0	25.0	88
N98-7334	25.0	17.5	91
N98-7251	28.3	15.0	91
N98-7299	28.3	15.0	91
N98-7210	25.2	17.5	94
N98-7224	25.2	17.5	94
N98-7288	25.2	17.5	94
N98-7313	21.7	22.5	94
N98-7230	28.6	15.0	98
N98-7231	23.6	20.0	99
N98-7254	29.3	15.0	100
N98-7173	25.7	17.5	101
N98-7252	25.7	17.5	101
N98-7236	21.4	25.0	103
N98-7239	27.1	17.5	104
N98-7244	30.7	15.0	104
N98-7250	27.4	17.5	106
N98-7208	24.8	20.0	107
N98-7307	31.2	15.0	107
N98-7294	31.7	15.0	109
N98-7221	22.4	25.0	109
N98-7235	23.3	22.5	111
N98-7316	21.9	27.5	112
N98-7178	29.3	17.5	113
N98-7279	29.3	17.5	113
N98-7290	22.1	30.0	115
N98-7220	26.4	20.0	116
N98-7314	29.5	17.5	117
N98-7199	24.8	22.5	117
N98-7227	27.6	20.0	119
N98-7213	23.3	27.5	120
N98-7228	31.0	17.5	120

N98-7263	22.6	32.5	122
N98-7315	23.8	27.5	123
N98-7286	32.9	17.5	124
N98-7195	29.3	20.0	125
N98-7329	27.6	22.5	126
N98-7325	31.2	20.0	127
N98-7297	26.7	25.0	128
N98-7300	26.0	27.5	129
N98-7246	30.7	22.5	130
N98-7312	28.6	30.0	131
N98-7277	33.6	22.5	131
Mean	23.3	16.3	
LSD ($\alpha = 0.05$)	8.2	10.9	

Table 2.S3. Candidate genes and their functional annotation identified using the Glyma2.1 gene models in SoyBase (www.soybase.org) within plus or minus 10 kb of SNPs significantly associated with canopy wilting from GWAS.

Locus	SNP	Gene Name	Annotation	Gene Ontology Terms
1	ss715578894	Glyma.01g001700	3-phosphoinositide-dependent protein kinase-1	Unknown
		Glyma.01g001800	Leucine-rich repeat protein kinase family protein	Protein kinase activity; protein binding; ATP binding; protein phosphorylation
		Glyma.01g001900	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain	Protein kinase activity; ATP binding; protein phosphorylation; response to stress
		Glyma.01g002000	Unknown protein	Unknown
		Glyma.01g002100	Transcriptional factor B3 family protein / auxin-responsive factor AUX/IAA-related	DNA binding; nucleus; regulation of transcription, DNA-templated; response to hormone
2	ss715580284	Glyma.01g192600	LOB domain-containing protein 38	Unknown
3	ss715580569	Glyma.01g225000	Integrase-type DNA-binding superfamily protein	Sequence-specific DNA binding transcription factor activity; regulation of transcription, DNA-templated
		Glyma.01g225100	Highly ABA-induced PP2C gene 3	Catalytic activity
4	ss715583580	Glyma.02g058600	Protein phosphatase 2C family protein	Catalytic activity
		Glyma.02g058700	GRAS family transcription factor	Unknown
5	ss715582774	Glyma.02g227800	No annotation	Unknown
		Glyma.02g227900	PHD finger transcription factor	Protein binding
		Glyma.02g228000	Unknown protein	Unknown
6	ss715583400	Glyma.02g294800	F-box/RNI-like superfamily protein	Protein binding
		Glyma.02g294900	Trigger factor type chaperone family protein	Protein folding; protein transport
		Glyma.02g295000	Phosphatidate cytidylyltransferase family protein	Membrane; transferase activity, transferring phosphorus-containing groups
7	ss715585532	Glyma.03g134300	Mitochondrial glycoprotein family protein	Mitochondrial matrix
		Glyma.03g134400	No annotation	Unknown
8	ss715586200	Glyma.03g203000	Nuclear factor Y, subunit A10	Sequence-specific DNA binding transcription factor activity; regulation of transcription, DNA-templated
		Glyma.03g203100	No annotation	Unknown
		Glyma.03g203200	No annotation	Unknown
		Glyma.03g203300	AGAMOUS-like 92	DNA binding; protein dimerization activity
9	ss715592087	Glyma.05g210400	Unknown protein	Membrane; transferase activity, transferring glycosyl groups
		Glyma.05g210500	No annotation	Unknown

		Glyma.05g210600	Zinc finger (CCCH-type/C3HC4-type RING finger) family protein	Metal ion binding
		Glyma.05g210700	No annotation	Unknown
9	ss715592085	Glyma.05g210400	Unknown protein	Membrane; transferase activity, transferring glycosyl groups
		Glyma.05g210500	No annotation	Unknown
		Glyma.05g210600	Zinc finger (CCCH-type/C3HC4-type RING finger) family protein	Metal ion binding
		Glyma.05g210700	No annotation	Unknown
10	ss715592929	Glyma.06g016900	12-oxophytodienoate reductase 1	FMN binding; oxidoreductase activity; oxidation-reduction process
		Glyma.06g017000	FMN-linked oxidoreductases superfamily protein	FMN binding; oxidoreductase activity; oxidation-reduction process
		Glyma.06g017100	High mobility group B2	Unknown
		Glyma.06g017200	No annotation	Unknown
		Glyma.06g017300	DNA-binding bromodomain-containing protein	Protein binding
11	ss715595359	Glyma.06g073800	No annotation	Unknown
		Glyma.06g073900	Syntaxin of plants 32	Protein binding; membrane
		Glyma.06g074000	Syntaxin of plants 32	Protein binding; membrane
12	ss715593709	Glyma.06g193700	Mini zinc finger	Unknown
13	ss715593748	Glyma.06g195100	Atypical CYS HIS rich thioredoxin 4	Cell redox homeostasis
14	ss715593787	Glyma.06g200000	Glucuronidase 2	Membrane; hydrolase activity, acting on glycosyl bonds
		Glyma.06g200100	Soluble N-ethylmaleimide-sensitive factor adaptor protein 33	Protein binding
15	ss715598799	Glyma.07g089200	ATP binding	Unknown
		Glyma.07g089300	Alpha/beta-hydrolases superfamily protein	Unknown
16	ss715596504	Glyma.07g136000	NAD(P)-binding Rossmann-fold superfamily protein	Metabolic process; oxidoreductase activity
		Glyma.07g136100	NAD(P)-binding Rossmann-fold superfamily protein	Metabolic process; oxidoreductase activity
17	ss715597558	Glyma.07g202300	Cytochrome P450, family 93, subfamily D, polypeptide 1	Iron ion binding; electron carrier activity; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; heme binding; oxidation-reduction process
		Glyma.07g202400	Cytochrome P450, family 705, subfamily A, polypeptide 5	Iron ion binding; electron carrier activity; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; heme binding; oxidation-reduction process
		Glyma.07g202500	Histone superfamily protein	DNA binding
		Glyma.07g202600	HAT dimerisation domain-containing protein	Unknown
18	ss715597793	Glyma.07g214100	PLC-like phosphodiesterases superfamily protein	Unknown

19	ss715601594	Glyma.08g048400	THO2	Unknown
		Glyma.08g048500	No annotation	Unknown
		Glyma.08g048600	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	Unknown
		Glyma.08g048700	Transmembrane kinase 1	Protein kinase activity; ATP binding; protein phosphorylation
20	ss715599729	Glyma.08g195200	Protein phosphatase 2A, regulatory subunit PR55	Unknown
21	ss715599836	Glyma.08g203700	Isochorismatase family protein	Catalytic activity; metabolic process
		Glyma.08g203800	Isochorismatase family protein	Catalytic activity; metabolic process
		Glyma.08g203900	OLIGOPEPTIDE TRANSPORTER-RELATED	Catalytic activity; metabolic process
22	ss715601359	No gene models	N/A	N/A
23	ss715602239	Glyma.08g329100	DEK domain-containing chromatin associated protein	Unknown
		Glyma.08g329200	DEK domain-containing chromatin associated protein	Unknown
24	ss715605197	Glyma.09g061800	Auxin efflux carrier family protein	Integral component of membrane; transmembrane transport
25	ss715603000	Glyma.09g089700	CBL-interacting protein kinase 23	Protein kinase activity; ATP binding; protein phosphorylation; signal transduction
26	ss715606025	Glyma.10g031000	No annotation	Unknown
27	ss715606644	Glyma.10g151500	RING/U-box superfamily protein	Protein binding; zinc ion binding; metal ion binding
		Glyma.10g151600	No annotation	Unknown
		Glyma.10g151700	No annotation	Unknown
		Glyma.10g151800	Translationally controlled tumor protein	Unknown
		Glyma.10g151900	Regulatory particle triple-A ATPase 5A	ATP binding; DNA repair; DNA recombination; four-way junction helicase activity; ATPase activity
28	ss715610566	Glyma.11g248100	Lojap-related protein	Unknown
		Glyma.11g248200	No annotation	Unknown
29	ss715613249	Glyma.12g079500	Heavy metal transport/detoxification superfamily protein	Metal ion transport; metal ion binding
30	ss715613671	Glyma.12g103200	No annotation	Unknown
30	ss715613672	Glyma.12g103300	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	Metabolic process; methyltransferase activity
		Glyma.12g103400	Protein with RNI-like/FBD-like domains	Protein binding
31	ss715612794	Glyma.12g217900	Met-10+ like family protein	RNA binding; tRNA (guanine-N2-)-methyltransferase activity; cytoplasm; protein methylation; tRNA processing; protein methyltransferase activity; transferase activity

		Glyma.12g218000	CYCLIN D1;1	Unknown
		Glyma.12g218100	Ribonuclease II family protein	Unknown
32	ss715614812	Glyma.13g181800	WAPL (Wings apart-like protein regulation of heterochromatin) protein	Unknown
		Glyma.13g181900	Cytochrome P450, family 71, subfamily A, polypeptide 25	Iron ion binding; electron carrier activity; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; heme binding; oxidation-reduction process
		Glyma.13g182000	Alpha/beta-Hydrolases superfamily protein	Unknown
33	ss715616411	Glyma.13g338200	ZRT/IRT-like protein 2	Membrane; metal ion transport; metal ion transmembrane transporter activity; transmembrane transport
		Glyma.13g338300	ZRT/IRT-like protein 2	Membrane; metal ion transport; metal ion transmembrane transporter activity; transmembrane transport
		Glyma.13g338400	RING-H2 finger C2A	Protein binding; zinc ion binding; metal ion binding
		Glyma.13g338500	No annotation	Unknown
		Glyma.13g338600	Chaperone DnaJ-domain superfamily protein	Unknown
34	ss715618174	Glyma.14g148900	EamA-like transporter family protein	Membrane; transmembrane transport
		Glyma.14g149000	No annotation	Unknown
		Glyma.14g149100	ARM repeat superfamily protein	Unknown
35	ss715623086	Glyma.15g101100	F-box/RNI-like superfamily protein	Unknown
		Glyma.15g101200	Mo25 family protein	Unknown
		Glyma.15g101300	BSD domain-containing protein	Unknown
		Glyma.15g101400	Bax inhibitor-1 family protein	Unknown
36	ss715620317	Glyma.15g139100	Prefoldin 3	Protein folding; prefoldin complex; unfolded protein binding
37	ss715620339	Glyma.15g141600	No annotation	Unknown
		Glyma.15g141700	No annotation	Unknown
		Glyma.15g141800	No annotation	Unknown
		Glyma.15g141900	No annotation	Unknown
		Glyma.15g142000	No annotation	Unknown
		Glyma.15g142100	No annotation	Unknown
38	ss715622189	No gene models	N/A	N/A
39	ss715622244	No gene models	N/A	N/A
40	ss715628209	Glyma.17g098800	HVA22 homologue A	Unknown
		Glyma.17g098900	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	Unknown
		Glyma.17g099000	No annotation	Unknown
		Glyma.17g099100	TEOSINTE BRANCHED, cycloidea and PCF (TCP) 14	Unknown

41	ss715627773	Glyma.17g252500	ADPGLC-PPase large subunit	Biosynthetic process; nucleotidyltransferase activity
		Glyma.17g252600	Ribosomal protein L31	Structural constituent of ribosome; intracellular; ribosome; translation
		Glyma.17g252700	Jasmonic acid carboxyl methyltransferase	Methyltransferase activity
42	ss715629399	Glyma.18g024400	Lateral organ boundaries (LOB) domain family protein	Unknown
		Glyma.18g024500	No annotation	Unknown
		Glyma.18g024600	Ferritin 4	Cellular iron ion homeostasis; ferric iron binding
43	ss715633017	Glyma.19g010900	Potassium transporter family protein	Potassium ion transmembrane transporter activity; membrane; potassium ion transmembrane transport
		Glyma.19g011000	NHL domain-containing protein	Unknown
44	ss715635643	Glyma.19g219700	Pectin lyase-like superfamily protein	Polygalacturonase activity; carbohydrate metabolic process
		Glyma.19g219800	NAD(P)-binding Rossmann-fold superfamily protein	Metabolic process; oxidoreductase activity
45	ss715638586	Glyma.20g212700	Unknown protein	Unknown
		Glyma.20g212800	Leucine-rich repeat protein kinase family protein	Protein kinase activity; protein binding; ATP binding; protein phosphorylation

Table 2.S4. Candidate genes and their functional annotation identified using the Glyma2.1 gene models in SoyBase (www.soybase.org) within plus or minus 10 kb of the SNP with highest LOD peak for canopy wilting QTLs from composite interval mapping.

QTL Name	SNP	Gene Name	Annotation	Gene Ontology Terms
Wilt-GA-16-RIL	Gm13_29481243_C_A	Glyma.13g181400	Protein kinase superfamily protein	Protein kinase activity; ATP binding; protein phosphorylation
		Glyma.13g181500	Clathrin light chain protein	Structural molecule activity; intracellular protein transport; vesicle-mediated transport; clathrin coat of trans-Golgi network vesicle; clathrin coat of coated pit
		Glyma.13g181600	Tetratricopeptide repeat (TPR)-like superfamily protein	Unknown
		Glyma.13g181700	C-terminal domain phosphatase-like 1	Protein binding
Wilt-KS-16-RIL	Gm16_7851145_G_A	Glyma.16g076700	Xyloglucanase 113	Unknown
		Glyma.16g076800	No annotation	Unknown

CHAPTER 3

UNRAVELING THE GENETIC ARCHITECTURE FOR CARBON AND NITROGEN RELATED TRAITS AND LEAF HYDRAULIC CONDUCTANCE IN SOYBEAN USING GENOME-WIDE ASSOCIATION ANALYSES[1]

[1] C.J. Steketee, T.R. Sinclair, M.K. Riar, W.T. Schapaugh, T.E. Carter Jr., and Z. Li. To be submitted to *BMC Genomics*.

Abstract

Drought stress is a major productivity limiting factor of soybean [*Glycine max* (L.) Merr.] around the world. Soybean plants can ameliorate this stress with improved water use efficiency (WUE), sustained N_2 fixation during water deficits, and/or limited leaf hydraulic conductance. In this study, phenotyping for WUE was conducted by measuring carbon isotope composition ($\delta^{13}C$). Additionally, nitrogen isotope composition ($\delta^{15}N$) and nitrogen concentration that relate to nitrogen fixation were evaluated. Decrease in transpiration rate (DTR) of de-rooted soybean shoots in a silver nitrate ($AgNO_3$) solution compared to deionized water under high vapor pressure deficit (VPD) conditions was used as a surrogate measurement for limited leaf hydraulic conductance. A panel of over 200 genetically diverse soybean accessions genotyped with the SoySNP50K iSelect BeadChips was evaluated for the carbon and nitrogen related traits in two field environments (Athens, GA in 2015 and 2016) and for transpiration response to $AgNO_3$ in a growth chamber. Thirty two, 23, 26, and 10 loci for $\delta^{13}C$, $\delta^{15}N$, nitrogen concentration, and transpiration response to $AgNO_3$, respectively, were significantly associated with these traits using a genome-wide association analysis approach. Accessions with favorable breeding values were also identified for these traits. The genomic regions and germplasm identified in this study can be used by breeders to understand the genetic architecture for these traits and to improve soybean drought tolerance.

Keywords: soybean, *Glycine max*, drought tolerance, carbon isotope composition, nitrogen concentration, nitrogen isotope composition, aquaporin, genome-wide association study (GWAS)

Introduction

Soybean [*Glycine max* (L.) Merr.] is an important source of protein and oil for a range of applications. Drought stress is the most important abiotic factor affecting soybean production, and can cause large decreases in yield (Specht et al., 1999). Use of irrigation during drought stress could ameliorate this issue; however, less than 10% of U.S. soybean hectares are irrigated (USDA-NASS, 2012). Therefore, the development of soybean cultivars that can withstand periods of drought stress is necessary to protect yield when water resources are limited.

Certain morphological and physiological traits could allow soybean plants to better tolerate drought stress. One such adaptation is higher water use efficiency (WUE), which can be defined as the ratio of grain yield to water used during growth at the crop production level, amount of biomass produced per unit of water transpired at the plant level, or photosynthetic carbon gained per unit of water transpired at the leaf level (Angus and van Herwaarden, 2001; Gilbert et al., 2011). Increasing and selecting for higher WUE could be a useful way to improve soybean yields in environments experiencing water deficit stress (Condon et al., 2004; Sinclair, 2012). However, direct selection for actual WUE in a large number of genotypes is difficult, due to the time- and labor-consuming nature of the measurements. Carbon isotope composition has been previously identified as a useful screening method that closely correlates with WUE, especially in C3 plant species (O'Leary, 1981; Farquhar et al., 1982; Hubick and Farquhar, 1989; Condon et al., 1990, 1993). C3 plants prefer the ^{12}C isotope of carbon for photosynthetic purposes, and therefore discriminate against the heavier ^{13}C isotope, which constitutes only around 1% of the atmosphere (O'Leary, 1981). Carbon isotope composition can be expressed as either carbon isotope discrimination ($\Delta^{13}C$, CID) which has a negative relationship with WUE, or carbon isotope ratio ($\delta^{13}C$) which has a positive relationship with WUE (Condon et al., 1990;

Ehleringer et al., 1991; Ismail and Hall, 1992; Rebetzke et al., 2002). Due to this correlation with WUE, carbon isotope composition can be used as an indirect method for selection of genotypes with improved WUE and productivity in water-stressed environments.

The use of carbon isotope composition as a selection tool has been reported in several C3 plant species (Condon et al., 1990; Condon and Richards, 1992; White, 1993; Wright et al., 1994; Specht et al., 2001; Rebetzke et al., 2008). Additionally, previous genome-wide association studies (GWAS) and QTL mapping studies have identified genomic regions that control carbon isotope composition and WUE in soybean. In one of these studies, 373 diverse maturity group (MG) IV soybean genotypes were grown in four environments and 39 SNPs were identified with GWAS that had significant association with $\delta^{13}C$ in at least two environments (Dhanapal et al., 2015a). Another study using the same set of accessions and phenotypic data, but with ~20,000 additional SNP markers and a different GWAS model, found 54 environment-specific SNPs tagging 46 putative loci for $\delta^{13}C$ (Kaler et al., 2017a). Previous QTL mapping in soybean identified five loci controlling CID (Specht et al., 2001), and nine loci controlling WUE (Mian et al., 1996, 1998).

Soybean is a legume which uses a symbiotic association with bradyrhizobia to fix N_2 from the atmosphere. This nitrogen fixation provides a supply of nitrogen (N) to the plant that it can use for its growth and development, as well as help with nitrogen supply for subsequent crops when soybean is used in a crop rotation. However, symbiotic N_2 fixation can be affected by limited water availability, and certain soybean genotypes are more sensitive than others in regards to N_2 fixation during drought stress (Sall and Sinclair, 1991; Serraj and Sinclair, 1997; King and Purcell, 2006; Sinclair et al., 2007; King et al., 2014). A previous simulation study that investigated the benefits of altered soybean drought traits found that sustained N_2 fixation during

water deficits had the most consistent and greatest yield advantage compared to four other traits using 50 years of weather data across U.S. soybean growing regions (Sinclair et al., 2010). Using a three-stage screening process, Sinclair et al. (2000) identified eight soybean genotypes which have superior N_2 fixation during water deficits. In addition, PI 471938 has been reported to have tolerant N_2 fixation as soil dries (Devi and Sinclair, 2013). Differences in the amount of N present in leaf tissue has previously been used as a way to determine a soybean genotype's sensitivity to N_2 fixation during drought conditions, with lower foliar N concentrations having superior fixation during water deficits (Sall and Sinclair, 1991; King and Purcell, 2006; King et al., 2014). This could be due to genotypes with higher N concentrations under well-watered conditions being closer to a threshold affecting N_2 fixation in response to drought stress as a result of feedback of nitrogen compounds from the shoot to nodules, whereas genotypes with lower N concentrations may continue to fix nitrogen during water deficits due to a lack of this feedback and to meet the plant's nitrogen requirements. Four QTLs for foliar N concentration were previously identified on Chr 13, 16, and 17 using a 'KS4895' × 'Jackson' RIL population (Hwang et al., 2013).

Nitrogen isotope composition ($\delta^{15}N$) could be a useful evaluation tool given that ^{15}N is present at much greater levels in soil compared to the atmosphere (Amarger et al., 1979; Shearer et al., 1980; Houngnandan et al., 2008). The fraction of ^{15}N found in a soybean plant would be reduced if it is actively fixing N_2 from the atmosphere, and could be an indicator of how much nitrogen fixation is affected by drought stress (Dhanapal et al., 2015b). A previous association mapping study using 373 soybean genotypes in MG IV found 19 and 17 SNP markers significantly associated with N concentration and the fraction of N derived from the atmosphere

(Ndfa), respectively, that were found in at least two of the four environments tested (Dhanapal et al., 2015b).

Leaf hydraulic conductance is defined as the water flux through the leaf per unit water potential driving force, and is a measure of water flow efficiency through the leaf (Sack and Holbrook, 2006). Limited leaf hydraulic conductance is a trait related to soybean drought tolerance that results in conserved soil moisture for use during water deficits (Sinclair et al., 2008). According to previous research, reduced hydraulic conductance allows certain soybean plants, namely PI 416937, to express a slow canopy wilting phenotype in the field after extended periods with little to no precipitation (Sinclair et al., 2008). Additionally, it was hypothesized that differences in hydraulic conductance were a result of different populations of aquaporins, water-conducting membrane proteins that are involved in water movement through cell membranes. Further, is was suggested that these aquaporin populations are sensitive to exposure to certain chemical inhibitors (Sadok and Sinclair, 2010b). Subjecting de-rooted soybean shoots to a silver nitrate ($AgNO_3$) solution resulted in some genotypes expressing a decreased transpiration rate (Sadok and Sinclair, 2010a), and it was hypothesized that this decrease in transpiration is a result of silver nitrate blocking silver-sensitive aquaporins. PI 416937, a slow wilting genotype with low hydraulic conductance, exhibited an insensitivity to silver nitrate by not decreasing its transpiration rate when subjected to the inhibitor solution (Sadok and Sinclair, 2010a). Given the possible relationship of the transpiration response to silver nitrate and hydraulic conductance, soybean genotypes could be characterized using this procedure as a way to potentially differentiate aquaporin populations and identify drought tolerant germplasm. A previous QTL mapping study identified four QTLs explaining 17.7 to 24.7% of the phenotypic

variation for the limited leaf hydraulic conductance trait using transpiration response to silver nitrate as the measurement for the trait (Carpentieri-Pipolo et al., 2011).

Unraveling the genetic architecture for drought tolerance related traits in soybean is aided by extensive germplasm and genomic resources. A reference genome for soybean has been previously published (Schmutz et al., 2010), and the SoySNP50K BeadChips allow for the identification of high quality SNP markers distributed across the genome to enable genome-wide association analyses (Song et al., 2013). Identification of new drought tolerant germplasm for a number of different traits could allow soybean breeders to understand the genetic architecture of these traits and to improve stress tolerance by incorporating beneficial germplasm and/or alleles into their breeding programs.

In this study, a genetically diverse panel of over 200 soybean genotypes was evaluated for $\delta^{13}C$, $\delta^{15}N$, and nitrogen concentration from leaf samples collected in two field environments. Additionally, this panel was evaluated for transpiration response to silver nitrate under high vapor pressure deficit (VPD) conditions in a growth chamber. The objectives of this study were to identify genomic regions controlling these traits using genome-wide association analyses, validate genomic loci for these traits across environments or studies, and identify genotypes in the panel which have favorable breeding values for these traits.

Materials and methods

Soybean populations

An association panel of 211 genetically diverse soybean genotypes was evaluated for transpiration response to a silver nitrate solution. The panel was previously described in Chapter 2, but with the addition of two lines and replacement of 10 other lines that did not produce

enough seed for the field evaluations of drought tolerance related traits described in Chapter 2 and also in this chapter. This panel was selected based on SoySNP50K genotype data to be genetically diverse, consisted mostly of maturity group (MG) VI-VIII plant introductions, and included drought tolerant and susceptible check genotypes. One hundred ninety five and 205 of the soybean genotypes described in Chapter 2 were evaluated in 2015 and 2016 in Athens, GA, respectively, for carbon and nitrogen related traits. The majority of these lines had not previously been evaluated for drought tolerance related traits, and are later maturing lines than those previously tested (MG IV) and used for association mapping of these traits.

Isotope analysis and sample collection

Leaf samples were collected from field plots of the association panel grown in Athens, GA in 2015 (GA-15) and 2016 (GA-16) and used for stable isotope analysis. More information about sowing dates, row spacing, and management of these plots can be found in Chapter 2. Based on soil sample testing, no fertilizer was added to the field in 2015, and a 4-15-30 fertilizer was applied at a rate of 392 kg ha^{-1} in 2016 prior to planting. These plots were grown under rain-fed conditions and experienced intermittent drought stress periods in both years. In 2015, the leaf samples were collected on 23 September and on 12 September in 2016. The majority of the soybean genotypes in the panel were in early reproductive growth stages (R1-R4) at the time of sample collection. Five leaves were randomly selected from each of the two-row plots at the third trifoliolate leaves below the top of the plants. These leaves were placed in seed envelopes, and stored in a -20^0 C freezer until they could be processed at a later date. In increments of 100-150 samples, the leaf samples were transferred to 50 ml Falcon tubes and placed in a lyophilizer for 2 days to freeze dry. The samples were then ground to a fine powder by placing 4.5 mm zinc

plated BBs in the tubes and grinding them using a Geno/Grinder (SPEX SamplePrep, Metuchen, New Jersey, USA). Immediately before using this ground leaf tissue for isotope analysis, the tubes were placed in a drying oven to ensure all residual moisture was removed. In an effort to further keep out moisture, the Falcon tube caps were wrapped with Parafilm immediately after this second drying step.

Stable isotope analysis was then performed using a Carlo Erba NA1500 CHN combustion analyzer coupled to a Delta V isotope ratio mass spectrometer via the Conflo III open split interface. Three experimental replications of the dry leaf tissue of each genotype were analyzed at the Center for Applied Isotope Studies, University of Georgia, Athens, GA. A detailed protocol for the procedure can be found at http://sisbl.uga.edu/ratio.html. The quantity of ^{13}C in the leaf samples was compared to a reference standard Pee Dee Belemnite (PDB), and these $\delta^{13}C$ values were used for further analyses. $\delta^{13}C$ was expressed in units per mil (‰) using the following equations (O'Leary, 1981):

$$R = {}^{13}CO_2/{}^{12}CO_2$$

$$\delta^{13}C\ (‰) = 1000\ (R_{sample} - R_{standard})/R_{standard}$$

The quantity of ^{15}N in the leaf samples was compared to air and expressed in units per mil (‰) according to the following equations:

$$R = {}^{15}N/{}^{14}N$$

$$\delta^{15}N\ (‰) = \ 1000\ (R_{sample} - R_{air\ N2})/R_{air\ N2}$$

Nitrogen concentration was expressed as g kg^{-1}.

Evaluation of response to silver nitrate inhibitor

Soybean plants for evaluation of transpiration response to silver nitrate were grown in a greenhouse at the University of Georgia in Athens, GA, USA. Three seeds of each genotype were sown in 32 oz. styrofoam cups using a Fafard 2B soil media (Sun Gro Horticulture, Agawam, MA, USA). Approximately 1.5 weeks after seedling emergence, the plants were thinned to one plant per cup and maintained under well-watered conditions. Once the soybean plants reached the V3-V4 growth stage (approximately 4 weeks after planting), the tests for response to the silver nitrate inhibitor began (Sadok and Sinclair, 2010b).

The tests were conducted over 2 days. In the afternoon of the first day, the soybean plants were removed from their growing media in the greenhouse and de-rooted using clippers. A second cut on the stem was then made underwater using a razor blade. The remaining shoot was then placed in a 250 ml Erlenmeyer flask filled with deionized (DI) water and the mouth of the flask was sealed with Parafilm to avoid water evaporation. Plants in flasks were then placed in a walk-in growth chamber at approximately 20^0 C and 60% relative humidity overnight in dark conditions.

In the morning of day 2, the lights were turned on, temperature was raised to 30^0 C, and relative humidity was decreased to 30% to obtain a chamber vapor pressure deficit (VPD) of approximately 2.3 kPa. The plants were allowed to acclimate to the high VPD condition for 60 min. Then, each flask/soybean was weighed inside the growth chamber using a balance with a resolution of 0.001 g in order by flask number. Sixty min after the first weighing, they were weighed again in the same order to determine the transpiration rate in water (TR_W). The soybean shoots were then transferred to a 60 mL amber glass bottle containing a 200 µM solution of silver nitrate ($AgNO_3$) under semi-dark conditions. This $AgNO_3$ solution concentration was

previously shown to best differentiate the transpiration response of drought tolerant versus susceptible soybean plants in Sadok and Sinclair (2010). Parafilm was again used to seal the mouth of the amber bottles to avoid evaporation and spilling of any chemical. Then, the plants were returned to the growth chamber and allowed to acclimate to the inhibitor treatment for 60 min. The amber bottles with shoots were then weighed for their initial weight in order by bottle number. After approximately 120-160 min, the bottles were reweighed in bottle order to determine the transpiration response to the silver nitrate inhibitor (TR_I). Differences in the amount of time that elapsed between weight measurements were accounted for in the TR_W and TR_I calculations by changing the denominator in increments of minutes. Decrease in transpiration rate (DTR, %) was then calculated as follows:

$$\mathrm{DTR} = 100 \times \frac{(\mathrm{TR_W} - \mathrm{TR_I})}{\mathrm{TR_W}}$$

Due to limitations in the size of the walk-in growth chamber and ability to weigh the flasks/bottles in an orderly and timely fashion, eight separate replications of this experiment were conducted for the association panel. Each replication consisted of the entire panel of 211 soybean genotypes, and the flask/bottle order was randomized for each replication. DTR values were normalized by the lowest DTR within each replication to help reduce variation across replications using the following equation:

Normalized DTR (NDTR) within Each Replication = $DTR_{Genotype}/DTR_{Genotype\ with\ Lowest\ DTR}$

A full summary of the measurement dates and average VPD for each experimental replicate are included as a supplemental table (Table 3.S1).

Genotype data and quality control

The association panel was genotyped with the SoySNP50K iSelect BeadChip. DNA extraction and genotyping procedures for this panel were conducted as described in Chapter 2. A total of 42,079 genome-wide SNP markers resulted from the genotyping effort. However, markers with minor allele frequencies (MAF) lower than 0.05 were eliminated leaving 35,262 SNP markers for the association analysis of transpiration response to silver nitrate. For the carbon and nitrogen related traits, 35,234 (Both), 35,101 (GA-15), and 35,219 (GA-16) markers were used after eliminating markers with MAF lower than 0.05. The number of markers varied, because certain SNPs with a MAF close to 0.05 were either included or excluded depending on the number of entries at the given environment. Physical positions are based on the Glyma.Wm82.a2 version of the soybean genome.

Statistical analyses

Analyses of variance (ANOVA) was conducted using PROC GLM in SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). For the response variables relating to carbon and nitrogen traits, genotype was treated as a fixed effect, and environment, genotype by environment interaction, and replication within environment were random effects. For transpiration response to silver nitrate, a model was created with genotype as a fixed effect and replication as a random effect, with NDTR as the response variable. Broad-sense heritability was calculated on an entry-mean basis according to Holland et al. (2002) with the variance components being calculated with PROC MIXED of SAS 9.4 using a model where all variables were treated as random.

Best linear unbiased predictors (BLUPs) were calculated from both across and within environments and used as the phenotype values for subsequent GWAS analyses. The BLUP

calculations for carbon and nitrogen related traits across both environments were performed using JMP Pro (JMP®, Version 13, SAS Institute Inc., Cary, NC, USA). The model was built by treating genotype, environment, genotype by environment, and replication within environment as random variables using the Standard Least Squares personality and REML method. For individual environments for carbon and nitrogen related traits and transpiration response to silver nitrate, genotype and replication were used as variables and treated as random to calculate BLUPs.

Genome-wide association analyses

Fixed and random model Circulating Probability Unification (FarmCPU) was used to perform the genome-wide association analyses for all traits evaluated (Liu et al., 2016). FarmCPU is an R package that implements a multiple loci linear mixed model (MLMM) incorporating a modified mixed linear model (MLM) that includes the most significant markers as covariates. It uses fixed and random effect models iteratively to help reduce potential confounding between the markers and kinship. This model has previously been successfully utilized in soybean genome-wide association analyses to identify genomic regions controlling canopy wilting (Chapter 2, Kaler et al., 2017), carbon and oxygen isotope ratios (Kaler et al., 2017a), and resistance to *Sclerotinia sclerotiorum* (Wei et al., 2017).

After removing markers with MAF below 0.05, 35,101-35,234 SNP markers were used for the analysis of carbon and nitrogen related traits, and 35,262 SNP markers were used for transpiration response to silver nitrate. Manhattan plots were visualized with the 'qqman' (Turner, 2014) and 'CMplot' R packages using the p-values generated from the FarmCPU output. The significance threshold ($p < 0.0001$; $-\log_{10}(P) > 4$) was used to determine if SNPs

were significantly associated with the traits of interest. This threshold is less stringent than a Bonferroni-corrected threshold, but is more stringent than many other soybean GWAS studies using 50K SNP genotyping data (Kaler et al., 2017a; b, 2018; Zeng et al., 2017). It is also near the point at which the p-values deviated from the linear expected p-values in the quantile-quantile (QQ) plots. Days to flowering (DTF) was recorded in both field environments as the number of days from sowing until 50% of the plants in a plot reached the first bloom (R1) growth stage. The carbon and nitrogen related traits evaluated had relatively strong correlations (data not shown) with DTF in both environments, so DTF was used as a fixed effect covariate, along with the first four genetic principal coordinates, in the GWAS to account for this correlation and population structure, respectively.

Haploview version 4.2 software (Barrett et al., 2005) was used to calculate pairwise estimates of D′ and r^2 and estimate linkage disequilibrium (LD) blocks. Using D′ > 0.8 to extend the spine, LD blocks were identified by chromosome with the Solid Spine of LD option. These LD blocks were used to determine if significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs that are physically close (less than 1 Mb) were at the same locus (genomic region) controlling the trait of interest. Allelic effects were calculated by taking the mean difference in phenotypic values for the trait between the two alleles at a particular SNP, and were provided as part of the FarmCPU output. A negative effect value indicates that an individual possessing the second nucleotide alphabetically for this SNP would have lower phenotypic values, whereas a positive effect value would have higher phenotypic values. The direction, negative or positive, of the effect is based on how the genotype data was converted from HapMap to numerical format using GAPIT (Tang et al., 2016) prior to conducting the GWAS with the numerically formatted genotype data in FarmCPU. Since BLUP values were used as the phenotype in the GWAS, the allelic effects

reported are based on these BLUP values rather than the original raw data. Phenotypic variation explained (R^2) by significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs was calculated using a linear regression in R. For individual SNPs, the model lm(BLUP ~ SNP) was used, whereas lm(BLUP ~ SNP_1 + SNP_2 + …) was used to determine the total amount of phenotypic variation explained by all significant SNPs for a given trait in a particular environment.

Breeding values for the traits were calculated by summing the allelic effects for all significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs in each individual environment and with the across environments BLUPs. Breeding values across the individual environments were also summed and used for comparisons. Allelic effects for a given SNP were considered negative if the allele contributed to lower phenotypic values, and positive if it increased phenotypic values. Heterozygous and missing allele calls were not included in the breeding value calculation.

Identification of gene models at significant SNPs and with aquaporin functional annotation

Using SoyBase (Grant et al., 2010), candidate genes along with their functional annotation and gene ontologies were identified for significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs from GWAS explaining more than 15% of the phenotypic variation for each of the carbon and nitrogen related traits. Glyma2.1 gene models within plus or minus 10 kb of the SNP physical position were recorded and further investigated. The median distance between SNP markers used in the GWAS was 9 kb, and the mean distance was 26 kb. Although identifying all gene models in LD with significant SNPs would be ideal, we focused our efforts on models in close proximity (within plus or minus 10 kb), which approximately spans this distance between markers.

Given the hypothesized relationship between transpiration response to silver nitrate and sensitivity of aquaporin populations in soybean (Sadok and Sinclair, 2010a; Carpentieri-Pipolo et

al., 2011; Sinclair et al., 2017), a search for the term "aquaporin" was performed in Phytozome v12.1 for the *Glycine max* Wm82.a2.v1 version of the soybean genome. This identified 88 gene models which had "aquaporin" in their functional annotation. In comparison, 82 of these gene models were also found when searching for "aquaporin" on the SoyBase website (www.soybase.org). The physical locations of the full list of 88 gene models having an aquaporin annotation from Phytozome were used to make comparisons between the significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs identified for transpiration response to silver nitrate from the GWAS results to see if any aquaporin genes were in or near these regions.

Results

$\delta^{13}C$, $\delta^{15}N$, and N concentration

Carbon isotope composition ($\delta^{13}C$), nitrogen isotope composition ($\delta^{15}N$), and foliar nitrogen (N) concentration were evaluated in two field environments (GA-15 and GA-16). Based on the ANOVA, genotypes, environments, and their interaction were statistically significant ($p < 0.05$) for all carbon and nitrogen related traits (Table 3.1). Genotype mean values within environments of $\delta^{13}C$ ranged from -29.97 to -25.14‰ (Figure 3.1), and had a correlation of $r = 0.74$ between environments. Broad-sense heritability of $\delta^{13}C$ on an entry-mean basis for each environment was 61% (GA-15), 72% (GA-16), and 62% across both environments (Both) (Table 3.2). $\delta^{15}N$ had a correlation of $r = 0.28$ between environments, and ranged from -1.23 to 4.50‰ based on mean genotype values within environments (Figure 3.1). Heritability for $\delta^{15}N$ was lower than for all other carbon and nitrogen related traits at 24% (GA-15), 40% (GA-16), and 17% across both environments (Both) (Table 3.2). The range of nitrogen concentrations observed for genotype means within environments was from 16.67 to 55.45 g kg^{-1}, and the correlation between the two

environments was r = 0.73. Broad-sense heritability for N concentration was between 63-73% (Table 3.2).

In general, these carbon and nitrogen related traits had fairly strong relationships with one another. Using BLUP values calculated from across both environments, correlations between the carbon and nitrogen related traits were from r = -0.52 to 0.71 (Table 3.3). The most negative correlation (r = -0.52) was between $\delta^{13}C$ and $\delta^{15}N$, and the most positive correlation (r = 0.71) was observed between $\delta^{13}C$ and N concentration (Table 3.3).

Higher (less negative) values of $\delta^{13}C$ is often assumed to indicate higher WUE in most environments, and therefore positive breeding values were considered favorable for this trait. PI 398823, a MG IV accession had the highest breeding value for $\delta^{13}C$ using the sum across the two individual environments (Table 3.S3). In addition, PI 416937, a slow wilting check genotype, had a relatively high breeding value for this trait and ranked 15th out of the over 200 genotypes tested (Table 3.S3). Other genotypes with positive breeding values for $\delta^{13}C$ could be sources of improvement for WUE in soybean.

Low leaf N concentration could be beneficial given that lower foliar N is correlated with sustained N_2 fixation during drought stress (King et al., 2014). A MG VI accession from China, PI 567377B, had the most negative (favorable) breeding value for N concentration using the sum across both individual environments (Table 3.S3). PI 471938, which was previously identified as a genotype possessing nitrogen fixation drought tolerance (Devi and Sinclair, 2013; Riar et al., 2018), had the 40th lowest breeding value for N concentration (Table 3.S3). Only 20 of the genotypes tested had negative breeding values for N concentration, and could be considered sources of nitrogen fixation drought tolerance improvement.

For $\delta^{15}N$, lower values would indicate that more nitrogen fixation from the atmosphere is occurring, and therefore potentially more negative breeding values could be advantageous (Dhanapal et al., 2015b). Forty four of the genotypes evaluated in the panel had negative breeding values for $\delta^{15}N$, with PI 567386, a MG VI accession from China, having the most negative breeding value.

Transpiration response to silver nitrate aquaporin inhibitor

Transpiration response to the silver nitrate aquaporin inhibitor was evaluated with the association panel. Decrease in transpiration response (DTR) was normalized by the lowest DTR (NDTR) in each separate experimental replication, so the maximum value the NDTR could be is 1. NDTR values ranged from -2.18 to 1.00 within individual replications (Figure 3.2), and from -0.57 to 0.64 based on genotype means. Genotype effects were statistically significant ($p < 0.05$) (Table 3.1), and broad-sense heritability on an entry-mean basis was 17% (Table 3.2). Using BLUP values across replications and environments, the relationships between NDTR to $AgNO_3$ and the carbon and nitrogen related traits were also evaluated (Table 3.3). Low correlations ($r = -0.04$ to -0.01) were observed for all comparisons between silver nitrate NDTR and the previously described carbon and nitrogen related traits.

Low or negative DTR to silver nitrate values (transpiration less affected by $AgNO_3$) have been previously correlated with limited leaf hydraulic conductance, which is a beneficial trait in certain drought stress environments (Sadok and Sinclair, 2010b). However, given that the DTR values in the current study were normalized by the lowest DTR value in each replication to calculate NDTR, accessions with higher NDTR values and high positive breeding values could be potential sources of improvement for this trait given that DTR and NDTR have an inverse

relationship. Nine out of the 15 accessions with the most positive breeding values originated from China (Table 3.S3). Interestingly, the four accessions with the most favorable breeding values were all MG IV plant introductions that originated from China (Table 3.S3). PI 416937 was previously identified as a genotype with a transpiration response that is relatively insensitive to silver nitrate (Sadok and Sinclair, 2010a), and ranked 88th based on NDTR breeding values. Genotypes with favorable breeding values could be targets for improvement of the limited leaf hydraulic conductance trait in soybean.

GWAS of carbon and nitrogen related traits

A total of 35 unique SNPs tagging 32 loci were identified either in individual environments or when using the BLUP calculated across both environments for $\delta^{13}C$ (Figure 3.S1 and Table 3.4). Two SNPs for $\delta^{13}C$ (ss715587736 and ss715587739) on Chr 4 were in the same genomic region, and were found in GA-15 and across both environments, respectively (Table 3.4). Of all other SNPs identified for $\delta^{13}C$, each SNP tagged a single genomic region, with the exception of two SNPs identified on Chr 4 and 16. The allelic effects across all significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs ranged from -0.19 to 0.13 (Table 3.4), and mean phenotypic variation explained (R^2) for individual SNPs was 3.4% with all significant SNPs explaining a total of 29-44% of the variation, depending on the environment (Table 3.4).

For $\delta^{15}N$, 23 loci were identified in the GWAS (Figure 3.S2 and Table 3.4). Depending on the environment, 36 to 51% of the phenotypic variation for $\delta^{15}N$ was explained by the significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs. On average, a single significant SNP had a R^2 of 5.9%. The allelic effects ranged from -0.14 to 0.11 for the SNPs significantly associated with $\delta^{15}N$ (Table 3.4). One SNP (ss715635458) was found for $\delta^{15}N$ both in GA-16 and using the

across both environments BLUPs (Table 3.4). All other SNPs identified tagged a single genomic region.

Twenty seven SNPs tagging 26 loci were identified in the GWAS for nitrogen concentration (Figure 3.S3 and Table 3.4). One SNP (ss715610522) was identified in both an individual environment (GA-15) and with the BLUP value from across both environments (Table 3.4). All other SNPs tagged a single genomic region, except for two SNPs (locus 17) on Chr 13. Allelic effects for nitrogen concentration ranged from -1.33 to 1.46 (Table 3.4). On average, 4.4% of the phenotypic variation for N concentration was explained by a significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNP, with R^2 of 50%, 35%, and 21% for GA-15, GA-16, and across both environments (Both), respectively, for all significant SNPs.

GWAS for transpiration response to silver nitrate aquaporin inhibitor

Ten SNPs tagging 10 loci were significantly ($p < 0.0001$; $-\log_{10}(P) > 4$) associated with NDTR to the silver nitrate aquaporin inhibitor solution (Figure 3.3 and Table 3.5). Thirty two percent of the phenotypic variation for the trait was explained by these 10 SNPs, with individual SNPs having a mean R^2 of 3.6%. The allelic effects for these significant SNPs ranged from -0.03 to 0.03 (Table 3.5).

Candidate genes for drought tolerance related traits

A total of six SNPs across the carbon and nitrogen related traits explained greater than 15% of the phenotypic variation (R^2) for their respective trait in a given environment. One, four, and four gene models for $\delta^{13}C$, $\delta^{15}N$, and N concentration respectively, were identified within plus or

minus 10 kb (approximately spans the mean distance between all markers) of these SNPs which had a wide range of annotations and ontologies (Table 3.S2).

Drought tolerant germplasm based on breeding values

Germplasm identified with favorable breeding values for drought tolerance related traits could be utilized by breeders to improve soybean drought tolerance. Breeding values from canopy wilting in Chapter 2 were included in the supplemental material, so that comparisons could be made between all drought tolerance related traits evaluated in this dissertation (Table 3.S3). As is evident is this table, accessions rarely possess favorable breeding values for all of the traits evaluated (Table 3.S3). As a reference point, PI 416937, a genotype previously identified as possessing the slow canopy wilting trait (Sloane et al., 1990), was ranked as the 65th best accession tested based on an overall median rank across breeding value ranks for canopy wilting, carbon isotope composition, nitrogen concentration, nitrogen isotope composition, and NDTR to silver nitrate (Table 3.S3). It ranked 69th best for canopy wilting and 15th best for carbon isotope composition, but ranked 189th for nitrogen concentration, 140th for nitrogen isotope composition, and 88th best for transpiration response to silver nitrate (Table 3.S3). Sixty-four accessions with overall median ranks lower than PI 416937 were identified in this research (Table 3.S3). These accessions show potential as parents for future genetic studies, as well as possess favorable alleles that could be introgressed into elite germplasm to improve many traits simultaneously. However, several other accessions could be targets for specific traits given that they possess more favorable breeding values for the particular trait.

Discussion

Rationale for trait evaluation

In this study, a genetically diverse panel of over 200 soybean genotypes was evaluated for $\delta^{13}C$, $\delta^{15}N$, and nitrogen concentration from leaf samples collected in two different field environments. In addition, this panel was also evaluated for transpiration response to silver nitrate under high vapor pressure deficit conditions in a growth chamber. Using genome-wide association mapping, genomic regions were identified controlling each of these different drought tolerance related traits and the results were compared to previous mapping studies for these traits. In addition, genotypes in the panel were identified which possessed favorable breeding values for these drought tolerance related traits.

Carbon isotope composition has been previously shown to be correlated in some cases with WUE and yield improvement in C3 plant species such as wheat, and therefore could be a useful surrogate measurement for WUE in soybean (Hall et al., 1994). Nitrogen fixation can be highly sensitive to drought stress (Sall and Sinclair, 1991; Serraj and Sinclair, 1997; Sinclair et al., 2007), and above-ground measurements such as nitrogen concentration and nitrogen isotope composition can be evaluated to help understand how these traits relate to fixation and soybean drought tolerance (Amarger et al., 1979; King and Purcell, 2006; Houngnandan et al., 2008). The amount of ^{15}N found in a soybean plant would be reduced if it is actively fixing N_2 from the atmosphere, and lower N concentrations have been shown to correlate with superior fixation during water deficits. However, given the high protein content of soybean, and the amount of nitrogen required to produce protein in seed, lower N concentrations would likely in general be a poor trait for a soybean genotype to possess. Water-transporting proteins called aquaporins are involved in water movement through cell membranes (Chaumont and Tyerman, 2017), and

populations of aquaporins in soybean lines can vary as detected by transpiration response to chemical inhibitors such as silver nitrate (Sadok and Sinclair, 2010a; Carpentieri-Pipolo et al., 2011; Devi et al., 2016). It is hypothesized that insensitivity to silver nitrate is correlated with the limited leaf hydraulic conductance trait, a beneficial trait for improving drought tolerance in certain environments (Sinclair et al., 2008; Sadok and Sinclair, 2010b). All of these traits were evaluated in the current study in order to help provide a better understanding of the genetic architecture of these drought tolerance related traits and identify germplasm with favorable breeding values for the traits.

$\delta^{13}C$, $\delta^{15}N$, and N concentration

Values for $\delta^{13}C$ were in a similar range to those observed in two previous carbon isotope association mapping studies (Dhanapal et al., 2015a; Kaler et al., 2017a) (Figure 3.1). The range of values observed for nitrogen concentration was wider and concentrations were higher compared to those observed in a previous study (Dhanapal et al., 2015b). Direct comparisons to Dhanapal et al. (2015a) were not able to be made for $\delta^{15}N$ due to differences in the units used for these measurements. Analyses of variance showed that genotype, environment, and their interaction were statistically significant ($p < 0.05$) for all carbon and nitrogen related traits evaluated with the association panel (Table 3.1). Although these genotype by environment interactions were significant ($p < 0.05$), correlations were generally high between the two environments. Correlations for $\delta^{13}C$ and nitrogen concentration were all above $r = 0.70$ between the two environments tested, and the lowest correlation was for $\delta^{15}N$ at $r = 0.28$, indicating the genotypes performed similarly across environments.

Heritability for $\delta^{13}C$ and nitrogen concentration in single environments and using across both environments data were above 61% (Table 3.2). Heritability for $\delta^{15}N$ was substantially lower and ranged from 17-40% (Table 3.2). This lower heritability for $\delta^{15}N$ could potentially be explained by the fact that we did not adjust our values to a non-nodulating reference crop, and that these values are also affected by field variation in soil nitrogen concentration (Evans, 2001). However, heritability estimates for all of these carbon and nitrogen related traits are comparable to the values observed in other studies (Dhanapal et al., 2015a; b; Kaler et al., 2017a).

Transpiration response to $AgNO_3$

Normalized DTR (NDTR) values for the current study were calculated as described in Carpentieri-Pipolo et al. (2011) by dividing each DTR by the lowest DTR value in each experimental replication, so the maximum NDTR value observed was 1. The genotype with the lowest DTR value in each replication was negative, so all positive NDTR values reported had negative non-normalized DTR values while all negative NDTR values reported had positive non-normalized DTR values due to how NDTR was calculated (negative divided by negative versus positive divided by negative). Given the hypothesis that silver nitrate blocks aquaporins and reduces transpiration, and that most previously reported DTR values were positive, we observed an unexpected distribution of NDTR values given that many of the genotypes we tested had negative non-normalized DTR (positive NDTR). This could indicate that silver nitrate blocked some aquaporins as expected, but in some genotypes this blockage resulted in a stimulus in the number or activity of other silver-insensitive aquaporins. However, this theory would need further experimental verification, and should only be considered a possible explanation at this point. Analyses of variance found that genotype effects were statistically significant ($p < 0.05$)

(Table 3.1), and heritability for this trait was 17% (Table 3.2). This low heritability estimate could have been a result of a technical issue or that this phenotyping method may not be a reliable proxy for limited leaf hydraulic conductance, and would make it difficult for soybean breeders to make effective selection for this trait.

Comparison to previous mapping results for C and N traits

Thirty two, 23, and 27 loci were identified for $\delta^{13}C$, $\delta^{15}N$, and nitrogen concentration, respectively, to be significantly associated with these traits using a genome-wide association mapping approach (Figures 3.S1-4 and Table 3.4). Given that FarmCPU uses the most significant markers as covariates in the GWAS model, SNPs are seldom identified within the same LD block for an environment-specific dataset. However, two genomic regions were found both in individual environments and when using the across both environments BLUP data for these carbon and nitrogen related traits. Significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs for carbon and nitrogen related traits were found on all 20 soybean chromosomes (Table 3.4).

QTLs for carbon isotope discrimination and WUE are numbered with their approximate physical positions on the SoyBase website (www.soybase.org). Locus 32 identified with GWAS for $\delta^{13}C$ in the current study is found within the CID 1-5 QTL on Chr. 19 identified in Specht et al. (2001). Loci 28 and 29 on Chr. 18 identified for $\delta^{13}C$ are within the QTL interval for WUE 1-1, and locus 25 on Chr. 16 is co-located with WUE 1-4, 1-5, and 1-6 identified in Mian et al. (1996) (Table 3.4). A comparison of SNPs significantly associated with $\delta^{13}C$ from two previous association mapping studies (Dhanapal et al., 2015a; Kaler et al., 2017a) and the current study was conducted (Figure 3.4A). Two SNPs on Chr. 6 and 11 from the current study are near

significant markers identified in Kaler et al. (2017), and one SNP on Chr. 13 and another SNP on 18 were found near the Dhanapal et al. (2015) significant SNPs for $\delta^{13}C$.

No QTLs for $\delta^{15}N$ identified with linkage mapping are reported on the SoyBase website. One previous linkage mapping study for foliar nitrogen concentration identified four QTLs, of which one QTL on Chr 16 was 256 kb away from locus 21 identified in the current study (Hwang et al., 2013). A comparison of SNPs identified for nitrogen related traits in a previous association mapping study (Dhanapal et al., 2015a) and the current study was also performed (Figure 3.4B). SNPs on Chr. 9 and 15 were found in common for $\delta^{15}N$ in the current study and nitrogen derived from the atmosphere (Ndfa) in Dhanapal et al. (2015). No SNPs were within 1 Mb of previously identified genomic regions for nitrogen concentration. Additionally, when making comparisons only across studies and different nitrogen related traits, only two regions on Chr. 15 and 16 had common SNPs within 1 Mb of each other. Within the current study only, two regions contained N related significant ($p < 0.0001$; $-\log_{10}(P) > 4$) SNPs within 1 Mb of each other on Chr. 13 and 20 (Table 3.4). The consistent QTLs and genomic regions across environments, studies, and traits, along with SNPs explaining a high amount of phenotypic variation in the current study could be useful breeding targets for these carbon and nitrogen drought tolerance related traits.

Genetic mapping for transpiration response to $AgNO_3$ and proximity of identified regions to aquaporin gene models

For NDTR to silver nitrate with the association panel, 10 SNPs tagging 10 loci were identified on Chr. 2, 4, 7, 8, 9, 11, 16, and 19 (Figure 3.3 and Table 3.5). This is the first report of association mapping for this trait to the authors' knowledge in any crop species. A previous QTL

mapping study for limited leaf hydraulic conductance identified QTLs on Chr. 3, 5, 10, and 12 (Carpentieri-Pipolo et al., 2011). None of the loci identified in the current study or in this previous study are located near one another. This could be due to differences in the populations utilized for the mapping, and could also be affected by the low heritability for this trait (Table 3.2). However, a search on Phytozome for gene models with a functional annotation which contained the word "aquaporin" was also conducted given the hypothesized relationship between this limited leaf hydraulic conductance trait and aquaporins, and found 88 gene models. The physical locations of these gene models and the loci identified in the current study with association and linkage mapping were compared (Figure 3.4C). Three SNPs identified in the GWAS were within 0.4 Mb of four gene models with an aquaporin functional annotation. These regions could be further investigated to see how this trait relates to aquaporins.

Candidate genes at identified genomic regions

A total of nine gene models were identified near the six SNPs which explained greater than 15% of the phenotypic variation (R^2) for a trait in a certain environment for the carbon and nitrogen related traits. The only gene model (Glyma.10g103000) identified near locus 12 for $\delta^{13}C$ is an elongation factor family protein with gene ontology terms of GTPase activity and GTP-binding (Table 3.S2). In a study where rice plants were subjected to extreme drought stress conditions, small GTP-binding proteins were significantly up-regulated (Mirzaei et al., 2012). Therefore, small G-proteins could play a role in drought stress response of plants. Near locus 8 identified for $\delta^{15}N$, a gene model (Glyma.09g225400) for cytokinin oxidase/dehydrogenase is located (Table 3.S2). Tobacco plants with ectopically enhanced cytokinin oxidase/dehydrogenase gene expression had enhanced drought tolerance, indicating that modulating cytokinin levels in plants

could have an effect on improving drought stress tolerance (Macková et al., 2013). These gene models could be potential targets for understanding and improving these drought tolerance related traits given their relationship with drought stress tolerance response or enhancement.

Relationship between drought tolerance related traits

Correlations between the carbon and nitrogen related traits evaluated ranged from r = -0.52 between $\delta^{13}C$ and $\delta^{15}N$ to r = 0.71 for $\delta^{13}C$ and nitrogen concentration when using BLUP values across both environments (Table 3.3). A high correlation for $\delta^{13}C$ and nitrogen concentration was also observed in a greenhouse study of six soybean genotypes subjected to low and high soil moisture treatments (Miles Ingwers, personal communication), and whether the relationship was positive or negative was determined by the soil moisture level. Another measurement related to soybean drought tolerance, canopy wilting, was added to the correlation matrix using data from Chapter 2 (Table 3.3). This additional data provides another trait to compare to carbon and nitrogen related traits and NDTR to silver nitrate. Canopy wilting and normalized DTR to silver nitrate had relatively low correlations with each of the other traits evaluated and with one another. Ries et al. (2012) also found that there was not a consistent relationship among genotypes within slow or fast canopy wilting groups and carbon isotope discrimination. Given that the carbon and nitrogen related traits evaluated are time integrative measurements, whereas canopy wilting and response to silver nitrate are more instantaneous, it is somewhat expected that these correlations are low and relationships not consistent. They are different drought tolerance traits governed by different genetic loci. Drought tolerance is a complex, quantitative trait, so it is expected that multiple different traits and loci are responsible for soybeans ability to withstand water deficit stress.

Breeding implications

Many different genotypes were identified in the current study with favorable breeding values for drought tolerance related traits and could be utilized by breeders to improve soybean drought tolerance or be used as parents to create mapping populations. However, the challenge as a breeder would be to determine which trait(s) to target given the quantitative nature of the genetic architecture for many traits that could lead to soybean drought tolerance improvement, and some of these traits could be associated with poor agronomic performance. Accessions in the current study often had favorable breeding values for certain traits, but then also had less favorable breeding values for other traits (Table 3.S3). To make selections based on multiple traits an index accounting for trait heritability, economic importance, and genetic and phenotypic correlations between the traits would likely need to be employed. Ultimately, a breeder may need to weight traits according to which would provide the best drought tolerance in their given target environment, and then utilize the germplasm and genomic regions identified for that specific trait.

Conclusions

Genome-wide association analyses were conducted for $\delta^{13}C$, $\delta^{15}N$, and nitrogen concentration from two environments using over 200 genetically diverse soybean genotypes. Thirty two, 23, and 26 loci were identified for $\delta^{13}C$, $\delta^{15}N$, and nitrogen concentration, respectively. Four loci detected with the GWAS for $\delta^{13}C$ were co-located with previously identified QTLs for CID or WUE, and four SNPs were near SNPs found in previous association mapping studies. Two SNPs for $\delta^{15}N$ were found in the GWAS near genomic regions identified in an association mapping

study for nitrogen related traits. Ten SNPs tagging 10 loci were identified with a GWAS approach for normalized DTR to silver nitrate, and three of the SNPs identified were found near four aquaporin related gene models. Breeding values calculated with the significant SNPs from the GWAS enabled the identification of accessions which possess favorable combinations of alleles for these drought tolerance related traits. The genomic regions and germplasm identified in this study, especially those found in common across environments, studies, and traits, can be used to understand the genetic architecture for these traits and by soybean breeders to improve drought tolerance.

References

Amarger, N., A. Mariotti, F. Mariotti, J.C. Durr, C. Bourguignon, and B. Lagacherie. 1979. Estimate of symbiotically fixed nitrogen in field grown soybeans using variations in 15N natural abundance. Plant Soil 52(2): 269–280.

Angus, J.F., and A.F. van Herwaarden. 2001. Increasing water use and water use efficiency in dryland wheat. Agron. J. 93: 290–298.

Barrett, J.C., B. Fry, J. Maller, and M.J. Daly. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21(2): 263–265.

Carpentieri-Pipolo, V., A.E. Pipolo, H. Abdel-Haleem, H.R. Boerma, and T.R. Sinclair. 2011. Identification of QTLs associated with limited leaf hydraulic conductance in soybean. Euphytica 186: 679–686.

Chaumont, F., and S. Tyerman (Eds). 2017. Plant aquaporins: From transport to signaling. Springer International Publishing.

Condon, A.G., G.D. Farquhar, and R.A. Richards. 1990. Genotypic variation in carbon isotope discrimination and transpiration efficiency in wheat. Leaf gas-exchange and whole plant studies. Aust. J. Plant Physiol. 17: 9–22.

Condon, A.G., and R.A. Richards. 1992. Broad sense heritability and genotype × environment interaction for carbon isotope discrimination in field-grown wheat. Aust. J. Agric. Res. 43(5): 921–934.

Condon, A.G., R.A. Richards, and G.D. Farquhar. 1993. Relationships between carbon isotope discrimination, water use efficiency and transpiration efficiency for dryland wheat. Aust. J. Agric. Res. 44(8): 1693–1711.

Condon, A.G., R.A. Richards, G.J. Rebetzke, and G.D. Farquhar. 2004. Breeding for high water-

use efficiency. J. Exp. Bot. 55(407): 2447–2460.

Devi, M.J., and T.R. Sinclair. 2013. Nitrogen fixation drought tolerance of the slow-wilting soybean PI 471938. Crop Sci. 53(5): 2072–2078.

Devi, M.J., T.R. Sinclair, and E. Taliercio. 2016. Silver and zinc inhibitors influence transpiration rate and aquaporin transcript abundance in intact soybean plants. Environ. Exp. Bot. 122: 168–175.

Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C. Andy King, P.B. Cregan, Q. Song, and F.B. Fritschi. 2015a. Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [Glycine max (L.) Merr.] genotypes. Theor. Appl. Genet. 128(1): 73–91.

Dhanapal, A.P., J.D. Ray, S.K. Singh, V. Hoyos-Villegas, J.R. Smith, L.C. Purcell, C.A. King, and F.B. Fritschi. 2015b. Genome-wide association analysis of diverse soybean genotypes reveals novel markers for nitrogen traits. Plant Genome 8(3): 1–15.

Ehleringer, J.R., S. Klassen, C. Clayton, D. Sherrill, M. Fullerholbrook, Q.N. Fu, and T.A. Cooper. 1991. Carbon isotope discrimination and transpiration efficiency in common bean. Crop Sci. 31(6): 1611–1615.

Evans, R.D. 2001. Physiological mechanisms influencing plant nitrogen isotope composition. Trends Plant Sci. 6(3): 121–126.

Farquhar, G., M. O'Leary, and J. Berry. 1982. On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves. Aust. J. Plant Physiol. 9(2): 121–137.

Gilbert, M.E., M.A. Zwieniecki, and N.M. Holbrook. 2011. Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use

efficiency in 11 soybean genotypes before and during mild drought. J. Exp. Bot. 62: 2875–2887.

Grant, D., R.T. Nelson, S.B. Cannon, and R.C. Shoemaker. 2010. SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res. 38: 843–846.

Hall, A.E., A. Richard, A.G. Condon, G.C. Wright, and G. Farquhar. 1994. Carbon isotope discrimination and plant breeding. p. 81–113. *In* Plant Breeding Reviews.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2010. Estimating and interpreting heritability for plant breeding: An update. *In* Janick, J. (ed.), Plant Breeding Reviews. John Wiley and Sons, Inc., Oxford, UK.

Houngnandan, P., R.G.H. Yemadje, S.O. Oikeh, C.F. Djidohokpin, P. Boeckx, and O. Van Cleemput. 2008. Improved estimation of biological nitrogen fixation of soybean cultivars (Glycine max L. Merril) using 15N natural abundance technique. Biol. Fertil. Soils 45(2): 175–183.

Hubick, K., and G. Farquhar. 1989. Carbon isotope discrimination and the ratio of carbon gained to water lost in barley cultivars. Plant. Cell Environ. 12: 795–804.

Hwang, S., C.A. King, M.K. Davies, J.D. Ray, P.B. Cregan, and L.C. Purcell. 2013. QTL analysis of shoot ureide and nitrogen concentrations in soybean [Glycine max (L.) Merr.]. Crop Sci. 53(6): 2421–2433.

Ismail, A.M., and A.E. Hall. 1992. Correlation between water-use efficiency and carbon isotope discrimination in diverse cowpea genotypes and isogenic lines. Crop Sci. 32(3): 7–12.

Kaler, A.S., A.P. Dhanapal, J.D. Ray, C.A. King, F.B. Fritschi, and L.C. Purcell. 2017a. Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. Crop Sci. 57: 1–16.

Kaler, A.S., J.D. Ray, W.T. Schapaugh, M.K. Davies, C.A. King, and L.C. Purcell. 2018. Association mapping identifies loci for canopy coverage in diverse soybean genotypes. Mol. Breed. 38: 50.

Kaler, A.S., J.D. Ray, W.T. Schapaugh, C.A. King, and L.C. Purcell. 2017b. Genome-wide association mapping of canopy wilting in diverse soybean genotypes. Theor. Appl. Genet. 130(10): 2203–2217.

King, C.A., and L.C. Purcell. 2006. Genotypic variation for shoot N concentration and response to water deficits in soybean. Crop Sci. 46(6): 2396–2402.

King, C.A., L.C. Purcell, A. Bolton, and J.E. Specht. 2014. A possible relationship between shoot N concentration and the sensitivity of N2 fixation to drought in soybean. Crop Sci. 54(2): 746–756.

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

Macková, H., M. Hronková, J. Dobrá, V. Turečková, O. Novák, Z. Lubovská, V. Motyka, D. Haisel, T. Hájek, I.T. Prášil, A. Gaudinová, H. Štorchová, E. Ge, T. Werner, T. Schmülling, and R. Vanková. 2013. Enhanced drought and heat stress tolerance of tobacco plants with ectopically enhanced cytokinin oxidase/dehydrogenase gene expression. J. Exp. Bot. 64(10): 2805–2815.

Mian, M.A.R., D.A. Ashley, and H.R. Boerma. 1998. An additional QTL for water use efficiency in soybean. Crop Sci. 38(2): 390–393.

Mian, M.A.R., T.E. Carter, W.A. Parrott, R. Wells, M.A. Bailey, D.A. Ashley, and H.R. Boerma. 1996. Molecular markers associated with water use efficiency and leaf ash in soybean. Crop

Sci. 36: 1252–1257.

Mirzaei, M., D. Pascovici, B.J. Atwell, and P.A. Haynes. 2012. Differential regulation of aquaporins, small GTPases and V-ATPases proteins in rice leaves subjected to drought stress and recovery. Proteomics 12(6): 864–877.

O'Leary, M.H. 1981. Carbon isotope fractionation in plants. Phytochemistry 20(4): 553–567.

Rebetzke, G.J., A.G. Condon, G.D. Farquhar, R. Appels, and R.A. Richards. 2008. Quantitative trait loci for carbon isotope discrimination are repeatable across environments and wheat mapping populations. Theor. Appl. Genet. 118(1): 123–137.

Rebetzke, G.J., A.G. Condon, R.A. Richards, and G.D. Farquhar. 2002. Selection for reduced carbon isotope discrimination increases aerial biomass and grain yield of rainfed bread wheat. Crop Sci. 42(3): 739–745.

Riar, M.K., P. Cerezini, A. Manandhar, T.R. Sinclair, Z. Li, and T.E. Carter. 2018. Expression of drought-tolerant N fixation in heterogeneous inbred families derived from PI471938 and Hutcheson soybean. Crop Sci. 58: 364–369.

Ries, L.L., L.C. Purcell, T.E. Carter, J.T. Edwards, and C.A. King. 2012. Physiological traits contributing to differential canopy wilting in soybean under drought. Crop Sci. 52: 272–281.

Sack, L., and N.M. Holbrook. 2006. Leaf hydraulics. Annu. Rev. Plant Biol. 57(1): 361–381.

Sadok, W., and T.R. Sinclair. 2010a. Genetic variability of transpiration response of soybean [Glycine max (L.) Merr.] shoots to leaf hydraulic conductance inhibitor AgNO3. Crop Sci. 50(4): 1423–1430.

Sadok, W., and T.R. Sinclair. 2010b. Transpiration response of "slow-wilting" and commercial soybean (Glycine max (L.) Merr.) genotypes to three aquaporin inhibitors. J. Exp. Bot.

61(3): 821–829.

Sall, K., and T.R. Sinclair. 1991. Soybean genotypic differences in sensitivity of symbiotic nitrogen fixation to soil dehydration. Plant Soil 133(1): 31–37.

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng, D. Xu, U. Hellsten, G.D. May, Y. Yu, T. Sakurai, T. Umezawa, M.K. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, B. Abernathy, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.-C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, and S.A. Jackson. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463(7278): 178–183.

Serraj, R., and T.R. Sinclair. 1997. Variation among soybean cultivars in dinitrogen fixation response to drought. Agron. J. 89: 963–969.

Shearer, G., D.H. Kohl, and J.E. Harper. 1980. Distribution of N among plant parts of nodulating and nonnodulating isolines of soybeans. Plant Physiol. 66: 57–60.

Sinclair, T. 2012. Is transpiration efficiency a viable plant trait in breeding for crop improvement? Funct. Plant Biol. 39: 359–365.

Sinclair, T.R., J. Devi, A. Shekoofa, S. Choudhary, W. Sadok, V. Vadez, M. Riar, and T. Rufty. 2017. Limited-transpiration response to high vapor pressure deficit in crop species. Plant Sci. 260: 109–118.

Sinclair, T.R., C.D. Messina, A. Beatty, and M. Samples. 2010. Assessment across the United States of the benefits of altered soybean drought traits. Agron. J. 102(2): 475–482.

Sinclair, T.R., L.C. Purcell, C.A. King, C.H. Sneller, P. Chen, and V. Vadez. 2007. Drought tolerance and yield increase of soybean resulting from improved symbiotic N2 fixation. F.

Crop. Res. 101: 68–71.

Sinclair, T.R., L.C. Purcell, V. Vadez, R. Serraj, C.A. King, and R. Nelson. 2000. Identification of soybean genotypes with N fixation tolerance to water deficits. Crop Sci. 40: 1803–1809.

Sinclair, T.R., M.A. Zwieniecki, and N.M. Holbrook. 2008. Low leaf hydraulic conductance associated with drought tolerance in soybean. Physiol. Plant. 132(4): 446–451.

Sloane, R.J., R.P. Patterson, and T.E. Carter. 1990. Field drought tolerance of a soybean plant introduction. Crop Sci. 30: 118–123.

Song, Q., D.L. Hyten, G. Jia, C. V Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS One 8(1): e54985.

Specht, J.E., K. Chase, M. Macrander, G.L. Graef, J. Chung, J.P. Markwell, M. Germann, J.H. Orf, and K.G. Lark. 2001. Soybean response to water: A QTL analysis of drought tolerance. Crop Sci. 41: 493–509.

Specht, J., D. Hume, and S. Kumudini. 1999. Soybean yield potential—A genetic and physiological perspective. Crop Sci. 39: 1560–1570.

Tang, Y., X. Liu, J. Wang, M. Li, Q. Wang, F. Tian, Z. Su, Y. Pan, D. Liu, A.E. Lipka, E.S. Buckler, and Z. Zhang. 2016. GAPIT version 2: An enhanced integrated tool for genomic association and prediction. Plant Genome 9(2): 1–9.

Turner, S.D. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv Prepr.

USDA-NASS. 2012. USDA-National Agricultural Statistics Service (USDA-NASS). Census Agric.: Table 37.

Wei, W., A.C.O. Mesquita, A. de A. Figueiró, X. Wu, S. Manjunatha, D.P. Wickland, M.E.

Hudson, F.C. Juliatti, and S.J. Clough. 2017. Genome-wide association mapping of resistance to a Brazilian isolate of Sclerotinia sclerotiorum in soybean genotypes mostly from Brazil. BMC Genomics 18: 849.

White, J.W. 1993. Implications of carbon isotope discrimination studies for breeding common bean under water deficits. p. 387–398. *In* Ehleringer, J.R., Hall, A.E., Farquhar, G.D. (eds.), Stable Isotopes and Plant Carbon-water Relations. Academic Press, San Diego.

Wright, G.C., R.C. Rao, and G.D. Farquhar. 1994. Water-use efficiency and carbon isotope discrimination in peanut under water deficit conditions. Crop Sci. 34: 92–97.

Zeng, A., P. Chen, K. Korth, F. Hancock, A. Pereira, K. Brye, C. Wu, and A. Shi. 2017. Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. Mol. Breed. 37: 30.

Figures and tables

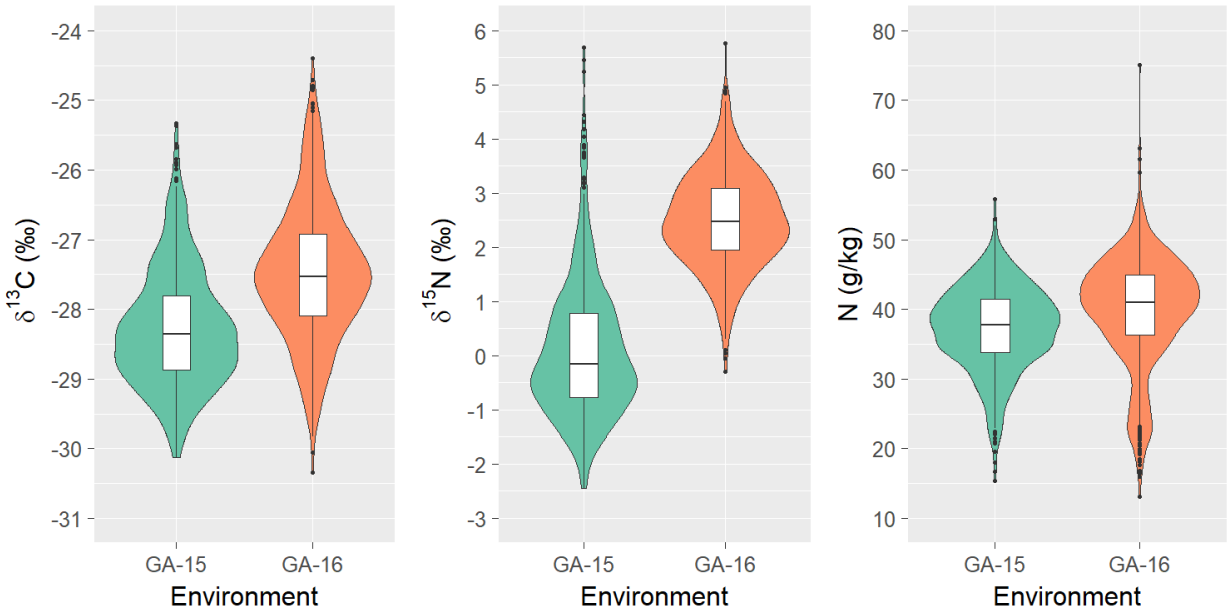


Figure 3.1. Violin plots with boxplots inside showing distributions of individual plot data for carbon and nitrogen related traits evaluated in two environments with association panel.

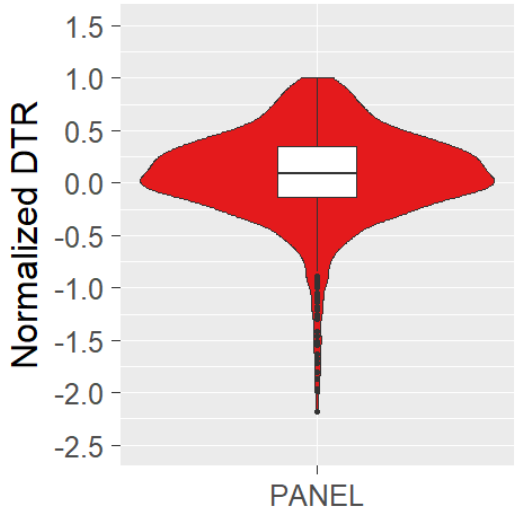


Figure 3.2. Violin plot with boxplot inside showing distribution of individual observations of normalized decrease in transpiration response (NDTR) to silver nitrate inhibitor for the association panel across eight experimental replications. DTR values were normalized by the lowest DTR value in each separate experimental replication in both populations to calculate NDTR.

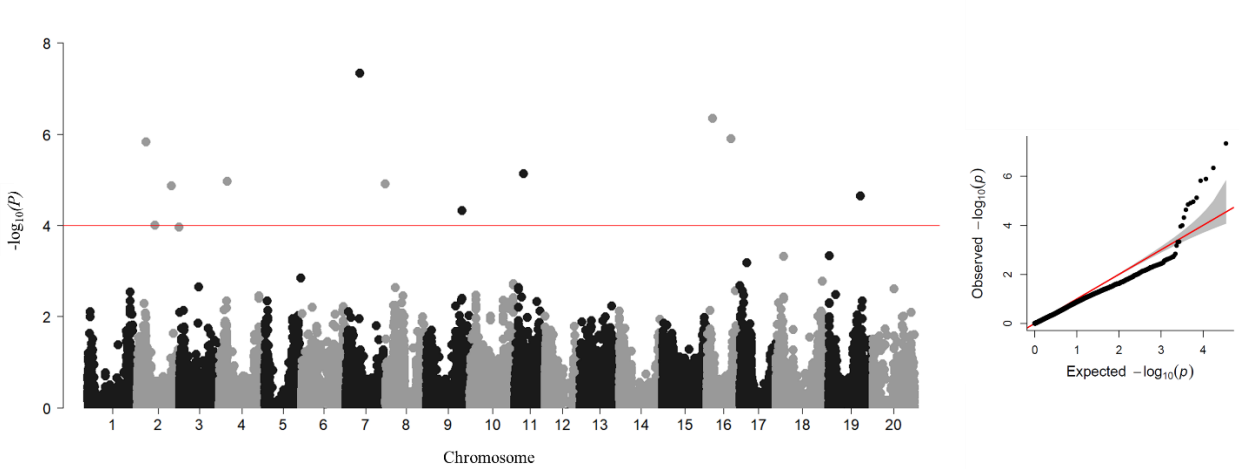
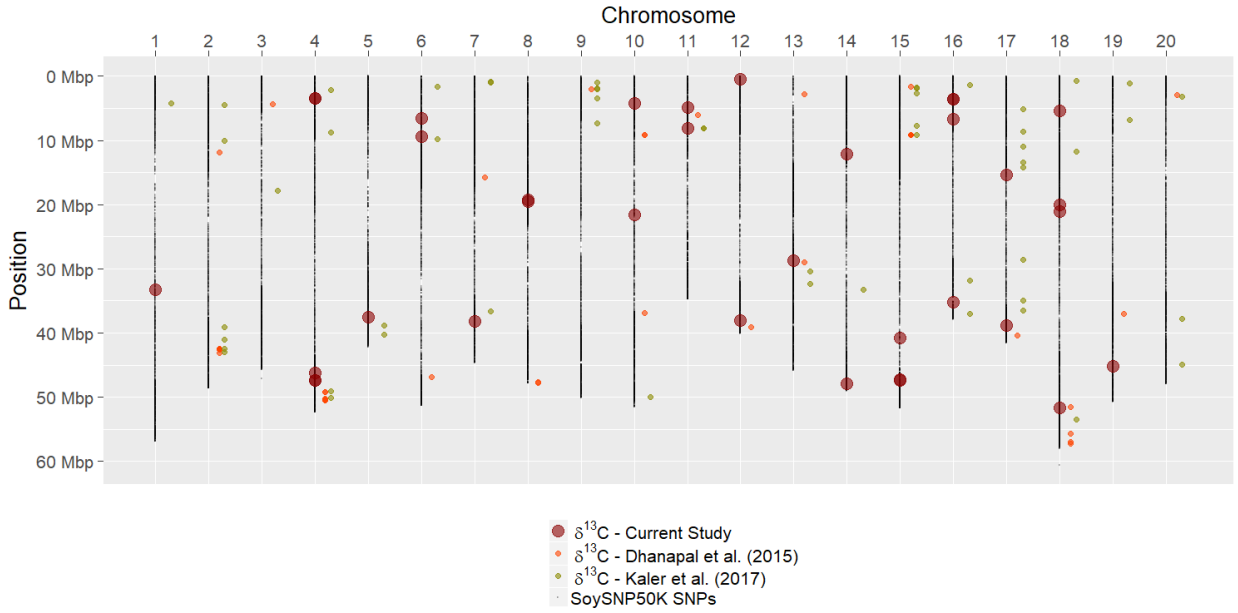
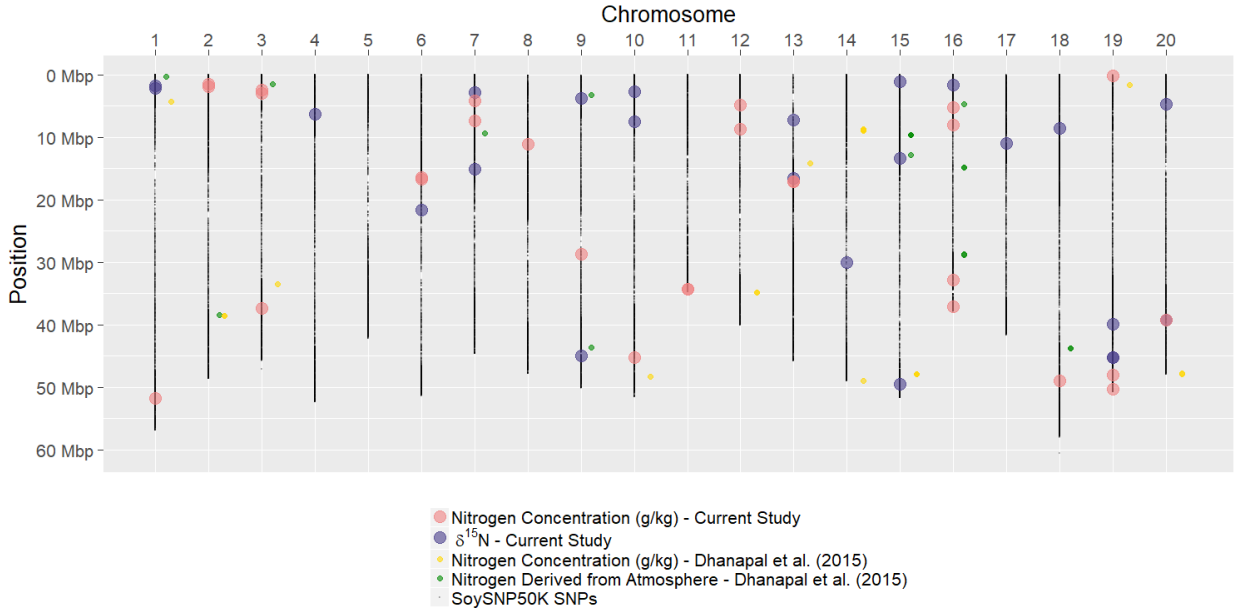


Figure 3.3. Genome-wide Manhattan plot for normalized decrease in transpiration rate (NDTR) to silver nitrate. The X-axis is the genomic position of SNPs by chromosome across the soybean genome, and the Y-axis is the $-\log_{10}$ of the p-values obtained from the GWAS model. Significance threshold $-\log_{10}(P) > 4$ (red line). The quantile-quantile (QQ) plot to the right of the Manhattan plot shows the expected versus observed p-values of each SNP tested in the GWAS model.

A


Chromosome
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Position
0 Mbp
10 Mbp
20 Mbp
30 Mbp
40 Mbp
50 Mbp
60 Mbp
δ13C - Current Study
δ13C - Dhanapal et al. (2015)
δ13C - Kaler et al. (2017)
SoySNP50K SNPs

B


Chromosome
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Position
0 Mbp
10 Mbp
20 Mbp
30 Mbp
40 Mbp
50 Mbp
60 Mbp
Nitrogen Concentration (g/kg) - Current Study
δ15N - Current Study
Nitrogen Concentration (g/kg) - Dhanapal et al. (2015)
Nitrogen Derived from Atmosphere - Dhanapal et al. (2015)
SoySNP50K SNPs

C

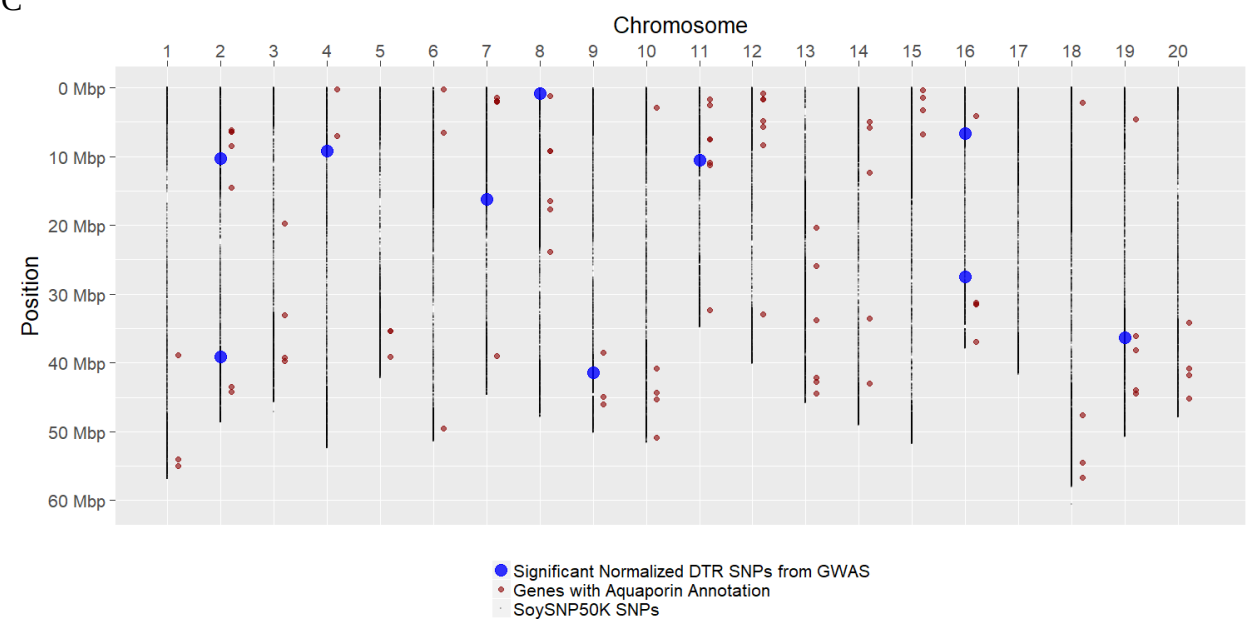


Figure 3.4. Location and comparison of SNPs significantly associated with drought tolerance related traits. Physical positions are based on the Glyma.Wm82.a2 version of the soybean genome. SNPs identified in GWAS from current study that met $-\log_{10}(P) > 4$ significance threshold are shown as larger circles for A) carbon isotope composition ($\delta^{13}C$), B) Nitrogen concentration and nitrogen isotope composition ($\delta^{15}N$), and C) normalized decrease in transpiration rate (NDTR) to silver nitrate aquaporin inhibitor. Smaller circles represent SNPs identified in A) Dhanapal et al. (2015b) or Kaler et al. (2017a), B) Dhanapal et al. (2015a) that were converted from version 1 to 2 physical positions of the soybean genome assembly, and C) location of gene models with the term "aquaporin" in their functional annotation from Phytozome v12.1. BARC_1.01_Gm20_46575262_G_A identified for nitrogen concentration in Dhanapal et al. (2015a) does not have a perfect match in the version 2 assembly, and therefore was excluded from this comparison.

Table 3.1. Summary of analyses of variance (ANOVA) for each trait evaluated.

Carbon Isotope Composition ($\delta^{13}C$)				Nitrogen Isotope Composition ($\delta^{15}N$)			
Source	DF	F Value	P > F	Source	DF	F Value	P > F
Genotype (G)	208	12.1	<0.0001	Genotype (G)	208	3.1	<0.0001
Environment (E)	1	834.3	<0.0001	Environment (E)	1	2440.1	<0.0001
G × E	194	1.6	<0.0001	G × E	194	1.6	<0.0001
Nitrogen Concentration [N]				Normalized DTR to $AgNO_3$			
Source	DF	F Value	P > F	Source	DF	F Value	P > F
Genotype (G)	208	12.4	<0.0001	Genotype (G)	210	2.1	<0.0001
Environment (E)	1	284.0	<0.0001				
G × E	194	1.7	<0.0001				

Table 3.2. Broad-sense heritability on an entry-mean basis for carbon and nitrogen related traits, and normalized decrease in transpiration rate (NDTR) to silver nitrate.

Trait	Both	GA-15	GA-16
	---------- Heritability (%) ----------		
Carbon Isotope Composition ($\delta^{13}C$)	62	61	72
Nitrogen Isotope Composition ($\delta^{15}N$)	17	24	40
Nitrogen Concentration [N]	64	63	73

Trait	Panel
	Heritability (%)
Normalized DTR to Silver Nitrate	17

Table 3.3. Correlations among all traits evaluated in the current study as well as canopy wilting data from Chapter 2 using best linear unbiased predictions (BLUPs) from across all replications and environments.

	$\delta^{13}C$	$\delta^{15}N$	[N]	NDTR to AgNO3	Canopy Wilting
$\delta^{13}C$	1.00				
$\delta^{15}N$	-0.52	1.00			
[N]	0.71	-0.50	1.00		
NDTR to AgNO3	-0.03	-0.04	-0.01	1.00	
Canopy Wilting	-0.08	-0.02	0.08	-0.02	1.00

Table 3.4. SNPs identified in a single environment or when using the BLUPs from both environments for carbon and nitrogen related traits that met the significance threshold level of $-\log_{10}(P) > 4$.

Carbon Isotope Composition								
Locus[a]	Chr.[b]	Pos.[c]	SNP	$-\log_{10}(P)$	MAF[d]	Effect[e]	R^2 (%)	Env[f]
1	1	33203133	ss715578992	7.11	0.27	-0.12	0.15	Both
2	4	3418112	ss715587736	5.36	0.44	-0.09	0.65	GA-15
	4	3425900	ss715587739	6.90	0.46	-0.10	1.01	Both
3	4	46166265	ss715588297	4.40	0.45	0.06	0.20	GA-15
4	4	47373969	ss715588481	4.23	0.19	0.10	0.03	GA-15
	4	47376582	ss715588482	4.11	0.19	0.10	0.01	GA-15
5	5	37563155	ss715591464	5.44	0.42	-0.09	0.33	GA-16
6	6	6576054	ss715595435	5.61	0.47	0.08	0.13	Both
7	6	9451023	ss715595676	7.29	0.09	-0.18	0.00	GA-15
8	7	38213845	ss715597738	4.96	0.19	0.10	10.54	Both
9	8	19267914	ss715600198	6.65	0.19	-0.13	3.04	GA-15
10	8	19518756	ss715600277	5.87	0.32	0.11	0.04	GA-15
11	10	4260367	ss715607234	8.65	0.36	0.10	0.24	Both
12	10	21586075	ss715605850	5.49	0.47	0.10	15.20	GA-16
13	11	4875880	ss715610795	5.67	0.41	-0.11	13.33	GA-16
14	11	8151411	ss715611206	5.42	0.13	0.13	12.31	GA-16
15	12	458748	ss715613097	5.79	0.45	-0.09	1.24	GA-16
16	12	38049740	ss715612828	6.25	0.21	-0.12	1.09	GA-16
17	13	28776094	ss715614695	4.89	0.36	-0.07	0.17	GA-15
18	14	12079082	ss715617567	4.45	0.47	-0.07	1.35	GA-15
19	14	47854709	ss715619453	5.15	0.43	-0.09	14.58	GA-15
20	15	40841088	ss715621829	4.25	0.12	-0.10	2.68	GA-15
21	15	47257859	ss715622121	6.99	0.07	-0.19	1.69	Both
22	15	47349730	ss715622149	7.93	0.12	-0.16	10.25	GA-15
23	16	3557974	ss715624794	4.71	0.33	0.08	0.07	Both
	16	3566872	ss715624799	5.15	0.37	0.09	0.45	Both
24	16	6706066	ss715625333	5.68	0.48	-0.08	0.00	Both
25	16	35166856	ss715624733	8.21	0.23	-0.12	13.30	Both
26	17	15380811	ss715626252	4.86	0.24	-0.12	0.04	GA-15
27	17	38826185	ss715627535	4.06	0.10	-0.12	0.04	GA-16
28	18	5429903	ss715631531	4.52	0.39	-0.08	1.14	GA-16
29	18	20093832	ss715629730	6.62	0.18	-0.12	0.04	GA-15
30	18	21021784	ss715629903	4.07	0.19	-0.12	0.56	GA-16
31	18	51704746	ss715631722	6.61	0.42	-0.09	0.43	Both
32	19	45240169	ss715635451	6.39	0.26	0.10	13.82	Both

Nitrogen Isotope Composition								
Locus	Chr.	Pos.	SNP	$-\log_{10}(P)$	MAF	Effect	R^2 (%)	Env
1	1	1756948	ss715578613	7.63	0.50	-0.05	4.33	Both
2	1	2126801	ss715578694	7.97	0.30	-0.10	2.23	GA-16
3	4	6329113	ss715589139	6.84	0.28	0.05	1.35	GA-15
4	6	21606676	ss715593886	5.99	0.12	0.07	17.81	GA-15
5	7	2811470	ss715597004	5.31	0.29	-0.05	5.05	Both
6	7	15036339	ss715596324	5.19	0.14	0.10	18.82	GA-16
7	9	3732795	ss715603834	8.26	0.22	-0.06	12.19	GA-15
8	9	45017460	ss715604529	7.87	0.32	-0.13	16.85	GA-16
9	10	2699011	ss715606028	5.10	0.08	-0.14	0.22	GA-16
10	10	7565702	ss715608519	6.18	0.20	-0.06	6.20	GA-15
11	13	7212966	ss715617100	5.54	0.07	0.08	1.57	GA-15
12	13	16630119	ss715616751	5.09	0.15	0.05	0.01	Both
13	14	30072552	ss715618124	9.22	0.11	0.11	14.60	Both
14	15	1121373	ss715620300	4.46	0.09	0.05	10.71	GA-15
15	15	13304091	ss715620571	4.71	0.23	-0.05	1.83	Both
16	15	49446994	ss715622476	6.61	0.33	0.04	0.01	GA-15
17	16	1675623	ss715623543	5.40	0.14	0.07	3.65	Both
18	17	11003712	ss715625747	4.24	0.34	-0.03	4.23	GA-15
19	18	8504254	ss715632791	4.11	0.20	-0.08	1.28	GA-16
20	19	39924653	ss715634905	7.68	0.49	-0.04	1.94	GA-15
21	19	45292930	ss715635458	4.44	0.37	0.08	9.51	GA-16
	19	45292930	ss715635458	5.35	0.37	0.06	6.86	Both
22	20	4645190	ss715638934	5.01	0.35	0.04	0.06	Both
23	20	39218472	ss715638011	4.19	0.24	-0.06	0.81	GA-16

Nitrogen Concentration								
Locus	Chr.	Pos.	SNP	$-\log_{10}(P)$	MAF	Effect	R^2 (%)	Env
1	1	51706358	ss715580153	5.08	0.09	-1.14	0.00	GA-16
2	2	1482658	ss715581317	4.16	0.24	0.62	4.92	GA-16
3	2	1864987	ss715581422	4.93	0.05	1.46	6.31	GA-16
4	3	2475142	ss715584877	5.60	0.45	-0.61	4.60	GA-15
5	3	2945818	ss715585023	8.03	0.15	-1.10	5.22	Both
6	3	37336737	ss715585803	8.25	0.27	-0.94	8.10	GA-15
7	6	16421594	ss715593518	9.09	0.28	-1.09	4.83	GA-15
8	6	16706510	ss715593613	4.49	0.27	-0.63	0.21	GA-15
9	7	4232510	ss715598067	7.40	0.15	1.16	16.59	GA-15
10	7	7433625	ss715598611	4.01	0.38	-0.56	2.45	GA-16

11	8	11126044	ss715599253	4.64	0.45	-0.58	1.06	GA-15
12	9	28739753	ss715603408	5.06	0.22	-0.79	12.89	GA-16
13	10	45301855	ss715607477	4.01	0.22	0.63	0.06	GA-15
14	11	34311552	ss715610522	4.50	0.33	0.58	3.77	GA-15
	11	34311552	ss715610522	4.14	0.33	0.53	1.88	Both
15	12	4839133	ss715613118	8.74	0.24	1.01	13.75	Both
16	12	8677962	ss715613605	5.53	0.10	1.05	5.95	GA-16
17	13	17044187	ss715616699	8.69	0.49	0.95	0.12	GA-16
	13	17062998	ss715616695	4.69	0.42	-0.55	0.10	Both
18	16	5223380	ss715625193	5.37	0.30	0.64	0.31	Both
19	16	8026492	ss715625562	5.89	0.17	-0.91	0.29	Both
20	16	32772447	ss715624545	7.06	0.46	0.99	7.57	GA-16
21	16	37144699	ss715624944	5.31	0.11	-1.17	0.00	GA-16
22	18	49000413	ss715631434	5.19	0.18	-0.84	1.35	GA-15
23	19	179420	ss715635206	6.10	0.45	0.70	0.01	Both
24	19	48043481	ss715635757	4.57	0.38	0.68	1.21	Both
25	19	50357367	ss715636025	4.67	0.36	-0.66	0.00	GA-16
26	20	39264651	ss715638016	9.63	0.14	-1.33	19.69	GA-15

[a] If multiple SNPs were identified in the same linkage disequilibrium (LD) block they were deemed part of the same locus (genomic region).

[b] Chromosome.

[c] Glyma.Wm82.a2 physical position.

[d] Minor allele frequency.

[e] Allelic effects were calculated by taking the difference in mean phenotypic value between the two alleles at a particular SNP, and the direction, negative or positive, of the allelic effect estimates are relative to the alphabetical order of the nucleotides at each particular marker.

[f] Environment written as location-year.

Table 3.5. SNPs that met significance level of $-\log_{10}(P) > 4$ for normalized decrease in transpiration rate (NDTR) to silver nitrate aquaporin inhibitor.

Locus[a]	Chr.[b]	Pos.[c]	SNP	$-\log_{10}(P)$	MAF[d]	Effect[e]	R^2 (%)
1	2	10320993	ss715580906	5.82	0.40	-0.02	7.72
2	2	39085373	ss715582355	4.86	0.31	-0.02	1.82
3	4	9244048	ss715589614	4.96	0.14	0.02	0.00
4	7	16199323	ss715596527	7.34	0.17	0.03	2.64
5	8	797007	ss715602728	4.91	0.18	0.02	4.56
6	9	41392225	ss715604054	4.32	0.30	0.01	3.03
7	11	10575472	ss715608766	5.12	0.07	-0.03	0.42
8	16	6643454	ss715625324	6.34	0.43	0.02	8.62
9	16	27510603	ss715623885	5.90	0.23	-0.02	1.46
10	19	36365868	ss715634460	4.64	0.20	0.02	6.01

[a] If multiple SNPs were identified in the same linkage disequilibrium (LD) block they were deemed part of the same locus (genomic region).

[b] Chromosome.

[c] Glyma.Wm82.a2 physical position.

[d] Minor allele frequency.

[e] Allelic effects were calculated by taking the difference in mean phenotypic value between the two alleles at a particular SNP, and the direction, negative or positive, of the allelic effect estimates are relative to the alphabetical order of the nucleotides at each particular marker.

Appendix A: Supplemental figures

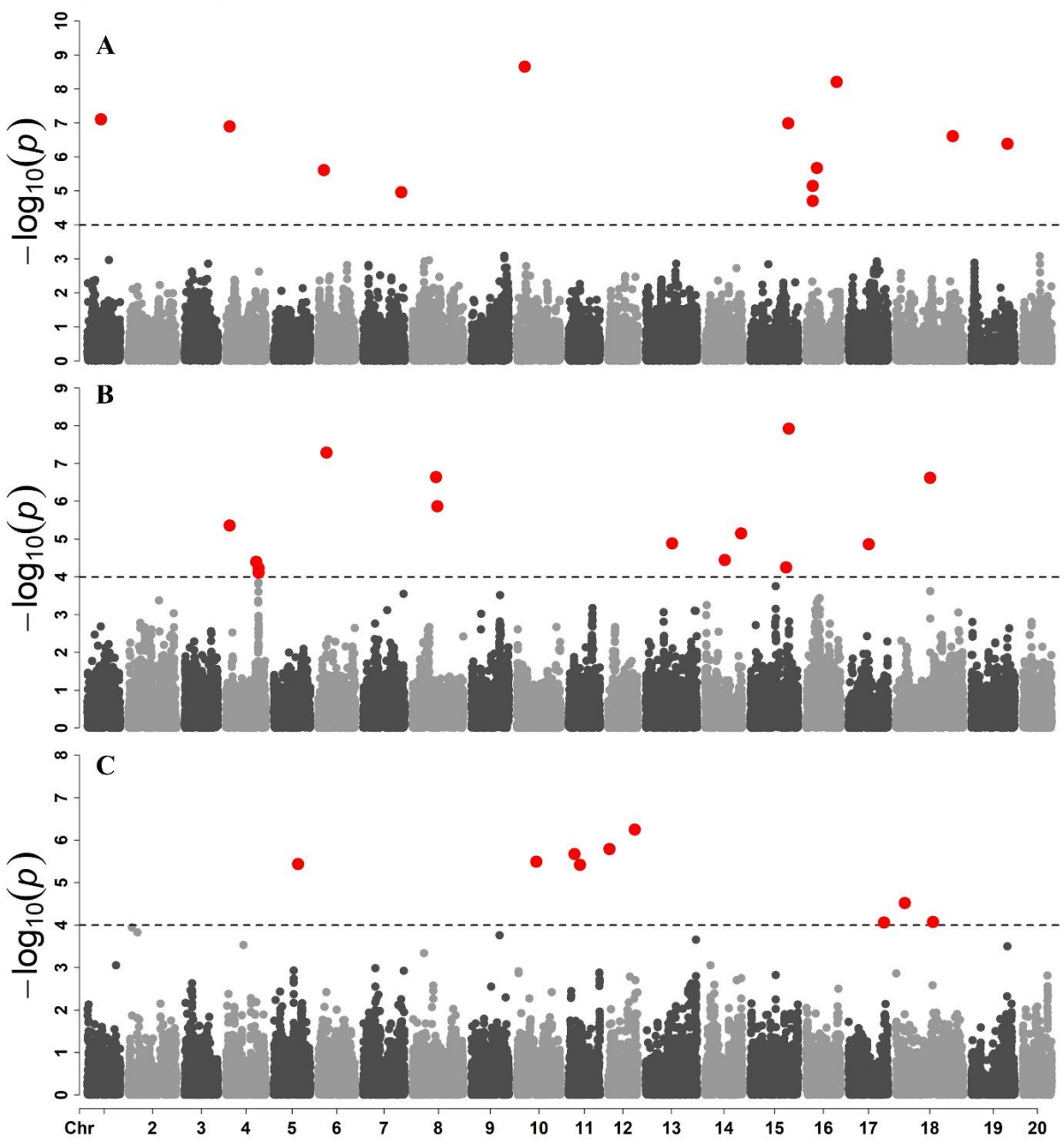


Figure 3.S1. Manhattan plots for carbon isotope composition ($\delta^{13}C$) for A) Both, B) GA-15, and C) GA-16 environments. The X-axis is the genomic position of SNPs across the soybean genome by chromosome, and the Y-axis is the -$\log_{10}$ of the p-values obtained from the GWAS model. SNPs that were above significance threshold ($-\log_{10}(P) > 4$) are colored in red and enlarged.

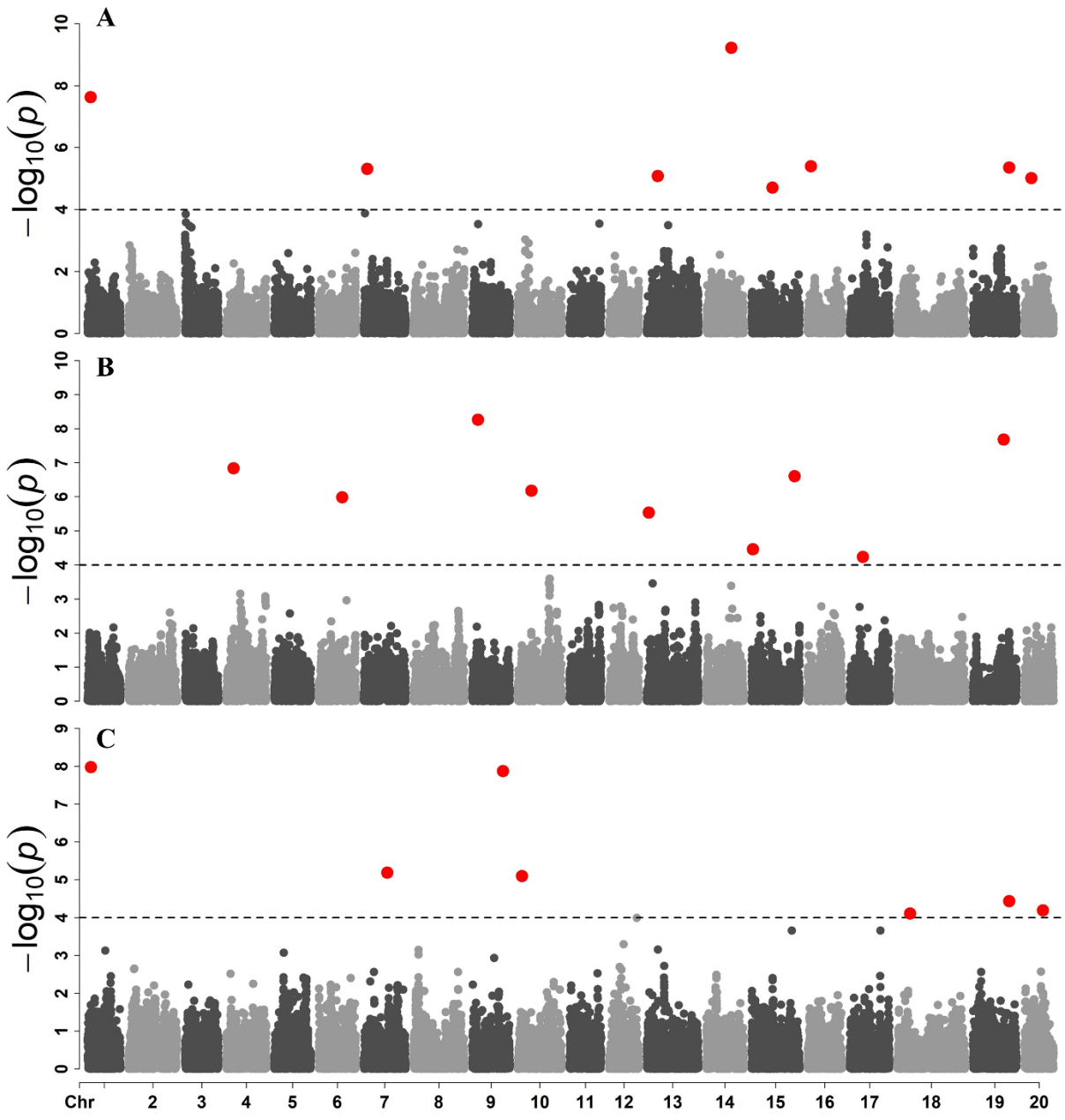


Figure 3.S2. Manhattan plots for nitrogen isotope composition ($\delta^{15}N$) for A) Both, B) GA-15, and C) GA-16 environments. The X-axis is the genomic position of SNPs across the soybean genome by chromosome, and the Y-axis is the $-\log_{10}$ of the p-values obtained from the GWAS model. SNPs that were above significance threshold ($-\log_{10}(P) > 4$) are colored in red and enlarged.

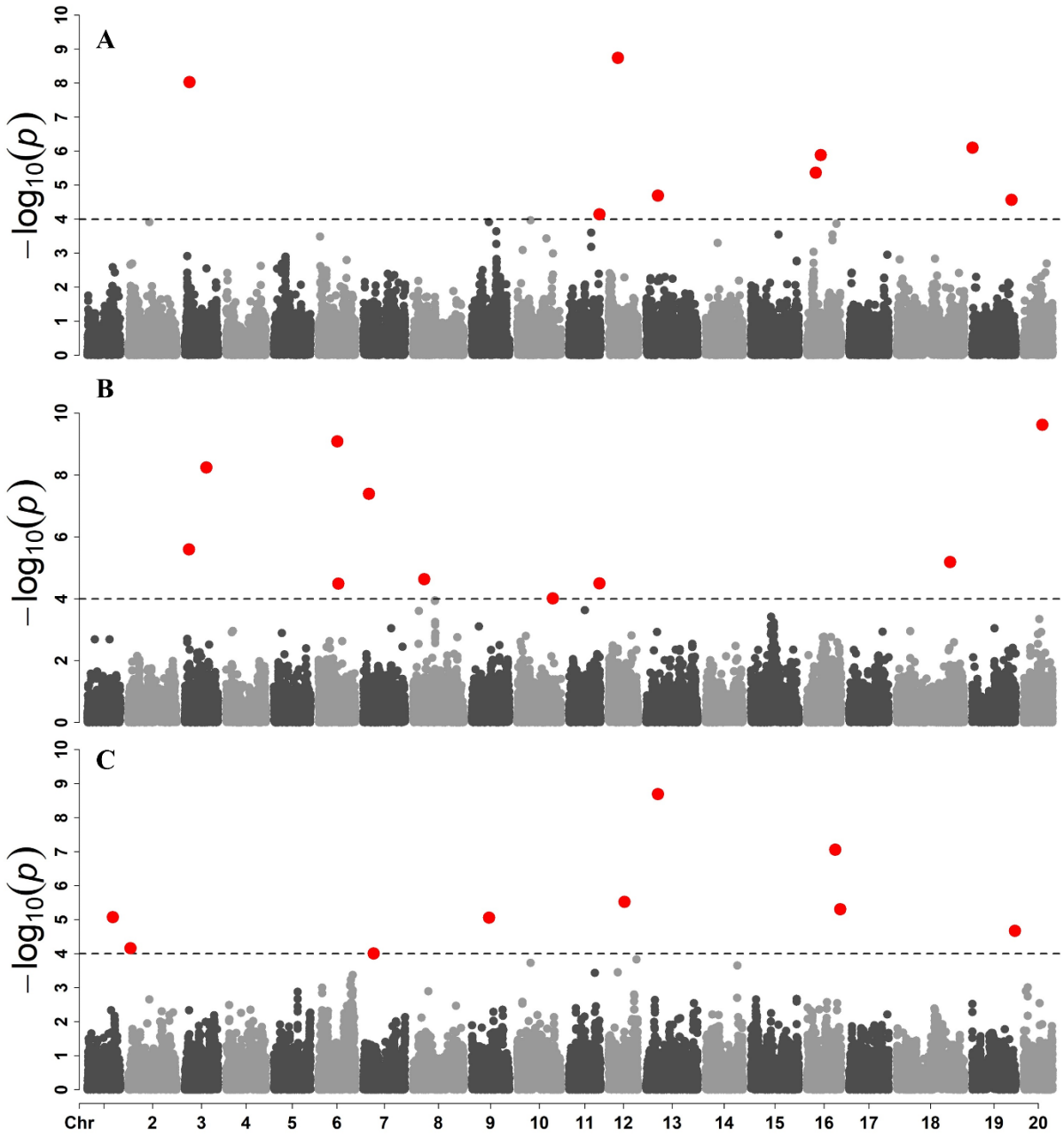


Figure 3.S3. Manhattan plots for nitrogen concentration for A) Both, B) GA-15, and C) GA-16 environments. The X-axis is the genomic position of SNPs across the soybean genome by chromosome, and the Y-axis is the -$\log_{10}$ of the p-values obtained from the GWAS model. SNPs that were above significance threshold (-$\log_{10}(P)$ > 4) are colored in red and enlarged.

Appendix B: Supplemental tables

Table 3.S1. Summary of transpiration response to silver nitrate experiments for the association panel. Temperature and relative humidity were measured during these experiments with two data loggers. Vapor pressure deficit (VPD) is shown as the average VPD from the time of the first weighing of the de-rooted shoots in deionized water to the final weighing of the de-rooted shoots in silver nitrate solution, and was averaged between the two data loggers.

Replicate	Measurement Date	Average VPD
1	4/1/2015	2.33
2	4/10/2015	1.92
3	4/15/2015	1.91
4	4/22/2015	2.24
5	6/23/2015	1.63
6	6/24/2015	1.56
7	10/20/2015	2.29
8	3/31/2016	1.74

Table 3.S2. Candidate genes and their functional annotation identified using the Glyma2.1 gene models in SoyBase (www.soybase.org) within plus or minus 10 kb of SNPs significantly associated with carbon and nitrogen related traits from GWAS which explained greater than 15% of the phenotypic variation for the trait.

Carbon Isotope Composition				
Locus	SNP	Gene Name	Annotation	Gene Ontology Terms
12	ss715605850	Glyma.10g103000	Elongation factor family protein	GTPase activity; GTP binding; intracellular; small GTPase mediated signal transduction

Nitrogen Isotope Composition				
Locus	SNP	Gene Name	Annotation	Gene Ontology Terms
4	ss715593886	No gene models	N/A	N/A
6	ss715596324	Glyma.07g125800	Protein of unknown function (DUF1005)	N/A
		Glyma.07g125900	HAD superfamily, subfamily IIIB acid phosphatase	Acid phosphatase activity
		Glyma.07g126000	Leucine-rich repeat (LRR) family protein	Protein binding
8	ss715604529	Glyma.09g225400	Cytokinin oxidase/dehydrogenase 6	UDP-N-acetylmuramate dehydrogenase activity; cytokinin metabolic process; oxidoreductase activity; cytokinin dehydrogenase activity; flavin adenine dinucleotide binding; oxidation-reduction process

Nitrogen Concentration				
Locus	SNP	Gene Name	Annotation	Gene Ontology Terms
9	ss715598067	Glyma.07g049700	Transducin family protein / WD-40 repeat family protein	Intracellular protein transport; vesicle-mediated transport
		Glyma.07g049800	TTF-type zinc finger protein with HAT dimerisation domain	Protein dimerization activity
26	ss715638016	Glyma.20g153600	Phosphoglucomutase	Carbohydrate metabolic process; intramolecular transferase activity, phosphotransferases
		Glyma.20g153700	K-box region and MADS-box transcription factor family protein	DNA binding; sequence-specific DNA binding transcription factor activity; nucleus; regulation of transcription, DNA-templated; protein dimerization activity

Table 3.S3. Breeding value ranks for accessions tested for five different traits: canopy wilting (from Chapter 2), carbon isotope composition ($\delta^{13}C$), nitrogen concentration, nitrogen isotope composition ($\delta^{15}N$), and normalized decrease in transpiration (NDTR) rate to silver nitrate ($AgNO_3$). Only genotypes that were tested for all five traits are shown in this table.

Accession	Name	Country	MG	Canopy Wilting	$\delta^{13}C$	[N]	$\delta^{15}N$	NDTR to $AgNO_3$	Median Rank
				----Breeding Value Rank[a]----					
Fendou78	-	China	IV	11	93	14	68	5	14
PI574484	Jin dou No. 6	China	IV	2	176	20	76	1	20
PI548983	Tracy	United States	VI	10	20	140	42	17	20
PI170886	-	South Africa	VI	174	16	174	13	23	23
PI398823	-	South Korea	IV	24	1	5	110	68	24
PI567036	-	Morocco	IX	26	4	175	36	10	26
PI592939	Jin dou 16	China	IV	6	3	106	70	26	26
PI090499	Black and white	China	VI	152	28	6	197	23	28
PI341248	Sangalo	Tanzania	IX	29	29	84	18	148	29
Jindou19	-	China	IV	3	95	30	180	7	30
PI603513A	Xiao niu mao huang	China	VIII	22	203	199	32	29	32
PI322694	Hernnon	Zimbabwe	VI	51	61	18	30	36	36
PI578495	Jin dou No. 4	China	IV	37	58	33	142	1	37
PI567316B	(Hong huang dou)	China	VI	38	40	109	29	120	40
PI495016	Nuwara Eliya Local	Sri Lanka	IX	148	43	158	33	21	43
PI170890	-	South Africa	VI	121	45	176	14	8	45
PI603566	Jin dou No. 4	China	III	20	100	46	78	11	46
PI306702A	3H/1	Kenya	IX	126	18	9	122	49	49
PI603540A	Hei huang dou	China	VII	50	37	185	129	39	50
PI592937	Jin dou 14	China	IV	169	140	19	25	54	54
PI423927	Tousan 93	Japan	IV	54	51	36	71	67	54
PI159093	34S51	South Africa	VII	177	54	29	46	79	54
PI322692	Max C.P1159A8	Australia	IX	53	86	123	21	55	55
PI429328	-	Nigeria	VIII	21	98	32	126	56	56
PI574483	Jin dou No. 5	China	IV	57	75	10	97	11	57
PI212605	-	Afghanistan	VI	120	191	57	38	14	57
PI578494A	Jin dou No. 1	China	IV	46	58	94	142	1	58
PI567356	Zao bai huang dou	China	VI	5	27	181	59	151	59
PI603588	Jing si dou	China	V	61	127	48	15	192	61

PI374219	Blyvoor	South Africa	VI	187	48	120	17	63	63
Fendou65	-	China	IV	1	94	53	104	64	64
PI556949	Ke feng No. 1	China	IV	193	173	3	12	64	64
PI381662	Hernon 49	Uganda	VI	180	65	127	35	35	65
PI639576	-	Burundi	VIII	79	107	34	65	36	65
PI210349	Jubiltan 65	Mozambique	VIII	65	139	79	28	50	65
PI360846	Shiroge-9	Japan	IV	66	82	58	56	108	66
PI531068	Stonewall	United States	VII	67	22	136	199	6	67
PI567326B	(Huang dou)	China	VI	45	154	67	4	192	67
PI341241A	Seminole	Israel	IX	114	118	24	69	29	69
PI341244B	(Yellow Kedele)	Tanzania	IX	39	106	69	53	180	69
PI269518B	(Koolat)	Pakistan	VI	84	174	71	26	14	71
PI398276	Chirpan 90 (Bulgaria)	South Korea	IV	129	71	35	169	36	71
PI603538C	(Wan dou zao)	China	VIII	108	70	177	60	72	72
PI567758	Pei xian tu shan da ping ding huang	China	IV	48	72	2	134	118	72
PI430737	Oribi	Zimbabwe	VII	203	57	37	72	146	72
PI219698	Kulat	Pakistan	VI	72	89	49	23	169	72
PI639572	-	Ghana	VIII	44	74	114	103	27	74
PI462312	Ankur	India	VIII	95	19	64	74	120	74
PI374180	-	India	VIII	49	119	55	75	147	75
PI574485	Jin dou No. 9	China	IV	46	100	76	78	1	76
PI279081	Masterpiece	South Africa	VII	204	76	137	3	8	76
PI603517A	Lao shu pi	China	VI	25	77	135	81	72	77
PI574486	Jin dou 13	China	III	168	178	76	78	22	78
N05-7432	-	United States	VIII	8	79	143	128	78	79
PI171441	Mud-bean	China	VI	202	199	45	6	79	79
PI381661	Bukalasa 6	Uganda	VIII	144	80	16	108	31	80
N04-9646	-	United States	VII	82	81	128	125	71	82
PI567205	GL2671/89	Georgia	VI	161	83	145	39	41	83
PI603534B	(Da niu mao huang)	China	VIII	12	204	165	49	83	83
PI417561	48.S.103 DL/63/180	South Africa	VI	75	32	83	114	88	83
PI090406	-	China	VI	133	24	81	109	84	84
PI486330	Macs-75	India	VIII	43	84	68	160	132	84
PI548980	Hood	United States	VI	4	141	41	84	160	84
PI434981	Indo 226	Central African Republic	VIII	76	87	117	7	141	87

PI416937	Houjaku Kuwazu	Japan	VI	69	15	189	146	88	88
PI381680	S7	Uganda	VII	85	23	188	88	143	88
PI346300	-	India	VII	30	192	86	88	163	88
PI603537D	(Niu yan jing quan zi)	China	VII	89	41	80	140	182	89
PI381657	3H55 F4/9/2	Uganda	VIII	158	45	92	77	90	90
PI548975	Centennial	United States	VI	56	39	90	198	114	90
PI603506	Xiao ke zao huang dou	China	VI	90	136	89	16	176	90
PI595645	Benning	United States	VII	91	92	39	184	19	91
PI615694	N7001	United States	VII	23	47	204	186	91	91
WOODRUFF	-	United States	VII	16	49	179	194	91	91
G00-3213	-	United States	VII	42	49	179	172	91	91
PI602597	Boggs	United States	VI	140	67	59	200	91	91
PI548659	Braxton	United States	VII	73	169	42	190	91	91
PI548989	Ransom	United States	VII	15	78	150	195	91	91
PI518664	Hutcheson	United States	V	92	160	28	184	19	92
PI567683B	(Zheng zhou niu yao qi)	China	VI	157	88	15	92	139	92
PI429330	-	Nigeria	VIII	93	130	54	81	158	93
PI210350	Jubiltan 67	Mozambique	IX	183	137	47	94	56	94
PI567314	Hei you huang dou	China	VI	142	206	52	95	31	95
PI205384	-	Pakistan	VI	94	96	156	153	17	96
PI090768	-	China	VI	172	91	96	202	85	96
PI522236	Thomas	United States	VII	96	73	42	190	114	96
PI567334	Jiang dou zi	China	VI	113	36	56	96	205	96
PI247678	Herman	Zaire	VIII	143	149	97	9	77	97
PI458517	Xiao Wuyie	China	III	98	148	21	127	82	98
Fendou56	-	China	IV	9	30	98	149	120	98
PI159095	41S31	South Africa	VII	141	14	172	99	39	99
PI592756	Dillon	United States	VI	99	150	84	177	51	99
PI567405	Wei zi dou	China	VI	60	99	118	43	104	99
PI567207	-	Georgia	VI	127	102	107	40	34	102
PI548660	Bragg	United States	VII	102	115	72	192	91	102
PI381683	S36	Uganda	VI	153	6	50	173	102	102
PI322695	Bicolor do Cuima	Angola	VI	128	159	99	44	102	102

PI089775	-	China	VI	164	42	169	66	105	105
PI322689	Improved	Angola	VII	181	53	105	62	152	105
PI341264	-	Liberia	VI	68	31	173	105	153	105
PI567412	Yi wo feng	China	VI	198	105	8	98	208	105
PI203406	-	South Africa	VIII	123	9	38	106	192	106
PI482601	-	Zimbabwe	IX	173	129	95	85	108	108
PI567394A	Jiu yue han	China	VI	105	108	190	41	176	108
PI341241B	(Seminole)	Israel	IX	109	162	115	53	66	109
PI324067	Hernon 237	Zimbabwe	VII	189	109	126	85	86	109
PI429329	-	Nigeria	VII	139	110	65	20	158	110
PI639574	-	Burundi	VIII	115	66	93	115	111	111
PI567315	Hong huang dou	China	VII	27	7	111	156	114	111
PI221715	-	South Africa	VII	74	90	162	111	192	111
PI603529	Hei huang dou	China	VIII	112	123	27	206	56	112
PI341242	Hernon 247	Tanzania	IX	104	103	146	112	160	112
PI553045	Cook	United States	VIII	28	169	113	179	106	113
N06-7543	-	United States	VII	59	114	129	181	41	114
PI648270	Osage	United States	V	155	85	104	203	114	114
PI603532	Hong li huang dou	China	VI	159	144	63	115	14	115
PI603519	Lu da dou	China	VI	34	117	142	115	56	115
PI505649B	-	Zambia	IX	205	202	25	115	68	115
PI603539A	Huang dou	China	VI	107	161	154	115	79	115
PI553039	Davis	United States	VI	116	155	82	154	99	116
PI341253	CMS	Sudan	IX	110	157	197	91	119	119
PI330634	-	South Africa	VII	55	179	119	73	132	119
PI603535	Hei zong huang dou	China	VIII	17	156	192	83	120	120
PI376069	DRO 9	Cameroon	VIII	184	10	159	36	120	120
PI603521	Huang dou	China	VIII	62	190	196	85	120	120
PI553046	Gasoy 17	United States	VII	83	175	72	120	192	120
NTCPR94-5157	-	United States	VI	7	124	153	121	25	121
PI171443	Tea-bean	China	VI	147	69	121	67	202	121
PI612157	Prichard	United States	VIII	149	122	100	188	41	122
PI486328	Birsa Soybean-1	India	VIII	41	62	122	192	207	122
PI567406B	(Wu se da dou)	China	VI	36	56	191	123	190	123
PI434980B	(Indo 180)	Central African Republic	IX	124	8	124	27	157	124
N06-7194	-	United States	VIII	31	133	125	183	101	125

PI407738	-	China	VI	125	198	23	63	176	125
PI603512	Jin man dou	China	VI	64	125	22	160	182	125
PI323570	-	India	VII	97	126	11	164	132	126
PI647085	N7002	United States	VII	13	128	194	168	41	128
PI603536	Hui huang dou	China	VIII	19	171	130	139	56	130
PI639573	-	Burundi	VIII	199	25	163	51	130	130
PI430736	Kudu	Zimbabwe	VI	160	55	110	165	130	130
PI508266	Young	United States	VI	40	189	111	130	170	130
PI548656	Lee	United States	VI	58	172	131	135	41	131
PI221717	-	South Africa	VI	146	131	51	173	91	131
PI265498	-	Zaire	VIII	131	166	148	55	120	131
PI341245	Avoyelles	Tanzania	IX	150	185	132	33	112	132
PI603509	Huang dou	China	VIII	81	200	200	132	120	132
PI221716	-	South Africa	VII	122	134	75	177	132	132
PI374221	Welkom	South Africa	VI	195	34	198	124	132	132
PI567206	GL2674/90	Georgia	VI	119	132	91	160	164	132
PI561375	Qi huang No. 1	China	V	186	163	4	133	53	133
PI567410A	Yang huang dou	China	VII	179	184	133	58	112	133
PI639575	-	Burundi	VIII	134	177	62	50	137	134
PI532458	Ba yue bao	China	VIII	18	135	161	175	87	135
PI567378	Ba yue zha	China	VI	135	193	17	24	143	135
PI330635	-	South Africa	VII	206	12	182	48	137	137
PI417562	54.S.30 DL/64/185	South Africa	VI	208	63	184	138	68	138
PI374220	Geduld	South Africa	VI	176	35	138	44	181	138
PI428691	-	India	VIII	138	121	88	160	189	138
PI567403A	Shuan huang dou	China	VII	156	26	139	90	150	139
PI567393	Jiu yue han	China	VII	170	167	141	131	141	141
PI471938	-	Nepal	V	33	141	40	141	160	141
PI438430	-	Israel	VII	52	141	116	187	191	141
PI341246	CNS	Tanzania	IX	178	21	144	10	171	144
PI322691	Jubiltan 109	Mozambique	IX	167	44	74	144	176	144
PI617045	NC-Roy	United States	VI	86	145	202	159	41	145
PI567329	Huang huang dou	China	VI	88	120	149	145	164	145
PI567350A	Shu pi huang dou	China	VI	207	146	66	61	211	146
FC003659	Da Wu Don	China	VI	101	68	193	147	200	147
PI567350B	(Shu pi huang dou)	China	VI	136	111	160	148	188	148
PI567332	Huo huang dou	China	VI	192	201	70	102	149	149
PI494851	-	Zambia	VI	162	150	61	154	99	150

PI555453	Hagood	United States	VII	99	150	87	171	174	150
PI567349B	(Shu pi huang dou)	China	VI	70	197	101	150	185	150
PI497967	-	India	VII	151	205	13	101	192	151
PI567345	Niu mao huang	China	VI	78	52	166	152	200	152
PI548657	Jackson	United States	VII	77	153	44	205	192	153
PI381663	Kakira 1	Uganda	VI	201	13	163	31	154	154
PI567493	Huang dou	China	IV	190	116	7	176	154	154
PI641156	NC-Raleigh	United States	VII	87	180	147	195	154	154
PI603520	Huang dou	China	VI	118	181	154	151	202	154
PI567295	Bian huang dou	China	VIII	154	164	78	19	208	154
PI424131	Buffalo	Zimbabwe	VII	191	158	185	136	108	158
PI567404B	(Wang shan hou)	China	VI	200	2	102	158	205	158
PI145079	Hernon No. 6	Zimbabwe	VII	163	17	168	201	164	164
PI323278	K-30	Pakistan	IX	165	188	167	47	107	165
PI603514	Ni ba dou	China	VI	103	165	157	170	174	165
PI603534A	Da niu mao huang	China	VII	32	195	203	166	120	166
PI159094	35S377	South Africa	VII	196	112	171	167	56	167
PI159096	41S77	South Africa	VII	197	97	170	99	171	170
PI330633	-	South Africa	VII	175	11	206	57	171	171
PI221714	-	South Africa	VI	171	168	60	203	192	171
NCC06-1090	-	United States	VI	182	186	151	182	76	182
PI603528	Hei ke zha	China	VII	14	187	195	137	182	182
PI567386	Huang da dou (1)	China	VI	185	182	26	2	186	182
PI599333	Musen	United States	VI	188	183	103	188	41	183
PI567377B	(Ba yue zha)	China	VI	209	196	1	113	202	196

[a] Breeding values were calculated for each trait within an individual environment, and then summed across environments. These summed breeding values for each trait were ranked in ascending (canopy wilting, nitrogen concentration, and $\delta^{15}N$) or descending ($\delta^{13}C$ and NDTR to $AgNO_3$) order based on whether negative or positive breeding values would be more favorable for the trait.

CHAPTER 4

SUMMARY

Improving drought tolerance in soybean would help ameliorate the impact of water deficit stress on its productivity. Soybean is a widely grown crop in the USA and throughout the world, and breeding soybean for drought tolerance can have a major impact on water usage and food production. Selecting for soybean yield under drought conditions is hampered because of its low heritability, interactions between the genotypes and environments, and the polygenic nature of drought tolerance. Therefore, evaluation of physiological drought tolerance related traits and associating them with genome-wide markers, can reveal genomic regions that can be targeted by breeders. Favorable combinations of alleles for these physiological traits found in soybean genotypes can then also be incorporated into elite germplasm by breeders.

In this research, five different traits related to soybean drought tolerance were evaluated: canopy wilting, carbon isotope composition ($\delta^{13}C$), nitrogen concentration, nitrogen isotope composition ($\delta^{15}N$), and leaf hydraulic conductance (transpiration response to $AgNO_3$). The primary objective of this research was to identify germplasm or genetic loci useful for improving our understanding and ability to improve soybean drought tolerance. Towards this aim, two different populations, a genetically diverse association panel of over 200 genotypes and 130 recombinant inbred lines (RILs) derived from a drought tolerant by susceptible cross, were evaluated for the aforementioned traits. Genome-wide association analyses and composite interval mapping using genome-wide SNP data for these populations revealed loci controlling

these traits, of which many explained a high amount of the phenotypic variation and could be exploited by breeders to improve drought tolerance.

In Chapter 2, genome-wide association and linkage mapping were used to identify genomic regions controlling canopy wilting. For the association panel, canopy wilting was visually scored in four environments after rain-fed field plots experienced drought stress, and in two environments for the RIL population. Association mapping revealed 47 unique SNPs that tagged 45 loci in at least one environment, and five regions were found both in a single environment and across all environments. Ten of the 47 SNPs were found in or near the genomic locations of previously reported canopy wilting QTLs from linkage mapping studies, and eight were near the genomic regions from a previous association mapping study. Three of the SNPs identified explained more than 20% of the phenotypic variation in a given environment. Many genotypes were identified which have more favorable breeding values for canopy wilting compared to the slow wilting check genotypes PI 416937 and PI 471938.

In Chapter 3, $\delta^{13}C$, nitrogen concentration, $\delta^{15}N$, and transpiration response to $AgNO_3$ were evaluated for the association panel. Stable isotope analyses were performed on leaf samples collected from rain-fed field plots in two environments to assess the carbon and nitrogen related traits. Normalized decrease in transpiration rate (NDTR) of de-rooted soybean shoots subjected to silver nitrate compared to water was measured in a growth chamber for eight replications of the panel. Using an association mapping approach, 32, 26, 23, and 10 loci were found to be significantly associated with $\delta^{13}C$, nitrogen concentration, $\delta^{15}N$, and NDTR to $AgNO_3$, respectively. Four of the $\delta^{13}C$ and two of the $\delta^{15}N$ loci detected were co-located with previously reported QTLs. Breeding values were calculated for all traits evaluated, which enabled the identification of accessions in the panel with favorable combinations of these drought tolerance

related traits. Ultimately, the new germplasm and genomic regions discovered in this research help provide additional understanding about the genetic architecture for these traits, revealed genetic loci to target in breeding efforts, and identified new parental genotypes to use in crosses and for integration of favorable alleles into elite germplasm.