Genome-wide association study of plant physiological parameters, yield, and nitrogen fixation-related traits in soybean under drought stress

by

Dilrukshi Kaushalya Dias Kombala Liyanage

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in
Plant Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Soybean is one of the world's leading legume crops. It fixes atmospheric nitrogen with the symbiotic relationship of rhizobia bacteria that inhabit root nodules. The global population is expected to exceed 9.7 billion by 2050, resulting in increased food demand, particularly for protein. In Canada, soybean is the third largest field crop in terms of farm cash receipts, and its production has recently grown significantly in the Canadian Prairies. The changing global climate is predicted to lead large areas of the world to experience extensive drought conditions over the next few decades. Drought will be a critical issue for western Canada as it has been experiencing frequent and severe droughts over the last few decades. Drought has negative impacts on soybean plant development, yield, and symbiotic nitrogen fixation. The main objective of this project is to identify the allelic variations associated with diverse short-season soybean varieties for plant physiological parameters, yield traits, and symbiotic nitrogen fixation under drought stress. A greenhouse pot experiment with 103 early-maturity Canadian soybean varieties was conducted to determine the above-mentioned traits. Seedlings were inoculated with *Bradyrhizobium japonicum* USDA 110, and the initial soil moisture content of the growing media was maintained at 80% field capacity (FC). Drought stress was applied after three weeks of plant growth, with half of the plants kept at 30% FC and the remaining at 80% FC until maturity. Plant physiological traits such as photosynthesis, transpiration, stomatal conductance, leaf chlorophyll, water-use efficiency, and intrinsic water use efficiency were collected during the flowering stage. At seed maturity, data on yield and nitrogen fixation-related traits, including number of pods, number of seeds, seed yield, 100-seed weight, seed nitrogen content, % nitrogen derived from the atmosphere (%Ndfa), total seed nitrogen fixed, and carbon isotope discrimination (CID), were collected. Drought stress reduced stomatal conductance and transpiration, resulting in greater water-use efficiency compared to well-watered plants at the flowering stage. In comparison to the well-watered treatment, drought-stressed soybean plants had lower seed yield, yield components (number of pods, number of seeds, 100-seed weight), seed nitrogen content, %Ndfa, and total fixed nitrogen. Specifically, drought reduced yield by 34.7%. %Ndfa by 13.4%, total seed nitrogen by 34.9%, and amount of seed nitrogen fixed by 42.1% compared to the well-watered treatment. Significant genotypic variability among soybean varieties was found for plant physiological parameters, yield parameters, nitrogen fixation traits, and water use efficiency. A Genome-Wide Association Study (GWAS) was conducted using 2.16M SNPs (2,164,465 SNPs) for above mentioned parameters for 30% FC, 80% FC and their relative performance (30% FC / 80% FC). In total, 13 quantitative trait locus (QTL) regions, including multiple candidate genes, were detected as significantly associated with different plant physiological traits, including photosynthesis, stomatal conductance, and water-use efficiency for 30% FC and relative performance. In terms of yield and nitrogen fixation-related characteristics, six QTL regions and candidate genes were identified as significantly correlated with %Ndfa and CID under drought stress and relative performance. These QTL regions may be useful in future breeding efforts to create drought-resistant soybean cultivars.

Keywords: soybean, drought, symbiotic nitrogen fixation, candidate gene, genome wide association study, quantitative trait locus

Preface

The technical apparatus referred to in chapter 2 was a collaboration of Dr. Malinda Sameera Thilakarathna, Boyd A. Mori, Ishan Chathuranga, and me (Dilrukshi Kombala Liyanage). Author Contributions are as follows, conceptualization, Dilrukshi Kombala Liyanage., Ishan Chathuranga. and Dr. Malinda Sameera Thilakarathna.; methodology, Dilrukshi Kombala Liyanage and Ishan Chathuranga; software, Ishan Chathuranga.; validation, Dilrukshi Kombala Liyanage and Ishan Chathuranga; formal analysis, Dilrukshi Kombala Liyanage. and Ishan Chathuranga.; investigation, Dilrukshi Kombala Liyanage.; resources, Malinda Sameera Thilakarathna.; data curation, Dilrukshi Kombala Liyanage; writing—original draft preparation, Dilrukshi Kombala Liyanage and Ishan Chathuranga.; writing— review and editing, Dilrukshi Kombala Liyanage, Ishan Chathuranga., Malinda Sameera Thilakarathna and Boyd A. Mori; visualization, Dilrukshi Kombala Liyanage, Ishan Chathuranga. and Malinda Sameera Thilakarathna.; supervision, Malinda Sameera Thilakarathna.; project administration, Malinda Sameera Thilakarathna.; funding acquisition, Malinda Sameera Thilakarathna and Boyd A. Mori.

Chapter 2 of this thesis has been published as, Liyanage, D. K., Chathuranga, I., Mori, B. A., and Thilakarathna, M. S. (2022). A simple, semi-automated, gravimetric method to simulate drought stress on plants. Agronomy 12, 349.

Acknowledgment

I'd like to take this opportunity to thank a few people who have helped me through thick and thin. First and foremost, I'd like to thank my supervisor, Dr. Malinda Thilakarathna, for his guidance and mentorship during my time at the –Plant-Microbe Lab at the University of Alberta. All of my accomplishments over the last two and half years have been made possible by his constant encouragement and guidance. I’m really grateful to be a part of his research group. My sincere thanks should go to Dr. Davoud Torkamaneh and Dr. François Belzile who collaborated in this research and guided me with my research especially on GWAS. I am grateful for their support in completing my research study. I am also sincerely thankful to my committee members Dr. Linda Gorim, Dr. Guanqun (Gavin) Chen for the advice, guidance, and enormous support given to me throughout the degree program. I would also like to thank all my teammates at the Plant microbe lab for helping me throughout the research study. My sincere gratefulness also goes to Kelley Dunfield and Shay Missian for their kind support throughout the greenhouse experiments. Last but not least, I'd like to thank my family for always being there for me in happy and sad times, and for always encouraging me to reach my goals.

Table of contents

List of tables

List of figures

List of abbreviations

ABA	Abscisic acid
ANK	Ankyrin repeat family
ANOVA	Analysis of variance
Asp	Asparagine
BNF	Biological nitrogen fixation
CID	Carbon isotope discrimination
CVD	Cardiovascular disease
EMMA	Efficient mixed model association
FarmCPU	Fixed and random model circulating probability unification
FC	Field capacity
FDR	False discovery rate
Fe	Iron
GBS	Genotyping by sequencing
GWAS	Genome wide association studies
IRMS	Isotope Ratio Mass Spectrometer
iWUE	Intrinsic water-use efficiency
Lb	Leghemoglobin
MAS	Marker-assisted selection
MDH	Malate dehydrogenase
MoFe	Molybdenum-iron
N_2	Dinitrogen
%Ndfa	Percentage of nitrogen derived from atmosphere
NH3	Ammonia
ODB	O_2-diffusion barrier
PBM	Peribacteroid membrane
PC	Principal component

PEP	Phosphoenolpyruvate
PEPC	Phosphoenolpyruvate carboxylase
QTL	Quantitative trait loci
SNF	Symbiotic nitrogen fixation
SNPs	Single Nucleotide Polymorphisms
TDR	Time-domain reflectometry
TDT	Time domain transmission
WGS	Whole-genome sequencing
WUE	Water-use efficiency

Chapter 1 – Literature Review

1.1 History of soybean

Soybean [*Glycine max* (L.) Merr.] is the leading legume crop in the world. The cultivated soybeans are proven to be domesticated from their wild annual progenitor (*Glycine soja* Sieb. and Zucc.) in China about 5000 years ago (Hymowitz, 2016). The identified centers of domestication for soybean include north-eastern (Fukuda, 1933) the yellow river valley of northern China (Li et al., 2010) and southern China (Gai et al., 2000). Then, it was introduced to North America in 1765 (Hymowitz, 2016). In 1881, the soybean was introduced to Ontario and is now considered one of the most valuable field crops in eastern Canada. Soybean is an important rotational crop in Ontario, often following corn (*Zea mays* L.) and preceding winter wheat (*Triticum aestivum* L.) (Bruce et al., 2019). In 2021, 5.32 million acres of agricultural land in Canada were cultivated with soybeans, yielding 6.27 MMT of grain (Soy Canada, 2022). With the introduction of new short-season soybean varieties, soybean is gradually expanding into Canadian prairies as well.

1.2 Nutritional value of soybean

Soybean is mainly grown as an oilseed crop and feed for livestock. It is a rich source of plant-based protein. The average protein content of commercial soybean varieties varies between 34%-48% but depends on the genotype, growing environment, and cultural practices (Singh, 2010). Soybean is also rich in important constituents such as isoflavonoids, tocopherols, lecithin, and biopeptides that can combat many diseases such as hormone-dependent cancers and cardiovascular disease (CVD), osteoporosis, menopausal problems and diabetes (Erdman and John, 2009; Singh, 2010; Levis et al., 2011; Yanti and Dafriani, 2021) Soybeans are high in dietary fibres and low in saturated fat (Mateos-Aparicio et al., 2008). Soybeans consist of isoflavone, which makes them unique among legumes (Mateos-Aparicio et al., 2008). Soybean is also rich in different amino acids, and the amino acid composition of soybean is listed in Table 1.1 (Liu, 1997).

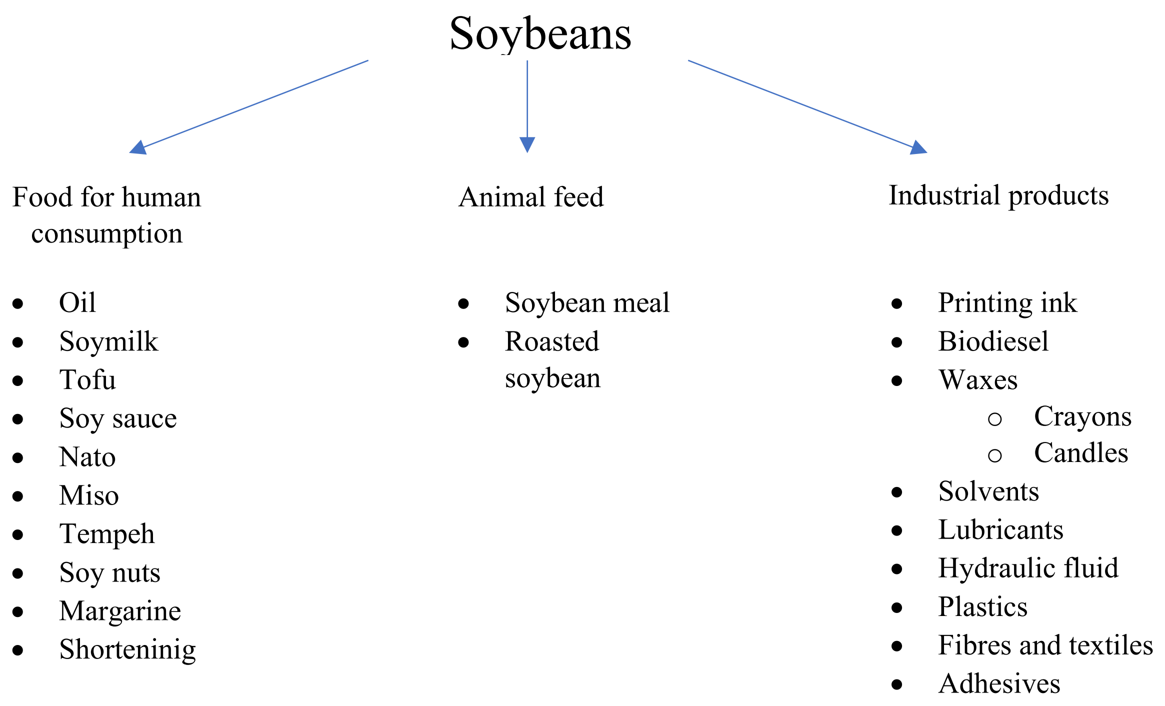
Table 1.1 Amino acid composition of soybean seed.

Amino Acid	**mg/g Protein**
Arginine	77.16
Alanine	40.23
Aspartic acid	68.86
Cystine	25.00
Glutamic acid	190.16
Glycine	36.72
Histidine	34.38
4-Hydroxiproline	1.40
Isoleucine	51.58
Leucine	81.69
Lysine	68.37
Methionine	10.70
Phenylalanine	56.29
Proline	52.91
Serine	54.05
Threonine	41.94
Tryptophan	12.73
Tyrosine	41.55
Valine	41.55

Source: (Liu, 1997)

1.3 Uses of soybean

Soybean can be identified as an important major legume with numerous uses in human consumption, animal consumption, and industrial uses. The main uses of soybean are oilseed crops and feed for livestock and fisheries. A general description of its uses is shown in Figure 1.1.



Soybeans

Food for human consumption

- Oil
- Soymilk
- Tofu
- Soy sauce
- Nato
- Miso
- Tempeh
- Soy nuts
- Margarine
- Shorteninig

Animal feed

- Soybean meal
- Roasted soybean

Industrial products

- Printing ink
- Biodiesel
- Waxes
  - Crayons
  - Candles
- Solvents
- Lubricants
- Hydraulic fluid
- Plastics
- Fibres and textiles
- Adhesives

Figure 1.1 General uses of soybean

(Source: Statistics Canada, Census of Agriculture, 2007)

1.4 World and Canadian soybean production

The world soybean production in the last five years is shown in Table 1.2. According to the 2021-2022 statistics, Brazil, the USA, Argentina, India, Paraguay, and Canada are the top seven soybean-producing countries. Furthermore, the world soybean production in 2021-2022 is represented in Figure 1.2. The USA and Brazil account for more than two-thirds (69%) of global soy production. The primary use of soybean is for feed production, where more than three-quarters (77%) of soy is used as livestock feed (Hannah and Roser, 2021). The nutritional quality (e.g. protein and lipid concentrations) of soybean can vary depending on the regions and countries (Grieshop and Fahey, 2001; Karr-Lilienthal et al., 2004). Furthermore, environmental conditions and processing conditions largely affect the nutrient quality and chemical composition of soybean (Grieshop and Fahey, 2001; Karr-Lilienthal et al., 2004).

Table 1.2 World soybean production trend in the last five years (in million metric tons).

Year	Brazil	USA	Argentina	China	India	Paraguay	Canada	Other
2021-2022	139	120.71	46.5	16.4	11.9	8.5	6.27	23.28
2020-2021	138	114.75	46.2	19.6	10.45	9.9	6.35	20.97
2019-2020	128.5	96.67	48.8	18.1	9.3	10.25	6.15	22.12
2018-2019	119.7	120.52	55.3	15.97	10.93	8.51	7.42	22.98
2017-2018	123.4	120.07	37.8	15.28	8.35	10.26	7.72	20.53

Source: Statista, Published by M. Shahbandeh, Feb 11, 2022

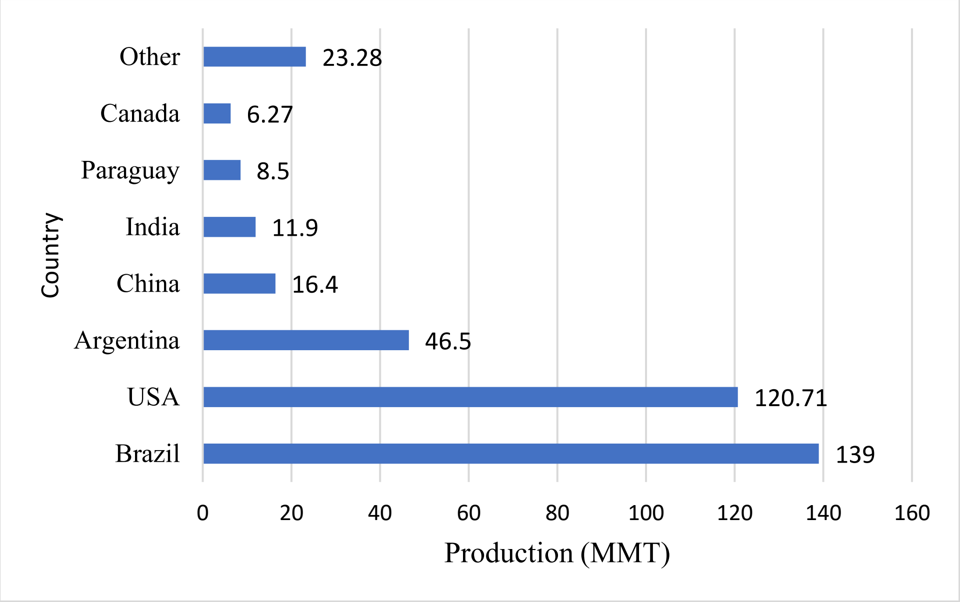


Figure 1.2 World Soybean production in 2021-2022

Source: Statista, Published by M. Shahbandeh, Feb 11, 2022

1.5 Soybean production in Canada

In Canada, soybean is the fourth-largest field crop in terms of farm cash receipts and its production has recently increased in the Canadian Prairies. According to Soy Canada, the total number of soybean producers in Canada during 2021 is 29,626. The total seeded area is 5.32 million acres. The production is 6.27 MMT and it is a 40% increase in production from 2011 and the average yield in 2021 is 38.2 bushels/acres (Soy Canada 2022). According to the distribution of soybean production among provinces, Ontario, Quebec, and Manitoba are in first, second, and third place, respectively.

Table 1.3. Provincial distribution of soybean production in 2021

Province	Production (MT)	Percentage of Total production (%)
Ontario	4,082,331	65.09
Quebec	1,101,708	17.57
Manitoba	963,764	15.37
Saskatchewan	50,935	0.81
Maritimes	72,643	1.16

Source: Statistics Canada. Table 001-0017 - Soybean Production in metric units, annual Updated: December 2021

1.6 Soybean plant growth stages

Soybean undergoes two main growth stages in its life cycle, they are the vegetative phase (V) and the reproductive phase (R). Vegetative stages are identified and differentiated on the number of fully developed trifoliate leaves. The reproductive phase consists of flowering, pod development, seed development, and maturation.

First cotyledons appear from the soil surface at VE stage. After the full exposure of cotyledons unifoliate leaves start to emerge (VC). In the V1 stage of growth the first trifoliate leaflets can be observed. During the V2 stage the second trifoliate leaf appears. Also, root nodules start to begin during this stage. When the plant is 6 to 8 inches taller, it begins to fix nitrogen and lateral roots start to develop. During the V3 to V5 stage the axillary buds mature into flower clusters. The lateral roots overlap at the V6 stage.

R1 is the beginning bloom stage. At least one flower can be observed in this stage. The secondary roots and root hairs start to grow throughout the R4 and R5 stages. During the R2 full bloom stage plant reaches approximately 50% of its mature height and a fully open flower emerge. The nitrogen fixation rapidly increases and at least one pod can be observed during the R3 beginning pod phase. The pods and seeds are rapidly forming during the R4 stage. Seed fill during the R5 growth stage demands large amounts of water and nutrients. At the beginning of maturity (R7) at least one brown color pod can be observed, and the seed dry matter is at the peak. At least one pod on the main stem reaches brown colour at the beginning of maturity (R7) and seed dry matter is at peak. Seeds and pods start to turn into a yellow to brown colour. Then at the full maturity stage (R8), at least 95% of pods turned brown colour and is ready to harvest (Purcell et al., 2014). Figure 1.3 illustrates the soybean growth stages from emergence to maturity stage. One of the most interesting features of soybean is, it has the ability to fix atmospheric nitrogen with the symbiosis of rhizobia bacteria reside in rood nodules.

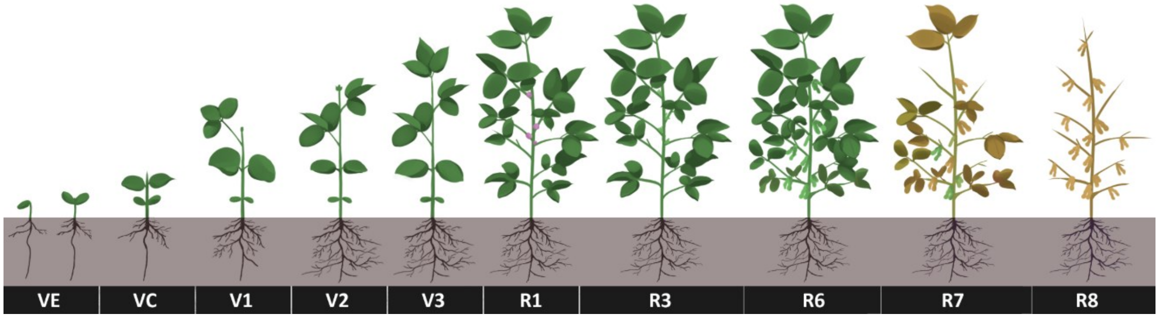


Figure 1.3 Soybean growth stages (Photo credit: Mitchel Tran)

1.7 Biological Nitrogen Fixation (BNF)

In terms of plants, nitrogen is one of the major macronutrients that plants need to complete their life cycle, growth, and development. Nitrogen (N_2) is an essential key element in chlorophyll, amino acids, ATP, and nucleic acids. Even though the atmosphere consists of 78% of N_2 it is the most limiting nutrient available in the soil for plants (Valentine et al., 2011). Some prokaryotes can fix atmospheric N_2 (diazotrophs) in the presence of an enzyme called nitrogenase where they reduced atmospheric N_2 into ammonia (NH_3) (Halbleib and Ludden, 2000). Nitrogenase is an enzyme complex with two components; the iron (Fe) protein (dinitrogenase reductase) and the molybdenum-iron (MoFe) protein (dinitrogenase) (Dixon and Kahn, 2004). Biological nitrogen fixation is an energy-consuming process that requires more reducing power and high-energy ATP (Adenosine Tri Phosphate) (Halbleib and Ludden, 2000). Biological nitrogen fixation under optimum conditions can be illustrated as below:

$$N_2 + 8\ e^- + 8\ H^+ + 16\ MgATP \rightarrow 2\ NH_3 + H_2 + 16\ MgADP + 16Pi$$

Rhizobia are gram-negative bacteria belonging to proteobacteria (Minamisawa et al., 1998; Franche et al., 2009). Also, they are called nitrogen-fixing heterotrophic bacteria (Minamisawa et al., 1998). In the past, it was believed that only *Bradyrhizobium japonicum* colonize the root system and form nodules in soybean (Rodríguez-Navarro et al., 2011). Later, it was found that not only *Bradyrhizobium* but also different species such as *Rhizobium, Mesorhizobium, and Sinorhizobium fredii* also can form nodules in soybean (Biate et al., 2014). Depending on the growth rate, there are two types of rhizobia; slow growing rhizobia (*Bradyrhizobium spp.*) and fast-growing rhizobia (*R. tropici, R. oryzae, M. tianshanense*) (Hungria et al., 2001; Biate et al., 2014). Rhizobial strains vary in terms of compatibility with various crop cultivars, nodulation ability, and SNF efficiency (Biate et al., 2014). There are 29 *Bradyrhizobium* species reported: *B. japonicum, B. elkanii, B. liaoningense, B. yuanmingense, B. betae, B. canariense, B. denitrificans, B. pachyrhizi, B. jicamae, B. iriomotense, B. cytisi, B. lablabi, B. daqingense, B. huanghuaihaiense, B. oligotrophicum, B. rifense, B. arachidis, B. retamae, B. diazoefficiens, B. ganzhouense, B. paxllaeri, B. icense, B. manausense, B. ingae, B. valentinum, B. neotropicale, B. ottawaense, B. erythrophlei, and B. ferriligni.*

1.7.1 Signal exchange in legume-rhizobia symbiotic biological nitrogen fixation

Legume-rhizobia symbiotic relationship is a signal exchange mechanism between both rhizobia and host plants (Hirsch and Fujishige, 2012). Legume plants release plant-specific flavonoids through their root system to attract rhizobia (Zhang et al., 2009), where they bind with nodD protein in rhizobia and activate nodulation genes (nod genes) (Hirsch and Fujishige, 2012). When the nod genes are activated, rhizobia to plant signal exchange happens by releasing nodulation factors/nod factors (lipo-chitooligosaccharides). Nod factors have different functions, such as rhizobia invasion and nodule formation (Geurts et al., 2005). When the plants receive nod factors, root hairs trap the rhizobia between cell walls (Gage, 2004). Then rhizobia enter the plant roots with the initiation of the infection thread. It grows and directs bacteria to the inner cortex of the roots (Jones et al., 2007). Nod factors trigger the root cortical cell division and the formation of nodule primordium (Gage, 2004). The infection thread releases bacteria into the cortical cells of the nodule primordium covered by the peribacteroid membrane (PBM) (Jones et al., 2007). These released bacteria transformed into bacteroids, and it surrounded by PBM is called symbiosome (Day et al., 2001). In return, the host soybean plant supplies photosynthate/carbon to the bacteroids, which make legume-rhizobia symbiotic relationship. A simple illustration of nitrogen fixation is shown in Figure 1.4.

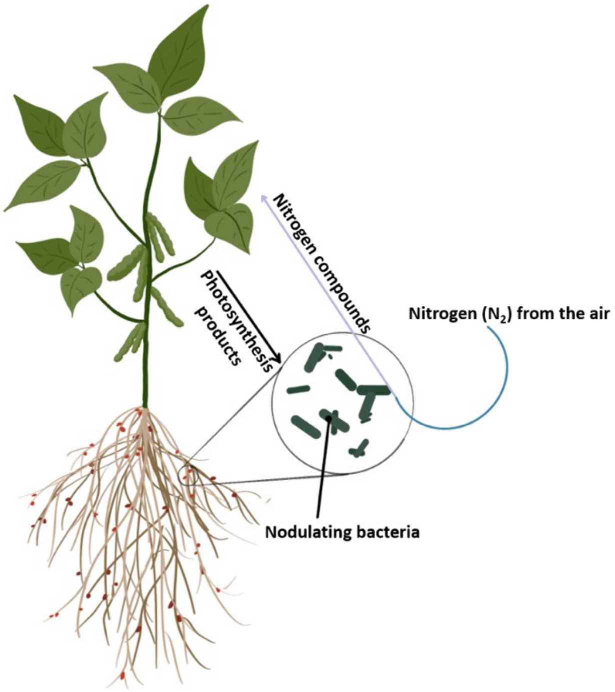


Figure 1.4 Legume-rhizobia symbiotic nitrogen fixation

Nitrogen fixation is catalyzed by the nitrogenase enzyme present in rhizobia. Nitrogenase is a complex enzyme that consists of two enzymes called dinitrogenase reductase and dinitrogenase. Dinitrogenase reductase binds ATP and transfers electrons to dinitrogenase. It is a dimeric Fe-protein with a 4Fe-4S cluster. Nitrogenase is a tetrameric FeMo protein that contains Fe-S clusters and binds N_2. It is important that these enzymes are irreversibly inactivated by oxygen hence, maintaining an anaerobic environment is critical. There are mechanisms in the nodules to control O_2 concentration to avoid oxidation of the nitrogenase enzyme complex while maintaining the bacteroid respiration at an optimum level (Hossain et al., 2017).

Leghemoglobin (Lb) is a monomeric heme-protein that acts as an oxygen carrier in legumes, similar to hemoglobin in animals (Downie, 2005). One of the interesting features of Lb is its ability to rapid O_2 association and relatively slow O_2 release rate (Rutten and Poole, 2019), which enables buffering of the free oxygen concentration at around 7-11nM (Downie, 2005). Lb is produced by the roots colonized rhizobia. It maintains the free oxygen in the cytoplasm of infected cells and aids the function of root nodules (Hossain et al., 2017). In the meantime, the bacteria need a high flux of oxygen for ATP synthesis (Downie, 2005). Hence, Lb is capable of keeping free oxygen concentration at a low level for the optimum function of nitrogenase whereas total oxygen concentration (free and bound to Lb) also at the required level for aerobic respiration (Downie, 2005).

The second mechanism that regulates low oxygen concentration in root nodules is the O_2-diffusion barrier (ODB) localized in the nodule parenchyma that is capable of limiting the flow of O_2 into the infected zone. It can adjust the permeability rapidly (Denison and Kinraide, 1995). This diffusion barrier allows only nanomolar concentrations of oxygen (Rutten and Poole, 2019). This layer helps the legume to control the O_2 inflow into the nodule, maintaining the internal oxygen concentration without any fluctuations due to environmental factors (Rutten and Poole, 2019). All these mechanisms collectively prepare the nodule environment favourable for rhizobial nitrogen fixation.

1.7.2 Importance of legume-rhizobia symbiotic nitrogen fixation

The use of synthetic nitrogen fertilizers causes both environmental and health issues. When nitrogen fertilizers are applied, the excess nitrogen is lost in either ionic or gaseous form through the processes of leaching, volatilization, and denitrification (Follett and Delgado, 2002). Part of the applied nitrate can be carried away by runoff water or leaches into the soil with water, and it creates environmental issues such as eutrophication, greenhouse effect and acid rains (Follett and Delgado, 2002). Consumption of contaminated groundwater will create health problems in humans as well as animals.

Biological nitrogen fixation is an interesting substitution for synthetic N fertilizer because legumes provide a sustainable N input via BNF for legumes as well as for companion non-legume crops in the cropping system (Thilakarathna et al., 2016a), especially under limited N conditions (Fujita et al., 1992). According to the statistics, annually 50-70 Tg of N can be derived via BNF in agricultural systems, and pasture and fodder legumes alone contribute 12-25 Tg N annually (Herridge et al., 2008). Nitrogen fixed by legumes is slowly released into the environment and is also less prone to leaching and volatilizing (Dixon and Kahn, 2004). Since nitrogen is released gradually through microbial decomposition, it reduces the environmental pollution acquainted with mineral N. Hence legume symbiotic nitrogen fixation is an important process in Agriculture.

1.8 Drought stress and drought tolerance mechanisms in soybean

Recent and possible future global temperature rises are likely to have an impact on the hydrologic cycle, including changes in precipitation and an increase in extreme occurrences like droughts (Sheffield and Wood, 2008). Drought stress is one of the most severe abiotic stresses that soybeans face, as drought reduces production significantly in many growing areas (Oya et al., 2004). Despite the fact that most parts of Canada experience drought on a periodical basis, the agricultural regions of the Canadian Prairies are the most vulnerable, owing to their reliance on agriculture for economic survival (Bonsal et al., 2013). Therefore, drought is expected to become a major issue for the Canadian Prairies in the future.

Plants use a variety of methods to combat the negative effects of drought, and the ability of crops to change using adaptive features is referred to as 'drought tolerance' (Basu et al., 2016). Drought-induced stomatal closure and reduced photosynthesis are common drought responses in plants (Liu et al., 2003; Mak et al., 2014). According to research, drought stress decreases relative leaf expansion rate, stomatal conductance, and leaf turgor while raising ABA concentration in the leaf and xylem (Liu et al., 2003). Soybeans are able to withstand drought by increasing leaf pubescence density, which limits transpiration during water constraints (Manavalan et al., 2009). Because highly pubescent lines are associated with greater root density and deeper root extension, this is a key adaptation characteristic of soybean (Manavalan et al., 2009). Under high irradiance, leaf pubescence density enhances reflectance from the leaf, resulting in lower leaf temperatures. Dense pubescence has been linked to lower leaf temperature, lower transpiration loss, and increased photosynthesis as a result of reduced canopy sunlight penetration (Manavalan et al., 2009; Specht and Williams, 2022). Furthermore, plants can also boost their drought resistance by producing cuticular wax on their aerial surface (Islam et al., 2009).

When tissue water potential drops, Osmotic adjustment is considered a valuable measure because it allows cells to retain turgor (Manavalan et al., 2009). The osmotic adjustment has been demonstrated to preserve stomatal conductance and photosynthesis at lower water potentials, delay leaf senescence and mortality, reduce flower abortion, boost root growth, and increase water extraction from the soil as water (Turner et al., 2001). Soybean plants respond to drought by growing a larger taproot that allows them to reach the deep soil layers and access water (Taylor et al., 1978). Furthermore, the fibrous root system is also beneficial when it uptake moisture and minerals like phosphorous. In addition, drought stress increases the biomass partitioning to the roots and increases the root: shoot ratio (Manavalan et al., 2009).

1.9 Drought stress on legume symbiotic nitrogen fixation

According to climate predictions, drought stress is a major threat to crop production in the near future (Kunert et al., 2016). Drought stress results in poor plant establishment, growth reduction, yield reduction, poor grain quality, and low nitrogen fixation in legumes (Farooq et al., 2017). In terms of low nitrogen fixation, drought stress reduces the nodulation and nitrogenase enzyme activity (Farooq et al., 2012). Drought stress causes the low symbiotic nitrogen fixation in nodules by three mechanisms; oxygen limitation in nodules, carbon shortage, and feedback inhibition of nitrogen fixation.

1.9.1 Oxygen limitation in nodules

Nitrogenase is the key enzyme involved in the symbiotic nitrogen fixation process. Yet, the nitrogenase enzyme is sensitive to oxygen, which irreversibly stopped functioning when exposed to high oxygen concentrations (Sulieman and Tran, 2016). Legume root nodules maintain a microaerobic environment within the nitrogen fixation zone. During drought stress, the lenticels of the soybean nodule collapse and limit the oxygen diffusion to the central zone of the nodule (Serraj et al., 1999; Serraj, 2003). As a result, the nitrogenase activity gets inhibited and negatively affects nitrogen fixation.

1.9.2 Carbon shortage

Reduction in nodule carbon flux relates to inhibition of nitrogen fixation under drought stress (Arrese-igorl et al., 1999). Sucrose is the primary carbon source supplied from shoots to bacteroids to fuel the symbiotic nitrogen fixation process. Sucrose synthase hydrolyzes sucrose into hexose, and then it catabolizes into phosphoenolpyruvate (PEP) through the glycolytic pathway, which is further converted to oxaloacetate by PEP carboxylase (PEPC) (Sulieman and Tran, 2016). This oxaloacetate reduces to malate by malate dehydrogenase (MDH) regenerating NAD+ (Vance and Gantt, 1992). During the drought stress, nodule sucrose synthase activity sharply declines (González et al., 1995), hence limiting the carbon flux required for the bacteroid respiration. Accumulation of sucrose synthase and malate depletion occurred in nodules as a result of sucrose synthase down-regulation (González et al., 1995).

1.9.3 Feedback inhibition of nitrogen fixation

In addition to the above two mechanisms, the N-feedback mechanism involves the N-status of the plant that regulates symbiotic nitrogen fixation. This is mostly considered in ureide-exporter tropical legumes (e.g., soybean). There are several N-related compounds that found to be acted as inhibitory signal molecules. They are ureides (Vadez et al., 2000), glutamine (Neo and Layzell, 1997), asparagine (Asp) (Vadez et al., 2000) and aspartate (King and Purcell, 2005). Furthermore, there is a restriction on the export of N-related compounds and their accumulation in nodules (Gil-Quintana et al., 2013). According to Serraj et al. (2001), there are two possible origins for the feedback inhibition of N fixation: 1) direct feedback within the nodules and 2) indirect feedback due to N compounds signals from the shoot level (Serraj et al., 2001). According to King and Purcell (2005), a combination of ureide and Asp levels in nodules and the transport of several amino acids from leaves might be involved in such a feedback inhibition mechanism in soybean (King and Purcell, 2005). It was found that prior to any N accumulation occurs at the shoot level, ureides accumulate in nodules, representing a local regulation of symbiotic nitrogen fixation (Gil-Quintana et al., 2013).

1.10 Genome-wide association study (GWAS)

The Genome-wide association study is an approach used to study hundreds of thousands to millions of genetic variants across the genomes of many individuals to see the association of variants with the trait of interest (Pantalião et al., 2016). Usually, GWAS studies focus on high-density single nucleotide polymorphisms, and it is an efficient and powerful tool in plant breeding to dissect the genetic architecture of complex traits (Rebolledo et al., 2015). Using GWAS analysis, we can identify Single Nucleotide Polymorphisms (SNPs) and candidate genes that are significantly associated with a particular trait and they can be used in Marker-Assisted selection to develop new cultivars in plant breeding (He et al., 2014; Zhang et al., 2014). Genome-wide association studies look for genotype-phenotype relationships by evaluating genetic variants throughout the genomes of many individuals (Tam et al., 2019). When compared to Quantitative Trait Loci (QTL) association mapping, GWAS is more precise at identifying candidate genes (Zhu et al., 2008). Therefore, it is one of the most promising strategies for detecting QTLs (Contreras-Soto et al., 2017). GWAS has been conducted for different traits in plant breeding, such as yield (Turuspekov et al., 2017; Ward et al., 2019), disease resistance (Boudhrioua et al., 2020; Kato et al., 2021), different agronomic traits such as N fixation (Torkamaneh et al., 2020), plant height (Sonah et al., 2015; Zhang et al., 2015; Li et al., 2016a), seed oil (Hwang et al., 2014; Sonah et al., 2015; Zhang et al., 2019), seed protein (Hwang et al., 2014; Sonah et al., 2015; Zhang et al., 2019), and root characteristics (Seck et al., 2020).

Previous literature reports number of GWAS related to different plant physiological parameters, yield, and nitrogen fixation-related traits in soybean. In terms of plant physiological traits, GWAS have been conducted for photosynthesis (Wang et al., 2020), chlorophyll content (Hao et al., 2012a; Dhanapal et al., 2016), carbon isotope ratio (Dhanapal et al., 2015a; Kaler et al., 2017a; Steketee et al., 2019), and photosynthesis traits related to phosphorous efficiency (Li et al., 2016b; Lü et al., 2018). Furthermore, GWAS has also been reported in the past with soybean for 100 seed weight (Contreras-Soto et al.,2017; Hu et al., 2020; Adeboye et al., 2021; Li et al., 2021; Priyanatha et al., 2022), seed yield (Contreras-Soto et al., 2017; Hu et al., 2020; Adeboye et al., 2021; Ravelombola et al., 2021; Priyanatha et al., 2022), number of pods per plant (Hu et al., 2020; Adeboye et al., 2021), pod weight (Li et al., 2021), (Ravelombola et al., 2021; Zhang et al., 2021; Jo et al., 2022), yield stability (Quero et al., 2021), number of seeds per plant (Chen et al., 2020; Adeboye et al., 2021), and number of seeds per pod (Chen et al., 2020). However, limited numbers of GWAS have been conducted for SNF-related traits in soybean (Dhanapal et al., 2015b).

1.11 Research hypothesis

There is an allelic variation associated with drought-resistant nitrogen fixation among the short-season soybean varieties.

1.12 Research objectives

The main objective of this research was to identify allelic variations associated with drought-resistant nitrogen fixation in a short season soybean diversity panel.

The specific objectives of this research are:

1. To develop a simple, Arduino microcontroller-based, semi-automated lysimeter system to adjust soil moisture content in pot experiments.
2. To identify allelic variation associated with different plant physiological parameters under drought stress at the flowering stage and identify the genomic regions controlling drought tolerance in short-season soybean varieties.
3. To identify allelic variation associated with different yield parameters and symbiotic nitrogen fixation in soybean under drought stress and identify the genomic regions controlling nitrogen fixation under drought stress in short-season soybean varieties.

Chapter 2 - A Simple, Semi-Automated, Gravimetric Method to Simulate Drought Stress on Plants

Abstract

Drought is a major constraint of global crop production. Given that drought-induced crop losses can threaten world food security, it has been and continues to be the focus of a large body of interdisciplinary research. Most drought experiments are conducted under controlled environmental conditions, where maintaining accurate soil moisture content is critical. In this study, we developed a simple, Arduino microcontroller-based, semi-automated, lysimeter that uses the gravimetric method to adjust soil moisture content in pot experiments. This method employs an Arduino microcontroller interfaced with a balance as part of a portable lysimeter and irrigation system which can weigh and record the mass of plants growing in pots, determine water loss due to evapotranspiration, and adjust soil moisture automatically to a desired relative soil water content. The system was validated with a greenhouse pot experiment using a panel of 50 early-maturity Canadian soybean varieties. Drought was induced in the experiment by adjusting soil moisture content to 30% field capacity while maintaining control pots at 80%. Throughout the experiment, the two moisture levels were efficiently maintained using the Arduino-based lysimeter. Plant physiological responses confirmed that plants in the drought treatment were under physiological stress. This semi-automated lysimeter is low-cost, portable, and easy to handle, which allows for high-throughput screening.

Keywords: drought; soil moisture; lysimeter; field capacity; Arduino microcontroller; moisture adjustment; water deficit

2.1 Introduction

Drought is the foremost abiotic stress that reduces plant growth and crop production throughout the world (Keyantash and Dracup, 2002; Bacelar et al., 2012). As the world population is predicted to surpass 9.5 billion (United Nations (2019). and drought risk and severity is predicted to increase due to climate change (Sheffield et al., 2012), future global crop production will be under significant pressure to keep pace with food demand. To supply this demand, research on drought stress is needed to improve crop resiliency and increase food production.

A wide-variety of interdisciplinary studies have shown that plants exhibit a vast array of mechanisms to tolerate drought stress (Earl and Davis, 2003; Ganjeali et al., 2011; Basu et al., 2016; Janiak et al., 2016; Jabbari et al., 2018; Wasaya et al., 2018; Tarawneh et al., 2020). However, to study these mechanisms it is crucial to accurately maintain soil moisture content (Marchin et al., 2020). As it is challenging to maintain precise soil moisture in the field, most drought experiments are conducted under controlled environmental conditions. Small pot or tube based laboratory experiments are commonly used to explore plant drought stress (Gorim and Vandenberg, 2017a, 2017b; Lehnert et al., 2018; Shirinbayan et al., 2019; Turner, 2019) and numerous methods have been developed to adjust soil moisture content including dual probe heat pulse, electromagnetic (e.g. time-domain reflectometry [TDR] and time-domain transmission [TDT] techniques) and gravimetric methods (Campbell et al., 2002; Earl, 2003; Robinson et al., 2003; Blonquist et al., 2005; Susha Lekshmi et al., 2014; Kojima et al., 2016; Bogena et al., 2017; Ortiz et al., 2018). However, the most popular, direct, and accurate method used to measure soil moisture content is the gravimetric method (Dobriyal et al., 2012; Susha Lekshmi et al., 2014). The usage of other indirect methods for measuring soil moisture content depends on accuracy, cost, response time, ease of installation, and durability of the instruments (Dobriyal et al., 2012).

Gravimetric-based soil moisture content is the ratio of the mass of the moisture present in a soil sample to the dried soil sample mass (Dobriyal et al., 2012). Gravimetric methods are usually more time-consuming and labor-intensive than other methods, but there is no need for expensive equipment (Susha Lekshmi et al., 2014), and drawbacks can be reduced by integrating computer-based automation techniques along with computational methods (Earl, 2003; Chard et al., 2010). Gravimetric-based methods have been used to measure evapotranspiration in pot experiments and to adjust soil moisture content to target levels (Earl, 2003; Chard et al., 2010; Wright et al., 2019; Gebre and Earl, 2020, 2021; McCauley et al., 2021). These methods involve frequently measuring the mass of pots, and replacing transpired water to maintain a targeted soil moisture content (Earl, 2003; Ortiz et al., 2018; Wright et al., 2019). However, the higher cost and complexity of the previous computer-based automated systems have limited their wider use for adjusting soil moisture levels in pot-based experiments.

Here, we developed a simple, low-cost, Arduino microcontroller-based lysimeter to gravimetrically adjust soil moisture content in pot experiments without the need for specialized facilities or equipment. The system measures soil moisture deficit and automatically adjusts the soil moisture content to a targeted level. We then demonstrate the effectiveness of this system with a drought experiment using a panel of 50 early maturity Canadian soybean varieties. Ultimately, this system will reduce costs and help researchers efficiently conduct drought-related experiment.

2.2 Materials and methods

2.2.1 Lysimeter system

2.2.1.1 Design and components

In this system (Figures 2.1 and 2.2; Table 2.1), water loss due to transpiration and evaporation in each pot is determined and recorded based on mass, and soil moisture content automatically adjusted to a targeted moisture level. A balance was made with two load cells (20 kg HX711AD pressure sensor modules; SZYT, Shenzhen, China) attached to a 10 cm diameter plastic tray, which served as a platform to place pots (Figures 2.1A and 2.2A). Two load cell amplifiers (HX711 load cell amplifiers; SZYT, Shenzhen, China), one per load cell, were used to amplify the signal generated from the load cells (Figure 2.2B). A standard breadboard (MT Technology Co., Ltd., Shenzhen, China) (Figure 2.2C) was used to connect the load cell amplifiers, a 1-channel 5 V relay module, and an Arduino R3 USB microcontroller (Arduino, A000066; Arduino SRL, Torino, Italy) (Figure 2.2D) which was used to control the irrigation system. The 1-channel 5 V relay module was used to connect the Arduino to the submersible water pump (ultra-quiet, 12 V, 4.2 W; ANSELF, Shenzhen, China) and power supply (1 A 12 V DC power adaptor with US plug type; ELECAPITAL, Shenzhen, China). Following the signal given by the Arduino, the relay (Figure 2.2E) connects or breaks the circuit, which turns on or turns off the water pump, respectively. A compact wire wiring connector (VENSTPOW, Shenzhen, China) was used to connect the submersible water pump and the power supply. This compact wire wiring connector was used to avoid manual soldering. The wiring diagram of the lysimeter system is illustrated in Figure 2.2. The submersible water pump (Figure 2.2H) was placed in a reservoir and pumped water via a flexible silicone hose (8 mm diameter; UXCELL, Shenzhen, China) (Figure 2.1D) to the pot following the signal given by the Arduino. A ring stand with a burette holder was used to direct the hose to the pot (Figure 2.1E). To avoid water damage, the Arduino, breadboard, and relay were placed in a water-resistant plastic container (Figure 2.1B). The system was connected via a USB cable (YCDC, Shenzhen, China) (Figure 2.1F) from the Arduino to a laptop computer (Figure

2.1C) to record the respective pot identification numbers and weights. For easy mobility, the entire system can be placed on a trolley and moved between locations within and between greenhouses (Figure 2.1). All components including their specifications and sources are listed in Table 2.1 and Appendix File A.1, respectively.

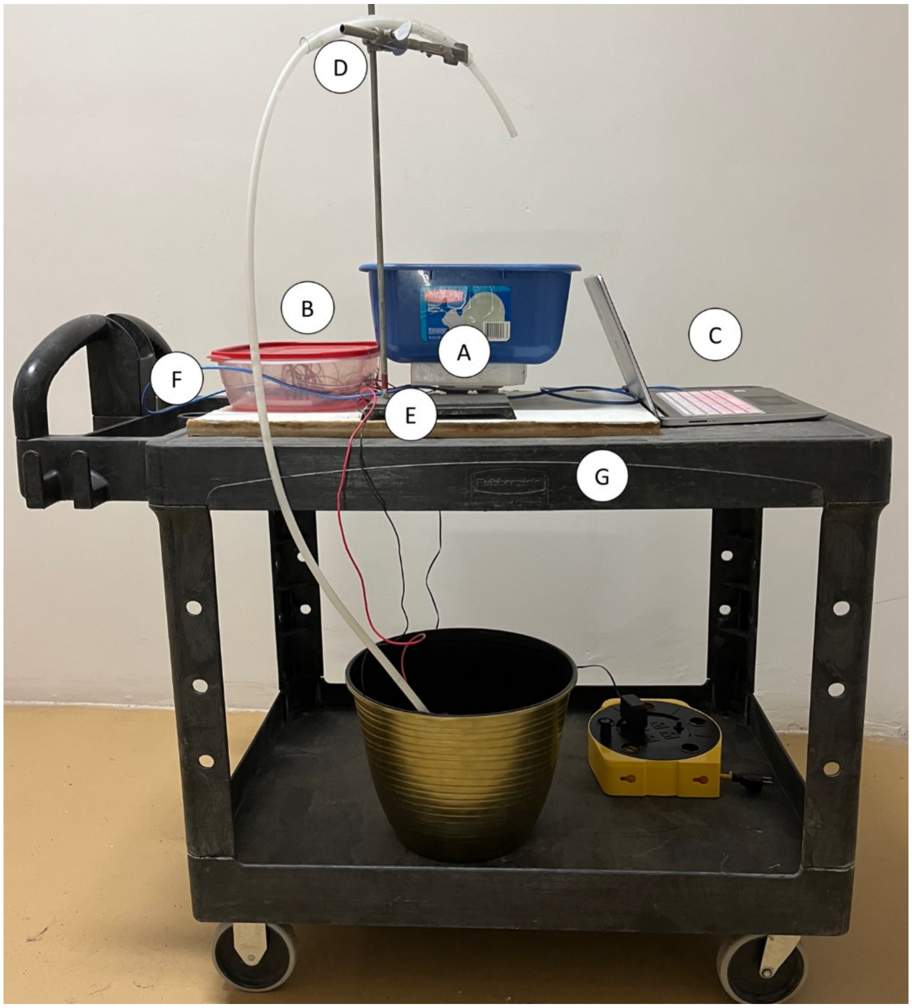


Figure 2.1 Arduino-based lysimeter. (A) Load cell and the pot holding tray, (B) Plastic water-resistant container holding the circuitry, (C) Laptop computer, (D) Hose from water reservoir, (E) Ring stand with burette holder, (F) USB cable to connect the system to connect the system to the laptop, (G) trolley.

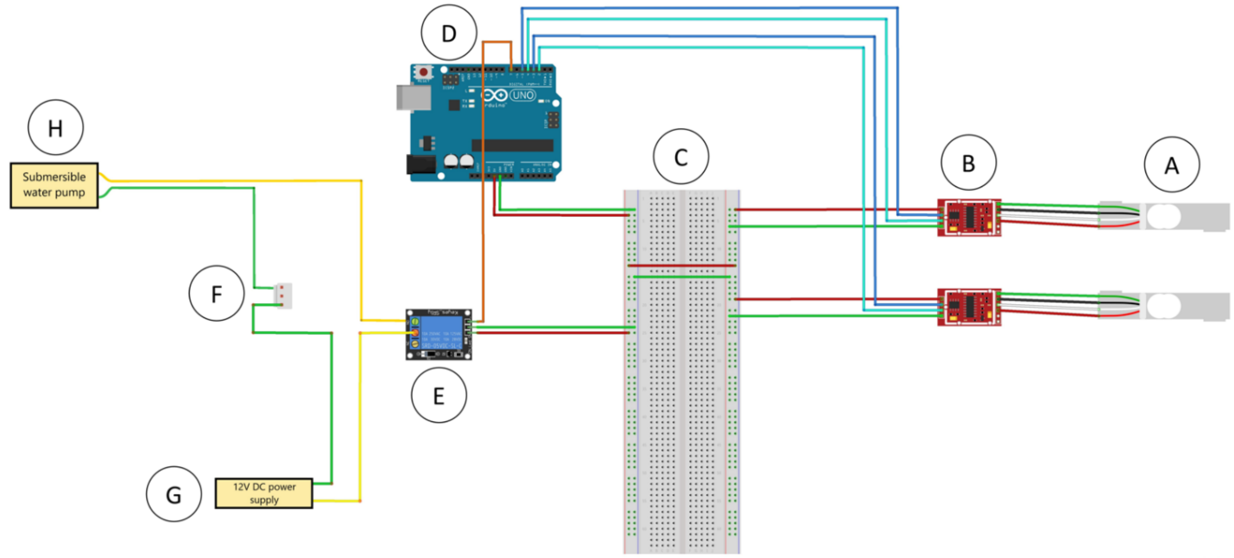
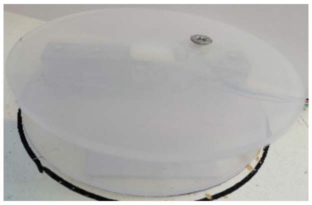
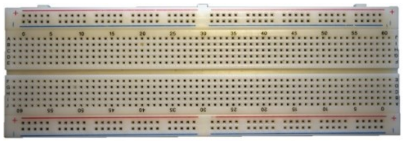


Figure 2.2 Wiring diagram of the Arduino-based lysimeter. (A) Two 20 kg load cells, (B) HX711 load cell amplifier, (C) Breadboard, (D) Arduino Uno microcontroller board, (E) One channel 5 V relay, (F) Compact wire wiring connector with lever, (G) 12 V DC power supply, (H) Submersible water pump (quiet, 12 V, 4.2 W).

Table 2.1. Lysimeter components and specifications.

Component	Component Image	Specifications	Number of Units Required
Arduino Uno R3 USB Microcontroller		Microcontroller: ATmega328 Operating Voltage: 5 V Input Voltage (recommended): 7–12 V Input Voltage (limits): 6–20 V Digital I/O Pins: 14 (of which 6 provide PWM output) Analog Input Pins: 6 DC Current per I/O Pin: 40 mA DC Current for 3.3 V Pin: 50 mA Flash Memory: 32 KB (ATmega328) of which 0.5 KB used by the bootloader SRAM: 2 KB (ATmega328) EEPROM: 1 KB (ATmega328) Clock Speed: 16 MHz	1
Submersible water pump		Ultra-quiet DC12 V 4.2 W Power: Electric Pressure: Low Pressure Structure: Submersible Pump Theory: Brushless Submersible pump	1

Power Adaptor		1 A Output Voltage: 12 V Plug Type: US Plug	1
Relay Module for Arduino		1 Channel 5 V	1
Flexible Hose		Material: Silicon Main Color: Clear Inner Dia (ID): 8 mm Outer Dia (OD): 10 mm; Wall Thickness: 1.2 mm Length: 1 m	1
Compact Wire Wiring Connector with lever		Model Number: 222–412 413 415 Material of Insulation: modified nylon (PA66) Contact material: phosphor copper Wire Cross section: 0.08~2.5 mm^2 (single hardwire), 0.08~4 mm^2 (multi soft wire) Rated Current: 32 A Rated power: 7 KW	1

		Gauge: 28~12 AWG Strip length: 9–10 mm	
Transparent Box Case Shell for Arduino UNO R3		Material: Acrylic Locking mechanism: Screwless locking Color: Transparent	1
USB cable		Material: Metal and environmental PVC Connectors: 4-pin USB type A (male) to 4-pin USB type B (male) Shield: Metal woven mesh + aluminum foil Transmission Rate: Up to 480 Mbps Approval: RoHS Color: Black, light blue Length: 1 m/3.28 ft	1
20 kg pressure sensor + HX711AD module weighing scale		20 kg Tray diameter 10 cm Operating voltage DC 5 V HX711AD (24-bit conversion)	2
Solderless Breadboard		400 tie points in total, 100 in 2 power rails, 300 in a 30 × 10 matrix Transparent plastic, with black legend. Colored power rails For wires 21 to 26 AWG 2-sided peelable adhesive tape	1

Dupont Jumper Wire		Length: 30 cm Package: 20 wires per each category	1
Lab Trolly		Material: Hard Plastic Functions: portable Weight: Lightweight	1
Burette holder/clamp		Material: Metal Functions: Fully adjustable	1
Laptop		Operating system: Windows/Apple operating systems	1

Power extension code		Length: As per your requirement	1
Water reservoir		Any plastic container to hold sufficient amount of water	1

2.2.1.2. Software and code

To set up the system, Arduino IDE (https://www.arduino.cc/en/Guide/Windows, accessed on 21 November 2021) and PuTTy (https://www.putty.org/, accessed on 21 November 2021) were installed on a laptop computer. PuTTy is used as the console as it has the capability of recording the data from the Arduino directly into a text (.txt) file. Data can also be saved to the cloud by configuring PuTTy if needed.

The Arduino Uno board was coded using C++. A specific algorithm was created to carry out the moisture adjustment. The HX711_ADC library was used with the algorithm to calibrate and measure weight data from the load cells. The given calibration example code was used to find the calibration values of the load cells, and those values were included in the main algorithm at the coding stage. As this system consisted of two load cells, the pot weight was determined by adding the two load cell readings. Before starting the watering process, the user should upload the algorithm to the Arduino by selecting the correct code and pressing the upload button in the software. After uploading it once, the Arduino will keep the algorithm in its memory for all the other trials. The relevant C++ codes are listed in Appendix File A.2.

2.2.2 Operating the lysimeter system

2.2.2.1 Determining field capacity

Before experiments begin, soil field capacity needs to be determined. Here, we used field capacity as a proxy for soil field capacity. Field capacity is defined as the amount of soil moisture or water content retained in the soil when all excess water has drained away (Ahuja et al., 2008). For this experiment, 6.52 L plastic pots were filled with a mixture of sand (QUIKRETE® Premium Play Sand) and growing mix (Sunshine Mix #4 Gro Professional) (1:3 volume basis) until a final constant weight was reached (e.g., 4500 g). The bottom of each pot was lined with a coffee filter (12" Mother Parkers Coffee Filters) to prevent soil loss. The initial weight of the dry soil (D_w) was measured after drying the soil in an oven at 80 °C until a constant weight was reached (Earl, 2003). The pots were watered slowly until soil was saturated and water drained out from the holes in the bottom. The top of the pots were covered with aluminum foil and then kept for 24 hrs until water no longer drained from the bottom. The final saturated weights of the pots were recorded (S_w). Field capacity was calculated as $FC = S_w – D_w$. From here, treatment weights can be calculated for a well-watered (W) and a drought (D) treatment. These treatment values (W, D) are then supplied to the algorithm using the Arduino IDE prior to the start of the experiment.

2.2.2.2. Applying soil water treatments

To initiate the process, the correct Arduino code file has to be opened using the Arduino IDE (integrated development environment). Once the upload button is pressed in the Arduino IDE, it will upload the algorithm to the Arduino board. Then the plastic tray has to be placed on the load cells and the PuTTy console software opened on the laptop. This will run the algorithm on the Arduino Board and will open a monitor on the laptop screen which will be used to show the outputs and to send the input values to the Arduino board, also a pre-saved data entry text (.txt) file is opened in the laptop. The algorithm will then initiate the load cells and tare the reading with the weight of the tray. Once this operation is completed, the system will consider the weight of the tray as zero or the tare point and will ask the user to place the first pot on the tray. Once the pot is on the tray, the algorithm will start monitoring the weight readings of the load cells and will identify the peak value readings by determining the point at which the load cell readings will increase and, subsequently, slightly decrease. A minimum threshold weight value was included in the algorithm to improve the accuracy of this operation. If the user accidentally touches the pot or hits the trolley while keeping the pot on the tray, it may affect the load cells and cause the algorithm to read an incorrect peak value. However, the threshold will prevent the algorithm from reading such false peak values before it hits the actual peak weight. Once the algorithm successfully determines the initial weight of the pot, it will ask the user to enter the pot identification number on the laptop. The algorithm will identify whether the user has started typing the number of the pot or not, and it will wait until the user starts typing the pot number. For example, the pot number will denote as 010W4 or 010D4, wherein the first three numbers represent the numerical number given to the pot and the W or D character represents whether it is a well-watered or a drought conditioned pot, respectively. The coding can be modified to assign more than two moisture treatments as well (Appendix File A.3). Finally, the last number will represent the replicate number of the pot. Once the pot identification number is typed on the laptop screen, the algorithm will count the number of characters in the given pot number, then run a loop to search whether there is W or D character in the pot number. The algorithm will identify both capital and simple representations of W and D as valid characters. If it finds a W character, then it will input the prescribed weight corresponding to a well-water conditioned pot in the algorithm. Similarly, if the algorithm finds a D character, it will input the prescribed weight corresponding to the specific drought condition pot in the algorithm. If more than two treatments are needed, the algorithm is available in Appendix - A (Appendix-A File A.4). Additionally, if the user mistakenly entered any other character or forgot to enter any character in the pot number, the algorithm will show a notification to check the pot number and will give space to re-enter it. This loop will run until the user enters a valid pot number.

Once a valid number is entered, the algorithm will check whether the pot weight is below the prescribed weight value (e.g., 5000 g for the well-watered condition and 3000 g for the drought condition). If the weight is equal to or higher than the prescribed value, the program will ask the user to remove the pot, and the data will be saved in a database. If the weight of the pot is below the prescribed value, the program will switch on the water pump. The algorithm will continue monitoring the weight of the pot, where once the pot weight reaches the prescribed value it will switch off the water pump. Once the water pump is switched off, a portion of the residual water in the tube will fall into the pot, each user must determine the weight of this residual water prior to using this system (e.g., We measured this as 40 g) and update in the coding (Appendix - A File A.4). The program will terminate the water pump as soon as the weight scale hits the value of the prescribed weight minus the residual water weight (e.g., 40 g). The laptop display will show the initial weight and the final weight of the pot, and the data will be saved in the text (.txt) file. It is important to note that the weight of the residual water (e.g., 40 g) needs to be added to the final weight as the software only records the weight before adding this residual water to the pot. At the next step, the algorithm will ask the user to remove the pot and wait until the system is ready for the next iteration. This step will take around 1–2 s. However, if ≥100 g of water spilled on the tray in the previous iteration, the system will pause this step until the tray is cleaned and placed on the load cells. If <100 g of water spilled into the tray, the user can either remove the spilled water from the tray before going to the next iteration or the system will automatically tare the scale with a new tare point accounting for the weight of the spilled water. Thus, it is only necessary to clean the tray after each iteration if ≥100 g of water spilled on the tray. The algorithm is also equipped with the capacity of the water reservoir. It is advised to fill the water reservoir completely prior to the start of data collection. The algorithm will notify the user when the water reservoir is almost empty (i.e., when 2 kg of water remains), which will prevent interruptions in the water supply while the pot is being watered. This will improve the life span of the pump as it will prevent the pump from running without water in the reservoir. A demonstration of the system operating is provided in Appendix Video A.1.

2.2.3. Validation of Arduino-based lysimeter

The lysimeter was tested using a greenhouse pot experiment with two soil moisture treatments maintaining the final soil moisture content at 80% and 30% field capacity (FC). The soil was mixed as above, and field capacity determined. Based on the FC data, final weights corresponding to 80% FC (well-watered) and 30% FC (drought) were calculated as 5363.4 g and 4025.9 g, respectively.

The greenhouse pot experiment was conducted using a variety panel of 50 early maturity Canadian soybean varieties. Pots were prepared as described above. Surface sterilized soybean seeds were planted at a depth of 2.5 cm in each pot (3 seeds per pot; 2 pots per variety) and supplied with sufficient water to germinate. Plants were grown at 24 ± 2 °C with supplemental lighting (range: 500–600 µmol m^{-2} s^{-1} at the top of the canopy, Fortimo LED Line, High Flux VO), maintaining a photoperiod of 16:8 h light:dark. One week after germination, seedlings were thinned, leaving one per pot. Seedlings were inoculated with 2 ml of *Bradyrhizobium japonicum* USDA 110, adjusted to a rhizobial density of $OD_{600} = 0.1$ (Thilakarathna et al., 2017). *Bradyrhizobium japonicum* form root nodules in soybean and fix atmospheric nitrogen, and make it available to host soybean plants. Plants were supplied with quarter strength N-free Hoagland's nutrient solution (HOP03-50LT, Caisson Labs, UT) weekly (100 ml per pot). After three weeks of plant growth, two treatments were applied (80% FC (well-watered) and 30% FC (drought)) using the lysimeter system. One pot of each of 50 lines was used in each treatment (n = 50 per treatment). The drought treatment was induced by withholding water until pots reached 30% FC. Moisture adjustment in all the pots was carried out based on the gravimetric method using the lysimeter until six weeks of plant growth (i.e. 3 weeks after treatments were applied).

Cumulative evapotranspiration during the drought-imposed period was calculated based on the moisture deficit (g) between the consecutive irrigation events (amount of water supplied). In addition, plant physiological traits, including photosynthesis, stomatal conductance, and transpiration, were measured using a LI-6400XT portable infrared gas analyzer (LI-COR, Inc., Lincoln, NE, USA). To measure the physiological traits, a photon flux density of 500 µmol m^{-2} s^{-}

^{1} and CO_2 concentration of 400 mol m^{-2} s^{-1} inside the chamber were maintained. Data were collected on a single fully expanded young soybean leaf on each plant in each treatment (n = 50 per treatment).

All statistical analyses were performed using GraphPad Prism Software (v9, GraphPad Software, San Diego, CA, USA). We compared the leaf photosynthesis, stomatal conductance, transpiration, and evapotranspiration between the drought and well-watered treatments with paired t-tests.

2.3 Results and discussion

The semi-automated lysimeter accurately maintained soil moisture levels in both the 80% and 30% FC treatments (Figure 2.3 and Appendix Figure A.1). Over the course of 3 weeks, pots were weighed every 2-5 days depending on the rate of evapotranspiration, and soil moisture content adjusted back to the targeted FC levels (Figure 2.3) based on the pot weight (Appendix Figure A.2). The error variance under each irrigation event was due to the differences in plant size, where the larger genotypes depleted soil water more quickly than the smaller plants. Plants in the 30% FC treatment had significantly lower stomatal conductance (Figure 2.4B), transpiration (Figure 2.4C), and evapotranspiration (Figure 2.4D). However, there was no significant difference in leaf photosynthesis between the 30% and 80% FC treatments (Figure 2.4A). Soybean plants' response to drought stress in terms of leaf stomatal conductance, transpiration, and evapotranspiration is well studied and our results corroborate previous findings (Liu et al., 2003, 2005; Ohashi et al., 2006; Medina and Gilbert, 2016; He et al., 2017). Our results confirmed that the semi-automated lysimeter was effective in inducing and maintaining drought and well-watered soil conditions. Therefore, this system could be applied to other studies aimed at examining plant response to different soil moisture levels.

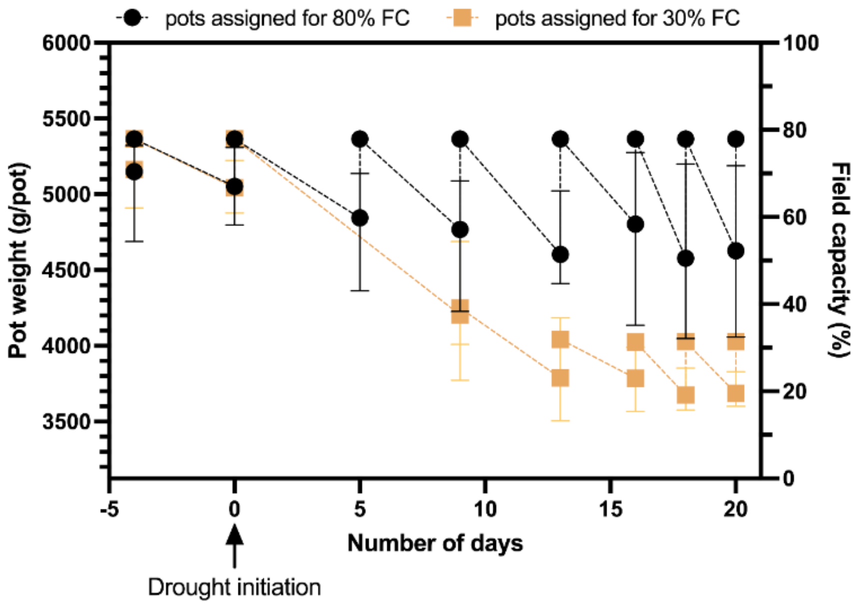


Figure 2.3 Total pot weight and soil water content as a percentage of field capacity (FC) in the well-watered (80% FC) and drought (30% FC) treatments over the soybean plant growth period measured using the Arduino-based lysimeter. The black color represents the 80% FC treatment pot weights, and the orange color represents the 30% FC pot weights. Data are means and ± range (*n* = 50).

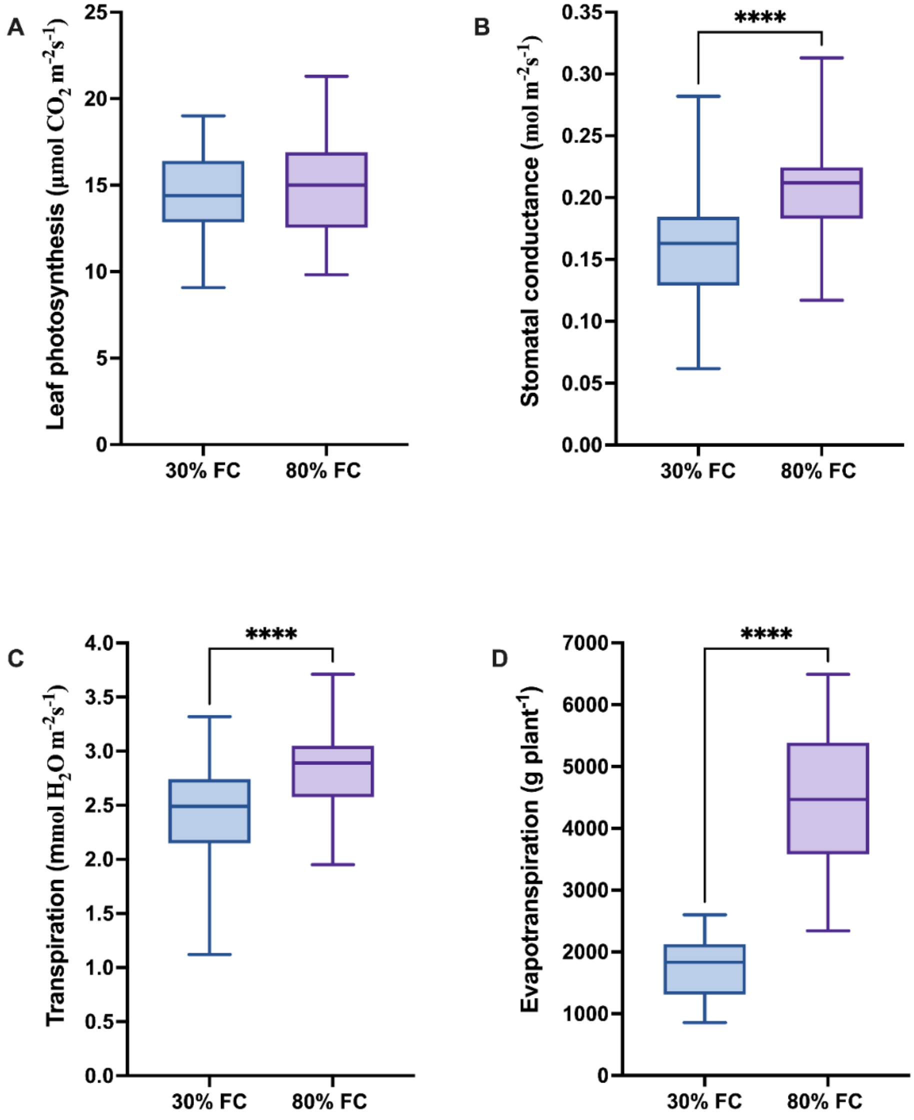


Figure 2.4 Plant physiological parameters and evapotranspiration under drought (30% field capacity) and well-watered (80% field capacity) conditions in soybean. (A) Leaf photosynthesis, (B) Stomatal conductance, (C) Transpiration, (D) Evapotranspiration. Box plots show the median (horizontal line), first and third quartiles (the lower and upper bounds respectively), and whiskers above and below the box plot indicate the range. FC; field capacity. **** indicates $p < 0.0001$.

Mini-lysimeters are portable, accurate, and effective in measuring evapotranspiration in pots (Misra et al., 2011; Meena et al., 2015). Although other mini-lysimeters have been developed to simulate drought stress in pot-based experiments (Earl, 2003; Chard et al., 2010), this newly described system is more portable and economical to build. All the components can be purchased for less than $200 USD, and it will run with a standard laptop that most research groups already have. The data collected can be automatically saved in the cloud, making it easier to handle and access. This semi-automated system is very user-friendly and does not require high technical competence to setup and operate. In the current experiment, it took ca. 1 minute to complete a single pot, which included time to bring the pot to the lysimeter, enter the pot identification number on the laptop, supply the water to the pot, and to return the pot to the greenhouse bench. This time will vary depending on different greenhouse arrangements but is quite efficient. One limiting factor can be the time required to fill the reservoir when the water level is low, but this can be overcome by using a larger reservoir and having a water supply close at hand. On average, we were able to adjust the moisture content in 50 pots per hour, so in 5 working hours in the greenhouse, ca. 250 pots can be adjusted. The current system was created to implement two treatments (well-watered, and drought) only; however, we have supplied the necessary code (Appendix - A File A.4) to increase the number of treatments. Users can setup any number of soil moisture treatments while using the Arduino RAM memory efficiently.

There are some limitations associated with this Arduino-based lysimeter. This semi-automated system is not fully water-resistant as its housing was made using a commonly available plastic box. It will withstand small water splashes, but to make the system fully water-resistant, a custom-made housing compartment with water barrier passages for the wires to pass through could be made. Another alternative solution would be to use two separate and independent systems for the weight measuring load cell section and the data processing Arduino section. However, in this setup, power must be supplied separately to the two units and wireless technology would be needed to transmit the weight data to the Arduino circuitry. This would minimize the complexity of making the whole system water-resistant; however, it would increase the complexity of the system as wireless connectivity would be essential on both devices.

In the validation experiment, we did not consider the weight of the plants for adjusting the FC as the plants grew over time. To make the FC values more accurate, extra pots of plants can be grown and shoot biomass destructively measured at each water adjustment period and added to the final target weight (e.g. 80% FC = 5363.4 g + shoot biomass; and 30% FC = 4025.9 g + shoot biomass) (Earl, 2003). In the future, it is possible to increase the efficiency and functionality of this system. The laptop could be replaced with a small LCD display and a wireless keyboard which could make the system more user-friendly (but potentially increase the technical competence to setup the system). A barcode reader or a QR code reader could also be attached to the system to identify the pot identification numbers, improving the accuracy of the data collection and making the process more efficient (Walden-Coleman et al., 2013). To make the system fully portable, the current wall power connector could be replaced with a rechargeable battery module

2.4 Conclusions

Maintaining accurate soil moisture content is critical in drought experiments. This semi-automated Arduino-based, lysimeter, irrigation system is an economical and high-throughput system for moisture adjustment in pot experiments. It can be further developed to minimize human errors and to reduce the cycle time, which will increase productivity.

References

Adeboye, K. A., Akhter Bhat, J., Ganie, S. A., Varshney, R. K., and Yu, D. (2021). Genome-Wide Association Analysis of yield-related traits of soybean using haplotype-based framework. *Biol. Life Sci. Forum* 11, 49.

Ahuja, L. R., Nachabe, M. H., and Rockiki, R. (2008). Soils: field capacity. *Encycl. Water Sci. Second Ed.* 2, 1128–1131. doi:10.1081/E-EWS--120010264.

Arrese-Igor, C., Esther González, B. M., Daniel Marino, B., Rubén Ladrera, B., Estíbaliz Larrainzar, B., and Erena Gil-Quintana, B. (2011). Physiological responses of legume nodules to drought. *Plant Stress* 5, 24–31.

Arrese-igorl, C., Gonzalezl, E. M., Anthony, J., Minchin, F. R., Galvezl, L., Cabrerizol, P. M., et al. (1999). Sucrose synthase and nodule nitrogen fixation under drought and other environmental stresses. *Symbiosis* 27, 189–212.

Arya, H., Singh, M. B., and Bhalla, P. L. (2021). Towards developing drought-smart soybeans. *Front. Plant Sci.* 12, 750664. doi:10.3389/fpls.2021.750664.

Ashraf, M., and Foolad, M. R. (2007). Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environ. Exp. Bot.* 59, 206–216. doi:10.1016/j.envexpbot.2005.12.006.

Bacelar, E. L. V. A., Moutinho-Pereira, J. M., Gonçalves, B. M. C., Brito, C. V. Q., Gomes-Laranjo, J., Ferreira, H. M. F., et al. (2012). "Water use strategies of plants under drought conditions," in *Plant responses to drought stress*, ed. Aroca R. (Springer, Berlin, Heidelberg.), 145–170. doi:10.1007/978-3-642-32653-0_6.

Basal, O., Szabó, A., and Veres, S. (2020). Physiology of soybean as affected by PEG-induced drought stress. *Curr. Plant Biol.* 22, 100135. doi:10.1016/j.cpb.2020.100135.

Basu, S., Ramegowda, V., Kumar, A., and Pereira, A. (2016). Plant adaptation to drought stress. *F1000Research* 5, 1554. doi:10.12688/F1000RESEARCH.7678.1.

Bellaloui, N., Hu, Y., Mengistu, A., Kassem, M. A., and Abel, C. A. (2013). Effects of foliar boron application on seed composition, cell wall boron, and seed δ15N and δ13C isotopes in water-stressed soybean plants. *Front. Plant Sci.* 4, 270. doi:10.3389/fpls.2013.00270.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate : a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300.

Biate, D. L., Kumar, L. V., Ramadoss, D., Kumari, A., Naik, S., Reddy, K. K., et al. (2014). *Genetic diversity of soybean root nodulating bacteria*. doi:10.1007/978-3-319-05936-5_5.

Blonquist, J. M., Jones, S. B., and Robinson, D. A. (2005). Standardizing characterization of electromagnetic water content sensors: Part 2. evaluation of seven sensing systems. *Vadose Zo. J.* 4, 1059–1069. doi:10.2136/vzj2004.0141.

Bogena, H. R., Huisman, J. A., Schilling, B., Weuthen, A., and Vereecken, H. (2017). Effective calibration of low-cost soil water content sensors. *Sensors* 17, 208. doi:10.3390/s17010208.

Bonsal, B., Liu, Z., Wheaton, E., and Stewart, R. (2020). Historical and projected changes to the stages and other characteristics of severe canadian prairie droughts. *Water (Switzerland)* 12. doi:10.3390/w12123370.

Bonsal, B. R., Aider, R., Gachon, P., and Lapp, S. (2013). An assessment of Canadian prairie drought: past, present, and future. *Clim. Dyn.* 41, 501–516. doi:10.1007/s00382-012-1422-0.

Bonsal, B. R., Wheaton, E. E., Chipanshi, A. C., Lin, C., Sauchyn, D. J., and Wen, L. (2011). Drought research in Canada: A review. *Atmos. - Ocean* 49, 303–319. doi:10.1080/07055900.2011.555103.

Bonsal, B., and Regier, M. (2007). Historical comparison of the 2001/2002 drought in the Canadian Prairies. *Clim. Res.* 33, 229–242. doi:10.3354/cr033229.

Boudhrioua, C., Bastien, M., Torkamaneh, D., and Belzile, F. (2020). Genome-wide association mapping of *Sclerotinia sclerotiorum* resistance in soybean using whole-genome resequencing data. *BMC Plant Biol.* 20, 1–9. doi:10.1186/s12870-020-02401-8.

Bruce, R. W., Grainger, C. M., Ficht, A., Eskandari, M., and Rajcan, I. (2019). Trends in soybean trait improvement over generations of selective breeding. *Crop Sci.* 59, 1870–1879. doi:10.2135/cropsci2018.11.0664.

Campbell, D. I., Laybourne, C. E., and Ian J. Blair (2002). Measuring peat moisture content using the dual-probe heat pulse technique. *Aust. J. Soil Sci.* 40, 177–190.

Carter, T. E., Hymowitz, T., and Nelson, R. L. (2004). "Biogeography, local adaptation, vavilov, and genetic diversity in soybean," in *Biological Resources and Migration* (Springer, Berlin, Heidelberg.), 47–59. doi:10.1007/978-3-662-06083-4.

Chamarthi, S. K., Kaler, A. S., Abdel-Haleem, H., Fritschi, F. B., Gillman, J. D., Ray, J. D., et al. (2021). Identification and confirmation of loci associated with canopy wilting in soybean using genome-wide association mapping. *Front. Plant Sci.* 12, 698116. doi:10.3389/fpls.2021.698116.

Chard, J., van Iersel, M. W., and Bugbee, B. (2010). Mini-lysimeters to monitor transpiration and control drought stress : system design and unique applications.

Chen, L., Fang, Y., Li, X., Zeng, K., Chen, H., Zhang, H., et al. (2020). Identification of soybean drought-tolerant genotypes and loci correlated with agronomic traits contributes new candidate genes for breeding. *Plant Mol. Biol.* 102, 109–122. doi:10.1007/s11103-019-00934-7.

Ciampitti, I. A., and Salvagiotti, F. (2018). New insights into soybean biological nitrogen fixation. *Agron. J.* 110, 1185–1196. doi:10.2134/agronj2017.06.0348.

Collier, R., and Tegeder, M. (2012). Soybean ureide transporters play a critical role in nodule development, function and nitrogen export. *Plant J.* 72, 355–367. doi:10.1111/j.1365-313X.2012.05086.x.

Contreras-Soto, R. I., Mora, F., De Oliveira, M. A. R., Higashi, W., Scapim, C. A., and Schuster, I. (2017). A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-Based haplotype analysis. *PLoS One* 12, 0171105. doi:10.1371/journal.pone.0171105.

Copley, T. R., Duceppe, M. O., and O'Donoughue, L. S. (2018). Identification of novel loci associated with maturity and yield traits in early maturity soybean plant introduction lines. *BMC Genomics* 19, 167. doi:10.1186/s12864-018-4558-4.

Cornic, G., and Briantais, J. M. (1991). Partitioning of photosynthetic electron flow between CO2 and O2 reduction in a C3 leaf (Phaseolus vulgaris L.) at different CO_2 concentrations and during drought stress. *Planta* 183, 178–184. doi:10.1007/BF00197786.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nat. Clim. Chang.* 3, 52–58. doi:10.1038/nclimate1633.

Day, D. A., Poole, P. S., Tyerman, S. D., and Rosendahl, L. (2001). Ammonia and amino acid transport across symbiotic membranes in nitrogen-fixing legume nodules. *Cell. Mol. Life Sci.* 58, 61–71.

Dayoub, E., Lamichhane, J. R., Schoving, C., Debaeke, P., and Maury, P. (2021). Early-stage phenotyping of root traits provides insights into the drought tolerance level of soybean cultivars. *Agronomy* 11. doi:10.3390/agronomy11010188.

Denison, R. F., and Kinraide, T. B. (1995). Oxygen-induced membrane depolarizations in legume root nodules possible evidence for an osmoelectrical mechanism controlling nodule gas permeability. *Plant Physiol.* 108, 235–240.

Desclaux, D., Huynh, T. T., and Roumet, P. (2000). Identification of soybean plant characteristics that indicate the timing of drought stress. *Crop Sci.* 40, 716–722. doi:10.2135/cropsci2000.403716x.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015a). Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. *Theor. Appl. Genet.* 128, 73–91. doi:10.1007/s00122-014-2413-9.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2016). Genome-wide association mapping of soybean chlorophyll traits based on canopy spectral reflectance and leaf extracts. *BMC Plant Biol.* 16, 1–15. doi:10.1186/s12870-016-0861-x.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015b). Genome-Wide Association analysis of diverse soybean genotypes reveals novel markers for nitrogen traits. *Plant Genome* 8, 1–15. doi:10.3835/plantgenome2014.11.0086.

Dixon, R., and Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nat. Rev.* 2, 621–631. doi:10.1038/nrmicro954.

Dobriyal, P., Qureshi, A., Badola, R., and Hussain, S. A. (2012). A review of the methods available for estimating soil moisture and its implications for water resource management. *J. Hydrol.* 458, 110–117. doi:10.1016/j.jhydrol.2012.06.021.

Dong, S., Jiang, Y., Dong, Y., Wang, L., Wang, W., Ma, Z., et al. (2019). A study on soybean responses to drought stress and rehydration. *Saudi J. Biol. Sci.* 26, 2006–2017. doi:10.1016/j.sjbs.2019.08.005.

Downie, J. A. (2005). Legume haemoglobins : symbiotic nitrogen fixation needs bloody nodules. *Curr. Biol.* 15, 196–198. doi:10.1016/j.cub.2005.03.006.

Du, Y., Zhao, Q., Chen, L., Yao, X., and Xie, F. (2020). Effect of drought stress at reproductive stages on growth and nitrogen metabolism in soybean. *Agronomy* 10, 302. doi:10.3390/agronomy10020302.

Dumschott, K., Dechorgnat, J., and Merchant, A. (2019). Water deficit elicits a transcriptional response of genes governing d-pinitol biosynthesis in soybean (*Glycine max*). *Int. J. Mol. Sci.* 20, 1–11. doi:10.3390/ijms20102411.

Earl, H. J. (2003). A precise gravimetric method for simulating drought stress in pot experiments. *Crop Sci.* 43, 1868–1873. doi:10.2135/cropsci2003.1868.

Earl, H. J., and Davis, R. F. (2003). Effect of drought stress on leaf and whole canopy radiation use efficiency and yield of maize. *Agron. J.* 95, 688–696. doi:10.2134/agronj2003.6880.

Ehleringer, J. R., and Cooper, T. A. (1992). On the role of orientation in reducing photoinhibitory damage in photosynthetic-twig desert shrubs. *Plant. Cell Environ.* 15, 301–306. doi:10.1111/j.1365-3040.1992.tb00977.x.

Einsle, O., and Rees, D. C. (2020). Structural enzymology of nitrogenase enzymes. *Chem. Rev.* 120, 4969–5004. doi:10.1021/acs.chemrev.0c00067.

Erdman, J., and John, W. (2009). Soy protein and cardiovascular disease: a statement for healthcare professionals from the nutrition committee of the AHA. *Circulation* 102, 2555–

2559. doi:10.1111/j.1751-7117.2009.00028.x.

Farooq, M., Gogoi, N., Barthakur, S., Baroowa, B., Bharadwaj, N., Alghamdi, S. S., et al. (2017). Drought stress in grain legumes during reproduction and grain filling. *J. Agron. Crop Sci.* 203, 81–102. doi:10.1111/jac.12169.

Farooq, M., Wahid, A., and Siddique, K. H. M. (2012). Micronutrient application through seed treatments - a review. *J. Soil Sci. Plant Nutr.* 12, 125–142. doi:10.4067/S0718-95162012000100011.

Follett, R. F., and Delgado, J. A. (2002). Nitrogen fate and transport in agricultural systems. *J. Soil Water Conserv.* 57, 402–408.

Foyer, C. H., Lam, H. M., Nguyen, H. T., Siddique, K. H. M., Varshney, R. K., Colmer, T. D., et al. (2016). Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants* 2, 1–10. doi:10.1038/NPLANTS.2016.112.

Franche, C., Lindström, K., and Claudine, E. (2009). Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant Soil* 321, 35–59. doi:10.1007/s11104-008-9833-8.

Frederick, J. R., Camp, C. R., and Bauer, P. J. (2001). Crop ecology, production & management: drought-stress effects on branch and mainstem seed yield and yield components of determinate soybean. *Crop Sci.* 41, 759–763. doi:10.2135/cropsci2001.413759x.

Fujita, K., Ofosu-Budu, K. G., and Ogata, S. (1992). Biological nitrogen fixation in mixed legume . cereal cropping systems. *Plant Soil* 141, 155–175.

Fukuda, Y. (1933). Cytogenetical studies on the wild and cultivated Manchurian soybeans (*Glycine* L.). *Japanese J. Bot.* 6, 489–506. Available at: https://ci.nii.ac.jp/naid/10012526143/.

Gadallah, M. A. A. (2000). Effects of indole-3-acetic acid and zinc on the growth, osmotic potential and soluble carbon and nitrogen components of soybean plants growing under water deficit. *J. Arid Environ.* 44, 451–467. doi:10.1006/jare.1999.0610.

Gage, D. J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol. Mol. Biol. Rev.* 68, 280–300. doi:10.1128/MMBR.68.2.280.

Gai, J., Xu, D., Gao, Z., Shimamoto, Y., Abe, J., Fukushi, H., et al. (2000). Studies on the evolutionary relationship among eco-types of *G. max* and *G. soja* in China. *Zuo wu xue bao* 26, 513—520. Available at: http://europepmc.org/abstract/CBA/339184.

Ganjeali, A., Porsa, H., and Bagheri, A. (2011). Assessment of Iranian chickpea (*Cicer arietinum* L.) germplasms for drought tolerance. *Agric. Water Manag.* 98, 1477–1484. doi:10.1016/j.agwat.2011.04.017.

Gebre, M. G., and Earl, H. J. (2020). Effects of growth medium and water stress on soybean [*Glycine max* (L.) Merr.] growth, soil water extraction and rooting profiles by depth in 1-m rooting columns. *Front. Plant Sci.* 11, 487. doi:10.3389/fpls.2020.00487.

Gebre, M. G., and Earl, H. J. (2021). Soil water deficit and fertilizer placement effects on root biomass distribution, soil water extraction, water use, yield, and yield components of soybean [*Glycine max* (L.) Merr.] grown in 1-m rooting columns. *Front. Plant Sci.* 12, 581127. doi:10.3389/fpls.2021.581127.

Geurts, R., Elena, F., and Bisseling, T. (2005). Nod factor signaling genes and their function in the early stages of *Rhizobium* infection. *Plant Biol.* 8, 346–352. doi:10.1016/j.pbi.2005.05.013.

Gil-Quintana, E., Larrainzar, E., Seminario, A., Díaz-Leal, J. L., Alamillo, J. M., Pineda, M., et al. (2013). Local inhibition of nitrogen fixation and nodule metabolism in drought-stressed soybean. *J. Exp. Bot.* 64, 2171–2182. doi:10.1093/jxb/ert074.

González, E. M., Gordon, A. J., James, C. L., and Arrese-lgor, C. (1995). The role of sucrose synthase in the response of soybean nodules to drought. *J. Exp. Bot.* 46, 1515–1523. doi:10.1093/jxb/46.10.1515.

Gordon, A. J., Minchin, F. R., Skøt, L., and James, C. L. (1997). Stress-induced declines in soybean N_2 fixation are related to nodule sucrose synthase activity. *Plant Physiol.* 114, 937–946. doi:10.1104/pp.114.3.937.

Gorim, L. Y., and Vandenberg, A. (2017a). Evaluation of wild lentil species as genetic resources to improve drought tolerance in cultivated lentil. *Front. Plant Sci.* 8, 1129. doi:10.3389/fpls.2017.01129.

Gorim, L. Y., and Vandenberg, A. (2017b). Root traits, nodulation and root distribution in soil for five wild lentil species and *Lens culinaris* (Medik.) grown under well-watered conditions. *Front. Plant Sci.* 8, 1632. doi:10.3389/fpls.2017.01632.

Grant, D., Nelson, R. T., Cannon, S. B., and Shoemaker, R. C. (2009). SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* 38, 843–846. doi:10.1093/nar/gkp798.

Grieshop, C. M., and Fahey, G. C. (2001). Comparison of quality characteristics of soybeans from Brazil, China, and the United States. *J. Agric. Food Chem.* 49, 2669–2673. doi:10.1021/jf0014009.

Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H., et al. (2010). A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): Implications from microsatellites and nucleotide sequences. *Ann. Bot.* 106, 505–514. doi:10.1093/aob/mcq125.

Halbleib, C. M., and Ludden, P. W. (2000). Recent advances in nutritional sciences. *Am. Soc. Nutr. Sci.* 136, 1081–1084.

Hannah, R., and Roser, M. (2021). Forests and deforestation. *OurWorldInData.org*.

Hao, D., Chao, M., Yin, Z., and Yu, D. (2012a). Genome-wide association analysis detecting significant single nucleotide polymorphisms for chlorophyll and chlorophyll fluorescence parameters in soybean (*Glycine max*) landraces. *Euphytica* 186, 919–931. doi:10.1007/s10681-012-0697-x.

Hao, D., Cheng, H., Yin, Z., Cui, S., Zhang, D., Wang, H., et al. (2012b). Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. *Theor. Appl. Genet.* 124, 447–458. doi:10.1007/s00122-011-1719-0.

Hartman, G. L., West, E. D., and Herman, T. K. (2011). Crops that feed the World 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. *Food Secur.* 3, 5–17. doi:10.1007/s12571-010-0108-x.

Hatfield, J. L., and Dold, C. (2019). Water-use efficiency: Advances and challenges in a changing climate. *Front. Plant Sci.* 10, 103. doi:10.3389/fpls.2019.00103.

He, C., Gao, H., Wang, H., Guo, Y., He, M., Peng, Y., et al. (2021). GSK3-mediated stress signaling inhibits legume–rhizobium symbiosis by phosphorylating *GmNSP1* in soybean. *Mol. Plant* 14, 488–502. doi:10.1016/j.molp.2020.12.015.

He, J., Du, Y. L., Wang, T., Turner, N. C., Yang, R. P., Jin, Y., et al. (2017). Conserved water use improves the yield performance of soybean (*Glycine max* (L.) Merr.) under drought. *Agric. Water Manag.* 179, 236–245. doi:10.1016/j.agwat.2016.07.008.

He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H. K., and Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant Sci.* 5, 1–8. doi:10.3389/fpls.2014.00484.

Herridge, D. F., Peoples, M. B., and Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant Soil* 311, 1–18. doi:10.1007/s11104-008-9668-3.

Herritt, M., Dhanapal, A. P., Purcell, L. C., and Fritschi, F. B. (2018). Identification of genomic loci associated with 21chlorophyll fluorescence phenotypes by genome-wide association analysis in soybean. *BMC Plant Biol.* 18, 1–19. doi:10.1186/s12870-018-1517-9.

Hirsch, A. M., and Fujishige, N. A. (2012). Molecular signals and receptors : communication between nitrogen-fixing bacteria and their plant hosts. *Biocommunication of Plants* 14, 255–280.

Höhner, R., Pribil, M., Herbstová, M., Lopez, L. S., Kunz, H. H., Li, M., et al. (2020). Plastocyanin is the long-range electron carrier between photosystem II and photosystem i in plants. *Proc. Natl. Acad. Sci. U. S. A.* 117, 15354–15362. doi:10.1073/pnas.2005832117.

Hossain, M. A., Burritt, D. J., Fujita, M., Munne-Bosch, S., and Pedro, D.-V. (2017). *Ascorbic Acid in Plant Growth , Development and Stress Tolerance*. Springer Publishers doi:10.1007/978-3-319-74057-7.

Hu, D., Zhang, H., Du, Q., Hu, Z., Yang, Z., Li, X., et al. (2020). Genetic dissection of yield-related traits via genome-wide association analysis across multiple environments in wild soybean (*Glycine soja* Sieb. and Zucc.). *Planta* 251, 39. doi:10.1007/s00425-019-03329-6.

Hufstetler, E. V., Boerma, H. R., Carter, T. E., and Earl, H. J. (2007). Genotypic variation for three physiological traits affecting drought tolerance in soybean. *Crop Sci.* 47, 25–35. doi:10.2135/cropsci2006.04.0243.

Hungria, M., and Bohrer, T. R. J. (2000). Variability of nodulation and dinitrogen fixation capacity among soybean cultivars. *Biol. Fertil. Soils* 31, 45–52. doi:10.1007/s003740050622.

Hungria, M., Campo, R. J., Chueire, L. M. O., Grange, L., and Megías, M. (2001). Symbiotic effectiveness of fast-growing rhizobial strains isolated from soybean nodules in Brazil. *Biol. Fertil. Soils* 33, 387–394. doi:10.1007/s003740100338.

Hungria, M., and Vargas, M. A. T. (2000). Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. *F. Crop. Res.* 65, 151–164. doi:10.1016/S0378-4290(99)00084-2.

Hwang, E. Y., Song, Q., Jia, G., Specht, J. E., Hyten, D. L., Costa, J., et al. (2014). A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics* 15, 1–12. doi:10.1186/1471-2164-15-1.

Hwang, S., King, C. A., Ray, J. D., Cregan, P. B., Chen, P., Carter, T. E., et al. (2015). Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. *Theor. Appl. Genet.* 128, 2047–2065. doi:10.1007/s00122-015-2566-1.

Hymowitz, T. (2016). “Speciation and Cytogenetics,” in *Soybeans: Improvement, Production, and Uses* (John Wiley & Sons, Ltd), 97–136. doi:10.2134/agronmonogr16.3ed.c4.

Hymowitz, T., and Harlan, J. R. (1983). Introduction of soybean to North America by Samuel Bowen in 1765. *Econ. Bot.* 37, 371–379. doi:10.1007/BF02904196.

Inamullah, and Isoda, A. (2005). Adaptive responses of soybean and cotton to water stress: I. Transpiration changes in relation to stomatal area and stomatal conductance. *Plant Prod.*

Sci. 8, 16–26. doi:10.1626/pps.8.16.

Islam, M. A., Du, H., Ning, J., Ye, H., and Xiong, L. (2009). Characterization of *Glossy1*-homologous genes in rice involved in leaf wax accumulation and drought resistance. *Plant Mol. Biol.* 70, 443–456. doi:10.1007/s11103-009-9483-0.

Jabbari, M., Fakheri, B. A., Aghnoum, R., Nezhad, N. M., and Ataei, R. (2018). GWAS analysis in spring barley (*Hordeum vulgare* L.) for morphological traits exposed to drought. *PLoS One* 13, e0204952. doi:10.1371/journal.pone.0204952.

Janiak, A., Kwaśniewski, M., and Szarejko, I. (2016). Gene expression regulation in roots under drought. *J. Exp. Bot.* 67, 1003–1014. doi:10.1093/jxb/erv512.

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017). WRKY transcription factors in plant responses to stresses. *J. Integr. Plant Biol.* 59, 86–101. doi:10.1111/jipb.12513.

Jo, H., Lee, J. Y., and Lee, J. D. (2022). Genome-Wide Association mapping for seed weight in soybean with black seed coats and green cotyledons. *Agronomy* 12, 250. doi:10.3390/agronomy12020250.

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. (2007). How rhizobial symbionts invade plants : the *Sinorhizobium – Medicago* model. *Nat. Rev.* 5, 619–633. doi:10.1038/nrmicro1705.

Jumrani, K., and Bhatia, V. S. (2019). Identification of drought tolerant genotypes using physiological traits in soybean. *Physiol. Mol. Biol. Plants* 25, 697–711. doi:10.1007/s12298-019-00665-5.

Kaler, A. S., Dhanapal, A. P., Ray, J. D., King, C. A., Fritschi, F. B., and Purcell, L. C. (2017a). Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. *Crop Sci.* 57, 3085–3100. doi:10.2135/cropsci2017.03.0160.

Kaler, A. S., Ray, J. D., Schapaugh, W. T., King, C. A., and Purcell, L. C. (2017b). Genome-wide association mapping of canopy wilting in diverse soybean genotypes. *Theor. Appl. Genet.* 130, 2203–2217. doi:10.1007/s00122-017-2951-z.

Kantolic, A. G., and Slafer, G. A. (2005). Reproductive development and yield components in indeterminate soybean as affected by post-flowering photoperiod. *F. Crop. Res.* 93, 212–222. doi:10.1016/j.fcr.2004.10.001.

Karr-Lilienthal, L. K., Grieshop, C. M., Merchen, N. R., Mahan, D. C., and Fahey, G. C. (2004). Chemical composition and protein quality comparisons of soybeans and soybean meals from five leading soybean-producing countries. *J. Agric. Food Chem.* 52, 6193–6199. doi:10.1021/jf049795+.

Kato, S., Samanfar, B., Morrison, M. J., Bekele, W. A., Torkamaneh, D., Rajcan, I., et al. (2021). Genome-wide association study to identify soybean stem pushing resistance and lodging resistance loci. *Can. J. Plant Sci.* 101, 663–670. doi:10.1139/cjps-2020-0187.

Keyantash, J., and Dracup, J. A. (2002). The quantification of drought: an evaluation of drought indices. *Bull. Am. Meteorol. Soc.* 83, 1167–1180.

Kim, M. Y., Van, K., Kang, Y. J., Kim, K. H., and Lee, S. H. (2012). Tracing soybean domestication history: From nucleotide to genome. *Breed. Sci.* 61, 445–452. doi:10.1270/jsbbs.61.445.

King, C. A., and Purcell, L. C. (2005). Inhibition of N_2 fixation in soybean is associated with elevated ureides and amino acids. *Plant Physiol.* 137, 1389–1396. doi:10.1104/pp.104.056317.1.

Kojima, Y., Shigeta, R., Miyamoto, N., Shirahama, Y., Nishioka, K., Mizoguchi, M., et al. (2016). Low-cost soil moisture profile probe using thin-film capacitors and a capacitive touch sensor. *Sensors* 16, 1292. doi:10.3390/s16081292.

Kunert, K. J., Vorster, B. J., Fenta, B. A., Kibido, T., Giuseppe, D., and Foyer, C. H. (2016). Drought stress responses in soybean roots and nodules. *Front. Plant Sci.* 7, 1015. doi:10.3389/fpls.2016.01015.

Larcher, W. (2000). Temperature stress and survival ability of mediterranean sclerophyllous plants. *Plant Biosyst.* 134, 279–295. doi:10.1080/11263500012331350455.

Le, D. T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Ham, L. H., et al. (2012). Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by Genome-Wide Transcriptome analysis. *PLoS One* 7, e495. doi:10.1371/journal.pone.0049522.

Lehnert, H., Serfling, A., Friedt, W., and Ordon, F. (2018). Genome-wide association studies reveal genomic regions associated with the response of wheat (*Triticum aestivum* L.) to mycorrhizae under drought stress conditions. *Front. Plant Sci.* 9, 1728. doi:10.3389/fpls.2018.01728.

Levis, S., Strickman-Stein, N., Ganjei-Azar, P., Xu, P., Doerge, D. R., and Krischer, J. (2011). Soy isoflavones in the prevention of menopausal bone loss and menopausal symptoms: A randomized, double-blind trial. *Arch. Intern. Med.* 171, 1363–1369. doi:10.1001/archinternmed.2011.330.

Li, F., Chen, B., Xu, K., Gao, G., Yan, G., Qiao, J., et al. (2016a). A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*). *Plant Sci.* 242, 169–177. doi:10.1016/j.plantsci.2015.05.012.

Li, H., Yang, Y., Zhang, H., Chu, S., Zhang, X., Yin, D., et al. (2016b). A genetic relationship between phosphorus efficiency and photosynthetic traits in soybean as revealed by QTL analysis using a high-density genetic map. *Front. Plant Sci.* 7, 1–16. doi:10.3389/fpls.2016.00924.

Li, X., Zhou, Y., Bu, Y., Wang, X., Zhang, Y., Guo, N., et al. (2021). Genome-wide association analysis for yield-related traits at the R6 stage in a Chinese soybean mini core collection. *Genes and Genomics* 43, 897–912. doi:10.1007/s13258-021-01109-9.

Li, Y. H., Li, W., Zhang, C., Yang, L., Chang, R. Z., Gaut, B. S., et al. (2010). Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. *New Phytol.* 188, 242–253. doi:10.1111/j.1469-8137.2010.03344.x.

Li, Y., Li, H., Li, Y., and Zhang, S. (2017). Improving water-use efficiency by decreasing stomatal conductance and transpiration rate to maintain higher ear photosynthetic rate in drought-resistant wheat. *Crop J.* 5, 231–239. doi:10.1016/j.cj.2017.01.001.

Liu, F., Andersen, M. N., Jacobsen, S. E., and Jensen, C. R. (2005). Stomatal control and water use efficiency of soybean (*Glycine max* L. Merr.) during progressive soil drying. *Environ. Exp. Bot.* 54, 33–40. doi:10.1016/j.envexpbot.2004.05.002.

Liu, F., Jensen, C. R., and Andersen, M. N. (2003). Hydraulic and chemical signals in the control of leaf expansion and stomatal conductance in soybean exposed to drought stress. *Funct. Plant Biol.* 30, 65–73. doi:10.1071/FP02170.

Liu, K. (1997). *Soybeans: chemistry, technology and utilization'*. International Thomson Publishing Asia.

Liu, K. S. (2008). "Food use of whole soybeans," in *Soybeans: Chemistry, Production, Processing, and Utilization* (AOCS Press), 441–481. doi:10.1016/B978-1-893997-64-6.50017-2.

Liu, S., Zhang, M., Feng, F., and Tian, Z. (2020). Toward a "Green Revolution" for Soybean. *Mol. Plant* 13, 688–697. doi:10.1016/j.molp.2020.03.002.

Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS One* 12, e1005767. doi:10.1186/1471-2156-13-100.

Liyanage, D. K., Chathuranga, I., Mori, B. A., and Thilakarathna, M. S. (2022). A simple , semi-automated , gravimetric method to simulate drought stress on plants. *Agronomy* 12, 349.

Lü, H., Yang, Y., Li, H., Liu, Q., Zhang, J., Yin, J., et al. (2018). Genome-wide association studies of photosynthetic traits related to phosphorus efficiency in soybean. *Front. Plant Sci.* 9. doi:10.3389/fpls.2018.01226.

Ma, J., Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., et al. (2022). Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4 .0 , r10285). *Geosci. Model Dev.* 15, 815–839.

Mak, M., Babla, M., Xu, S. C., O'Carrigan, A., Liu, X. H., Gong, Y. M., et al. (2014). Leaf mesophyll K^+, H+ and Ca^{2+} fluxes are involved in drought-induced decrease in photosynthesis and stomatal closure in soybean. *Environ. Exp. Bot.* 98, 1–12. doi:10.1016/j.envexpbot.2013.10.003.

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

Marchin, R. M., Ossola, A., Leishman, M. R., and Ellsworth, D. S. (2020). A simple method for simulating drought effects on plants. *Front. Plant Sci.* 10, 1715. doi:10.3389/fpls.2019.01715.

Marcolino-Gomes, J., Rodrigues, F. A., Fuganti-Pagliarini, R., Bendix, C., Nakayama, T. J., Celaya, B., et al. (2014). Diurnal oscillations of soybean circadian clock and drought responsive genes. *PLoS One* 9. doi:10.1371/journal.pone.0086402.

Mateos-Aparicio, I., Cuenca, A. R., Villanueva-Suárez M.J., and Zapata-Revilla, M. A. (2008). Soybean , a promising health source. *Nutr. Hosp.* 23, 305–312.

McCauley, D., Levin, A., and Nackley, L. (2021). Reviewing mini-lysimeter controlled irrigation in container crop systems. *Hort Technol.*, 1–8. doi:10.21273/horttech04826-21.

Medina, V., and Gilbert, M. E. (2016). Physiological trade-offs of stomatal closure under high evaporative gradients in field grown soybean. *Funct. Plant Biol.* 43, 40–51. doi:10.1071/FP15304.

Meena, H. M., R.K.Singh, and P.Santra (2015). Design and development of a load-cell based cost effective mini- lysimeter. *J. Agric. Phys.* 15, 1–6.

Minamisawa, K., Isawa, T., Nakatsuka, Y., and Ichikawa, N. (1998). New *Bradyrhizobium japonicum* strains that possess high copy numbers of the repeated sequence RSα. *Appl. Environ. Microbiol.* 64, 1845–1851. doi:10.1128/aem.64.5.1845-1851.1998.

Misra, R. K., Padhi, J., and Payero, J. O. (2011). A calibration procedure for load cells to improve accuracy of mini-lysimeters in monitoring evapotranspiration. *J. Hydrol.* 406, 113–118. doi:10.1016/j.jhydrol.2011.06.009.

Mustofa, A., Zubaidah, S., and Kuswantoro, H. (2021). Correlation and path analysis on yield and yield components in segregating populations. *Genetika* 53, 157–166. doi:10.2298/GENSR2101157R.

Mutava, R. N., Prince, S. J. K., Syed, N. H., Song, L., Valliyodan, B., Chen, W., et al. (2015). Understanding abiotic stress tolerance mechanisms in soybean: a comparative evaluation of soybean response to drought and flooding stress. *Plant Physiol. Biochem.* 86, 109–120.

Mwenye, O. J., van Rensburg, L., van Biljon, A., and van der Merwe, R. (2016). The role of proline and root traits on selection for drought-stress tolerance in soybeans: a review. *South African J. Plant Soil* 33, 245–256. doi:10.1080/02571862.2016.1148786.

Naderi, A., Naseri, R., Fathi, A., Bahamin, S., and Maleki, R. (2013). Physiological performance of soybean cultivars under drought stress. *Bull. Environ. Pharmacol. Life Sci.* 2, 38–44.

Naya, L., Ladrera, R., Ramos, J., González, E. M., Arrese-Igor, C., Minchin, F. R., et al. (2007). The response of carbon metabolism and antioxidant defenses of alfalfa nodules to drought stress and to the subsequent recovery of plants. *Plant Physiol.* 144, 1104–1114. doi:10.1104/pp.107.099648.

Neo, H. H., and Layzell, D. B. (1997a). Phloem glutamine and the regulation of O_2 diffusion in legume nodules. *Plant Physiol.* 113, 259–267. doi:10.1104/pp.113.1.259.

Neo, H. H., and Layzell, D. B. (1997b). Phloem glutamine and the regulation of oxygen , diffusion in legume nodules ’. *Plant Physiol.* 113, 259–267.

Nouri, M. Z., Moumeni, A., and Komatsu, S. (2015). Abiotic stresses: Insight into gene regulation and protein expression in photosynthetic pathways of plants. *Int. J. Mol. Sci.* 16, 20392–20416. doi:10.3390/ijms160920392.

Ohashi, Y., Nakayama, N., Saneoka, H., and Fujita, K. (2006). Effects of drought stress on photosynthetic gas exchange, chlorophyll fluorescence and stem diameter of soybean plants. *Biol. Plant.* 50, 138–141. doi:10.1007/s10535-005-0089-3.

Ortiz, D., Litvin, A. G., and Salas Fernandez, M. G. (2018). A cost-effective and customizable automated irrigation system for precise high-throughput phenotyping in drought stress studies. *PLoS One* 13, e0198546. doi:10.1371/journal.pone.0198546.

Oya, T., Nepomuceno, A. L., Neumaier, N., Farias, J. R. B., Tobita, S., and Ito, O. (2004). Drought tolerance characteristics of Brazilian soybean cultivars - evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field. *Plant Prod. Sci.* 7, 129–137. doi:10.1626/pps.7.129.

Pantalião, G. F., Narciso, M., Guimarães, C., Castro, A., Colombari, J. M., Breseghello, F., et al. (2016). Genome wide association study (GWAS) for grain yield in rice cultivated under water deficit. *Genetica* 144, 651–664. doi:10.1007/s10709-016-9932-z.

Patinni, I. R. G., de Andrade, C. A., Campos, C. N. S., Teodoro, L. P. R., Andrade, S. M., Roque, C. G., et al. (2020). Agronomic performance and water-use efficiency of F3 soybean populations grown under contrasting base saturation. *J. Agron. Crop Sci.* 206, 806–814. doi:10.1111/jac.12413.

Pejić, B., Maksimović, L., Cimpeanu, S., Bucur, D., Milić, S., and Ćupina, B. (2011). Response of soybean to water stress at specific growth stages. *J. Food, Agric. Environ.* 9, 280–284.

Phang, T.-H., Li, M.-W., Cheng, C.-C., Wong, F.-L., Chan, C., and Lam, H.-M. (2011). Molecular Responses to Osmotic Stresses in Soybean. *Soybean - Mol. Asp. Breed.* doi:10.5772/15155.

Pitumpe Arachchige, P. S., Rosso, L. H. M., Hansel, F. D., Ramundo, B., Torres, A. R., Asebedo, R., et al. (2020). Temporal biological nitrogen fixation pattern in soybean inoculated with *Bradyrhizobium*. *Agrosystems, Geosci. Environ.* 3, 1–11. doi:10.1002/agg2.20079.

Priyanatha, C., Torkamaneh, D., and Rajcan, I. (2022). Genome-Wide Association Study of soybean germplasm derived from Canadian × Chinese crosses to mine for novel alleles to improve seed yield and seed quality traits. *Front. Plant Sci.* 13. doi:10.3389/fpls.2022.866300.

Purcell, L. C., Salmeron, M., and Ashlock, L. (2014). “Soybean growth and development,” in *Arkansas soybean production handbook* (University of Arkansas), 1–8.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795.

Quero, G., Simondi, S., Ceretta, S., Otero, Á., Garaycochea, S., Fernández, S., et al. (2021). An integrative analysis of yield stability for a GWAS in a small soybean breeding population. *Crop Sci.* 61, 1903–1914. doi:10.1002/csc2.20490.

R Core Team (2021). R: a language and environment for statistical computing. Available at: https://www.r-project.org/.

Raj, A., Stephens, M., and Pritchard, J. K. (2014). FastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. doi:10.1534/genetics.114.164350.

Ravelombola, W., Qin, J., Shi, A., Song, Q., Yuan, J., Wang, F., et al. (2021). Genome-wide association study and genomic selection for yield and related traits in soybean. *PLoS One* 16, e0255761. doi:10.1371/journal.pone.0255761.

Rebolledo, M. C., Dingkuhn, M., Courtois, B., Gibon, Y., Clément-Vidal, A., Cruz, D. F., et al. (2015). Phenotypic and genetic dissection of component traits for early vigour in rice using plant growth modelling, sugar content analyses and association mapping. *J. Exp. Bot.* 66, 5555–5566. doi:10.1093/jxb/erv258.

Robinson, D. A., Jones, S. B., Wraith, J. M., Or, D., and Friedman, S. P. (2003). A Review of advances in dielectric and electrical conductivity measurement in soils using time domain reflectometry. *Vadose Zo. J.* 2, 444–475. doi:10.2136/vzj2003.4440.

Rodríguez-Navarro, D. N., Margaret Oliver, I., Albareda Contreras, M., and Ruiz-Sainz, J. E. (2011). Soybean interactions with soil microbes, agronomical and molecular aspects. *Agron. Sustain. Dev.* 31, 173–190. doi:10.1051/agro/2010023.

Rutten, P. J., and Poole, P. S. (2019). *Oxygen regulatory mechanisms of nitrogen fixation in rhizobia*. Elsevier Ltd doi:10.1016/bs.ampbs.2019.08.001.

Saeidi, M., and Abdoli, M. (2015). Effect of drought stress during grain filling on yield and its components, gas exchange variables, and some physiological traits of wheat cultivars. *J. Agric. Sci. Technol.* 17, 885–898.

Santachiara, G., Salvagiotti, F., and Rotundo, J. L. (2019). Nutritional and environmental effects on biological nitrogen fixation in soybean: A meta-analysis. *F. Crop. Res.* 240, 106–115.

doi:10.1016/j.fcr.2019.05.006.

Sebastián, D. I., Fernando, F. D., Raúl, D. G., and Gabriela, G. M. (2020). Overexpression of *Arabidopsis* aspartic protease APA1 gene confers drought tolerance. *Plant Sci.* 292, 110406. doi:10.1016/j.plantsci.2020.110406.

Seck, W., Torkamaneh, D., and Belzile, F. (2020). Comprehensive genome-wide association analysis reveals the genetic basis of root system srchitecture in soybean. *Front. Plant Sci.* 11, 590740. doi:10.3389/fpls.2020.590740.

Serraj, R. (2003). Effects of drought stress on legume symbiotic nitrogen fixation: physiological mechanisms. *Indian J. Exp. Biol.* 41, 1136–1141.

Serraj, R. S., Adez, V. V, and Inclair, T. R. S. (2001). Feedback regulation of symbiotic N_2 fixation under drought stress. *Agronomie* 21, 621–626.

Serraj, R., Sinclair, T. R., and Purcell, L. C. (1999a). Symbiotic N_2 fixation response to drought. *J. Exp. Bot.* 50, 143–155. doi:10.1093/jxb/50.331.143.

Serraj, R., Sinclair, T. R., and Purcell, L. C. (1999b). Symbiotic N_2 fixation response to drought. *J. Exp. Bot.* 50, 143–155. doi:10.1093/jxb/50.331.143.

Sheffield, J., and Wood, E. F. (2008). Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Clim. Dyn.* 31, 79–105. doi:10.1007/s00382-007-0340-z.

Sheffield, J., Wood, E. F., and Roderick, M. L. (2012). Little change in global drought over the past 60 years. *Nature* 491, 435–438. doi:10.1038/nature11575.

Shi, W. Y., Du, Y. T., Ma, J., Min, D. H., Jin, L. G., Chen, J., et al. (2018). The WRKY transcription factor *GmWRKY12* confers drought and salt tolerance in soybean. *Int. J. Mol. Sci.* 19, 4087. doi:10.3390/ijms19124087.

Shibata, M., Breuer, C., Kawamura, A., Clark, N. M., Rymen, B., Braidwood, L., et al. (2018). GTL1 and DF1 regulate root hair growth through transcriptional repression of ROOT HAIR DEFECTIVE 6-LIKE 4 in Arabidopsis. *Dev.* 145, 1–11. doi:10.1242/dev.159707.

Shirinbayan, S., Khosravi, H., and Malakouti, M. J. (2019). Alleviation of drought stress in maize

(*Zea mays*) by inoculation with *Azotobacter* strains isolated from semi-arid regions. *Appl. Soil Ecol.* 133, 138–145. doi:10.1016/j.apsoil.2018.09.015.

Shurtleff, W., and Aoyagi, A. (2010). *History of soybeans and soyfood in Canada (1831-2010): Extensively annotated bibliography and sourcebook.* Soyinfo Center,Lafayette, CA.

Siddiqui, M. H., Al-Khaishany, M. Y., Al-Qutami, M. A., Al-Whaibi, M. H., Grover, A., Ali, H. M., et al. (2015a). Response of different genotypes of faba bean plant to drought stress. *Int. J. Mol. Sci.* 16, 10214–10227. doi:10.3390/ijms160510214.

Siddiqui, Z. S., Cho, J. Il, Kwon, T. R., Ahn, B. O., Lee, K. S., Jeong, M. J., et al. (2014). Physiological mechanism of drought tolerance in transgenic rice plants expressing *Capsicum annuum* methionine sulfoxide reductase B2 (CaMsrB2) gene. *Acta Physiol. Plant.* 36, 1143–1153. doi:10.1007/s11738-014-1489-9.

Siddiqui, Z. S., Cho, J. Il, Park, D. B., Lee, G. S., Ryu, T. H., Shahid, H., et al. (2015b). Field assessment of CaMsrB2 transgenic lines in a drought stress environment. *Turk. J. Botany* 39, 973–981. doi:10.3906/bot-1502-18.

Silva, A. F., Sediyama, T., Silva, F. C. S., Bezerra, A. R. G., and Ferreira, L. V. (2015). Correlation and path analysis of yield components in soybean varieties. *Turkish J. F. Crop.* 10, 177–179.

Silvente, S., Sobolev, A. P., and Lara, M. (2012). Metabolite adjustments in drought tolerant and sensitive soybean genotypes in response to water stress. *PLoS One* 7, e38554. doi:10.1371/journal.pone.0038554.

Singh, G. (2010). *The soybean: botany, production and uses.* CABI doi:10.1016/s0016-0032(32)91432-x.

Singh, J., and Verma, P. K. (2021). NSP1 allies with GSK3 to inhibit nodule symbiosis. *Trends Plant Sci.* 26, 999–1001. doi:10.1016/j.tplants.2021.07.001.

Son, S., Oh, C. J., Bae, J. H., Lee, H., and An, C. S. (2015). GmREM1.1 and GmREM2.1, which encode the remorin proteins in soybean, have distinct roles during root nodule development. *J. Plant Biol.* 58, 17–25. doi:10.1007/s12374-014-0273-0.

Sonah, H., O'Donoughue, L., Cober, E., Rajcan, I., and Belzile, F. (2015). Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. *Plant Biotechnol. J.* 13, 211–221. doi:10.1111/pbi.12249.

Specht, J. E., and Williams (2022). Breeding for drought and heat resistance: Prerequisites and examples. in *World Soybean Research Conference III: Proceedings* (CRC Press), 468–475.

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

Steketee, C. J., Sinclair, T. R., Riar, M. K., Schapaugh, W. T., and Li, Z. (2019). Unraveling the genetic architecture for carbon and nitrogen related traits and leaf hydraulic conductance in soybean using genome-wide association analyses. *BMC Genomics* 20, 811. doi:10.1186/s12864-019-6170-7.

Sulieman, S., and Tran, L. S. P. (2016). *Legume nitrogen fixation in a changing environment: achievements and challenges*. doi:10.1007/978-3-319-06212-9.

Susha Lekshmi, S. U., Singh, D. N., and Shojaei Baghini, M. (2014). A critical review of soil moisture measurement. *J. Int. Meas. Confed.* 54, 92–105. doi:10.1016/j.measurement.2014.04.007.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., et al. (2002). Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J.* 29, 417–426. doi:10.1046/j.0960-7412.2001.01227.x.

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484. doi:10.1038/s41576-019-0127-1.

Tarawneh, R. A., Alqudah, A. M., Nagel, M., and Börner, A. (2020). Genome-wide association mapping reveals putative candidate genes for drought tolerance in barley. *Environ. Exp. Bot.* 180, 104237. doi:10.1016/j.envexpbot.2020.104237.

Taylor, H. M., Burnett, E., and Booth, G. D. (1978). Taproot elongation rates of soybeans.

Zeitschrift fuer Acker und Pflanzenbau 146, 33–39.

Thilakarathna, M. S., McElroy, M. S., Chapagain, T., Papadopoulos, Y. A., and Raizada, M. N. (2016a). Belowground nitrogen transfer from legumes to non-legumes under managed herbaceous cropping systems. A review. *Agron. Sustain. Dev.* 36. doi:10.1007/s13593-016-0396-4.

Thilakarathna, M. S., Moroz, N., and Raizada, M. N. (2017). A biosensor-based leaf punch assay for glutamine correlates to symbiotic nitrogen fixation measurements in legumes to permit rapid screening of *rhizobia* inoculants under controlled conditions. *Front. Plant Sci.* 8, 1714. doi:10.3389/fpls.2017.01714.

Thilakarathna, M. S., Papadopoulos, Y. A., Rodd, A. V., Grimmett, M., Fillmore, S. A. E., Crouse, M., et al. (2016b). Nitrogen fixation and transfer of red clover genotypes under legume–grass forage based production systems. *Nutr. Cycl. Agroecosystems* 106, 233–247. doi:10.1007/s10705-016-9802-1.

Thilakarathna, M. S., and Raizada, M. N. (2017). A meta-analysis of the effectiveness of diverse rhizobia inoculants on soybean traits under field conditions. *Soil Biol. Biochem.* 105, 177–196. doi:10.1016/j.soilbio.2016.11.022.

Thilakarathna, M. S., Torkamaneh, D., Bruce, R. W., Rajcan, I., Chu, G., Grainger, C. M., et al. (2021). Testing whether pre-pod-fill symbiotic nitrogen fixation in soybean is subject to drift or selection over 100 years of soybean breeding. *Front. Agron.* 3, 725813. doi:10.3389/fagro.2021.725813.

Tombesi, S., Nardini, A., Frioni, T., Soccolini, M., Zadra, C., Farinelli, D., et al. (2015). Stomatal closure is induced by hydraulic signals and maintained by ABA in drought-stressed grapevine. *Nature* 5, 12449. doi:10.1038/srep12449.

Torkamaneh, D., Chalifour, F. P., Beauchamp, C. J., Agrama, H., Boahen, S., Maaroufi, H., et al. (2020). Genome-wide association analyses reveal the genetic basis of biomass accumulation under symbiotic nitrogen fixation in African soybean. *Theor. Appl. Genet.* 133, 665–676. doi:10.1007/s00122-019-03499-7.

Torkamaneh, D., Laroche, J., Tardivel, A., O'Donoughue, L., Cober, E., Rajcan, I., et al. (2018).

Comprehensive description of genomewide nucleotide and structural variation in short-season soya bean. *Plant Biotechnol. J.* 16, 749–759. doi:10.1111/pbi.12825.

Turner, N. C. (2019). Imposing and maintaining soil water deficits in drought studies in pots. *Plant Soil* 439, 45–55. doi:10.1007/s11104-018-3893-1.

Turner, N. C., Wright, G. C., and Siddique, K. H. M. (2001). *Adaptation of grain legumes (pulses) to water-limited environments.*

Turuspekov, Y., Baibulatova, A., Yermekbayev, K., Tokhetova, L., Chudinov, V., Sereda, G., et al. (2017). GWAS for plant growth stages and yield components in spring wheat (*Triticum aestivum* L.) harvested in three regions of Kazakhstan. *BMC Plant Biol.* 17. doi:10.1186/s12870-017-1131-2.

Tuteja, N., Gill, S. S., and Tuteja, R. (2011). Plant responses to abiotic stresses : shedding light on salt , drought , cold and heavy metals. *Omi. Plant Abiotic Stress Toler.* 1, 39–64.

United Nations (2019). Department of economics and social affairs. https://www.un.org/development/desa/en/news/population/world-population-prospects-2019.html Malinda Thilakarathna.

Vadez, V., Sinclair, T. R., and Serraj, R. (2000). Asparagine and ureide accumulation in nodules and shoots as feedback inhibitors of N_2 fixation in soybean. *Physiol. Plant.* 110, 215–223. doi:10.1034/j.1399-3054.2000.110211.x.

Valentine, A. J., A., B. V., and Yun, K. (2011). “Legume nitrogen fixation and soil abiotic stress : from physiology to genomics and beyond,” in *Nitrogen Metabolism in Plants in the Post-genomic Era* (Blackwell Publishing Ltd), 207–248.

Vance, C. P., and Gantt, S. J. (1992). Control of nitrogen and carbon metabolism in root nodules. *Physiol. Plant.* 85, 266–274.

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., et al. (2017). ePlant: Visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *Plant Cell* 29, 1806–1821. doi:10.1105/tpc.17.00073.

Walden-Coleman, A. E., Rajcan, I., and Earl, H. J. (2013). Dark-adapted leaf conductance, but not minimum leaf conductance, predicts water use efficiency of soybean (*Glycine max* L. Merr.). *Can. J. Plant Sci.* 93, 13–22. doi:10.4141/CJPS2012-178.

Wang, J., Chu, S., Zhang, H., Zhu, Y., Cheng, H., and Yu, D. (2016a). Development and application of a novel genome-wide SNP array reveals domestication history in soybean. *Nature* 6, 20728. doi:10.1038/srep20728.

Wang, L., Yang, Y., Zhang, S., Che, Z., Yuan, W., and Yu, D. (2020). GWAS reveals two novel loci for photosynthesis-related traits in soybean. *Mol. Genet. Genomics* 295, 705–716. doi:10.1007/s00438-020-01661-1.

Wang, W., Wang, C., Pan, D., Zhang, Y., Luo, B., and Ji, J. (2018). Effects of drought stress on photosynthesis and chlorophyll fluorescence images of soybean (*Glycine max*) seedlings. *Int. J. Agric. Biol. Eng.* 11, 196–201. doi:10.25165/j.ijabe.20181102.3390.

Wang, X., Oh, M. W., and Komatsu, S. (2016b). Characterization of S-adenosylmethionine synthetases in soybean under flooding and drought stresses. *Biol. Plant.* 60, 269–278. doi:10.1007/s10535-016-0586-6.

Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS One* 14, 0208217. doi:10.1371/journal.pone.0208217.

Wasaya, A., Zhang, X., Fang, Q., and Yan, Z. (2018). Root phenotyping for drought tolerance: A review. *Agronomy* 8, 241. doi:10.3390/agronomy8110241.

Wei, Y., Jin, J., Jiang, S., Ning, S., and Liu, L. (2018). Quantitative response of soybean development and yield to drought stress during different growth stages in the Huaibei plain, China. *Agronomy* 8, 97. doi:10.3390/agronomy8070097.

Wilhite, D. A. (2000). Drought as a natural hazard: concepts and definitions. *Drought Mitig. Cent. Fac. Publ.*, 1–21.

Wilker, J., Navabi, A., Rajcan, I., Marsolais, F., Hill, B., Torkamaneh, D., et al. (2019). Agronomic performance and nitrogen fixation of heirloom and conventional dry bean varieties under low-nitrogen field conditions. *Front. Plant Sci.* 10, 952.

doi:10.3389/fpls.2019.00952.

Wilkinson, S., and Davies, W. J. (2002). ABA-based chemical signalling : the co-ordination of responses to stress in plants. *Plant. Cell Environ.* 25, 195–210.

Wright, D. E. J., Cline, J. A., and Earl, H. J. (2019). Physiological responses of four apple (*Malus × domestica* borkh.) rootstock genotypes to soil water deficits. *Can. J. Plant Sci.* 99, 510–524. doi:10.1139/cjps-2018-0276.

Wu, C., Ma, Q., Yam, K. M., Cheung, M. Y., Xu, Y., Han, T., et al. (2006). In situ expression of the GmNMH7 gene is photoperiod-dependent in a unique soybean (*Glycine max* [L.] Merr.) flowering reversion system. *Planta* 223, 725–735. doi:10.1007/s00425-005-0130-y.

Wu, T. ting, Li, J. yu, Wu, C. xiang, Sun, S., Mao, T. ting, Jiang, B. jun, et al. (2015). Analysis of the independent- and interactive-photo-thermal effects on soybean flowering. *J. Integr. Agric.* 14, 622–632. doi:10.1016/S2095-3119(14)60856-X.

Yan, L., Hofmann, N., Li, S., Ferreira, M. E., Song, B., Jiang, G., et al. (2017). Identification of QTL with large effect on seed weight in a selective population of soybean with genome-wide association and fixation index analyses. *BMC Genomics* 18, 529. doi:10.1186/s12864-017-3922-0.

Yang, W., Wu, T., Zhang, X., Song, W., Xu, C., Sun, S., et al. (2019). Critical photoperiod measurement of soybean genotypes in different maturity groups. *Crop Sci.* 59, 2055–2061. doi:10.2135/cropsci2019.03.0170.

Yang, Y., Wang, L., Zhang, D., Cheng, H., Wang, Q., Yang, H., et al. (2020). GWAS identifies two novel loci for photosynthetic traits related to phosphorus efficiency in soybean. *Mol. Breed.* 40. doi:10.1007/s11032-020-01112-0.

Yang, Y., Zhu, X., Cui, R., Wang, R., Li, H., Wang, J., et al. (2021). Identification of soybean phosphorous efficiency QTLs and genes using chlorophyll fluorescence parameters through GWAS and RNA-seq. *Planta* 254, 1–13. doi:10.1007/s00425-021-03760-8.

Yanti, E., and Dafriani, P. (2021). The effect of soybean power on blood sugar and cholesterol levels in type 2 diabetes mellitus patients in the work area of the Puskesmas Andalas

Padang. *Adv. Heal. Sci. Res.* 39, 340–344.

Ye, H., Song, L., Schapaugh, W. T., Ali, M. L., Sinclair, T. R., Riar, M. K., et al. (2020). The importance of slow canopy wilting in drought tolerance in soybean. *J. Exp. Bot.* 71, 642–652. doi:10.1093/jxb/erz150.

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, Proteomics Bioinforma.* 19, 619–628. doi:10.1016/j.gpb.2020.10.007.

Yuan, S., Li, X., Li, R., Wang, L., Zhang, C., Chen, L., et al. (2018). Genome-wide identification and classification of soybean C2H2 Zinc Finger proteins and their expression analysis in legume-rhizobium symbiosis. *Front. Microbiol.* 9, 126. doi:10.3389/fmicb.2018.00126.

Zapata, F., Danso, F., Hardarson, G., and Fried, M. (1987). Nitrogen fixation and translocation in field-grown fababean 1. *Agron. J.* 79, 505–509. doi:10.2134/agronj1987.00021962007900030020x.

Zargar, S. M., Gupta, N., Nazir, M., Mahajan, R., Malik, F. A., Sofi, N. R., et al. (2017). Impact of drought on photosynthesis: molecular perspective. *Plant Gene* 11, 154–159. doi:10.1016/j.plgene.2017.04.003.

Zhang, J., Song, Q., Cregan, P. B., and Jiang, G. L. (2016). Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). *Theor. Appl. Genet.* 129, 117–130. doi:10.1007/s00122-015-2614-x.

Zhang, J., Song, Q., Cregan, P. B., Nelson, R. L., Wang, X., Wu, J., et al. (2015). Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics* 16, 1–11. doi:10.1186/s12864-015-1441-4.

Zhang, J., Subramanian, S., Stacey, G., and Yu, O. (2009). Flavones and flavonols play distinct critical roles during nodulation of *Medicago truncatula* by *Sinorhizobium meliloti*. *Plant J.* 57, 171–183. doi:10.1111/j.1365-313X.2008.03676.x.

Zhang, L. X., Kyei-Boahen, S., Zhang, J., Zhang, M. H., Freeland, T. B., Watson, C. E., et al.

(2007). Modifications of optimum adaptation zones for soybean maturity groups in the USA. *Crop Manag.* 6, 1–11. doi:10.1094/cm-2007-0927-01-rs.

Zhang, T., Wu, T., Wang, L., Jiang, B., Zhen, C., Yuan, S., et al. (2019). A combined linkage and GWAS analysis identifies QTLs linked to soybean seed protein and oil content. *Int. J. Mol. Sci.* 20, 5915. doi:10.3390/ijms20235915.

Zhang, W., Xu, W., Zhang, H., Liu, X., Cui, X., Li, S., et al. (2021). Comparative selective signature analysis and high-resolution GWAS reveal a new candidate gene controlling seed weight in soybean. *Theor. Appl. Genet.* 134, 1329–1341. doi:10.1007/s00122-021-03774-6.

Zhang, Z., Ober, U., Erbe, M., Zhang, H., Gao, N., He, J., et al. (2014). Improving the accuracy of whole genome prediction for complex traits using the results of genome wide association studies. *PLoS One* 9, e93017. doi:10.1371/journal.pone.0093017.

Zhao, J. Y., Lu, Z. W., Sun, Y., Fang, Z. W., Chen, J., Zhou, Y. Bin, et al. (2020). The ankyrin-repeat gene *GmANK114* confers drought and salt tolerance in *Arabidopsis* and soybean. *Front. Plant Sci.* 11, 584167. doi:10.3389/fpls.2020.584167.

Zhao, Z., and Assmann, S. M. (2011). The glycolytic enzyme, phosphoglycerate mutase, has critical roles in stomatal movement, vegetative growth, and pollen production in *Arabidopsis thaliana*. *J. Exp. Bot.* 62, 5179–5189. doi:10.1093/jxb/err223.

Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and prospects of association mapping in plants. *Plant Genome* 1, 5–20. doi:10.3835/plantgenome2008.02.0089.

Chapter 3 - The Genotypic Variability Among Short-Season Soybean Cultivars for Photosynthesis and Water-Use Efficiency Under Drought Stress

Abstract

Soybean (*Glycine max* L.) is an important and widely grown legume crop globally. It is predicted that drought is going to be a critical issue for crop production in the future. Therefore, the identification of drought-tolerant soybean varieties is critical for maintaining soybean production in the future. The objective of this project is to identify the allelic variation associated with diverse short-season soybean varieties for photosynthesis, transpiration, stomatal conductance, leaf chlorophyll, and water-use efficiency. A greenhouse pot experiment with 103 early-maturity Canadian soybean varieties was conducted to determine different plant physiological traits under drought stress. Seedlings were inoculated with *Bradyrhizobium japonicum* USDA 110, and the initial soil moisture content was maintained at 80% field capacity (FC). The drought stress was imposed after 3 weeks of plant growth, where half of the plants were maintained at 30% FC and the rest at 80% FC until maturity. Under drought stress, soybean plants showed lower stomatal conductance and transpiration, leading to more water-use efficiency compared to the well-watered plants at the flowering stage. The leaf chlorophyll content was higher in drought-stressed plants compared to the well-watered plants, possibly due to nitrogen accumulation in leaves under drought stress. Significant genotypic variability among soybean varieties was found for stomatal conductance, transpiration, leaf chlorophyll content, water-use efficiency, photosynthesis, and intrinsic water-use efficiency. A Genome-Wide Association Study (GWAS) using 2.16M SNPs (2,164,465 SNPs) was performed for different plant physiological parameters for 30% and 80% FC and their relative performance. In total, 13 quantitative trait locus (QTL) regions, including multiple candidate genes, were detected as significantly associated with different plant physiological traits, including photosynthesis, stomatal conductance, and water-use efficiency for 30% FC and relative performance. These genes can serve to accelerate future breeding efforts on developing drought-resistant soybean varieties.

Keywords: soybean, drought tolerance, water-use efficiency, candidate gene, quantitative trait locus, genome wide association study

3.1 Introduction

The cultivated soybean (*Glycine max* (L.) Merr.) was domesticated from its wild progenitor *Glycine soja* 6000-9000 years ago in China (Carter et al., 2004; Kim et al., 2012). It was introduced to North America in 1765 (Hymowitz and Harlan, 1983) and has become one of the dominant crops in North America. Soybean was introduced to Ontario, Canada in 1881 (Shurtleff and Aoyagi, 2010), where 2.15 million ha were cultivated with soybean, producing 6.27 MMT in 2021 (Soy Canada, 2022). Soybean is primarily used for producing vegetable oil, animal feed, and serves as a rich source of protein in the human diet. Additionally, it is used in industrial products such as cosmetics, hygiene products, paint removers, plastics and biofuels (Liu, 2008; Hartman et al., 2011).

Photoperiod is a critical factor in soybean flowering, seed set, and yield (Kantolic and Slafer, 2005; Wu et al., 2006, 2015). Soybean cultivars are categorized into several maturity groups as per their response to photoperiod. In North America, soybean cultivars are classified into 13 maturity groups (MG)(000 to X) (Zhang et al., 2007). Early maturity cultivars range from group 0 to III, and the late maturity cultivars start from group VI (Zhang et al., 2007; Yang et al., 2019). The soybean cultivars with 000, 00, and 0 groups are considered as earliest maturity and are mainly adapted to production areas in Canada, and maturity groups I and II are typically grown in the northern US and some parts of southern Ontario (Zhang et al., 2007). Therefore, in this study, we were mostly interested in early maturity soybean cultivars. An interesting feature of soybean is that it can fix atmospheric nitrogen with the symbiotic rhizobia bacteria that reside in root nodules. Rhizobia can convert atmospheric nitrogen to fixed ammonia, and in return, the host plant provides photosynthetic products for rhizobia metabolism (Thilakarathna and Raizada, 2017). In general, soybean can fix about 68% of the plant's nitrogen demand through symbiotic nitrogen fixation (SNF) (Herridge et al., 2008). According to a recent study, the annual input of soybean symbiotic nitrogen fixation is estimated to be 11.6 ± 2.9 Tg N on a global scale (Ma et al., 2022).

Soybean production is threatened by the anticipated climate change with frequent droughts over many parts of the world (Dai, 2013; Foyer et al., 2016). Drought stress is one of the most severe abiotic stresses in soybean, as drought causes a significant yield reduction in many growing regions (Oya et al., 2004). In Western Canada, many regions have been experiencing frequent and severe droughts over the last few decades, and it is predicted that drought will be a critical issue for the Canadian Prairies in the future (Bonsal et al., 2013, 2020). Drought stress during the vegetative phase can decrease plant height and reduce seed number (Desclaux et al., 2000). Also, drought reduces the amount of biomass allocated to reproductive organs, reducing seed weight (Du et al., 2020). Furthermore, drought stress can reduce SNF in soybean nodules due to carbon shortage, oxygen limitation in nodules, feedback inhibition of N fixation, and nitrogenase enzyme disruption (Serraj et al., 1999; Downie, 2005; Arrese-Igor et al., 2011; Collier and Tegeder, 2012; Kunert et al., 2016).

Plants use diverse drought tolerance mechanisms. The earliest plant response during drought stress is the stomatal closure that results in decreased mesophyll CO_2 diffusion, transpiration rate, and photosynthesis (Nouri et al., 2015; Arya et al., 2021). Under drought stress, abscisic acid (ABA) is synthesized in plant roots and transported to guard cells of the leaves to induce stomatal closure (Wilkinson and Davies, 2002). Stomatal closure ultimately reduces growth and development (Cornic and Briantais, 1991; Ohashi et al., 2006; Mutava et al., 2015). In order to maintain cell turgidity during dehydration, plants synthesize osmoprotectants such as sugars, sugar alcohols, organic acids, and free amino acids (Dumschott et al., 2019). Additionally, plants show various adaptations to survive drought stress; changes in leaf architecture to minimize water losses by reducing light absorbance (Ehleringer and Cooper, 1992; Nouri et al., 2015), a dense trichome layer that increases reflectance (Larcher, 2000), and shedding of older leaves (Tuteja et al., 2011). High water-use efficiency is an important trait in drought tolerance, where the increase in water-use efficiency is associated with a reduction in transpiration and photosynthesis (Kaler et al., 2017a).

Drought tolerance also depends on the genotype, where genotypic variability among different soybean cultivars for drought tolerance has been identified in previous studies (Oya et al., 2004; Hufstetler et al., 2007; Naderi et al., 2013; Jumrani and Bhatia, 2019; Du et al., 2020; Dayoub et al., 2021). Genome-wide association studies (GWAS) can be used as a powerful tool to identify genomic regions and genes controlling drought tolerance. GWAS involves scanning the entire genome of many individuals with high-density markers to find genetic loci associated with traits of interest (Tam et al., 2019). When these traits exhibit a complex genetic architecture, such loci are termed Quantitative Trait Loci (QTL). Furthermore, different plant traits have been considered in previous drought-related GWAS in soybean, including carbon isotope ratio (Dhanapal et al., 2015a; Chamarthi et al., 2021), oxygen isotope ratio (Chamarthi et al., 2021), yield stability index (Quero et al., 2021), canopy wilting (Hwang et al., 2015; Kaler et al., 2017b; Steketee et al., 2020;

Ye et al., 2020; Chamarthi et al., 2021), chlorophyll and chlorophyll fluorescence (Hao et al., 2012a; Dhanapal et al., 2016; Herritt et al., 2018; Yang et al., 2021), net photosynthetic rate, stomatal conductance, intercellular carbon dioxide concentration, and transpiration rate (Wang et al., 2020), all of which are related to drought tolerance.

The main objective of this study was to identify allelic variation associated with different plant physiological parameters under drought stress at the flowering stage and to identify the genomic regions controlling drought tolerance in short-season soybean varieties. Here, a core set of 103 Canadian short-season soybean lines were phenotyped for different plant physiological parameters, including leaf chlorophyll content, photosynthesis, stomatal conductance, transpiration, water-use efficiency, and intrinsic water-use efficiency at the flowering stage under drought and well-watered conditions.

3.2 Materials and methods

3.2.1 Plant materials and plant growth conditions

A total of 103 Canadian short-season soybean genotypes representing a large genetic diversity (Sonah et al., 2015; Torkamaneh et al., 2018) was used in this study (Appendix Table B.1). Seeds were surface-sterilized with 70% ethanol for 2 minutes, followed by washing with five changes of autoclaved double distilled water (Thilakarathna et al. 2017). Three seeds were planted in 6.52-L pots (H.J.S. Wholesale Ltd, Winnipeg, Manitoba) and supplied with sufficient moisture to germinate. The soil media used was a professional growing mix (Sun Gro Horticultural Canada Ltd, Seba Beach, Canada) and sand (Target Products Ltd, Morinville, Alberta) mixed in a 3:1 ratio (v:v), respectively. Plants were grown at 26 ± 2°C (day) and 20 ± 2°C (night) temperatures with supplemental lighting (range: 500–600 µmol m^{-2} s^{-1} at the top of the canopy, Fortimo LED Line, High Flux VO), maintaining a photoperiod of 16 h/8 h light/dark cycles. Extra plants were removed after one week of plant growth, leaving one plant per pot. One-week-old seedlings were inoculated with 2 ml of *Bradyrhizobium japonicum* USDA 110 inoculum, where rhizobial density was adjusted to OD_{600} = 0.1 (Thilakarathna et al., 2017). Rhizobia inoculation was repeated one week after the first inoculation to ensure proper nodulation. Plants were supplied weekly with 100 ml of quarter-strength nitrogen-free Hoagland's nutrient solution (HOP03-50LT, Caisson Labs, UT).

3.2.2 Determination of soil field capacity

First, the field capacity of the pots was measured using a separate study. Briefly, plastic 6.52-L pots were filled with a professional growing mix (Sun Gro Horticultural Canada Ltd, Seba Beach, Canada) and sand (Target Products Ltd, Morinville, Alberta) mixed in a 3:1 ratio on volume basis. until a final constant weight was reached (e.g., 4500 g). The bottom of each pot was lined with a coffee filter (12" Mother Parkers Coffee Filters) to prevent soil loss. The initial weight of the dry soil (D_w) was measured after drying the soil in an oven at 80 °C until a constant weight was reached (Liyanage et al. 2022). The pots were slowly watered until the soil was saturated and water drained out from the holes in the bottom. The tops of the pots were covered with aluminum foil and then kept for 24 hrs until no water drained from the bottom. The final saturated weights of the pots were recorded (S_w). Field capacity (FC) was calculated as FC = Sw–Dw. Based on the FC data, the final weight of the well-watered treatment (80% FC) and drought treatments (30% FC) were calculated (Liyanage et al., 2022).

3.2.3 Initiation of drought treatment

After three weeks of plant growth, the soil moisture content of pots was maintained at 80% FC (well-watered) and 30% FC (drought) until seed maturity. The drought treatment was induced by withholding water until pots reached 30% FC. Moisture adjustment in all the pots was carried out using a semi-automated irrigation system and the system recorded the weight of each pot before and after soil moisture adjustments. The treatments (genotypes x moisture) were allocated in a randomized complete block design with four replicates. The main factor was the irrigation regime (80% or 30% FC) and the subfactor was the genotypes.

3.2.4 Data collection

Different plant physiological parameters were collected from seven-week-old soybean plants (flowering stage). Leaf photosynthesis, stomatal conductance, and transpiration data were collected using a LICOR-6400 photosynthesis system (LI-COR Biosciences, Lincoln, USA). A photon flux density of 500 μmol m^{-2} s^{-1} and CO_2 concentration of 400 mol m^{-2} s^{-1} inside the chamber were maintained to measure different physiological traits. Data were collected on a single fully expanded young soybean leaf on each plant. Water-use efficiency (WUE) was calculated as the ratio of photosynthesis to transpiration (Patinni et al., 2020), wherein intrinsic water-use efficiency (iWUE) was calculated as the ratio of photosynthesis to stomatal conductance (Hatfield and Dold, 2019). Additionally, leaf chlorophyll content was measured on single young fully expanded soybean leaves using a SPAD meter (KONICA-MINOLTA 502 Plus).

3.2.5 Statistical analysis

The effects of moisture treatments and soybean genotypes were analyzed using analysis of variance (ANOVA) set at $\alpha < 0.05$. The data were analyzed as a two-factor factorial design, where the main factor was soil moisture treatment, and the sub-factor was the soybean genotypes. The statistical analyses, including ANOVA, frequency distributions, and Pearson correlations of traits, were performed using R 3.5. (R Core Team, 2021)

3.2.6 Genotyping data

By combining approach of genotyping by sequencing (GBS) and whole-genome sequencing (WGS) approach a catalog of 2.18M SNPs was created for 137 accessions by Seck et al., 2020 and from that larger data set 103 accessions were extracted for this study with 2.16M SNPs. (2,164,465 SNPs) (Seck et al., 2020)

3.2.7 Population structure

A linkage disequilibrium-based pruning (r^2>0.5) was conducted on the previously generated 2.18M SNPs using PLINK (Purcell et al., 2007) and created a uniform panel of 14K markers. The population structure was characterized by the tested populations (K, 1-13) using fastSTRUCTURE (Raj et al., 2014). The ‘‘Choseek.py’’ python script was applied to assess the most suitable K value. This K value was based on the rate of change in LnP between the successive K values. Furthermore, a phylogenetic tree was built using the Neighbor-Joining method in TASSEL 5.0 (Appendix Figure B.1). The distribution of SNP markers across the soybean genome is shown in Appendix Figure B.2. The Appendix Figures B.3, B.4 and B.5 represented the quantile-quantile plots of p values for the association between SNP markers and different physiological traits at 30% FC, 80% FC and relative performance respectively. A scree plot (Appendix Figure B.6) was used to assess the most informative Principal Components (PCs) (Appendix Figures B.7 and B.8). A kinship matrix was determined using the efficient mixed-model association (EMMA).

3.2.8 Genome-wide association analysis on physiological traits

A genome-wide association study (GWAS) was performed with 2.16M SNPs (2,164,465 SNPs) using the fixed and random model circulating probability unification (FarmCPU) model (Liu et al., 2016) implemented in rMVP package (Yin et al., 2021) on Microsoft Open R. The Population Structure (Q) and Kinship (K) matrices were used as covariates to reduce false-positive signals. A genome-wide significance threshold of less than 0.05 was used to find the significant associations using the false discovery rate (FDR) test (Benjamini and Hochberg, 1995).

3.2.9 Candidate gene identification

Linkage Disequilibrium (D') was measured between the peak SNP and all the markers within 1-Mb on each side using PLINK (Purcell et al., 2007). The genes within that area were extracted from SoyBase (Grant et al., 2009). The "Gene expression and protein tools" (ePlant[2]) for soybeans were used to visualize the expression in tissues relevant to physiological traits (e.g. leaves, roots, nodules) in order to provide more information about probable candidate genes [based on the transcriptome data of (Waese et al., 2017)].

3.3 Results

3.3.1 Phenotypic variation for drought-related physiological parameters in soybean

A significant phenotypic variation among 103 soybean genotypes was found for leaf chlorophyll content, leaf photosynthesis, stomatal conductance, WUE, iWUE, and transpiration ($p<0.0001$). The soil moisture content had a significant effect on leaf chlorophyll content, stomatal conductance, WUE, and iWUE (Table 1). However, leaf photosynthesis and transpiration were not significantly different between the drought and well-watered treatments at $p<0.05$, yet they were very close to being significant. The leaf chlorophyll content, WUE, and iWUE were higher under 30% FC drought treatment compared to the 80% FC well-watered treatment. Conversely, stomatal conductance was lower under the drought treatment compared to the well-watered treatment. The summary statistics for different plant physiological parameters measured under 30% and 80% FC are shown in Table 3.1.

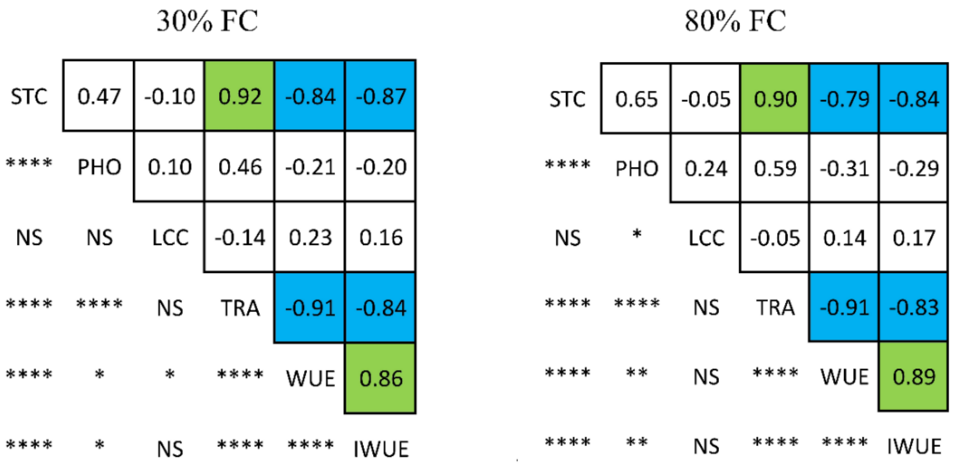
Table 3.1 Summary statistics of the six plant physiological traits of 103 soybean genotypes under two different soil moisture conditions.

Traits	80% FC	30% FC	p value
Leaf chlorophyll content	34.9 ± 0.4	39.5 ± 0.4	p<0.0001
Photosynthesis	15.5 ± 0.2	15.8 ± 0.1	p=0.0571
Stomatal conductance	0.30 ± 0.01	0.28 ± 0.01	p=0.0442
WUE	3.8 ± 0.1	4.2 ± 0.1	p<0.0001
iWUE	55.1 ± 1.1	59.1 ± 1.2	p=0.0003
Transpiration	4.6 ± 0.2	4.4 ± 0.1	p=0.0772

*Mean value of the trait ± SE; FC, field capacity; WUE, water-use efficiency; iWUE, intrinsic water-use efficie*ncy. *water-use efficiency = photosynthesis / transpiration. Intrinsic water-use efficiency = photosynthesis / stomatal conductance.*

3.3.2 Correlations among the measured physiological traits

Many significant correlations were observed under the 80% FC well-watered and 30% FC drought treatments (Figure 3.1). WUE and iWUE were negatively correlated with stomatal conductance and transpiration under both treatments. However, WUE and iWUE had weak negative and positive correlations with photosynthesis and relative leaf chlorophyll content, respectively. There was a strong positive correlation between transpiration and stomatal conductance under both moisture treatments. Furthermore, photosynthesis was positively correlated with transpiration and stomatal conductance under drought and well-watered conditions.



30% FC

STC	0.47	-0.10	0.92	-0.84	-0.87
****	PHO	0.10	0.46	-0.21	-0.20
NS	NS	LCC	-0.14	0.23	0.16
****	****	NS	TRA	-0.91	-0.84
****	*	*	****	WUE	0.86
****	*	NS	****	****	IWUE

80% FC

STC	0.65	-0.05	0.90	-0.79	-0.84
****	PHO	0.24	0.59	-0.31	-0.29
NS	*	LCC	-0.05	0.14	0.17
****	****	NS	TRA	-0.91	-0.83
****	**	NS	****	WUE	0.89
****	**	NS	****	****	IWUE

Figure 3.1 Correlations among physiological traits for the 103 soybean lines. Numbers above the diagonal correspond to Pearson's correlation coefficients (R). Green boxes highlight the positive values exceeding 0.65 and blue boxes highlight the negative values exceeding 0.65. Below the diagonal, we show the degree of significance of the corresponding correlations between traits (****P < 0.0001, ***P < 0.001, **P < 0.01, *P < 0.05, and NS: not significant). STC, stomatal conductance; PHO, photosynthesis; LCC, leaf chlorophyll content; TRA, transpiration; WUE,

water-use efficiency; IWUE, intrinsic water-use efficiency; WUE = photosynthesis/transpiration; IWUE = photosynthesis/stomatal conductance.

3.3.3 Genotypic data and population structure

In this study, both GBS and WGS genotyping approaches were used to cover the whole soybean genome as described in Seck et al., 2020 (Seck et al., 2020). Population structure was characterized by using a subset of 14K pruned SNPs. The optimum number of subpopulations (K) lies between 6 and 9. This was also conformed with PCA and phylogeny analysis (Appendix Figures B.7 and B.8).

3.3.4 Genome-wide association of physiological parameters

GWAS analysis was performed for six plant physiological parameters under 30% FC, 80% FC (Appendix Figure B.9) and their relative performance (30%/80% FC) using 2.16M SNPs (2,164,465 SNPs) and the FarmCPU statistical model. In total, 13 significantly associated SNPs were detected for 30% FC drought-induced treatment (Figure 3.2), and their relative performance (Figure 3.3). Three associated with 30% FC (photosynthesis qPH1, stomatal conductance qSC1 and water-use efficiency qWUE1, and 10 associated with the relative performance (photosynthesis qPH2-qPH3, leaf chlorophyll content qLCC1, stomatal conductance qSC2, intrinsic water-use efficiency qIWUE2-qIWUE6 and transpiration qTR1) (Table 3.2). No SNPs were detected as significantly associated with leaf chlorophyll content, intrinsic water-use efficiency and transpiration under 30% FC, and water-use efficiency under the relative performance.

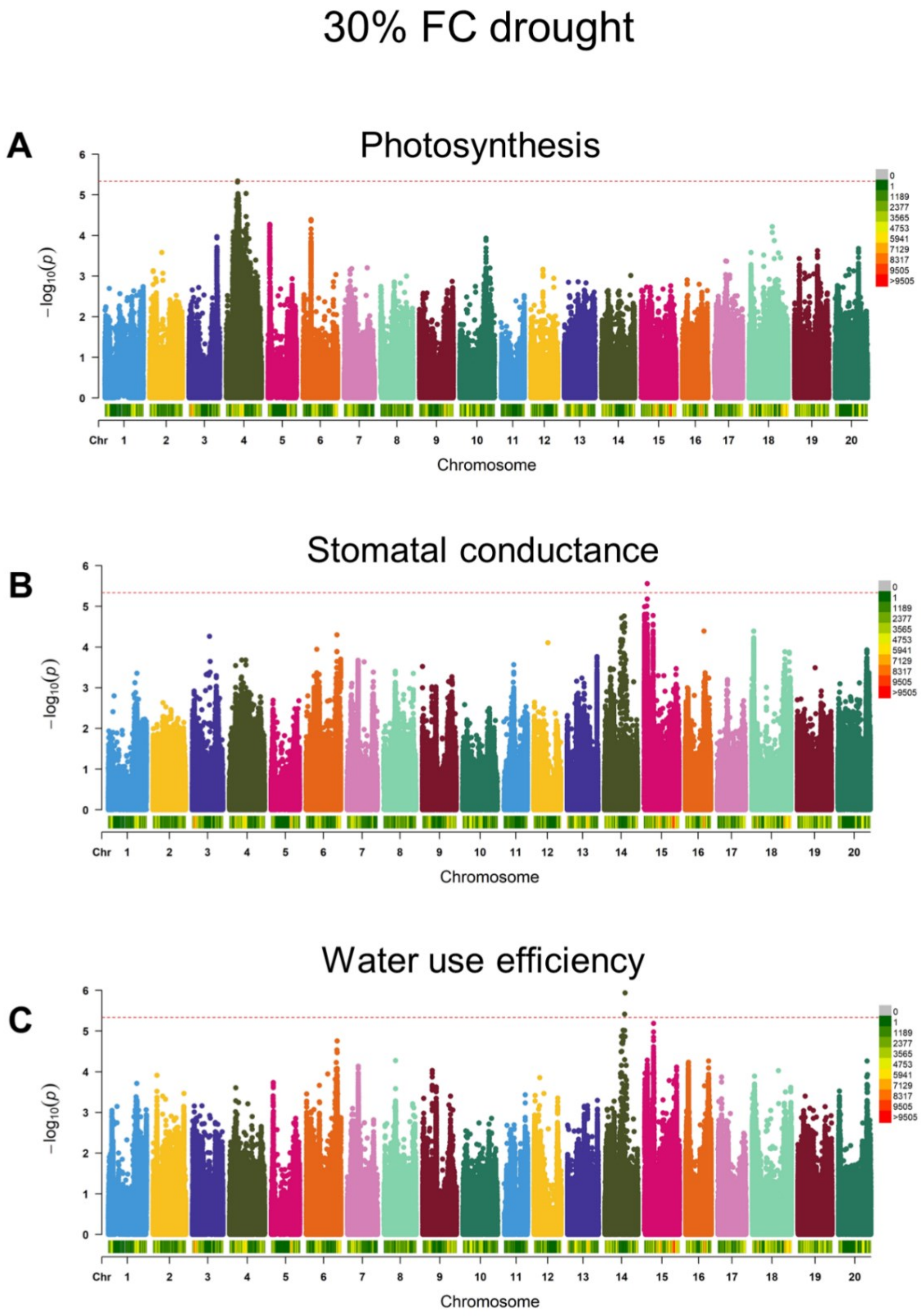


Figure 3.2 Manhattan plots of the genome-wide association results for different plant physiological traits measured at the flowering stage under drought (30% FC) conditions. (A) photosynthesis (B) stomatal conductance (C) water-use efficiency. Negative $\log_{10}$ (P-values) (y-axis) describing the strength of the association between each marker and trait are plotted against the physical position of each marker (x-axis). The pink dashed line indicates the significance threshold (FDR = 5%) and beyond that are considered as significant associations.

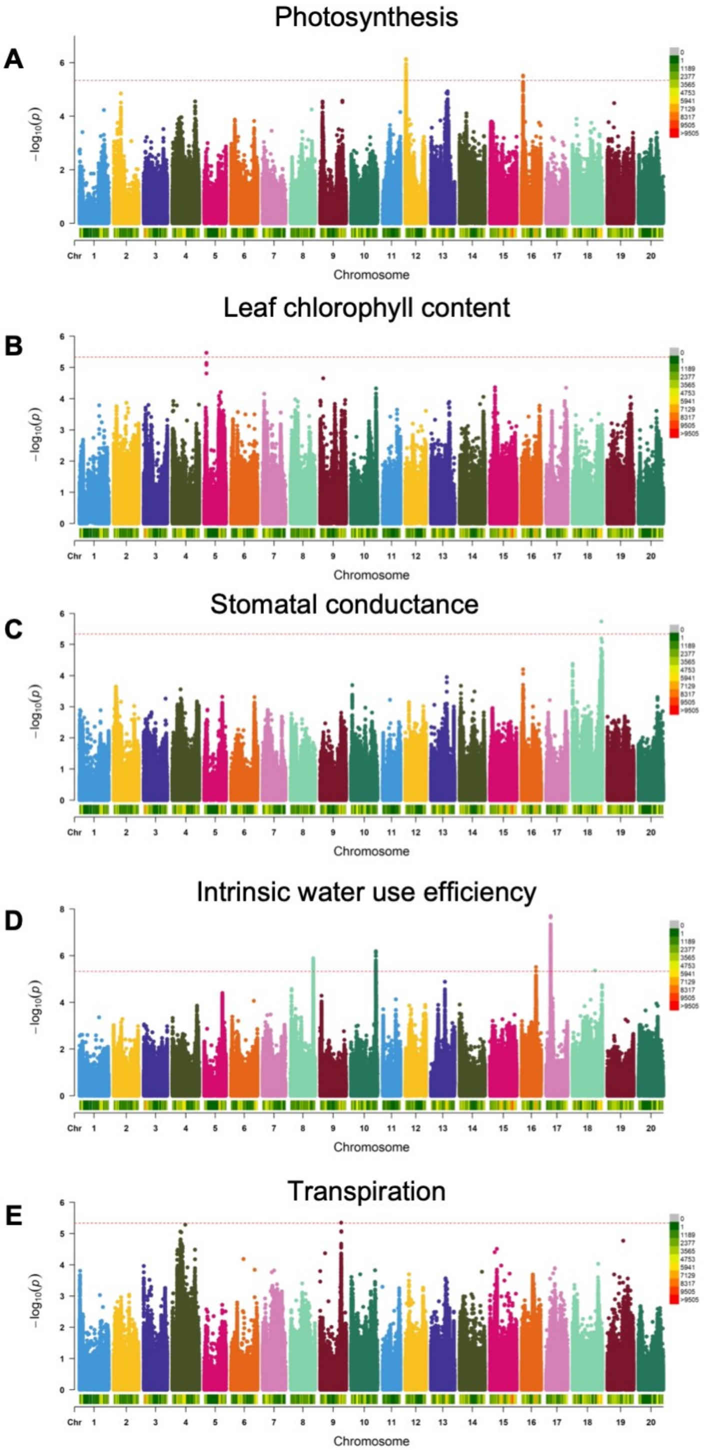


Figure 3.3 Manhattan plots of the genome-wide association results for relative performance. (A) photosynthesis, (B) leaf chlorophyll content, (C) stomatal conductance, (D) intrinsic water-use efficiency, (E) transpiration. Negative log10 (P-values) (y-axis) describing the strength of the association between each marker and trait are plotted against the physical position of each marker (x-axis). The pink dashed line indicates the significance threshold (FDR = 5%) and beyond that are considered as significant associations.

Table 3.2 List of quantitative trait loci (QTL) associated with 30% FC, and relative performance.

Treatment	Trait	Chr[a]	MSS[b] position	QTL ID	Major allele frequency / Minor allele frequency/ Average of phenotypic values	Effect
30% FC	Photosynthesis	4	17,183,396	qPH1	C:T/0.80:0.20/15.54:16.25	1.16
	Stomatal conductance	15	4,877,172	qSC1	C:T/0.77:0.22/0.30:0.21	-0.04
	[c]Water-use efficiency	14	29,320,423	qWUE1	A:C/ 0.68:0.31/3.91:4.66	0.86
		14	29,913,062		G:A/ 0.70:0.30/3.89:4.73	0.92
Relative performance	Photosynthesis	12	1,375,409	qPH2	T:G/ 0.79:0.21/1.06:0.93	-0.07
		12	1,546,067		C:T/ 0.74: 0.26/1.06:0.95	-0.06
		16	3,072,259	qPH3	T:A/0.84: 0.15/1.05:0.91	-0.08
	Leaf chlorophyll content	5	4,274,900	qLCC1	A:G/ 0.75:0.25/1.11:1.27	0.08
	Stomatal conductance	18	55,782,253	qSC2	A:G/ 0.62:0.38/0.92:1.04	0.08
	[d]Intrinsic water-use efficiency	8	44,111,118	qIWUE2	T:C/0.62: 0.38/1.02:1.17	0.07
		10	48,220,950	qIWUE3	A:G/ 0.84:0.16/1.10:0.96	-0.1
		10	48,320,395		T:C/ 0.79:0.21/1.11:0.97	-0.08
		16	28,085,463	qIWUE4	C:A/0.63: 0.37/1.03:1.17	0.08
		17	8,878,535	qIWUE5	C:T 0.87:0.13/1.05:1.25	0.1
		17	8,978,069		C:T/0.86:0.13/1.05:1.28	0.12
		17	9,019,014		G:T/0.85: 0.15/1.05:1.25	0.1
		18	43,418,751	qIWUE6	G:T/0.92: 0.08/1.06:1.30	0.16
	Transpiration	9	40,795,425	qTR1	G:A/ 0.93: 0.07/0.94:1.16	0.14

[a] *Chromosome number*

[b] *Most significant SNP*

[c] *water-use efficiency = photosynthesis/transpiration;* [d] *intrinsic-water use efficiency = photosynthesis/stomatal conductance*

3.3.5 Plant-physiology-related candidate genes

In order to find candidate genes, regions of interest for all 13 QTLs were defined as spanning from the leftmost to the rightmost marker in high LD (D’ ≥0.85) with the peak SNP. All the genes that are residing in whole or in part within the 13 QTLs of interest were extracted from SoyBase. Appendix Tables B.2, and B.3 provide the complete information of these genes, including their annotations. Based on their annotation, we identified some strong candidate genes separately for 30% FC, and their relative performance as mentioned below.

3.3.5.1 Candidate genes under drought stress

For stomatal conductance (qSC1), we found the gene *Glyma.15g060800* annotated for phosphoglycerate mutase family protein. The genes *Glyma.15g061300* and *Glyma.15g06581* encodes for soluble inorganic pyrophosphatase 1% 2C Chloroplastic-like and for proline/glycine/tyrosine-rich protein respectively. The gene *Glyma.15g06581* act as an indicator during drought stress. In addition, the gene *Glyma.15g064300* – plastid developmental protein DAG putative was also identified as an important candidate gene for light reaction in photosynthesis. The gene *Glyma.15g061100* encodes for methionine sulfoxide reductase B2 was found in qSC1. Appendix Table B.2 provides detailed information about the candidate genes associated under drought stress.

3.3.5.2 Candidate genes under relative performance

The candidate gene *Glyma.10g256000* annotated for photosystem I reaction center subunit N (PSAN or PSI-N) was identified for intrinsic water-use efficiency (qIWUE3). The gene *Glyma.10g257500* associated with intrinsic water-use efficiency (qIWUE3) and annotated for root hair defective 6-like 2 in *Arabidopsis thaliana* is an important transcription factor. Moreover, the candidate gene *Glyma.16g130200* annotated as vascular related NAC-domain protein 1 in qIWUE4 locus and *Glyma.17g113000* for early nodulin-like protein 1 in qIWUE5 were identified. Importantly, the published nodule gene in soybean, *Glyma.17g117100* encodes for early nodulin93 was recognized in qWUE5. The gene *Glyma.17g117200*, annotated as Early nodulin 93 ENOD93 protein in the qIWUE5 region, and *Glyma.16g030400,* a photosynthesis (qPH3) gene designated for chloroplast sulphur E, were also discovered. Appendix Table B.3 provides detailed information about the candidate genes associated with relative performance.

3.4 Discussion

3.4.1 Significant phenotypic variation of physiological traits in soybean

As mentioned earlier, we found significant genotypic variability for different plant physiological parameters at the flowering stage. Similarly, previous studies have also found phenotypic variability for different plant physiological parameters, including leaf epidermal conductance (Hufstetler et al., 2007; Wang et al., 2020), water-use efficiency (Hufstetler et al., 2007; Dhanapal et al., 2015a; Kaler et al., 2017a), chlorophyll content (Hao et al., 2012a; Dhanapal et al., 2016), photosynthesis (Wang et al., 2020), and transpiration (Kaler et al., 2017a; Wang et al., 2020). The existence of phenotypic variation within a germplasm pool for different plant physiological traits, including water-use efficiency, is important for plant breeders to make progress through selection. Soybean breeders need phenotypic and genotypic tools to improve drought stress tolerance.

Among the different soybean varieties evaluated, OT11-01, OAC Clinton, Roland, OAC Lauralain, and OAC Bayfield were shown to be the most water-efficient varieties during the drought, while Maple Ridge, AC Orford, AC Brant, Dundas, and 90B11 varieties performed poorly. OAC Petrel, PS36, S03-W4, Roland, and OAC Wallace performed well in terms of intrinsic water-use efficiency, whereas 90B11, Dundas, AC Brant, AC Orford, and Maple Ridge varieties had the lowest intrinsic water-use efficiency. In terms of photosynthesis, Dares, QS5030.46Bp, OAC Champion, OAC Ayton, OAC 07-26C had the highest photosynthesis, while Misty, McCall, Casino, AC Glengarry, and OAC Carman had the lowest. Relative leaf chlorophyll content was measured as SPAD units, where Lotus, Colby, SECAN 07-27C, Altesse, AC Proteina had the highest leaf chlorophyll content. In contrast, Naya, 90B11, OAC Avatar, AC Bravor and OAC 07-04C had the lowest. The highest stomatal conductance was observed in AC Brant, Maple Isle, Misty, AC Orford, and 90A01, whereas PS36, OAC Perth, Roland, OAC Avatar, and OAC Wallace had the lowest. For transpiration, the highest was observed in 90B11, AC Brant, Maple Isle, Maple Ridge, and Lotus, whereas the lowest was recorded in Roland, OAC 09-35C, OAC Clinton, OAC Ayton, and OAC Avatar.

3.4.2 High and significant correlations among physiological traits

We found that many physiological traits were highly and significantly correlated. These results are also in agreement with other reports in the literature. For example, it has been observed that transpiration exhibited a tight correlation with stomatal conductance in soybean (Inamullah and Isoda, 2005). This indicates that soybean reduces transpiration by reducing stomatal conductance. Moreover, we found a strong negative correlation between water-use efficiency *vs.* stomatal conductance and transpiration. During short-term drought stress conditions, plants tend to increase their water-use efficiency by reducing stomatal aperture and transpiration, whereas in the long run, plants reduce the maximum stomatal conductance by changing stomatal density and size (Li et al., 2017). Similarly, a previous study conducted with 23 soybean genotypes reported that WUE is negatively correlated with leaf epidermal conductance (Hufstetler et al., 2007).

3.4.3 Genome-Wide Association using whole genome data revealed 13 QTLs controlling physiological traits

In this study, a GWAS was performed using an exhaustive genome-wide set of SNPs (2,164,465 SNPs). Here, we uncovered a total of 13 genomic regions (QTL) contributing to drought stress, photosynthesis, stomatal conductance, water-use efficiency, intrinsic water-use efficiency, transpiration, and leaf chlorophyll content. Similarly, some previous GWAS in soybean reveal loci for photosynthesis related traits (Wang et al., 2020), soybean chlorophyll traits (Hao et al., 2012a; Dhanapal et al., 2016) carbon isotope ratio (Dhanapal et al., 2015a), and photosynthesis traits related to phosphorous efficiency (Li et al., 2016b; Lü et al., 2018). We found that previous GWAS studies on soybean physiological parameters revealed the same QTL regions as we discovered in our study; photosynthesis qPH2 (Lü et al., 2018; Yang et al., 2020) and transpiration qTR1 (Lü et al., 2018). We observed that most of the QTL regions detected in our work are novel. Moreover, most of the GWAS research studies in soybean were based on photosynthesis and leaf chlorophyll content, whereas a limited number of GWAS have been conducted for the other mentioned physiological traits. Especially, there has been no GWAS conducted for intrinsic water-use efficiency in soybean to date.

3.4.4 Candidate genes for physiological traits associated QTL

In this work, we considered genes to be candidate causal genes if three conditions were met : (1) they were residing in the same LD block containing the peak SNP associated with the physiological trait, (2) has a possible role in physiological trait and / or drought stress, (3) their highest expression, based on publicly available data on ePlant, is in the responsible organ such as for photosynthesis, stomatal conductance, leaf chlorophyll content, transpiration (e.g. leaves), and water-use efficiency and intrinsic water-use efficiency (e.g. roots and nodules).

3.4.5.1 Candidate genes under drought stress

The gene *Glyma.15g060800* associated with stomatal conductance (qSC1) is annotated for the Phosphoglycerate mutase family protein (Appendix Table B.2). It has been found that Phosphoglycerate mutase is an important enzyme in the glycolysis pathway that involves in the stomatal movement of *Arabidopsis thaliana* (Zhao and Assmann, 2011). The gene *Glyma.15g061300* annotated for soluble inorganic pyrophosphatase 1% 2C Chloroplastic-like is associated with stomatal conductance and it is previously functionally validated (Cheung et al., 2009) (Appendix Table B.2). The gene *Glyma15g064300* – plastid developmental protein DAG putative highly expressed in leaves is also an important candidate gene as it mediates electron transfer from photosystem II (PSII) to photosystem I (PSI) in light reaction of photosynthesis (Höhner et al., 2020) (Appendix Table B.2). The gene *Glyma.15g061100* annotated for methionine sulfoxide reductase B2 in leaves is known to provide tolerance to abiotic environmental stresses in rice (Siddiqui et al., 2014, 2015b) (Appendix Table B.2). Accumulation of proline in leaves, roots, root tips, and nodules is an indicator to monitor drought stress in soybean. Proline accumulation increase with the drought stress during the flowering stage (Silvente et al., 2012), and it exhibits the plants' ability to tolerate drought stress (Mwenye et al., 2016). Also, the accumulation of glycine protects the plant from osmotic adjustment and detoxification of reactive oxygen species (Ashraf and Foolad, 2007; Basu et al., 2016). The published gene *Glyma.15g06581* associated with stomatal conductance that is annotated for proline/glycine/tyrosine-rich protein is an important candidate to use as maker to identify drought tolerance and susceptible cultivars (Appendix Table B.2).

3.4.5.2 Candidate genes under relative performance

The gene *Glyma.10g 256000* annotated for photosystem I reaction center subunit N (PSAN or PSI-N) is an important candidate gene for qIWUE3 (Appendix Table B.3). The gene *Glyma.10g257500* associated with qIWUE3 and annotated for ROOT HAIR DEFECTIVE 6-LIKE 2, RSL2 in *Arabidopsis thaliana* is a transcription factor that promotes root hair growth (Shibata et al., 2018) (Appendix Table B.3). RSL2 was expressed concurrently with RSL4, and its expression was controlled by RHD6 and RSL1. This gene involves in root hair initiation and response to auxin stimulus as well. The candidate gene *Glyma.16g130200* annotated as vascular related NAC-domain protein 1 is highly expressed in roots (Appendix Table B.3). This gene encodes a NAC-domain transcription factor that is expressed in developing vessels and protoxylem in *Arabidopsis thaliana*. Furthermore, the candidate gene *Glyma.17g113000*, which encodes an early nodulin-like protein, is strongly expressed in roots and can be found in qIWUE5 (Appendix Table B.3). In qIWUE8, the genes *Glyma.17g117200* and *Glyma.17g117100* have a role in the Early nodulin 93 ENOD93 protein (Appendix Table B.3), where it is a crucial nitrogen fixation gene in soybean root nodules (Lauridsen et al., 1993). The chloroplast sulfur E gene, *Glyma.16g030400*, is related to photosynthesis (Appendix Table B.3). The chloroplast contains a large number of iron-sulfur proteins for photosynthesis and nitrogen and sulfur assimilation (Abdel-Ghany et al., 2005).

3.5 Conclusions

Overall, this study has revealed significant genotypic variability among soybean varieties for stomatal conductance, transpiration, leaf chlorophyll content, water-use efficiency, photosynthesis, and intrinsic water-use efficiency. This study identifies 13 QTLs for drought, and their relative performance. In addition, potential candidate genes were discovered to support the findings. The current study will contribute toward understanding the genetics underlying different plant physiological parameters, water-use efficiency, and drought tolerance in soybean.

References

Arrese-Igor, C., Esther González, B. M., Daniel Marino, B., Rubén Ladrera, B., Estíbaliz Larrainzar, B., and Erena Gil-Quintana, B. (2011). Physiological responses of legume nodules to drought. Plant Stress 5, 24–31.

Arya, H., Singh, M. B., and Bhalla, P. L. (2021). Towards developing drought-smart soybeans. Front. Plant Sci. 12, 750664. doi:10.3389/fpls.2021.750664.

Ashraf, M., and Foolad, M. R. (2007). Roles of glycine betaine and proline in improving plant abiotic stress resistance. Environ. Exp. Bot. 59, 206–216. doi:10.1016/j.envexpbot.2005.12.006.

Basu, S., Ramegowda, V., Kumar, A., and Pereira, A. (2016). Plant adaptation to drought stress. F1000Research 5, 1–10. doi:10.12688/F1000RESEARCH.7678.1.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. 57, 289–300.

Bonsal, B., Liu, Z., Wheaton, E., and Stewart, R. (2020). Historical and projected changes to the stages and other characteristics of severe Canadian prairie droughts. Water (Switzerland) 12. doi:10.3390/w12123370.

Bonsal, B. R., Aider, R., Gachon, P., and Lapp, S. (2013). An assessment of Canadian prairie drought: past, present, and future. Clim. Dyn. 41, 501–516. doi:10.1007/s00382-012-1422-0.

Carter, T. E., Hymowitz, T., and Nelson, R. L. (2004). “Biogeography, local adaptation, vavilov, and genetic diversity in soybean,” in Biological Resources and Migration (Springer, Berlin, Heidelberg.), 47–59. doi:10.1007/978-3-662-06083-4.

Chamarthi, S. K., Kaler, A. S., Abdel-Haleem, H., Fritschi, F. B., Gillman, J. D., Ray, J. D., et al. (2021). Identification and confirmation of loci associated with canopy wilting in soybean using genome-wide association mapping. Front. Plant Sci. 12, 698116. doi:10.3389/fpls.2021.698116.

Collier, R., and Tegeder, M. (2012). Soybean ureide transporters play a critical role in nodule development, function and nitrogen export. Plant J. 72, 355–367. doi:10.1111/j.1365-313X.2012.05086.x.

Cornic, G., and Briantais, J. M. (1991). Partitioning of photosynthetic electron flow between CO_2 and O_2 reduction in a C_3 leaf (*Phaseolus vulgaris* L.) at different CO_2 concentrations and during drought stress. Planta 183, 178–184. doi:10.1007/BF00197786.

Dai, A. (2013). Increasing drought under global warming in observations and models. Nat. Clim. Chang. 3, 52–58. doi:10.1038/nclimate1633.

Dayoub, E., Lamichhane, J. R., Schoving, C., Debaeke, P., and Maury, P. (2021). Early-stage phenotyping of root traits provides insights into the drought tolerance level of soybean cultivars. Agronomy 11. doi:10.3390/agronomy11010188.

Desclaux, D., Huynh, T. T., and Roumet, P. (2000). Identification of soybean plant characteristics that indicate the timing of drought stress. Crop Sci. 40, 716–722. doi:10.2135/cropsci2000.403716x.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015). Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. Theor. Appl. Genet. 128, 73–91. doi:10.1007/s00122-014-2413-9.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2016). Genome-wide association mapping of soybean chlorophyll traits based on canopy spectral reflectance and leaf extracts. BMC Plant Biol. 16, 1–15. doi:10.1186/s12870-016-0861-x.

Downie, J. A. (2005). Legume haemoglobins: symbiotic nitrogen fixation needs bloody nodules. Curr. Biol. 15, 196–198. doi:10.1016/j.cub.2005.03.006.

Du, Y., Zhao, Q., Chen, L., Yao, X., and Xie, F. (2020). Effect of drought stress at reproductive stages on growth and nitrogen metabolism in soybean. Agronomy 10, 302. doi:10.3390/agronomy10020302.

Dumschott, K., Dechorgnat, J., and Merchant, A. (2019). Water deficit elicits a transcriptional response of genes governing d-pinitol biosynthesis in soybean (*Glycine max*). Int. J. Mol. Sci. 20, 1–11. doi:10.3390/ijms20102411.

Ehleringer, J. R., and Cooper, T. A. (1992). On the role of orientation in reducing photoinhibitory damage in photosynthetic-twig desert shrubs. Plant. Cell Environ. 15, 301–306. doi:10.1111/j.1365-3040.1992.tb00977.x.

Foyer, C. H., Lam, H. M., Nguyen, H. T., Siddique, K. H. M., Varshney, R. K., Colmer, T. D., et al. (2016). Neglecting legumes has compromised human health and sustainable food production. Nat. Plants 2, 1–10. doi:10.1038/NPLANTS.2016.112.

Grant, D., Nelson, R. T., Cannon, S. B., and Shoemaker, R. C. (2009). SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res. 38, 843–846. doi:10.1093/nar/gkp798.

Hao, D., Chao, M., Yin, Z., and Yu, D. (2012). Genome-wide association analysis detecting significant single nucleotide polymorphisms for chlorophyll and chlorophyll fluorescence parameters in soybean (*Glycine max*) landraces. Euphytica 186, 919–931. doi:10.1007/s10681-012-0697-x.

Hartman, G. L., West, E. D., and Herman, T. K. (2011). Crops that feed the World 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. Food Secur. 3, 5–17. doi:10.1007/s12571-010-0108-x.

Hatfield, J. L., and Dold, C. (2019). Water-use efficiency: Advances and challenges in a changing climate. Front. Plant Sci. 10, 103. doi:10.3389/fpls.2019.00103.

Herridge, D. F., Peoples, M. B., and Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil 311, 1–18. doi:10.1007/s11104-008-9668-3.

Herritt, M., Dhanapal, A. P., Purcell, L. C., and Fritschi, F. B. (2018). Identification of genomic loci associated with 21chlorophyll fluorescence phenotypes by genome-wide association analysis in soybean. BMC Plant Biol. 18, 1–19. doi:10.1186/s12870-018-1517-9.

Höhner, R., Pribil, M., Herbstová, M., Lopez, L. S., Kunz, H. H., Li, M., et al. (2020). Plastocyanin is the long-range electron carrier between photosystem II and photosystem i in plants. Proc. Natl. Acad. Sci. U. S. A. 117, 15354–15362. doi:10.1073/pnas.2005832117.

Hufstetler, E. V., Boerma, H. R., Carter, T. E., and Earl, H. J. (2007). Genotypic variation for three physiological traits affecting drought tolerance in soybean. Crop Sci. 47, 25–35. doi:10.2135/cropsci2006.04.0243.

Hwang, S., King, C. A., Ray, J. D., Cregan, P. B., Chen, P., Carter, T. E., et al. (2015). Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. Theor. Appl. Genet. 128, 2047–2065. doi:10.1007/s00122-015-2566-1.

Hymowitz, T., and Harlan, J. R. (1983). Introduction of soybean to North America by Samuel Bowen in 1765. Econ. Bot. 37, 371–379. doi:10.1007/BF02904196.

Inamullah, and Isoda, A. (2005). Adaptive responses of soybean and cotton to water stress: I. Transpiration changes in relation to stomatal area and stomatal conductance. Plant Prod. Sci. 8, 16–26. doi:10.1626/pps.8.16.

Jumrani, K., and Bhatia, V. S. (2019). Identification of drought tolerant genotypes using physiological traits in soybean. Physiol. Mol. Biol. Plants 25, 697–711. doi:10.1007/s12298-019-00665-5.

Kaler, A. S., Dhanapal, A. P., Ray, J. D., King, C. A., Fritschi, F. B., and Purcell, L. C. (2017a). Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. Crop Sci. 57, 3085–3100. doi:10.2135/cropsci2017.03.0160.

Kaler, A. S., Ray, J. D., Schapaugh, W. T., King, C. A., and Purcell, L. C. (2017b). Genome-wide association mapping of canopy wilting in diverse soybean genotypes. Theor. Appl. Genet. 130, 2203–2217. doi:10.1007/s00122-017-2951-z.

Kantolic, A. G., and Slafer, G. A. (2005). Reproductive development and yield components in indeterminate soybean as affected by post-flowering photoperiod. F. Crop. Res. 93, 212–222. doi:10.1016/j.fcr.2004.10.001.

Kim, M. Y., Van, K., Kang, Y. J., Kim, K. H., and Lee, S. H. (2012). Tracing soybean domestication history: From nucleotide to genome. Breed. Sci. 61, 445–452. doi:10.1270/jsbbs.61.445.

Kunert, K. J., Vorster, B. J., Fenta, B. A., Kibido, T., Giuseppe, D., and Foyer, C. H. (2016). Drought stress responses in soybean roots and nodules. Front. Plant Sci. 7, 1015. doi:10.3389/fpls.2016.01015.

Larcher, W. (2000). Temperature stress and survival ability of mediterranean sclerophyllous plants. Plant Biosyst. 134, 279–295. doi:10.1080/11263500012331350455.

Li, H., Yang, Y., Zhang, H., Chu, S., Zhang, X., Yin, D., et al. (2016). A genetic relationship between phosphorus efficiency and photosynthetic traits in soybean as revealed by QTL analysis using a high-density genetic map. Front. Plant Sci. 7, 1–16. doi:10.3389/fpls.2016.00924.

Li, Y., Li, H., Li, Y., and Zhang, S. (2017). Improving water-use efficiency by decreasing stomatal conductance and transpiration rate to maintain higher ear photosynthetic rate in drought-resistant wheat. Crop J. 5, 231–239. doi:10.1016/j.cj.2017.01.001.

Liu, K. S. (2008). “Food use of whole soybeans,” in Soybeans: Chemistry, Production, Processing, and Utilization (AOCS Press), 441–481. doi:10.1016/B978-1-893997-64-6.50017-2.

Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. PLoS One 12, e1005767. doi:10.1186/1471-2156-13-100.

Lü, H., Yang, Y., Li, H., Liu, Q., Zhang, J., Yin, J., et al. (2018). Genome-wide association studies of photosynthetic traits related to phosphorus efficiency in soybean. Front. Plant Sci. 9. doi:10.3389/fpls.2018.01226.

Ma, J., Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., et al. (2022). Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4. 0, r10285). Geosci. Model Dev. 15, 815–839.

Mutava, R. N., Prince, S. J. K., Syed, N. H., Song, L., Valliyodan, B., Chen, W., et al. (2015). Understanding abiotic stress tolerance mechanisms in soybean: A comparative evaluation of soybean response to drought and flooding stress. Plant Physiol. Biochem. 86, 109–120.

Mwenye, O. J., van Rensburg, L., van Biljon, A., and van der Merwe, R. (2016). The role of proline and root traits on selection for drought-stress tolerance in soybeans: a review. South African J. Plant Soil 33, 245–256. doi:10.1080/02571862.2016.1148786.

Naderi, A., Naseri, R., Fathi, A., Bahamin, S., and Maleki, R. (2013). Physiological performance of soybean cultivars under drought stress. Bull. Environ. Pharmacol. Life Sci. 2, 38–44.

Nouri, M. Z., Moumeni, A., and Komatsu, S. (2015). Abiotic stresses: Insight into gene regulation and protein expression in photosynthetic pathways of plants. Int. J. Mol. Sci. 16, 20392–20416. doi:10.3390/ijms160920392.

Ohashi, Y., Nakayama, N., Saneoka, H., and Fujita, K. (2006). Effects of drought stress on photosynthetic gas exchange, chlorophyll fluorescence and stem diameter of soybean plants. Biol. Plant. 50, 138–141. doi:10.1007/s10535-005-0089-3.

Oya, T., Nepomuceno, A. L., Neumaier, N., Farias, J. R. B., Tobita, S., and Ito, O. (2004). Drought tolerance characteristics of Brazilian soybean cultivars - evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field. Plant Prod. Sci. 7, 129–137. doi:10.1626/pps.7.129.

Patinni, I. R. G., de Andrade, C. A., Campos, C. N. S., Teodoro, L. P. R., Andrade, S. M., Roque, C. G., et al. (2020). Agronomic performance and water-use efficiency of F3 soybean populations grown under contrasting base saturation. J. Agron. Crop Sci. 206, 806–814. doi:10.1111/jac.12413.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. doi:10.1086/519795.

Quero, G., Simondi, S., Ceretta, S., Otero, Á., Garaycochea, S., Fernández, S., et al. (2021). An integrative analysis of yield stability for a GWAS in a small soybean breeding population. Crop Sci. 61, 1903–1914. doi:10.1002/csc2.20490.

R Core Team (2021). R: a language and environment for statistical computing. available at: https://www.r-project.org/.

Raj, A., Stephens, M., and Pritchard, J. K. (2014). FastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197, 573–589. doi:10.1534/genetics.114.164350.

Seck, W., Torkamaneh, D., and Belzile, F. (2020). Comprehensive genome-wide association analysis reveals the genetic basis of root system architecture in soybean. Front. Plant Sci. 11, 590740. doi:10.3389/fpls.2020.590740.

Serraj, R., Sinclair, T. R., and Purcell, L. C. (1999). Symbiotic N_2 fixation response to drought. J. Exp. Bot. 50, 143–155. doi:10.1093/jxb/50.331.143.

Shibata, M., Breuer, C., Kawamura, A., Clark, N. M., Rymen, B., Braidwood, L., et al. (2018). GTL1 and DF1 regulate root hair growth through transcriptional repression of ROOT HAIR DEFECTIVE 6-LIKE 4 in Arabidopsis. Dev. 145, 1–11. doi:10.1242/dev.159707.

Shurtleff, W., and Aoyagi, A. (2010). History of soybeans and soyfood in Canada (1831-2010): Extensively annotated bibliography and sourcebook. Soyinfo Center, Lafayette, CA.

Siddiqui, Z. S., Cho, J. Il, Kwon, T. R., Ahn, B. O., Lee, K. S., Jeong, M. J., et al. (2014). Physiological mechanism of drought tolerance in transgenic rice plants expressing *Capsicum annuum* methionine sulfoxide reductase B2 (CaMsrB2) gene. Acta Physiol. Plant. 36, 1143–1153. doi:10.1007/s11738-014-1489-9.

Siddiqui, Z. S., Cho, J. Il, Park, D. B., Lee, G. S., Ryu, T. H., Shahid, H., et al. (2015). Field assessment of CaMsrB2 transgenic lines in a drought stress environment. Turk. J. Botany 39, 973–981. doi:10.3906/bot-1502-18.

Silvente, S., Sobolev, A. P., and Lara, M. (2012). Metabolite adjustments in drought tolerant and sensitive soybean genotypes in response to water stress. PLoS One 7, e38554. doi:10.1371/journal.pone.0038554.

Sonah, H., O’Donoughue, L., Cober, E., Rajcan, I., and Belzile, F. (2015). Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. Plant Biotechnol. J. 13, 211–221. doi:10.1111/pbi.12249.

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

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. Nat. Rev. Genet. 20, 467–484. doi:10.1038/s41576-019-0127-1.

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature 466, 707–713. doi:10.1038/nature09270.

Thilakarathna, M. S., Moroz, N., and Raizada, M. N. (2017). A biosensor-based leaf punch assay for glutamine correlates to symbiotic nitrogen fixation measurements in legumes to permit rapid screening of rhizobia inoculants under controlled conditions. Front. Plant Sci. 8, 1714. doi:10.3389/fpls.2017.01714.

Thilakarathna, M. S., and Raizada, M. N. (2017). A meta-analysis of the effectiveness of diverse rhizobia inoculants on soybean traits under field conditions. Soil Biol. Biochem. 105, 177–196. doi:10.1016/j.soilbio.2016.11.022.

Torkamaneh, D., Laroche, J., Tardivel, A., O'Donoughue, L., Cober, E., Rajcan, I., et al. (2018). Comprehensive description of genomewide nucleotide and structural variation in short-season soya bean. Plant Biotechnol. J. 16, 749–759. doi:10.1111/pbi.12825.

Tuteja, N., Gill, S. S., and Tuteja, R. (2011). Plant responses to abiotic stresses: shedding light on salt, drought, cold and heavy metals. Omi. Plant Abiotic Stress Toler. 1, 39–64.

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., et al. (2017). ePlant: Visualizing and exploring multiple levels of data for hypothesis generation in plant biology. Plant Cell 29, 1806–1821. doi:10.1105/tpc.17.00073.

Wang, L., Yang, Y., Zhang, S., Che, Z., Yuan, W., and Yu, D. (2020). GWAS reveals two novel loci for photosynthesis-related traits in soybean. Mol. Genet. Genomics 295, 705–716. doi:10.1007/s00438-020-01661-1.

Wilkinson, S., and Davies, W. J. (2002). ABA-based chemical signalling: the co-ordination of responses to stress in plants. Plant. Cell Environ. 25, 195–210.

Wu, C., Ma, Q., Yam, K. M., Cheung, M. Y., Xu, Y., Han, T., et al. (2006). In situ expression of the GmNMH7 gene is photoperiod-dependent in a unique soybean (*Glycine max* [L.] Merr.) flowering reversion system. Planta 223, 725–735. doi:10.1007/s00425-005-0130-y.

Wu, T. ting, Li, J. yu, Wu, C. xiang, Sun, S., Mao, T. ting, Jiang, B. jun, et al. (2015). Analysis of the independent- and interactive-photo-thermal effects on soybean flowering. J. Integr. Agric. 14, 622–632. doi:10.1016/S2095-3119(14)60856-X.

Yang, W., Wu, T., Zhang, X., Song, W., Xu, C., Sun, S., et al. (2019). Critical photoperiod measurement of soybean genotypes in different maturity groups. Crop Sci. 59, 2055–2061. doi:10.2135/cropsci2019.03.0170.

Yang, Y., Wang, L., Zhang, D., Cheng, H., Wang, Q., Yang, H., et al. (2020). GWAS identifies two novel loci for photosynthetic traits related to phosphorus efficiency in soybean. Mol. Breed. 40. doi:10.1007/s11032-020-01112-0.

Yang, Y., Zhu, X., Cui, R., Wang, R., Li, H., Wang, J., et al. (2021). Identification of soybean phosphorous efficiency QTLs and genes using chlorophyll fluorescence parameters through GWAS and RNA-seq. Planta 254, 1–13. doi:10.1007/s00425-021-03760-8.

Ye, H., Song, L., Schapaugh, W. T., Ali, M. L., Sinclair, T. R., Riar, M. K., et al. (2020). The importance of slow canopy wilting in drought tolerance in soybean. J. Exp. Bot. 71, 642–652. doi:10.1093/jxb/erz150.

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. Genomics, Proteomics Bioinforma. 19, 619–628. doi:10.1016/j.gpb.2020.10.007.

Zhang, L. X., Kyei-Boahen, S., Zhang, J., Zhang, M. H., Freeland, T. B., Watson, C. E., et al. (2007). Modifications of optimum adaptation zones for soybean maturity groups in the USA. Crop Manag. 6, 1–11. doi:10.1094/cm-2007-0927-01-rs.

Zhao, Z., and Assmann, S. M. (2011). The glycolytic enzyme, phosphoglycerate mutase, has critical roles in stomatal movement, vegetative growth, and pollen production in *Arabidopsis thaliana*. J. Exp. Bot. 62, 5179–5189. doi:10.1093/jxb/err223.

Chapter 4 - The Genotypic Variability Among Short-Season Soybean Cultivars for Grain Production and Nitrogen Fixation Under Drought Stress

Abstract

The global population is predicted to be reaching 9.7 billion by 2050, and there is an increasing food demand, with a specific need for protein. Soybean [*Glycine max* (L.) Merr.] is the leading commercial grain legume grown in the world. It fixes atmospheric nitrogen through the symbiotic rhizobia bacteria that inhabitant root nodules. Drought stress limits plant growth, yield, and symbiotic nitrogen fixation (SNF) in soybean. Notably, the changing global climate is predicted to lead large areas of the world to experience extensive drought conditions over the next few decades. Most importantly, many regions in North America have been experiencing frequent and severe droughts over the last few decades. Therefore, there is a great need to improve the drought tolerance and SNF in soybean to increase food production for the rising population in the world. We believe that there is genetic variation for drought tolerance and SNF in short-season soybean varieties. The main objective of this study is to identify allelic variations associated with yield, yield parameters, and SNF in short-season Canadian soybean varieties under drought stress. A diversity panel of 103 early maturity Canadian soybean varieties for which genome-wide genotypic data are available was used in this study. A greenhouse pot experiment was conducted to determine yield and symbiotic nitrogen fixation-related traits under drought stress. The drought was imposed after three weeks of plant growth, where plants were maintained at 30% field capacity - FC (drought) and 80% FC (well-watered) until seed maturity. Under drought stress, soybean plants had lower seed yield, yield components (number of pods, number of seeds, 100-seed weight), seed nitrogen content, % nitrogen derived from the atmosphere (%Ndfa), and total fixed nitrogen compared to the well-watered condition. Significant genotypic variability among soybean varieties was found for yield, yield parameters, nitrogen fixation traits, and water use efficiency (measured employing carbon isotope discrimination – CID). A Genome-Wide Association Study (GWAS) was conducted using 2.16M SNPs (2,164,465 SNPs) for different yield and nitrogen fixation related parameters for 30% FC, 80% FC and their relative performance (30% FC / 80% FC). In total, six quantitative trait locus (QTL) regions, including candidate genes, were detected as significantly associated with %Ndfa and CID under drought stress and relative performance. These genes can potentially aid in future breeding efforts to develop drought-resistant soybean varieties.

Keywords: soybean, drought, symbiotic nitrogen fixation, candidate gene, quantitative trait locus, genome wide association study

4.1 Introduction

Soybean (*Glycine max* (L) Merr) is the main grain legume grown in the world and is primarily used for oil, food, and feed production (Herridge et al., 2008). Soybean originated and domesticated in China 6000-9000 years ago from *Glycine soja* (Sieb. & Zucc.) (Guo et al., 2010; Wang et al., 2016a). Soybean production has increased significantly over the last five decades as a result of the expansion of the growing areas (Liu et al., 2020). In terms of Canada, according to 2021 statistics, the highest soybean production was recorded in Ontario, followed by Quebec, Manitoba, Maritimes, and Saskatchewan, where the total production accounts for 6.27 MMT (Soy Canada, 2022).

It is predicted that drought stress will be one of the world's costliest climatic problem in the near future (Wilhite, 2000). The agricultural regions of the Canadian Prairies are particularly vulnerable to frequent drought conditions (Bonsal and Regier, 2007). Drought stress limits soybean productivity and affects yield stability (Quero et al., 2021). Soybean yield can drop by more than half during droughts, resulting in significant losses for farmers (Wei et al., 2018). Drought stress reduces grain legume productivity in all the growth stages. However, drought stress during the reproductive and grain filling stages causes significant yield loss (Farooq et al., 2017). This yield loss reflects in different ways, such as a decrease in pod number, poorly developed pods, reduction in seed weight, reduction in seed number, and decline in seed quality (Arya et al., 2021). The plant growth stage and the duration of drought stress are important in determining the level of impact in soybean (Wei et al., 2018; Dong et al., 2019). Drought stress induced in early maturity soybean varieties at vegetative stages results in reduced plant height, decline in seed number in the early reproductive stages, and reduced seed weight in late reproductive stages (Desclaux et al., 2000). The drought condition between flowering and early seed filling stages can influence the vegetative growth of branches, leading to a decrease in branch seed number and a reduction in branch seed yield (Frederick et al., 2001). For example, research in China found that drought stress at the flowering and seed filling stages reduced yield by 73% to 82 percent (Wei et al., 2018). Another study reported that long-term drought stress in reproductive stages reduces biomass allocation to reproductive organs, resulting in lower seed weight in soybean (Du et al., 2020).

One of the most intriguing characteristics of soybean is its ability to create a symbiotic relationship with rhizobia bacteria found in root nodules. They can convert atmospheric nitrogen into ammonia inside the root nodules, and in return, the host plant provides photosynthesis products for rhizobia metabolism (Thilakarathna and Raizada, 2017). Symbiotic nitrogen fixation is sensitive to different biotic and abiotic factors (Hungria and Vargas, 2000; Thilakarathna and Raizada, 2017), where the percentage of nitrogen derived from the atmosphere (%Ndfa) can vary from 0 to 95% in soybean due to these biotic and abiotic factors (Hungria and Bohrer, 2000; Herridge et al., 2008; Thilakarathna and Raizada, 2017; Santachiara et al., 2019) Among the different abiotic factors, drought stress is a major factor that limits symbiotic nitrogen fixation (SNF) in soybean (Serraj et al., 1999b; Serraj, 2003; Thilakarathna and Raizada, 2017). Drought stress affects different stages of legume-rhizobia symbiosis, such as root hair infection, nodule growth and development, and nodule function (Arrese-igorl et al., 1999; Sulieman and Tran, 2016). Furthermore, drought stress inhibits nitrogenase activity which is the key enzyme in catalyzing the reduction of dinitrogen (N_2) to ammonia (NH_3) (Einsle and Rees, 2020). The reduction in SNF ultimately leads to a reduction in grain yield, seed nitrogen, and grain protein production. Drought stress reduces symbiotic nitrogen fixation in nodules through oxygen limitation in nodules, carbon scarcity, and nitrogen fixation feedback inhibition (Arrese-Igor et al., 2011; Collier and Tegeder, 2012; Kunert et al., 2016).

In the last decade, genome-wide association studies (GWAS) have become a feasible approach for discovering beneficial alleles from a genetic diversity panel. GWAS has achieved extensive use in crops, despite being a relatively new tool in the fields of plant breeding and molecular biology. Different yield and nitrogen fixation-related parameters have been considered in previous GWAS in soybean, including 100 seed weight (Contreras-Soto et al., 2017; Hu et al., 2020; Adeboye et al., 2021; Li et al., 2021; Priyanatha et al., 2022), seed yield (Contreras-Soto et al., 2017; Hu et al., 2020; Adeboye et al., 2021; Ravelombola et al., 2021; Priyanatha et al., 2022), number of pods (Hu et al., 2020; Adeboye et al., 2021), pod weight (Li et al., 2021), (Ravelombola et al., 2021; Zhang et al., 2021; Jo et al., 2022), yield stability (Quero et al., 2021), number of seeds per plant (Chen et al., 2020; Adeboye et al., 2021), number of seeds per pod (Chen et al., 2020), seed moisture content (Li et al., 2021), %Ndfa (Dhanapal et al., 2015b), and carbon isotope discrimination (CID) (Dhanapal et al., 2015a; Kaler et al., 2017a; Steketee et al., 2019). Specifically GWAS conducted in soybean during drought conditions for seed number per plant and number of seeds per pod was found only in one study (Chen et al., 2020). There is a research gap in identifying alleles involved in SNF in soybean.

The main objective of this study is to identify allelic variation associated with different yield parameters and SNF in soybean under drought stress and to identify the genomic regions controlling drought tolerance in short-season soybean varieties. A diverse panel of 103 Canadian short-season soybean cultivars was phenotyped for multiple yield and nitrogen fixation-related traits, including the number of pods per plant, number of seeds per plant, seed weight, 100-seed weight, %Ndfa, seed nitrogen, total nitrogen fixed, and water use efficiency under drought conditions.

4.2 Materials and methods

4.2.1 Germplasm, plant materials and growth conditions

A diverse Canadian short-season soybean panel consisting of 103 soybean genotypes was used in this study. First, seeds were surface sterilized using 70% ethanol for two minutes and washed with autoclaved double distilled water five times (Thilakarathna et al., 2017). The professional growing mix (Sun Gro Horticultural Canada Ltd, Seba Beach, Canada) and sand (Target Products Ltd, Morinville, Alberta) were mixed in a 3:1 ratio in 6.52 L pots (H.J.S. Wholesale Ltd, Winnipeg, Manitoba). Initially, three seeds were planted in a pot. Plants were maintained in a greenhouse with supplemental lighting (range: 500-600 mol m^{-2} s^{-1} at the top of the canopy, Fortimo LED Line, High Flux VO) at 26 ± 2 °C during the day and 20 ± 2 °C at night. The photoperiod was kept at 16/8 h light/dark cycles. Extra plants were removed, leaving one plant per pot after one week of plant growth. Seedlings were inoculated with 2 ml of *Bradyrhizobium japonicum* USDA 110 inoculum (rhizobial density OD_{600} = 0.1) (Thilakarathna et al., 2017). The same process was repeated one week after the first inoculation to ensure successful nodulation. Each week, plants received 100 ml of quarter-strength N-free Hoagland's nutrient solution (HOP03-50LT, Caisson Labs, UT). Plants were labelled with 25 ml of 0.5 mM $K_{15}NO_3$ solution (10 atom% 15N; 348481-25G; Sigma Aldrich, Oakville, ON, Canada) after two and three weeks of planting to measure SNF.

4.2.2 Soil field capacity

The determination of field capacity was done in a separate study. First, the bottom of the 6.52 L pots was covered with a coffee filter (12" Mother Parkers Coffee Filters) to avoid any soil leakage. Then, pots were filled with the media of sand (Target Products Ltd, Morinville, Alberta) and growing mix (Sun Gro Horticultural Canada Ltd, Seba Beach, Canada) on a 1:3 volume basis till a constant final weight was obtained (e.g., 4500 g). The initial dry weight of the soil (Dw) was measured after drying in an oven at 80 °C until obtained a constant weight (Earl, 2003). The pots were watered until saturation and drained out from the bottom. The top of the pots was covered using aluminum foil to avoid evaporation. The pots were kept for 24 hrs until observed no water drainage, and the final saturated weight was recorded (Sw). The Field Capacity (FC) was calculated as FC = S_W-D_W. Accordingly, final weights for the well-watered treatment (80% FC) and drought treatment (30% FC) were calculated (Liyanage et al., 2022).

4.2.3 Drought treatment and yield data collection

The soil moisture contents of the pots were maintained at 80% FC (well-watered) and 30% FC (drought) after three weeks of germination. The 30% FC was reached by withholding water until pots reached 30% FC. This moisture adjustment in all the pots was carried out using an Arduino-based semi-automated irrigation system throughout the greenhouse experiment (Liyanage et al., 2022). The main treatments were genotypes and moisture levels, and they were allocated according to the randomized complete block design with four replicates per treatment. Different yield parameters such as number of pods per plant, number of seeds per plant, 100 seed weight, and seed weight per plant were collected at seed maturity.

4.2.4 Determination of nitrogen fixation-related parameters

The %Ndfa was measured using the isotope dilution method (Thilakarathna et al., 2016b). Seeds were oven-dried at 60 °C for three days and ground to a coarse seed powder by using a coffee grinder. A sub-sample from each sample was further ground in a small Eppendorf tube along with a steel bead in a bead beater homogenizer (OMNI International, Kennesaw, GA, USA). Then, a 5 mg of soybean powder sample was measured into a small tin capsule (8 mm × 5 mm, D1008, Isomass Scientific Inc) using a microbalance. Samples were enveloped and compressed into a tiny pellet to make sure no air remained in it. The tin capsules were arranged in a 96 well plate and sent to the Stable Isotope Facility, Agri-Food Canada, Lethbridge to analyze 15N, 13C, and total N% (Thilakarathna et al., 2021).The encapsulated seed samples were analyzed with a Finnigan Delta V Plus (Thermo Electron, Bremen, Germany) Isotope Ratio Mass Spectrometer (IRMS) fitted with a Flash 2000 Elemental Analyzer (Thermo Fisher Scientific, Voltaweg, Netherlands) and a Conflo IV interface

(Thermo Fisher Scientific, Bremen, Germany) between the IRMS and the analyzer. The isotope standards were L-glutamic acid (USGA40) and L-glutamic acid enriched in 13C and 15N (USGA41A) (United States Geological Survey). The %Ndfa of the soybean was calculated using the following formula according to the isotope dilution technique:

$$\%Ndfa = \left(1 - \frac{\text{atom\%}\ {}^{15}N\ excess_{(fixing\ plant)}}{\text{atom\%}\ {}^{15}N\ excess_{(non-fixing\ plant)}}\right) x\ 100$$

where atom % ^{15}N excess = atom % ^{15}N soybean - 0.3663. The amount of seed nitrogen derived from nittrogen fixation was calculated based on the total seed nitrogen content and %Ndfa (seed N content x %Ndfa/100).

4.2.5 Carbon isotope discrimination

To evaluate plant water use efficiency, the carbon isotope discrimination (CID) was calculated as follows (Wilker et al., 2019).

$$\Delta C\ (CID) = \frac{\delta^{13}C_{air} - \delta^{13}C_{plant}}{1 + \delta^{13}C_{plant}}$$

where $\delta^{13}C_{air}$ is the C isotope composition of atmospheric CO_2 (−8‰) and $\delta^{13}C_{plant}$ is the C isotope composition of the soybean shoot samples.

4.2.6 Statistical analysis

Analysis of variance (ANOVA) set at $\alpha < 0.05$ was used to examine the effects of moisture treatments and soybean genotypes. A two-factor factorial design was used to analyze the data. The main factor was the soil moisture level with two levels of soil moisture: 80% FC (well-watered) and 30% FC (drought), and the subfactor was soybean genotypes. R 3.5 was used to perform the ANOVA, frequency distributions, and Pearson correlations (R Core Team, 2021).

4.2.7 Genotyping data

Seck et al., 2020 created a catalogue of 2.18M SNPs for 137 accessions by combining the genotyping by sequencing (GBS) and whole-genome sequencing (WGS) approaches, and from that larger data set 103 accessions with 2.16M SNPs were extracted for this study (2,164,465 SNPs) (Seck et al., 2020).

4.2.8 Population structure

In the panel of 2.18M SNPs, LD-based pruning (r^2>0.5) was done with PLINK (Purcell et al., 2007) to get reduced and uniformly distributed set of 14K markers. The fastSTRUCTURE algorithm (Raj et al., 2014) was used to characterize the population structure using tested subpopulations (K) from 1 to 13 and three independent runs of each. The python 'Choseek.py' script was used to find the most suitable K value based on the rate of change in LnP between the successive K values. In addition, a phylogenetic tree was constructed in TASSEL 5.0 using the Neighbor-Joining method (Appendix Figure B.1), and a scree plot (Appendix Figure C.1) was used to evaluate the most informative Principal Components (PCs) (Appendix Figures C.2 and C.3). Furthermore, a Kinship matrix was calculated using the efficient mixed-model association (EMMA) method.

4.2.9 Genome-wide association study

A genome-wide association study (GWAS) was carried out using 2.16M SNPs (2,164,465 SNPs) utilizing the fixed and random model circulation probability unification (FarmCPU) model (Liu et al., 2016) implemented in rMVP package on Microsoft open R (Yin et al., 2021). To reduce the false-positive signals, the population matrix (Q) and Kinship matrix (K) were measured and used as covariates. A genome-wide significance threshold level that is less than 0.05 was used to find the significant associations using the false discovery rate (FDR) test (Benjamini and Hochberg, 1995).

4.2.10 Candidate gene identification

The soybean public database (SoyBase 2020) and soybean reference genome annotation was used to identify candidate genes for the yield and nitrogen fixation-related parameters. The QTL flanking areas were set to 100kb on either side of the QTL peak to seek potential genes involved in yield parameters. The tool ePlant2 was used to get more information about the gene expression in different tissues (Waese et al., 2017).

4.3 Results

4.3.1 Phenotypic variation of yield and nitrogen fixation-related traits in soybean

A significant phenotypic variation among 103 soybean genotypes was found for the number of pods per plant, the number of seeds per plant, seed yield, 100-seed weight, %Ndfa, seed total nitrogen content, total nitrogen fixed, and CID ($p<0.0001$). The soil moisture content also had a significant effect on the number of pods per plant, the number of seeds per plant, total seed weight, 100-seed weight, %Ndfa, seed total nitrogen content, total nitrogen fixed ($p<0.0001$), and CID ($p=0.0227$) (Table 1). Drought stress significantly reduced yield and yield parameters compared to the well-watered condition. Similarly, %Ndfa, seed total nitrogen content, total nitrogen fixed, and CID were lower under 30% FC compared to 80% FC. The summary statistics for different plant physiological parameters measured under 30% and 80% FC are shown in Table 4.1.

Under drought stress, varieties such as DH 748, OAC Lakeview, Mario, OAC Madoc, and OAC Champion produced the most pods per plant, while Albions, 9004, Auriga, OAC-07-04C, and AC 2001 produced the fewest. In terms of the number of seeds per plant, DH 748, OAC Avatar, OAC Ginty, OAC Madoc, and OAC Wallace had the most, while Albinos, Auriga, 9004, Maple Donovan, and AC 2001 had the fewest under drought stress. Seed yield was higher in DH 748, OAC Avatar, OAC Ginty, OAC-09-35C, and OAC Lauralain, and lowest in Albions, Maple Donovan, 9004, Auriga, and Naya under drought. The 100-seed weight was greater in DH 618, Alta, Amasa, DH 420, and Ohgata under drought stress, where Maple presto, 90B11, Maple Donovan, 90A01, and OAC Madoc had the lowest.

In terms of nitrogen fixation-related traits, OAC Champion, DH 748, OAC Oxford, Toki, and OAC Wallace had the highest %Ndfa, while Maple Donovan, Costaud, Albinos, Gaillard, and Naya had the lowest under drought stress. The nitrogen content of the seed was highest in DH 748, OAC Ginty, OAC Avatar, OAC Oxford, and 91M10 under drought stress, where Maple Donovan, 9004, Albinos, Auriga, and AC 2001 had the lowest. Total nitrogen fixation was highest in DH 748, OAC Ginty, OAC Avatar, OAC Oxford, and OAC Stratford, where Maple Donovan, Costaud, Albinos, 9004, and Gaillard had the lowest under drought stress. In terms of water use efficiency, OAC Elora, OT11-01, Maple presto, OAC Carman, and 90A01 had higher water use efficiency, whereas the lowest found in OAC Champion, Krios, DH 530, OAC-08-22C, and OAC Avatar.

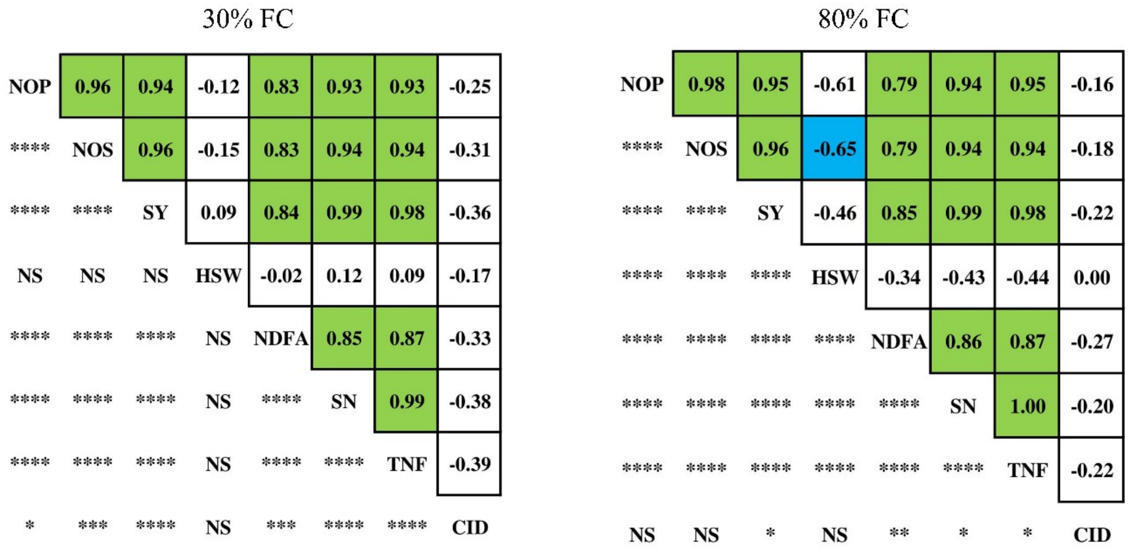
Table 4.1 Summary statistics of the grain yield, yield components, nitrogen fixation traits, and carbon isotope discrimination of 103 soybean genotypes under well-watered and drought conditions.

Traits (per plant)	80% FC	30% FC	p value
Number of pods	67.2 ± 3.2	39.6 ± 1.4	$p<0.0001$
Number of seeds	150.2 ± 7.5	87.9 ± 3.5	$p<0.0001$
Seed yield	24.7 ± 1.0	16.1 ± 0.6	$p<0.0001$
100-seed weight	17.5 ± 0.3	18.5 ± 0.2	$p<0.0001$
%Ndfa	84.4 ± 1.3	73.1 ± 1.9	$p<0.0001$
Seed total nitrogen	1.8 ± 0.1	1.2 ± 0.1	$p<0.0001$
Total nitrogen fixed	1.6 ± 0.1	0.9 ± 0.1	$p<0.0001$
CID	20.1 ± 0.1	20.2 ± 0.1	$p=0.0227$

Mean value of the trait ± SE; FC, Field capacity; %Ndfa, percentage of nitrogen derived from the atmosphere; CID, carbon isotope discrimination

4.3.2 Correlations among yield and nitrogen fixation-related traits in soybean

Significant correlations among different plant traits were found under the 80% FC and 30% FC treatments (Figure 4.1). The %Ndfa and total nitrogen fixation (g N plant^{-1}) were positively correlated with the number of pods, number of seeds, seed yield, and seed nitrogen content under both 80% FC and 30% FC treatments. Seed yield was positively correlated with the number of pods, number of seeds, %Ndfa, seed nitrogen content, and total nitrogen fixed under 80% FC and 30% FC treatments. Interestingly, seed yield, %Ndfa, and total nitrogen fixed were not correlated with 100-seed weight and CID under both irrigation treatments. The number of seeds was negatively correlated with 100-seed weight in 80% FC treatment.



30% FC

NOP	0.96	0.94	-0.12	0.83	0.93	0.93	-0.25
****	NOS	0.96	-0.15	0.83	0.94	0.94	-0.31
****	****	SY	0.09	0.84	0.99	0.98	-0.36
NS	NS	NS	HSW	-0.02	0.12	0.09	-0.17
****	****	****	NS	NDFA	0.85	0.87	-0.33
****	****	****	NS	****	SN	0.99	-0.38
****	****	****	NS	****	****	TNF	-0.39
*	***	****	NS	***	****	****	CID

80% FC

NOP	0.98	0.95	-0.61	0.79	0.94	0.95	-0.16
****	NOS	0.96	-0.65	0.79	0.94	0.94	-0.18
****	****	SY	-0.46	0.85	0.99	0.98	-0.22
****	****	****	HSW	-0.34	-0.43	-0.44	0.00
****	****	****	****	NDFA	0.86	0.87	-0.27
****	****	****	****	****	SN	1.00	-0.20
****	****	****	****	****	****	TNF	-0.22
NS	NS	*	NS	**	*	*	CID

Figure 4.1 Correlations among yield and nitrogen fixation related traits in soybean under drought and well-watered conditions. Numbers above the diagonal correspond to Pearson's correlation coefficients (r). Green boxes highlight the positive values exceeding 0.65 and blue boxes highlight the negative values exceeding or equal 0.65. Below the diagonal show the degree of significance of the corresponding correlations between traits (****P < 0.0001, ***P < 0.001, **P < 0.01, *P < 0.05, and NS: not significant). FC, field capacity; NOP, number of pods; NOS, number of seeds; SY, seed yield; HSW, 100-seed weight; %Ndfa, percentage of nitrogen derived from the atmosphere; SN, seed nitrogen; TNF, total nitrogen fixed; CID, carbon isotope discrimination.

4.3.3 Genotyping data and population structure

As reported in Seck et al., 2020, both GBS and WGS genotyping techniques were utilized in this study to cover the entire soybean genome (Seck et al., 2020). A subset of 14K trimmed SNPs was used to characterize population structure. The optimal number of subpopulations (K) was 6 to 9. This was also confirmed with PCA and phylogeny analysis.

4.3.4 Genome-wide association of yield and nitrogen fixation-related traits

GWAS analysis was performed for yield, yield parameters, and nitrogen fixation-related traits under 30% and 80% FC treatments (Appendix Figure C.4) and their relative performance using 2.16M SNPs (2,164,465 SNPs) and the FarmCPU statistical model. The Appendix Figure C.5 exhibited the distribution of SNP markers across the soybean genome. The Appendix Figures C.6, C.7 and C.8 represented quantile-quantile plots of p values for the association between SNP markers and different traits in 30% FC, 80% FC and relative performance respectively. In total, 6 SNPs were detected for 30% FC drought treatment (Figure 4.2) and their relative performance (Figure 4.3). Three SNPs were associated with 30% FC

(%Ndfa- qNDFA1, qNDFA2, qCID), and another three SNPs were associated with relative performance (%Ndfa – qNDFA3, qNDFA4, qNDFA5) (Table 4.2).

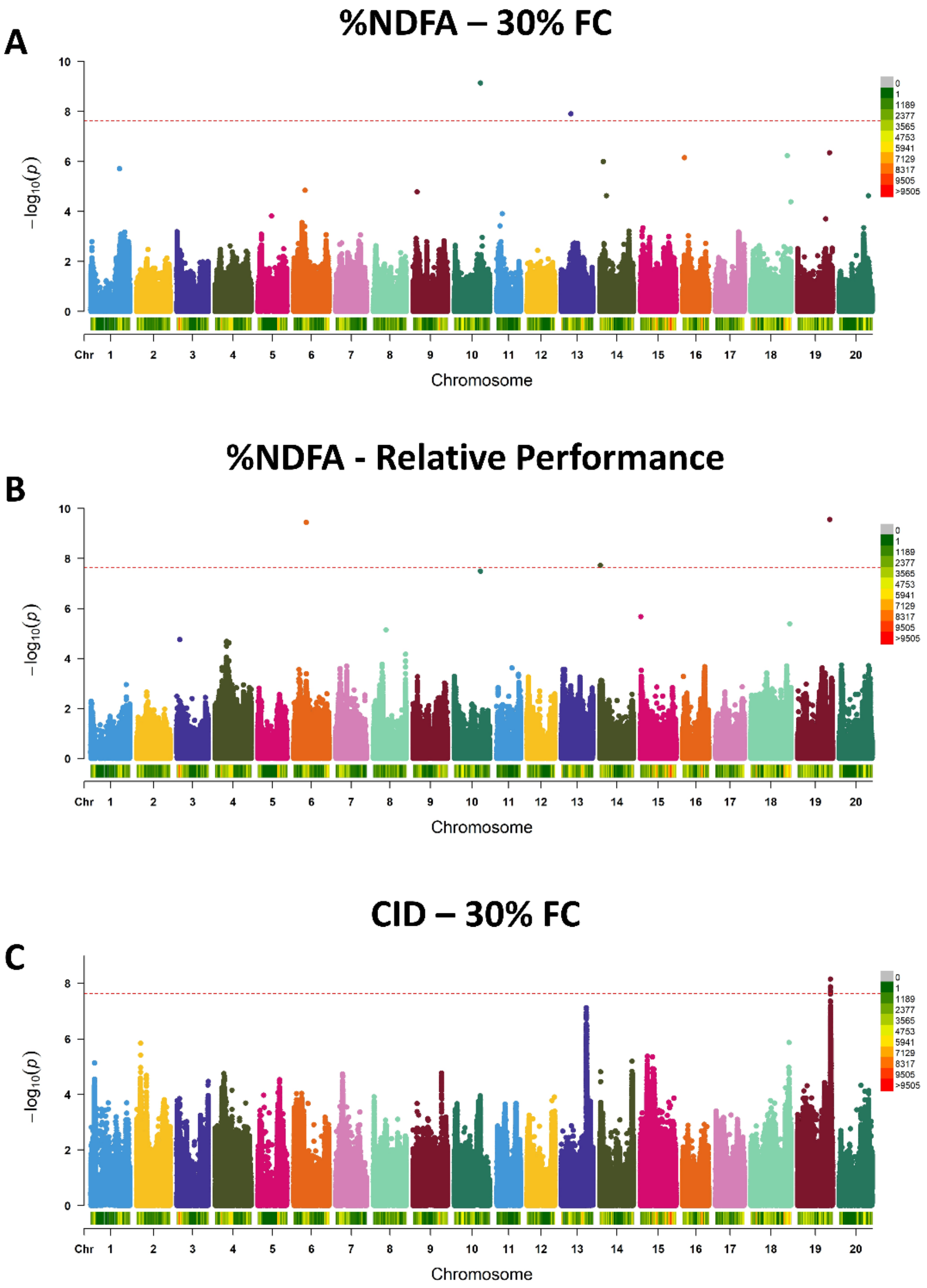


Figure 4.2 Manhattan plots of the genome-wide association results for % nitrogen derived from the atmosphere (%Ndfa) and carbon isotope discrimination (CID). (A) %Ndfa-30% field capacity (FC). (B) %Ndfa- relative performance. (C) Carbon isotope discrimination (CID) – 30% field capacity (FC). Relative performance was calculated as the ratio of %Ndfa under 30% Fc vs. 80% FC. Negative $\log_{10}$ (P-values, y-axis) describing the strength of the association between each marker and trait are plotted against the physical position of each marker (x-axis). The pink dashed line indicates the significance threshold (FDR = 5%) and beyond that are considered as significant associations.

Table 4.2 List of quantitative trait loci (QTL) associated with 30% field capacity (FC), and relative performance.

Moisture Effect	Traits	Chr[a]	MSS[b] position	QTL ID	Major allele frequency / Minor allele frequency/ Average of phenotypic values	Effect
30% FC	%Ndfa[⁑]	10	37995110	qNDFA1	G:A/0.81:0.19/77.18:56.63	-6.87
		13	13866995	qNDFA2	T:C/0.70:0.29/77.41:60.74	-4.79
	CID[¶]	19	47602841	qCID	G:A/0.60:0.40/19.60:21.13	0.75
Relative performance[c]	%Ndfa	6	18244365	qNDFA3	G:A/0.94:0.06/0.88:0.55	-0.12
		14	994141	qNDFA4	G:T/0.88:0.12/0.86:0.78	-0.05
		19	46575733	qNDFA5	C:T/0.89:0.11/0.88:0.65	-0.09

FC = field capacity
%Ndfa = percentage of nitrogen derived from the atmosphere
CID = carbon isotope discrimination
[a] *Chromosome number*
[b] *Most significant SNP*
[C] *Relative performance was calculated as the ratio of %Ndfa under 30% Fc vs. 80% FC*

4.3.5 Yield and nitrogen fixation traits related candidate genes

All the genes that are residing in whole or in part within the 6 QTLs of interest were extracted from SoyBase. Table 4.3 provides the complete information of these genes, including their annotations. Based on their annotation, we identified some strong candidate genes separately for 30% FC, and their relative performance as mentioned below.

4.3.5.1 Candidate genes under drought stress

For the trait %Ndfa, we found the gene *Glyma.10g144600* which is annotated for Glycogen synthase kinase-3 and Glycogen synthase kinase (qNDFA1) (Table 4.3). The gene *Glyma.10g145300* that encodes for galactinol synthase was also found as a candidate gene for qNDFA1. The gene *Glyma.10g144300* annotated for S-adenosylmethionine synthetase 2 was also found in this region (qNDFA1). For the CID, which estimates water use efficiency, we found the gene *Glyma.19g225500* that was annotated as aspartic proteinase A1 (qCID) (Table

4.3). The gene *Glyma.19g221700* was also identified for the qCID that was annotated as WRKY family transcription factor. Furthermore, the gene *Glyma.19g221900* encodes as Indole-3-acetic acid inducible 20 was found in this region.

4.3.5.2 Candidate genes under relative performance

For the relative performance of %Ndfa, we found the gene *Glyma.06g197700* which is annotated for putative endonuclease or glycosyl hydrolase with C2H2-type zinc finger domain (qNDFA3) (Table 4.3). The gene *Glyma.06g198600* encodes for Ankyrin repeat family protein identified in this region. Moreover, *Glyma.06g199700* gene encoding for Remorin family protein was recognized as an important candidate gene (qNDFA3). The gene *Glyma.19g212800* annotated as sucrose synthase 3 is another strong candidate gene for qNDFA5 (Table 4.3). The gene *Glyma.19g213900* encodes for Drought-responsive family protein was also identified for this trait.

Table 4.3 Candidate genes associated with drought stress (30% field capacity) and relative performance

Treatment	Trait	Chr	MSS position	REF/ALT	QTL ID	Candidate genes	Orthologous genes in *Arabidopsis*	Annotations
30% FC	%NDFA	10	37995110	G/A	qNDFA1	Glyma.10g144600	Gene Model: AT5G26751.1	Glycogen synthase kinase - FJ460228
						Glyma.10g144600	Gene Model: AT5G26751.1	Glycogen synthase kinase-3 - BT093874
						Glyma.10g145300	Gene Model: AT2G47180.1	Galactinol synthase 1- AK245720, AY126715
						Glyma.10g144300	Gene Model: AT4G01850.1	S-adenosylmethionine synthetase 2
	CID	19	47602841	G/A	qCID	Glyma.19g225500	Gene Model: AT1G11910.1	Aspartic proteinase A1
						Glyma.19g221700	Gene Model: AT4G04450.1	WRKY family transcription factor
						Glyma.19g221900	Gene Model: AT2G46990.1	Indole-3-acetic acid inducible 20
Relative Performance	%NDFA	6	18244365	G/A	qNDFA3	Glyma.06g197700	Gene Model: AT5G61190.1	Putative endonuclease or glycosyl hydrolase with C2H2-type zinc finger domain
						Glyma.06g198600	Gene Model: AT2G03430.1	Ankyrin repeat family protein
						Glyma.06g199700	Gene Model: AT5G61280.1	Remorin family protein
		19	46575733	C/T	qNDFA5	Glyma.19g212800	Gene Model: AT4G02280.1	Sucrose synthase 3
						Glyma.19g213900	Gene Model: AT4G02200.1	Drought-responsive family protein

FC, field capacity; %Ndfa, percentage of nitrogen derived from the atmosphere; CID, carbon isotope discrimination

4.4 Discussion

4.4.1 Significant phenotypic variation of yield and nitrogen fixation-related traits in soybean

As previously stated, we discovered significant phenotypic variability for various yield and nitrogen fixation-related parameters at the maturity stage. Our findings corroborate with previous findings, wherein phenotypic variability for different yield and nitrogen fixation-related traits were found for number of pods (Hu et al., 2020), number of seeds (Chen et al., 2020) seed weight (Yan et al., 2017; Ravelombola et al., 2021), 100-seed weight (Sonah et al., 2015; Zhang et al., 2016; Contreras-Soto et al., 2017; Copley et al., 2018; Hu et al., 2020), %Ndfa (Dhanapal et al., 2015b), and CID (Dhanapal et al., 2015a; Kaler et al., 2017a; Steketee et al., 2019). The presence of phenotypic variation within a germplasm pool for various yield and nitrogen fixation-related traits is critical for plant breeders to make selection for breeding.

4.4.2 Drought stress on symbiotic nitrogen fixation

In this study, drought stress had significant negative impact on symbiotic nitrogen fixation in soybean. Drought stress reduced %NDFA by 13.4%, total seed nitrogen by 34.9%, and the amount of seed nitrogen fixed by 42.1% compared to the well-watered plants. The reduction in nitrogen fixation under drought stress may be due to multiple plant responses. Drought stress can lead to decreases in carbon flux to nodules and a reduction in the performance of sucrose synthase activity. This can lead to reduced sucrose production, where the lack of sucrose in nodules inhibits SNF (King and Purcell, 2005; Kunert et al., 2016; Sulieman and Tran, 2016). Furthermore, a reduction in available oxygen in root nodules can disrupt rhizobia nitrogen fixation (Serraj 2003). Furthermore, feedback inhibition of nitrogen fixation can also be taken place due to the increase in ureides and free amino acids in soybean plant tissues (King and Purcell, 2005; Sulieman and Tran, 2016).

4.4.3 Correlations among yield and nitrogen fixation-related traits

Many yield and nitrogen fixation-related traits were found to be highly and significantly correlated. These findings are also consistent with the previous literature. For instance, it has been observed that the number of pods per plant has a significant correlation between the number of seeds per pod and seed yield (Hao et al., 2012b; Silva et al., 2015; Hu et al., 2020; Mustofa et al., 2021). We found a negative correlation between the number of seeds and 100-seed weight. The negative relationship could be that when the number of seeds per plant increases, the seed size decreases (Hao et al., 2012b; Silva et al., 2015). Soybean seeds contain a significant amount of plant nitrogen (~71%) (Ciampitti and Salvagiotti, 2018) and it mainly derives from SNF. The strong positive correlation between the %Ndfa and seed nitrogen content under drought and conditions indicates that symbiotic nitrogen fixation is crucial for seed nitrogen accumulation and protein production in soybean even under drought conditions. Surprisingly, we found that CID has no relationship with other metrics, implying that the water use efficiency does not correlate with yield or nitrogen fixation characteristics.

4.4.4 Genome-wide association using whole-genome data revealed 6 QTLs associated with %Ndfa and CID

In this study, a GWAS was performed utilizing a comprehensive genome-wide set of SNPs (2,164,465 SNPs). We discovered six genomic regions (QTL) that contribute to %Ndfa and CID under drought stress and relative performance. In the same manner, some previous GWAS studies in soybean also have revealed loci for %Ndfa (Dhanapal et al., 2015b) and CID (Dhanapal et al., 2015a; Kaler et al., 2017a; Steketee et al., 2019). We discovered that the majority of the QTL regions identified in our study are novel. Furthermore, the majority of GWAS research studies in soybean were based on seed yield because seed yield per plant is a key character that is the focus of crop improvement. A limited number of GWAS have been conducted for the %Ndfa and CID in soybean and specifically under drought conditions.

4.4.5 Candidate genes for %Ndfa and CID-associated QTLs

4.4.5.1 Candidate genes under drought stress

The gene *Glyma.10g144600* associated with %Ndfa (qNDFA1) is a strong candidate gene annotated for glycogen synthase kinase-3 and highly expressed in roots. (Table 4.3). Under salt stress, the glycogen synthase kinase 3 (GSK3)-like kinase plays an important role in inhibiting symbiotic signaling and nodule formation in soybean (Phang et al., 2011; He et al., 2021; Singh and Verma, 2021). The GSK3-like kinases directly regulate the activities of G. max Nodulation Signaling Pathway 1 (GmNSP1) proteins to facilitate legume–rhizobium symbiosis under salt stress (He et al., 2021). The galactinol and raffinose act as osmoprotectants for drought-stress tolerance in plants (Taji et al., 2002) and the gene *Glyma.10g145300* annotated for galactinol synthase and highly expressed in roots was also identified for qNDFA1 (Table 4.3). Overexpression of galactinol synthase (which catalyzes the first step in the biosynthesis of raffinose family oligosaccharides) results in increased galactinol and raffinose accumulation and improve drought tolerance in soybean (Taji et al., 2002; Bellaloui et al., 2013). It is found that the gene *Glyma.10g144300*-S-adenosylmethionine synthetase 2 is down-regulated in soybean root tips and roots under drought conditions, and it is highly expressed in soybean nodules (qNDFA1) (Wang et al., 2016b) (Table 4.3).

The gene *Glyma.19g225500* annotated as aspartic proteinase A1 was identified for qCID, which estimates the trait water use efficiency and is highly expressed in roots (Table 4.3). In Arabidopsis thaliana, drought tolerance is conferred by overexpression of the aspartic protease APA1 gene (Sebastián et al., 2020). The gene *Glyma.19g221700* annotated as WRKY family transcription factor and highly expressed in roots (qCID) (Table 4.3). The WRKY genes are involved in response to abiotic stress (Jiang et al., 2017). GmWRKY12, for example, is sensitive to drought, salt, ABA, and salicylic acid (Shi et al., 2018). GmWRKY12 overexpression improved drought and salt tolerance by increasing proline content, and lowering malondialdehyde level in transgenic soybean seedlings exposed to drought and salt (Shi et al., 2018). The gene *Glyma.19g221900* encodes as Indole-3-acetic acid inducible 20 (qCID) (Table 4.3). Indole acetic acid increases water transport in roots and works with abscisic acid to respond to turgor-affecting stressors like drought and salinity (Gadallah, 2000).

4.4.5.2 Candidate genes under relative performance

The gene *Glyma.06g197700* annotated for putative endonuclease or glycosyl hydrolase with C2H2-type zinc finger domain was found for the relative performance of qNDFA3 (Table 4.3). It is found that in soybean, C2H2 zinc finger proteins involve in nodule development, nodule function, and nodule signal transduction (Yuan et al., 2018). The gene *Glyma.06g198600* encodes for Ankyrin repeat family protein (ANK) and was found in the same region (qNDFA3) (Table 4.3). It is found that overexpression of GmANK114 improved the survival rate of transgenic soybean hairy roots under drought and salt stresses (Zhao et al., 2020). GmANK114 overexpression in soybean hairy root showed higher proline and lower malondialdehyde contents and also lower H2O2 and O2− contents in response to drought or salt stress (Zhao et al., 2020).

The gene *Glyma.06g199700* encodes for Remorin family protein and is highly expressed in soybean roots (qNDFA3) (Table 4.3). Remorin participates in a wide range of biotic and abiotic stresses and root nodule development (Son et al., 2015). GmREM1.1, for example, was found to be highly expressed in the nodule primordia and inner cortex region of root nodules. Furthermore, GmREM2.1 transcription was mostly found in rhizobia infected cells (Son et al., 2015). The gene *Glyma.19g212800* annotated as sucrose synthase 3 was found in qNDFA5 (Table 4.3). Sucrose synthase plays a key role in the regulation of nodule carbon metabolism (González et al., 1995; Gordon et al., 1997; Arrese-igorl et al., 1999). Therefore, it will directly affect SNF in soybean as photosynthesis and carbon supply hamper under the drought stress. The gene *Glyma.19g213900*, which encodes for Drought-responsive family protein, was found in the same QTL region (Table 4.3). These Drought-responsive family genes are differentially expressed upon exposure to drought stress (Le et al., 2012; Marcolino-Gomes et al., 2014). Importantly, Drought-responsive candidate genes contribute to the development of drought-tolerant soybean cultivars (Le et al., 2012).

4.5 Conclusions

This study found significant genotypic variability among soybean varieties in terms of the number of pods per plant, the number of seeds per plant, seed weight, 100-seed weight, %Ndfa, seed nitrogen, total nitrogen fixed, and CID. The GWAS conducted for the study revealed 6 QTLs for %Ndfa and CID under drought conditions and relative performance. Furthermore, strong candidate genes were discovered to support the findings. The current study will contribute toward understanding the genetics underlying different yield and nitrogen fixation-related traits and drought tolerance in soybean.

References

Adeboye, K. A., Akhter Bhat, J., Ganie, S. A., Varshney, R. K., and Yu, D. (2021). Genome-wide association analysis of yield-related traits of soybean using haplotype-based framework. Biol. Life Sci. Forum 11, 49.

Arrese-Igor, C., Esther González, B. M., Daniel Marino, B., Rubén Ladrera, B., Estíbaliz Larrainzar, B., and Erena Gil-Quintana, B. (2011). Physiological responses of legume nodules to drought. Plant Stress 5, 24–31.

Arrese-igorl, C., Gonzalezl, E. M., Anthony, J., Minchin, F. R., Galvezl, L., Cabrerizol, P. M., et al. (1999). Sucrose synthase and nodule nitrogen fixation under drought and other environmental stresses. Symbiosis 27, 189–212.

Arya, H., Singh, M. B., and Bhalla, P. L. (2021). Towards developing drought-smart soybeans. Front. Plant Sci. 12, 750664. doi:10.3389/fpls.2021.750664.

Bellaloui, N., Hu, Y., Mengistu, A., Kassem, M. A., and Abel, C. A. (2013). Effects of foliar boron application on seed composition, cell wall boron, and seed δ15N and δ13C isotopes in water-stressed soybean plants. Front. Plant Sci. 4, 270. doi:10.3389/fpls.2013.00270.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. 57, 289–300.

Bonsal, B., and Regier, M. (2007). Historical comparison of the 2001/2002 drought in the Canadian Prairies. Clim. Res. 33, 229–242. doi:10.3354/cr033229.

Chen, L., Fang, Y., Li, X., Zeng, K., Chen, H., Zhang, H., et al. (2020). Identification of soybean drought-tolerant genotypes and loci correlated with agronomic traits contributes new candidate genes for breeding. Plant Mol. Biol. 102, 109–122. doi:10.1007/s11103-019-00934-7.

Ciampitti, I. A., and Salvagiotti, F. (2018). New insights into soybean biological nitrogen fixation. Agron. J. 110, 1185–1196. doi:10.2134/agronj2017.06.0348.

Collier, R., and Tegeder, M. (2012). Soybean ureide transporters play a critical role in nodule development, function and nitrogen export. Plant J. 72, 355–367. doi:10.1111/j.1365-313X.2012.05086.x.

Contreras-Soto, R. I., Mora, F., De Oliveira, M. A. R., Higashi, W., Scapim, C. A., and Schuster, I. (2017). A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-Based haplotype analysis. PLoS One 12, 0171105. doi:10.1371/journal.pone.0171105.

Copley, T. R., Duceppe, M. O., and O'Donoughue, L. S. (2018). Identification of novel loci associated with maturity and yield traits in early maturity soybean plant introduction lines. BMC Genomics 19, 167. doi:10.1186/s12864-018-4558-4.

Desclaux, D., Huynh, T. T., and Roumet, P. (2000). Identification of soybean plant characteristics that indicate the timing of drought stress. Crop Sci. 40, 716–722. doi:10.2135/cropsci2000.403716x.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015a). Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. Theor. Appl. Genet. 128, 73–91. doi:10.1007/s00122-014-2413-9.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015b). Genome-wide association analysis of diverse soybean genotypes reveals novel markers for nitrogen traits. Plant Genome 8, 1–15. doi:10.3835/plantgenome2014.11.0086.

Dong, S., Jiang, Y., Dong, Y., Wang, L., Wang, W., Ma, Z., et al. (2019). A study on soybean responses to drought stress and rehydration. Saudi J. Biol. Sci. 26, 2006–2017. doi:10.1016/j.sjbs.2019.08.005.

Du, Y., Zhao, Q., Chen, L., Yao, X., and Xie, F. (2020). Effect of drought stress at reproductive stages on growth and nitrogen metabolism in soybean. Agronomy 10, 302. doi:10.3390/agronomy10020302.

Earl, H. J. (2003). A precise gravimetric method for simulating drought stress in pot experiments. Crop Sci. 43, 1868–1873. doi:10.2135/cropsci2003.1868.

Einsle, O., and Rees, D. C. (2020). Structural enzymology of nitrogenase enzymes. Chem. Rev. 120, 4969–5004. doi:10.1021/acs.chemrev.0c00067.

Farooq, M., Gogoi, N., Barthakur, S., Baroowa, B., Bharadwaj, N., Alghamdi, S. S., et al. (2017). Drought stress in grain legumes during reproduction and grain filling. J. Agron. Crop Sci. 203, 81–102. doi:10.1111/jac.12169.

Frederick, J. R., Camp, C. R., and Bauer, P. J. (2001). Crop ecology, production & management: Drought-stress effects on branch and mainstem seed yield and yield components of determinate soybean. Crop Sci. 41, 759–763. doi:10.2135/cropsci2001.413759x.

Gadallah, M. A. A. (2000). Effects of indole-3-acetic acid and zinc on the growth, osmotic potential and soluble carbon and nitrogen components of soybean plants growing under water deficit. J. Arid Environ. 44, 451–467. doi:10.1006/jare.1999.0610.

González, E. M., Gordon, A. J., James, C. L., and Arrese-lgor, C. (1995). The role of sucrose synthase in the response of soybean nodules to drought. J. Exp. Bot. 46, 1515–1523. doi:10.1093/jxb/46.10.1515.

Gordon, A. J., Minchin, F. R., Skøt, L., and James, C. L. (1997). Stress-induced declines in soybean N_2 fixation are related to nodule sucrose synthase activity. Plant Physiol. 114, 937–946. doi:10.1104/pp.114.3.937.

Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H., et al. (2010). A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): implications from microsatellites and nucleotide sequences. Ann. Bot. 106, 505–514. doi:10.1093/aob/mcq125.

Hao, D., Cheng, H., Yin, Z., Cui, S., Zhang, D., Wang, H., et al. (2012). Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. Theor. Appl. Genet. 124, 447–458. doi:10.1007/s00122-011-1719-0.

He, C., Gao, H., Wang, H., Guo, Y., He, M., Peng, Y., et al. (2021). GSK3-mediated stress signaling inhibits legume–rhizobium symbiosis by phosphorylating GmNSP1 in soybean. Mol. Plant 14, 488–502. doi:10.1016/j.molp.2020.12.015.

Herridge, D. F., Peoples, M. B., and Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil 311, 1–18. doi:10.1007/s11104-008-9668-3.

Hu, D., Zhang, H., Du, Q., Hu, Z., Yang, Z., Li, X., et al. (2020). Genetic dissection of yield-related traits via genome-wide association analysis across multiple environments in wild soybean (*Glycine soja* Sieb. and Zucc.). Planta 251, 39. doi:10.1007/s00425-019-03329-6.

Hungria, M., and Bohrer, T. R. J. (2000). Variability of nodulation and dinitrogen fixation capacity among soybean cultivars. Biol. Fertil. Soils 31, 45–52. doi:10.1007/s003740050622.

Hungria, M., and Vargas, M. A. T. (2000). Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. F. Crop. Res. 65, 151–164. doi:10.1016/S0378-4290(99)00084-2.

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017). WRKY transcription factors in plant responses to stresses. J. Integr. Plant Biol. 59, 86–101. doi:10.1111/jipb.12513.

Jo, H., Lee, J. Y., and Lee, J. D. (2022). Genome-wide association mapping for seed weight in soybean with black seed coats and green cotyledons. Agronomy 12, 250. doi:10.3390/agronomy12020250.

Kaler, A. S., Dhanapal, A. P., Ray, J. D., King, C. A., Fritschi, F. B., and Purcell, L. C. (2017). Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. Crop Sci. 57, 3085–3100. doi:10.2135/cropsci2017.03.0160.

King, C. A., and Purcell, L. C. (2005). Inhibition of N_2 fixation in soybean is associated with elevated ureides and amino acids. Plant Physiol. 137, 1389–1396. doi:10.1104/pp.104.056317.1.

Kunert, K. J., Vorster, B. J., Fenta, B. A., Kibido, T., Giuseppe, D., and Foyer, C. H. (2016). Drought stress responses in soybean roots and nodules. Front. Plant Sci. 7, 1015. doi:10.3389/fpls.2016.01015.

Le, D. T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Ham, L. H., et al. (2012). Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by Genome-wide transcriptome analysis. PLoS One 7, e495. doi:10.1371/journal.pone.0049522.

Li, X., Zhou, Y., Bu, Y., Wang, X., Zhang, Y., Guo, N., et al. (2021). Genome-wide association analysis for yield-related traits at the R6 stage in a Chinese soybean mini core collection. Genes and Genomics 43, 897–912. doi:10.1007/s13258-021-01109-9.

Liu, S., Zhang, M., Feng, F., and Tian, Z. (2020). Toward a “Green Revolution” for soybean. Mol. Plant 13, 688–697. doi:10.1016/j.molp.2020.03.002.

Liyanage, D. K., Chathuranga, I., Mori, B. A., and Thilakarathna, M. S. (2022). A simple, semi-automated, gravimetric method to simulate drought stress on plants. Agronomy 12, 349.

Marcolino-Gomes, J., Rodrigues, F. A., Fuganti-Pagliarini, R., Bendix, C., Nakayama, T. J., Celaya, B., et al. (2014). Diurnal oscillations of soybean circadian clock and drought responsive genes. PLoS One 9. doi:10.1371/journal.pone.0086402.

Mustofa, A., Zubaidah, S., and Kuswantoro, H. (2021). Correlation and path analysis on yield and yield components in segregating populations. Genetika 53, 157–166. doi:10.2298/GENSR2101157R.

Phang, T.-H., Li, M.-W., Cheng, C.-C., Wong, F.-L., Chan, C., and Lam, H.-M. (2011). Molecular responses to osmotic stresses in soybean. Soybean - Mol. Asp. Breed. doi:10.5772/15155.

Priyanatha, C., Torkamaneh, D., and Rajcan, I. (2022). Genome-wide association study of soybean germplasm derived from Canadian × Chinese crosses to mine for novel alleles to improve seed yield and seed quality traits. Front. Plant Sci. 13. doi:10.3389/fpls.2022.866300.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. doi:10.1086/519795.

Quero, G., Simondi, S., Ceretta, S., Otero, Á., Garaycochea, S., Fernández, S., et al. (2021). An integrative analysis of yield stability for a GWAS in a small soybean breeding population. Crop Sci. 61, 1903–1914. doi:10.1002/csc2.20490.

R Core Team (2021). R: a language and environment for statistical computing. available at: https://www.r-project.org/.

Raj, A., Stephens, M., and Pritchard, J. K. (2014). FastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics 197, 573–589. doi:10.1534/genetics.114.164350.

Ravelombola, W., Qin, J., Shi, A., Song, Q., Yuan, J., Wang, F., et al. (2021). Genome-wide association study and genomic selection for yield and related traits in soybean. PLoS One 16, e0255761. doi:10.1371/journal.pone.0255761.

Santachiara, G., Salvagiotti, F., and Rotundo, J. L. (2019). Nutritional and environmental effects on biological nitrogen fixation in soybean: a meta-analysis. F. Crop. Res. 240, 106–115. doi:10.1016/j.fcr.2019.05.006.

Sebastián, D. I., Fernando, F. D., Raúl, D. G., and Gabriela, G. M. (2020). Overexpression of *Arabidopsis* aspartic protease APA1 gene confers drought tolerance. Plant Sci. 292, 110406. doi:10.1016/j.plantsci.2020.110406.

Seck, W., Torkamaneh, D., and Belzile, F. (2020). Comprehensive genome-wide association analysis reveals the genetic basis of root system architecture in soybean. Front. Plant Sci. 11, 590740. doi:10.3389/fpls.2020.590740.

Serraj, R. (2003). Effects of drought stress on legume symbiotic nitrogen fixation: physiological mechanisms. Indian J. Exp. Biol. 41, 1136–1141.

Serraj, R., Sinclair, T. R., and Purcell, L. C. (1999). Symbiotic N_2 fixation response to drought. J. Exp. Bot. 50, 143–155. doi:10.1093/jxb/50.331.143.

Shi, W. Y., Du, Y. T., Ma, J., Min, D. H., Jin, L. G., Chen, J., et al. (2018). The WRKY transcription factor GmWRKY12 confers drought and salt tolerance in soybean. Int. J. Mol. Sci. 19, 4087. doi:10.3390/ijms19124087.

Silva, A. F., Sediyama, T., Silva, F. C. S., Bezerra, A. R. G., and Ferreira, L. V. (2015). Correlation and path analysis of yield components in soybean varieties. Turkish J. F. Crop. 10, 177–179.

Singh, J., and Verma, P. K. (2021). NSP1 allies with GSK3 to inhibit nodule symbiosis. Trends Plant Sci. 26, 999–1001. doi:10.1016/j.tplants.2021.07.001.

Son, S., Oh, C. J., Bae, J. H., Lee, H., and An, C. S. (2015). GmREM1.1 and GmREM2.1, which encode the remorin proteins in soybean, have distinct roles during root nodule development. J. Plant Biol. 58, 17–25. doi:10.1007/s12374-014-0273-0.

Sonah, H., O'Donoughue, L., Cober, E., Rajcan, I., and Belzile, F. (2015). Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. Plant Biotechnol. J. 13, 211–221. doi:10.1111/pbi.12249.

Steketee, C. J., Sinclair, T. R., Riar, M. K., Schapaugh, W. T., and Li, Z. (2019). Unraveling the genetic architecture for carbon and nitrogen related traits and leaf hydraulic conductance in soybean using genome-wide association analyses. BMC Genomics 20, 811. doi:10.1186/s12864-019-6170-7.

Sulieman, S., and Tran, L.-S. P. (2016). Legume nitrogen fixation in a changing environment. Springer Publishers doi:10.1007/978-3-319-06212-9.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., et al. (2002). Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. Plant J. 29, 417–426. doi:10.1046/j.0960-7412.2001.01227.x.

Thilakarathna, M. S., Moroz, N., and Raizada, M. N. (2017). A biosensor-based leaf punch assay for glutamine correlates to symbiotic nitrogen fixation measurements in legumes to permit rapid screening of rhizobia inoculants under controlled conditions. Front. Plant Sci. 8, 1714. doi:10.3389/fpls.2017.01714.

Thilakarathna, M. S., Papadopoulos, Y. A., Rodd, A. V., Grimmett, M., Fillmore, S. A. E., Crouse, M., et al. (2016). Nitrogen fixation and transfer of red clover genotypes under legume–grass forage based production systems. Nutr. Cycl. Agroecosystems 106, 233–247. doi:10.1007/s10705-016-9802-1.

Thilakarathna, M. S., and Raizada, M. N. (2017). A meta-analysis of the effectiveness of diverse rhizobia inoculants on soybean traits under field conditions. Soil Biol. Biochem. 105, 177–196. doi:10.1016/j.soilbio.2016.11.022.

Thilakarathna, M. S., Torkamaneh, D., Bruce, R. W., Rajcan, I., Chu, G., Grainger, C. M., et al. (2021). Testing whether pre-pod-fill symbiotic nitrogen fixation in soybean is subject to drift or selection over 100 years of soybean breeding. Front. Agron. 3, 725813. doi:10.3389/fagro.2021.725813.

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., et al. (2017). ePlant: Visualizing and exploring multiple levels of data for hypothesis generation in plant biology. Plant Cell 29, 1806–1821. doi:10.1105/tpc.17.00073.

Wang, J., Chu, S., Zhang, H., Zhu, Y., Cheng, H., and Yu, D. (2016a). Development and application of a novel genome-wide SNP array reveals domestication history in soybean. Nature 6, 20728. doi:10.1038/srep20728.

Wang, X., Oh, M. W., and Komatsu, S. (2016b). Characterization of S-adenosylmethionine synthetases in soybean under flooding and drought stresses. Biol. Plant. 60, 269–278. doi:10.1007/s10535-016-0586-6.

Wei, Y., Jin, J., Jiang, S., Ning, S., and Liu, L. (2018). Quantitative response of soybean development and yield to drought stress during different growth stages in the Huaibei plain, China. Agronomy 8, 97. doi:10.3390/agronomy8070097.

Wilhite, D. A. (2000). Drought as a natural hazard: concepts and definitions. Drought Mitig. Cent. Fac. Publ., 1–21.

Wilker, J., Navabi, A., Rajcan, I., Marsolais, F., Hill, B., Torkamaneh, D., et al. (2019). Agronomic performance and nitrogen fixation of heirloom and conventional dry bean varieties under low-nitrogen field conditions. Front. Plant Sci. 10, 952. doi:10.3389/fpls.2019.00952.

Yan, L., Hofmann, N., Li, S., Ferreira, M. E., Song, B., Jiang, G., et al. (2017). Identification of QTL with large effect on seed weight in a selective population of soybean with genome-wide association and fixation index analyses. BMC Genomics 18, 529. doi:10.1186/s12864-017-3922-0.

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. Genomics, Proteomics Bioinforma. 19, 619–628. doi:10.1016/j.gpb.2020.10.007.

Yuan, S., Li, X., Li, R., Wang, L., Zhang, C., Chen, L., et al. (2018). Genome-wide identification and classification of soybean C2H2 zinc finger proteins and their expression analysis in legume-rhizobium symbiosis. Front. Microbiol. 9, 126. doi:10.3389/fmicb.2018.00126.

Zhang, J., Song, Q., Cregan, P. B., and Jiang, G. L. (2016). Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). Theor. Appl. Genet. 129, 117–130. doi:10.1007/s00122-015-2614-x.

Zhang, W., Xu, W., Zhang, H., Liu, X., Cui, X., Li, S., et al. (2021). Comparative selective signature analysis and high-resolution GWAS reveal a new candidate gene controlling seed weight in soybean. Theor. Appl. Genet. 134, 1329–1341. doi:10.1007/s00122-021-03774-6.

Zhao, J. Y., Lu, Z. W., Sun, Y., Fang, Z. W., Chen, J., Zhou, Y. Bin, et al. (2020). The ankyrin-repeat gene GmANK114 confers drought and salt tolerance in *Arabidopsis* and soybean. Front. Plant Sci. 11, 584167. doi:10.3389/fpls.2020.584167.

Chapter 5 - General Discussion

Drought is a major abiotic stress that negatively affects crop growth and productivity. In terms of Canada, it is predicted that drought stress will be a critical issue in the near future, especially for the Canadian Prairies which we called the hub of legumes or protein baskets in Canada (Bonsal et al., 2011). Soybean is a drought sensitive plant, and the sensitivity depends on the growth stage of the plant (Pejić et al., 2011). Soil moisture availability is critical in three different growth stages of soybean plants; germination, emergence, and flowering-grain filling (Pejić et al., 2011). Soybean is the world leading legume crop and is primarily used as an animal feed, seed oil, and a vital source of plant-based protein in human diets (Singh, 2010). Because of the world's rapid population growth, there is a need to expand the grain protein supply, where soybean has become a significant protein source. Therefore, in this study, we assessed allelic variations associated with diverse short-season soybean varieties for various plant physiological parameters at flowering, yield, and symbiotic nitrogen fixation-related traits at maturity under drought stress which is critical for soybean breeding initiatives. The evaluation of drought tolerance, yield, and SNF-related traits and associating them with molecular markers can reveal genomic regions that plant breeders can focus on. This information can be used in Marker-assisted selection (MAS) in soybean breeding programs.

The majority of drought experiments are carried out in controlled environments where it is crucial to maintain proper soil moisture content. To facilitate our main objectives in this research study, we created a unique semi-automated lysimeter based on an Arduino microcontroller that uses the gravimetric technique to adjust soil moisture content in pot experiments (Liyanage et al. 2022). This irrigation system weighs and records the mass of plants growing in pots, calculates water loss due to evapotranspiration, and automatically adjusts soil moisture to a desired relative water content. A greenhouse pot experiment with a panel of 50 early maturity Canadian soybean varieties was used to validate the system. Drought was imposed in the experiment by lowering soil moisture content in the pots to 30% of FC while keeping control pots at 80% FC. This system efficiently maintained the moisture levels during the experiment. The measured plant physiological responses confirmed that plants in the drought treatment were under moisture stress. This innovative setup is fascinating since it is low-cost, portable, and simple to use in controlled environmental research.

There are different drought tolerance mechanisms in soybean, which are related to different plant physiological traits. In this research, six different plant traits related to soybean drought tolerance at the flowering stage were evaluated: photosynthesis, transpiration, stomatal conductance, leaf chlorophyll content, water use efficiency, and intrinsic water use efficiency. Drought stress generally reduces the net CO_2 assimilation rate, stomatal conductance, transpiration rate, and crop growth (Saeidi and Abdoli, 2015). The stomata get closed to decrease the transpiration and uptake of CO_2, which is mainly controlled by chemical signals such as Abscisic acid (ABA) production in dehydrating roots (Wilkinson and Davies, 2002; Tombesi et al., 2015). In our experiment, we observed the same trend that stomatal conductance was low in drought treatment compared to the well-watered treatment. Stomatal closure prevents water loss from transpiration (Tombesi et al., 2015), where transpiration is downregulated under drought stress. In terms of photosynthesis and transpiration, we found no significant moisture effect in the drought treatment compared to the well-watered treatment in this study. These findings contradict the fact that drought reduced net photosynthesis (Siddiqui et al., 2015a; Zargar et al., 2017; Wang et al., 2018). More specifically, the function of photosystem -II (PSII) is decreased and reduces the rate of photosynthesis (Basal et al., 2020). Possible explanations for the contradictory results include the fact that the LI-COR 6400 photosynthesis system readings we obtained were instantaneous and that plants may have adapted to drought conditions. Another possibility could be the drought condition we imposed (30% FC) was not intense enough to cause significant changes in photosynthesis and transpiration. Also, the data collection was performed during the flowering stage, where the timing of data collection may not be ideal for show a significant difference in photosynthesis as more than 60% of grain protein accumulation happens during the pod filling stage. The long-term measurement of photosynthesis is the accumulation of leaf chlorophyll, and we observed a significantly higher amount of nitrogen accumulation in leaves during the drought stress, which is consistent with previous literature (Vadez et al., 2000; King and Purcell, 2005; Sulieman and Tran, 2016). This scenario is called the feedback inhibition of nitrogen fixation that is associated with an increased in ureides and free amino acids in plant tissues such as leaves and nodules. These internal nitrogen compounds acts as signal molecules for inhibiting SNF under drought stress (Neo and Layzell, 1997; Vadez et al., 2000; Serraj et al., 2001; King and Purcell, 2005; Gil-Quintana et al., 2013). Furthermore, plants utilize water more efficiently during drought stress, which is associated with decreased transpiration and photosynthesis (Kaler et al., 2017a).

In our study, we also observed that plants use water more efficiently during drought stress compared to the well-watered treatment. The intrinsic water use efficiency (iWUE) trait also behaved in a similar manner.

A genome-wide association study was performed using 2.16M SNPs for the plant physiological characteristics, yielding 13 quantitative trait locus (QTL), including many candidate genes for 30% FC drought treatment and relative performance. Under drought stress, two QTL regions for photosynthesis and stomatal conductance were found on chromosomes 4 and 15, respectively, and two QTL for the water use efficiency on chromosome 14. For relative performance, two QTL regions on chromosomes 12 and 16 for photosynthesis, one QTL on chromosome 5 for leaf chlorophyll content, one QTL on chromosome 18 for stomatal conductance, and one QTL on chromosome 9 for transpiration were found in this study. For the relative performance, QTL regions for intrinsic water use efficiency were discovered in chromosomes 8, 10, 16, 17, and 18. To justify these findings number of candidate genes were identified.

Yield and yield parameters are the most important traits in crops, and these traits are directly linked with plant physiological parameters. In this study, we evaluated yield parameters such as number of pods per plant, number of seeds per plant, seed weight, and 100-seed weight during the seed maturity stage. According to the previous studies number of pods (Hu et al., 2020), number of seeds (Chen et al., 2020), seed weight (Yan et al., 2017; Ravelombola et al., 2021), 100-seed weight (Sonah et al., 2015; Zhang et al., 2016; Contreras-Soto et al., 2017; Copley et al., 2018; Hu et al., 2020) yield parameters reduced under drought stress. In our study, we observed the same trend where the seed yield was reduced by 34.7% compared to the well-watered condition. In terms of SNF-related traits, we measured %Ndfa, seed total nitrogen, and total seed nitrogen fixed. Drought causes negative impacts on SNF in legumes (King and Purcell, 2005), and we also discovered a reduction in %Ndfa and seed total nitrogen fixation in soybean under drought stress. The %Ndfa is highly variable in soybean, which can vary from 0 to 95% depending on biotic and abiotic conditions (Herridge et al., 2008). In our study, the %Ndfa varied from 18% - 94%, which may be due to the diverse 103 genotypes used in this study and the two soil moisture treatments used. There are different reasons that can have negative impacts on nitrogen fixation under drought stress, such as reduction in carbon flux to nodules, reduced nodule oxygen permeability, the decline in nodule sucrose synthase activity, and an increase in ureides and free amino acids as described above (King and Purcell, 2005; Sulieman and Tran, 2016). The decline in nodule sucrose synthase activity leads to a reduction in the substrate for rhizobia in the nodules to facilitate SNF (Kunert et al., 2016; Sulieman and Tran, 2016). Furthermore drought stress decreases nitrogenase enzyme activity, which converts atmospheric N_2 into ammonium, reducing the overall nitrogen fixation (Naya et al., 2007). All these factors collectively decrease the SNF in soybean during drought stress.

Similar to the plant physiological traits, a genome wide association study was conducted using 2.16M SNPs for the yield and nitrogen fixation-related traits and resulted in six QTL regions for drought treatment and relative performance. For the drought treatment, two QTL regions were found on chromosomes 10 and 13, whereas one QTL was found on chromosome 19 for CID. In terms of relative performance, three QTLs were found on chromosomes 6, 14, and 19 for %Ndfa. This knowledge is critical for producing new cultivars through plant breeding programs that can tolerate drought and maintain a high nitrogen fixation.

As future directions, we can design a similar kind of experiment to investigate how root and nodulation parameters (e.g., nodule number, nodule dry weight, nitrogenase activity, leghemoglobin) respond to drought conditions using a diverse soybean variety panel. Furthermore, molecular assays for testing how nodulation and nitrogen fixation-related genes behave under drought stress will address some unknown mechanisms in soybean under abiotic stress conditions. In this study, we evaluated the soybean plants under 30% FC conditions, which is a moderate drought condition. Therefore, in future studies, we can test soybean plants even under severe drought stress to evaluate how plant responses in terms of physiological parameters, yield, and SNF-related traits. Nitrogen fixation in soybean varies depending on the growth stage of the plant (Thilakarathna et al., 2021), where SNF is lowest at the vegetative stage, rapidly increases between the R1 and R3 development stages, and highest during the pod formation (R5 stage) (Zapata et al., 1987; Pitumpe Arachchige et al., 2020). Therefore, it will be more informative to conduct GWAS by imposing drought stress during the pod-filling stages and evaluate the subsequent effects on yield and SNF-related traits.

In summary, the present study provides evidence that drought causes significant adverse effects on plant physiological parameters at the flowering stage and yield parameters and symbiotic nitrogen fixation at the seed maturity. The identified QTLs for the aforementioned parameters aid in plant breeding programs to create new cultivars that can tolerate drought stress while maintaining high nitrogen fixation. Collectively, these findings enhance our understanding of the allelic variations associated with different plant physiological traits, yield, and nitrogen fixation-related traits under drought stress.

References

Adeboye, K. A., Akhter Bhat, J., Ganie, S. A., Varshney, R. K., and Yu, D. (2021). Genome-Wide Association Analysis of yield-related traits of soybean using haplotype-based framework. *Biol. Life Sci. Forum* 11, 49.

Ahuja, L. R., Nachabe, M. H., and Rockiki, R. (2008). Soils: field capacity. *Encycl. Water Sci. Second Ed.* 2, 1128–1131. doi:10.1081/E-EWS--120010264.

Arrese-Igor, C., Esther González, B. M., Daniel Marino, B., Rubén Ladrera, B., Estíbaliz Larrainzar, B., and Erena Gil-Quintana, B. (2011). Physiological responses of legume nodules to drought. *Plant Stress* 5, 24–31.

Arrese-igorl, C., Gonzalezl, E. M., Anthony, J., Minchin, F. R., Galvezl, L., Cabrerizol, P. M., et al. (1999). Sucrose synthase and nodule nitrogen fixation under drought and other environmental stresses. *Symbiosis* 27, 189–212.

Arya, H., Singh, M. B., and Bhalla, P. L. (2021). Towards developing drought-smart soybeans. *Front. Plant Sci.* 12, 750664. doi:10.3389/fpls.2021.750664.

Ashraf, M., and Foolad, M. R. (2007). Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environ. Exp. Bot.* 59, 206–216. doi:10.1016/j.envexpbot.2005.12.006.

Bacelar, E. L. V. A., Moutinho-Pereira, J. M., Gonçalves, B. M. C., Brito, C. V. Q., Gomes-Laranjo, J., Ferreira, H. M. F., et al. (2012). "Water use strategies of plants under drought conditions," in *Plant responses to drought stress*, ed. Aroca R. (Springer, Berlin, Heidelberg.), 145–170. doi:10.1007/978-3-642-32653-0_6.

Basal, O., Szabó, A., and Veres, S. (2020). Physiology of soybean as affected by PEG-induced drought stress. *Curr. Plant Biol.* 22, 100135. doi:10.1016/j.cpb.2020.100135.

Basu, S., Ramegowda, V., Kumar, A., and Pereira, A. (2016). Plant adaptation to drought stress. *F1000Research* 5, 1554. doi:10.12688/F1000RESEARCH.7678.1.

Bellaloui, N., Hu, Y., Mengistu, A., Kassem, M. A., and Abel, C. A. (2013). Effects of foliar boron application on seed composition, cell wall boron, and seed δ15N and δ13C isotopes in water-stressed soybean plants. *Front. Plant Sci.* 4, 270. doi:10.3389/fpls.2013.00270.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate : a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300.

Biate, D. L., Kumar, L. V., Ramadoss, D., Kumari, A., Naik, S., Reddy, K. K., et al. (2014). *Genetic diversity of soybean root nodulating bacteria*. doi:10.1007/978-3-319-05936-5_5.

Blonquist, J. M., Jones, S. B., and Robinson, D. A. (2005). Standardizing characterization of electromagnetic water content sensors: Part 2. evaluation of seven sensing systems. *Vadose Zo. J.* 4, 1059–1069. doi:10.2136/vzj2004.0141.

Bogena, H. R., Huisman, J. A., Schilling, B., Weuthen, A., and Vereecken, H. (2017). Effective calibration of low-cost soil water content sensors. *Sensors* 17, 208. doi:10.3390/s17010208.

Bonsal, B., Liu, Z., Wheaton, E., and Stewart, R. (2020). Historical and projected changes to the stages and other characteristics of severe canadian prairie droughts. *Water (Switzerland)* 12. doi:10.3390/w12123370.

Bonsal, B. R., Aider, R., Gachon, P., and Lapp, S. (2013). An assessment of Canadian prairie drought: past, present, and future. *Clim. Dyn.* 41, 501–516. doi:10.1007/s00382-012-1422-0.

Bonsal, B. R., Wheaton, E. E., Chipanshi, A. C., Lin, C., Sauchyn, D. J., and Wen, L. (2011). Drought research in Canada: A review. *Atmos. - Ocean* 49, 303–319. doi:10.1080/07055900.2011.555103.

Bonsal, B., and Regier, M. (2007). Historical comparison of the 2001/2002 drought in the Canadian Prairies. *Clim. Res.* 33, 229–242. doi:10.3354/cr033229.

Boudhrioua, C., Bastien, M., Torkamaneh, D., and Belzile, F. (2020). Genome-wide association mapping of *Sclerotinia sclerotiorum* resistance in soybean using whole-genome resequencing data. *BMC Plant Biol.* 20, 1–9. doi:10.1186/s12870-020-02401-8.

Bruce, R. W., Grainger, C. M., Ficht, A., Eskandari, M., and Rajcan, I. (2019). Trends in soybean trait improvement over generations of selective breeding. *Crop Sci.* 59, 1870–1879. doi:10.2135/cropsci2018.11.0664.

Campbell, D. I., Laybourne, C. E., and Ian J. Blair (2002). Measuring peat moisture content using the dual-probe heat pulse technique. *Aust. J. Soil Sci.* 40, 177–190.

Carter, T. E., Hymowitz, T., and Nelson, R. L. (2004). "Biogeography, local adaptation, vavilov, and genetic diversity in soybean," in *Biological Resources and Migration* (Springer, Berlin, Heidelberg.), 47–59. doi:10.1007/978-3-662-06083-4.

Chamarthi, S. K., Kaler, A. S., Abdel-Haleem, H., Fritschi, F. B., Gillman, J. D., Ray, J. D., et al. (2021). Identification and confirmation of loci associated with canopy wilting in soybean using genome-wide association mapping. *Front. Plant Sci.* 12, 698116. doi:10.3389/fpls.2021.698116.

Chard, J., van Iersel, M. W., and Bugbee, B. (2010). Mini-lysimeters to monitor transpiration and control drought stress : system design and unique applications.

Chen, L., Fang, Y., Li, X., Zeng, K., Chen, H., Zhang, H., et al. (2020). Identification of soybean drought-tolerant genotypes and loci correlated with agronomic traits contributes new candidate genes for breeding. *Plant Mol. Biol.* 102, 109–122. doi:10.1007/s11103-019-00934-7.

Ciampitti, I. A., and Salvagiotti, F. (2018). New insights into soybean biological nitrogen fixation. *Agron. J.* 110, 1185–1196. doi:10.2134/agronj2017.06.0348.

Collier, R., and Tegeder, M. (2012). Soybean ureide transporters play a critical role in nodule development, function and nitrogen export. *Plant J.* 72, 355–367. doi:10.1111/j.1365-313X.2012.05086.x.

Contreras-Soto, R. I., Mora, F., De Oliveira, M. A. R., Higashi, W., Scapim, C. A., and Schuster,

I. (2017). A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-Based haplotype analysis. *PLoS One* 12, 0171105. doi:10.1371/journal.pone.0171105.

Copley, T. R., Duceppe, M. O., and O'Donoughue, L. S. (2018). Identification of novel loci associated with maturity and yield traits in early maturity soybean plant introduction lines. *BMC Genomics* 19, 167. doi:10.1186/s12864-018-4558-4.

Cornic, G., and Briantais, J. M. (1991). Partitioning of photosynthetic electron flow between CO_2 and O_2 reduction in a C3 leaf (Phaseolus vulgaris L.) at different CO_2 concentrations and during drought stress. *Planta* 183, 178–184. doi:10.1007/BF00197786.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nat. Clim. Chang.* 3, 52–58. doi:10.1038/nclimate1633.

Day, D. A., Poole, P. S., Tyerman, S. D., and Rosendahl, L. (2001). Ammonia and amino acid transport across symbiotic membranes in nitrogen-fixing legume nodules. *Cell. Mol. Life Sci.* 58, 61–71.

Dayoub, E., Lamichhane, J. R., Schoving, C., Debaeke, P., and Maury, P. (2021). Early-stage phenotyping of root traits provides insights into the drought tolerance level of soybean cultivars. *Agronomy* 11. doi:10.3390/agronomy11010188.

Denison, R. F., and Kinraide, T. B. (1995). Oxygen-induced membrane depolarizations in legume root nodules possible evidence for an osmoelectrical mechanism controlling nodule gas permeability. *Plant Physiol.* 108, 235–240.

Desclaux, D., Huynh, T. T., and Roumet, P. (2000). Identification of soybean plant characteristics that indicate the timing of drought stress. *Crop Sci.* 40, 716–722. doi:10.2135/cropsci2000.403716x.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015a). Genome-wide association study (GWAS) of carbon isotope ratio (δ13C) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. *Theor. Appl. Genet.* 128, 73–91. doi:10.1007/s00122-014-2413-9.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2016). Genome-wide association mapping of soybean chlorophyll traits based on canopy spectral reflectance and leaf extracts. *BMC Plant Biol.* 16, 1–15. doi:10.1186/s12870-016-0861-x.

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015b). Genome-Wide Association analysis of diverse soybean genotypes reveals novel markers for nitrogen traits. *Plant Genome* 8, 1–15. doi:10.3835/plantgenome2014.11.0086.

Dixon, R., and Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nat. Rev.* 2, 621–631. doi:10.1038/nrmicro954.

Dobriyal, P., Qureshi, A., Badola, R., and Hussain, S. A. (2012). A review of the methods available for estimating soil moisture and its implications for water resource management. *J. Hydrol.* 458, 110–117. doi:10.1016/j.jhydrol.2012.06.021.

Dong, S., Jiang, Y., Dong, Y., Wang, L., Wang, W., Ma, Z., et al. (2019). A study on soybean responses to drought stress and rehydration. *Saudi J. Biol. Sci.* 26, 2006–2017. doi:10.1016/j.sjbs.2019.08.005.

Downie, J. A. (2005). Legume haemoglobins : symbiotic nitrogen fixation needs bloody nodules. *Curr. Biol.* 15, 196–198. doi:10.1016/j.cub.2005.03.006.

Du, Y., Zhao, Q., Chen, L., Yao, X., and Xie, F. (2020). Effect of drought stress at reproductive stages on growth and nitrogen metabolism in soybean. *Agronomy* 10, 302. doi:10.3390/agronomy10020302.

Dumschott, K., Dechorgnat, J., and Merchant, A. (2019). Water deficit elicits a transcriptional response of genes governing d-pinitol biosynthesis in soybean (*Glycine max*). *Int. J. Mol. Sci.* 20, 1–11. doi:10.3390/ijms20102411.

Earl, H. J. (2003). A precise gravimetric method for simulating drought stress in pot experiments. *Crop Sci.* 43, 1868–1873. doi:10.2135/cropsci2003.1868.

Earl, H. J., and Davis, R. F. (2003). Effect of drought stress on leaf and whole canopy radiation use efficiency and yield of maize. *Agron. J.* 95, 688–696. doi:10.2134/agronj2003.6880.

Ehleringer, J. R., and Cooper, T. A. (1992). On the role of orientation in reducing photoinhibitory damage in photosynthetic-twig desert shrubs. *Plant. Cell Environ.* 15, 301–306. doi:10.1111/j.1365-3040.1992.tb00977.x.

Einsle, O., and Rees, D. C. (2020). Structural enzymology of nitrogenase enzymes. *Chem. Rev.* 120, 4969–5004. doi:10.1021/acs.chemrev.0c00067.

Erdman, J., and John, W. (2009). Soy protein and cardiovascular disease: a statement for healthcare professionals from the nutrition committee of the AHA. *Circulation* 102, 2555–2559. doi:10.1111/j.1751-7117.2009.00028.x.

Farooq, M., Gogoi, N., Barthakur, S., Baroowa, B., Bharadwaj, N., Alghamdi, S. S., et al. (2017). Drought stress in grain legumes during reproduction and grain filling. *J. Agron. Crop Sci.* 203, 81–102. doi:10.1111/jac.12169.

Farooq, M., Wahid, A., and Siddique, K. H. M. (2012). Micronutrient application through seed treatments - a review. *J. Soil Sci. Plant Nutr.* 12, 125–142. doi:10.4067/S0718-95162012000100011.

Follett, R. F., and Delgado, J. A. (2002). Nitrogen fate and transport in agricultural systems. *J. Soil Water Conserv.* 57, 402–408.

Foyer, C. H., Lam, H. M., Nguyen, H. T., Siddique, K. H. M., Varshney, R. K., Colmer, T. D., et al. (2016). Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants* 2, 1–10. doi:10.1038/NPLANTS.2016.112.

Franche, C., Lindström, K., and Claudine, E. (2009). Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant Soil* 321, 35–59. doi:10.1007/s11104-008-9833-8.

Frederick, J. R., Camp, C. R., and Bauer, P. J. (2001). Crop ecology, production & management:

drought-stress effects on branch and mainstem seed yield and yield components of determinate soybean. *Crop Sci.* 41, 759–763. doi:10.2135/cropsci2001.413759x.

Fujita, K., Ofosu-Budu, K. G., and Ogata, S. (1992). Biological nitrogen fixation in mixed legume . cereal cropping systems. *Plant Soil* 141, 155–175.

Fukuda, Y. (1933). Cytogenetical studies on the wild and cultivated Manchurian soybeans (*Glycine* L.). *Japanese J. Bot.* 6, 489–506. Available at: https://ci.nii.ac.jp/naid/10012526143/.

Gadallah, M. A. A. (2000). Effects of indole-3-acetic acid and zinc on the growth, osmotic potential and soluble carbon and nitrogen components of soybean plants growing under water deficit. *J. Arid Environ.* 44, 451–467. doi:10.1006/jare.1999.0610.

Gage, D. J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol. Mol. Biol. Rev.* 68, 280–300. doi:10.1128/MMBR.68.2.280.

Gai, J., Xu, D., Gao, Z., Shimamoto, Y., Abe, J., Fukushi, H., et al. (2000). Studies on the evolutionary relationship among eco-types of *G. max* and *G. soja* in China. *Zuo wu xue bao* 26, 513—520. Available at: http://europepmc.org/abstract/CBA/339184.

Ganjeali, A., Porsa, H., and Bagheri, A. (2011). Assessment of Iranian chickpea (*Cicer arietinum* L.) germplasms for drought tolerance. *Agric. Water Manag.* 98, 1477–1484. doi:10.1016/j.agwat.2011.04.017.

Gebre, M. G., and Earl, H. J. (2020). Effects of growth medium and water stress on soybean [*Glycine max* (L.) Merr.] growth, soil water extraction and rooting profiles by depth in 1-m rooting columns. *Front. Plant Sci.* 11, 487. doi:10.3389/fpls.2020.00487.

Gebre, M. G., and Earl, H. J. (2021). Soil water deficit and fertilizer placement effects on root biomass distribution, soil water extraction, water use, yield, and yield components of soybean [*Glycine max* (L.) Merr.] grown in 1-m rooting columns. *Front. Plant Sci.* 12, 581127. doi:10.3389/fpls.2021.581127.

Geurts, R., Elena, F., and Bisseling, T. (2005). Nod factor signaling genes and their function in the early stages of *Rhizobium* infection. *Plant Biol.* 8, 346–352. doi:10.1016/j.pbi.2005.05.013.

Gil-Quintana, E., Larrainzar, E., Seminario, A., Díaz-Leal, J. L., Alamillo, J. M., Pineda, M., et al. (2013). Local inhibition of nitrogen fixation and nodule metabolism in drought-stressed soybean. *J. Exp. Bot.* 64, 2171–2182. doi:10.1093/jxb/ert074.

González, E. M., Gordon, A. J., James, C. L., and Arrese-lgor, C. (1995). The role of sucrose synthase in the response of soybean nodules to drought. *J. Exp. Bot.* 46, 1515–1523. doi:10.1093/jxb/46.10.1515.

Gordon, A. J., Minchin, F. R., Skøt, L., and James, C. L. (1997). Stress-induced declines in soybean N_2 fixation are related to nodule sucrose synthase activity. *Plant Physiol.* 114, 937–946. doi:10.1104/pp.114.3.937.

Gorim, L. Y., and Vandenberg, A. (2017a). Evaluation of wild lentil species as genetic resources to improve drought tolerance in cultivated lentil. *Front. Plant Sci.* 8, 1129. doi:10.3389/fpls.2017.01129.

Gorim, L. Y., and Vandenberg, A. (2017b). Root traits, nodulation and root distribution in soil for five wild lentil species and *Lens culinaris* (Medik.) grown under well-watered conditions. *Front. Plant Sci.* 8, 1632. doi:10.3389/fpls.2017.01632.

Grant, D., Nelson, R. T., Cannon, S. B., and Shoemaker, R. C. (2009). SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* 38, 843–846. doi:10.1093/nar/gkp798.

Grieshop, C. M., and Fahey, G. C. (2001). Comparison of quality characteristics of soybeans from Brazil, China, and the United States. *J. Agric. Food Chem.* 49, 2669–2673. doi:10.1021/jf0014009.

Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H., et al. (2010). A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): Implications from microsatellites and nucleotide sequences. *Ann. Bot.* 106, 505–514. doi:10.1093/aob/mcq125.

Halbleib, C. M., and Ludden, P. W. (2000). Recent advances in nutritional sciences. *Am. Soc. Nutr. Sci.* 136, 1081–1084.

Hannah, R., and Roser, M. (2021). Forests and deforestation. *OurWorldInData.org*.

Hao, D., Chao, M., Yin, Z., and Yu, D. (2012a). Genome-wide association analysis detecting significant single nucleotide polymorphisms for chlorophyll and chlorophyll fluorescence parameters in soybean (*Glycine max*) landraces. *Euphytica* 186, 919–931. doi:10.1007/s10681-012-0697-x.

Hao, D., Cheng, H., Yin, Z., Cui, S., Zhang, D., Wang, H., et al. (2012b). Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. *Theor. Appl. Genet.* 124, 447–458. doi:10.1007/s00122-011-1719-0.

Hartman, G. L., West, E. D., and Herman, T. K. (2011). Crops that feed the World 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. *Food Secur.* 3, 5–17. doi:10.1007/s12571-010-0108-x.

Hatfield, J. L., and Dold, C. (2019). Water-use efficiency: Advances and challenges in a changing climate. *Front. Plant Sci.* 10, 103. doi:10.3389/fpls.2019.00103.

He, C., Gao, H., Wang, H., Guo, Y., He, M., Peng, Y., et al. (2021). GSK3-mediated stress signaling inhibits legume–rhizobium symbiosis by phosphorylating *GmNSP1* in soybean. *Mol. Plant* 14, 488–502. doi:10.1016/j.molp.2020.12.015.

He, J., Du, Y. L., Wang, T., Turner, N. C., Yang, R. P., Jin, Y., et al. (2017). Conserved water use improves the yield performance of soybean (*Glycine max* (L.) Merr.) under drought. *Agric. Water Manag.* 179, 236–245. doi:10.1016/j.agwat.2016.07.008.

He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H. K., and Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant Sci.* 5, 1–8. doi:10.3389/fpls.2014.00484.

Herridge, D. F., Peoples, M. B., and Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant Soil* 311, 1–18. doi:10.1007/s11104-008-9668-3.

Herritt, M., Dhanapal, A. P., Purcell, L. C., and Fritschi, F. B. (2018). Identification of genomic loci associated with 21chlorophyll fluorescence phenotypes by genome-wide association analysis in soybean. *BMC Plant Biol.* 18, 1–19. doi:10.1186/s12870-018-1517-9.

Hirsch, A. M., and Fujishige, N. A. (2012). Molecular signals and receptors : communication between nitrogen-fixing bacteria and their plant hosts. *Biocommunication of Plants* 14, 255–280.

Höhner, R., Pribil, M., Herbstová, M., Lopez, L. S., Kunz, H. H., Li, M., et al. (2020). Plastocyanin is the long-range electron carrier between photosystem II and photosystem i in plants. *Proc. Natl. Acad. Sci. U. S. A.* 117, 15354–15362. doi:10.1073/pnas.2005832117.

Hossain, M. A., Burritt, D. J., Fujita, M., Munne-Bosch, S., and Pedro, D.-V. (2017). *Ascorbic Acid in Plant Growth , Development and Stress Tolerance*. Springer Publishers doi:10.1007/978-3-319-74057-7.

Hu, D., Zhang, H., Du, Q., Hu, Z., Yang, Z., Li, X., et al. (2020). Genetic dissection of yield-related traits via genome-wide association analysis across multiple environments in wild soybean (*Glycine soja* Sieb. and Zucc.). *Planta* 251, 39. doi:10.1007/s00425-019-03329-6.

Hufstetler, E. V., Boerma, H. R., Carter, T. E., and Earl, H. J. (2007). Genotypic variation for three physiological traits affecting drought tolerance in soybean. *Crop Sci.* 47, 25–35. doi:10.2135/cropsci2006.04.0243.

Hungria, M., and Bohrer, T. R. J. (2000). Variability of nodulation and dinitrogen fixation capacity among soybean cultivars. *Biol. Fertil. Soils* 31, 45–52. doi:10.1007/s003740050622.

Hungria, M., Campo, R. J., Chueire, L. M. O., Grange, L., and Megías, M. (2001). Symbiotic effectiveness of fast-growing rhizobial strains isolated from soybean nodules in Brazil. *Biol. Fertil. Soils* 33, 387–394. doi:10.1007/s003740100338.

Hungria, M., and Vargas, M. A. T. (2000). Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. *F. Crop. Res.* 65, 151–164. doi:10.1016/S0378-4290(99)00084-2.

Hwang, E. Y., Song, Q., Jia, G., Specht, J. E., Hyten, D. L., Costa, J., et al. (2014). A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics* 15, 1–12. doi:10.1186/1471-2164-15-1.

Hwang, S., King, C. A., Ray, J. D., Cregan, P. B., Chen, P., Carter, T. E., et al. (2015). Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. *Theor. Appl. Genet.* 128, 2047–2065. doi:10.1007/s00122-015-2566-1.

Hymowitz, T. (2016). "Speciation and Cytogenetics," in *Soybeans: Improvement, Production, and Uses* (John Wiley & Sons, Ltd), 97–136. doi:10.2134/agronmonogr16.3ed.c4.

Hymowitz, T., and Harlan, J. R. (1983). Introduction of soybean to North America by Samuel Bowen in 1765. *Econ. Bot.* 37, 371–379. doi:10.1007/BF02904196.

Inamullah, and Isoda, A. (2005). Adaptive responses of soybean and cotton to water stress: I. Transpiration changes in relation to stomatal area and stomatal conductance. *Plant Prod. Sci.* 8, 16–26. doi:10.1626/pps.8.16.

Islam, M. A., Du, H., Ning, J., Ye, H., and Xiong, L. (2009). Characterization of *Glossy1*-homologous genes in rice involved in leaf wax accumulation and drought resistance. *Plant Mol. Biol.* 70, 443–456. doi:10.1007/s11103-009-9483-0.

Jabbari, M., Fakheri, B. A., Aghnoum, R., Nezhad, N. M., and Ataei, R. (2018). GWAS analysis in spring barley (*Hordeum vulgare* L.) for morphological traits exposed to drought. *PLoS One* 13, e0204952. doi:10.1371/journal.pone.0204952.

Janiak, A., Kwaśniewski, M., and Szarejko, I. (2016). Gene expression regulation in roots under drought. *J. Exp. Bot.* 67, 1003–1014. doi:10.1093/jxb/erv512.

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017). WRKY transcription factors in plant responses to stresses. *J. Integr. Plant Biol.* 59, 86–101. doi:10.1111/jipb.12513.

Jo, H., Lee, J. Y., and Lee, J. D. (2022). Genome-Wide Association mapping for seed weight in soybean with black seed coats and green cotyledons. *Agronomy* 12, 250. doi:10.3390/agronomy12020250.

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. (2007). How rhizobial symbionts invade plants : the *Sinorhizobium – Medicago* model. *Nat. Rev.* 5, 619–633. doi:10.1038/nrmicro1705.

Jumrani, K., and Bhatia, V. S. (2019). Identification of drought tolerant genotypes using physiological traits in soybean. *Physiol. Mol. Biol. Plants* 25, 697–711. doi:10.1007/s12298-019-00665-5.

Kaler, A. S., Dhanapal, A. P., Ray, J. D., King, C. A., Fritschi, F. B., and Purcell, L. C. (2017a). Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. *Crop Sci.* 57, 3085–3100. doi:10.2135/cropsci2017.03.0160.

Kaler, A. S., Ray, J. D., Schapaugh, W. T., King, C. A., and Purcell, L. C. (2017b). Genome-wide association mapping of canopy wilting in diverse soybean genotypes. *Theor. Appl. Genet.* 130, 2203–2217. doi:10.1007/s00122-017-2951-z.

Kantolic, A. G., and Slafer, G. A. (2005). Reproductive development and yield components in indeterminate soybean as affected by post-flowering photoperiod. *F. Crop. Res.* 93, 212–222. doi:10.1016/j.fcr.2004.10.001.

Karr-Lilienthal, L. K., Grieshop, C. M., Merchen, N. R., Mahan, D. C., and Fahey, G. C. (2004). Chemical composition and protein quality comparisons of soybeans and soybean meals from five leading soybean-producing countries. *J. Agric. Food Chem.* 52, 6193–6199.

doi:10.1021/jf049795+.

Kato, S., Samanfar, B., Morrison, M. J., Bekele, W. A., Torkamaneh, D., Rajcan, I., et al. (2021). Genome-wide association study to identify soybean stem pushing resistance and lodging resistance loci. *Can. J. Plant Sci.* 101, 663–670. doi:10.1139/cjps-2020-0187.

Keyantash, J., and Dracup, J. A. (2002). The quantification of drought: an evaluation of drought indices. *Bull. Am. Meteorol. Soc.* 83, 1167–1180.

Kim, M. Y., Van, K., Kang, Y. J., Kim, K. H., and Lee, S. H. (2012). Tracing soybean domestication history: From nucleotide to genome. *Breed. Sci.* 61, 445–452. doi:10.1270/jsbbs.61.445.

King, C. A., and Purcell, L. C. (2005). Inhibition of N_2 fixation in soybean is associated with elevated ureides and amino acids. *Plant Physiol.* 137, 1389–1396. doi:10.1104/pp.104.056317.1.

Kojima, Y., Shigeta, R., Miyamoto, N., Shirahama, Y., Nishioka, K., Mizoguchi, M., et al. (2016). Low-cost soil moisture profile probe using thin-film capacitors and a capacitive touch sensor. *Sensors* 16, 1292. doi:10.3390/s16081292.

Kunert, K. J., Vorster, B. J., Fenta, B. A., Kibido, T., Giuseppe, D., and Foyer, C. H. (2016). Drought stress responses in soybean roots and nodules. *Front. Plant Sci.* 7, 1015. doi:10.3389/fpls.2016.01015.

Larcher, W. (2000). Temperature stress and survival ability of mediterranean sclerophyllous plants. *Plant Biosyst.* 134, 279–295. doi:10.1080/11263500012331350455.

Le, D. T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Ham, L. H., et al. (2012). Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by Genome-Wide Transcriptome analysis. *PLoS One* 7, e495. doi:10.1371/journal.pone.0049522.

Lehnert, H., Serfling, A., Friedt, W., and Ordon, F. (2018). Genome-wide association studies reveal genomic regions associated with the response of wheat (*Triticum aestivum* L.) to mycorrhizae under drought stress conditions. *Front. Plant Sci.* 9, 1728. doi:10.3389/fpls.2018.01728.

Levis, S., Strickman-Stein, N., Ganjei-Azar, P., Xu, P., Doerge, D. R., and Krischer, J. (2011). Soy isoflavones in the prevention of menopausal bone loss and menopausal symptoms: A randomized, double-blind trial. *Arch. Intern. Med.* 171, 1363–1369. doi:10.1001/archinternmed.2011.330.

Li, F., Chen, B., Xu, K., Gao, G., Yan, G., Qiao, J., et al. (2016a). A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*). *Plant Sci.* 242, 169–177. doi:10.1016/j.plantsci.2015.05.012.

Li, H., Yang, Y., Zhang, H., Chu, S., Zhang, X., Yin, D., et al. (2016b). A genetic relationship between phosphorus efficiency and photosynthetic traits in soybean as revealed by QTL analysis using a high-density genetic map. *Front. Plant Sci.* 7, 1–16. doi:10.3389/fpls.2016.00924.

Li, X., Zhou, Y., Bu, Y., Wang, X., Zhang, Y., Guo, N., et al. (2021). Genome-wide association analysis for yield-related traits at the R6 stage in a Chinese soybean mini core collection. *Genes and Genomics* 43, 897–912. doi:10.1007/s13258-021-01109-9.

Li, Y. H., Li, W., Zhang, C., Yang, L., Chang, R. Z., Gaut, B. S., et al. (2010). Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. *New Phytol.* 188, 242–253. doi:10.1111/j.1469-8137.2010.03344.x.

Li, Y., Li, H., Li, Y., and Zhang, S. (2017). Improving water-use efficiency by decreasing stomatal conductance and transpiration rate to maintain higher ear photosynthetic rate in drought-resistant wheat. *Crop J.* 5, 231–239. doi:10.1016/j.cj.2017.01.001.

Liu, F., Andersen, M. N., Jacobsen, S. E., and Jensen, C. R. (2005). Stomatal control and water use efficiency of soybean (*Glycine max* L. Merr.) during progressive soil drying. *Environ. Exp. Bot.* 54, 33–40. doi:10.1016/j.envexpbot.2004.05.002.

Liu, F., Jensen, C. R., and Andersen, M. N. (2003). Hydraulic and chemical signals in the control of leaf expansion and stomatal conductance in soybean exposed to drought stress. *Funct. Plant Biol.* 30, 65–73. doi:10.1071/FP02170.

Liu, K. (1997). *Soybeans: chemistry, technology and utilization'*. International Thomson Publishing Asia.

Liu, K. S. (2008). "Food use of whole soybeans," in *Soybeans: Chemistry, Production, Processing, and Utilization* (AOCS Press), 441–481. doi:10.1016/B978-1-893997-64-6.50017-2.

Liu, S., Zhang, M., Feng, F., and Tian, Z. (2020). Toward a "Green Revolution" for Soybean. *Mol. Plant* 13, 688–697. doi:10.1016/j.molp.2020.03.002.

Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS One* 12, e1005767. doi:10.1186/1471-2156-13-100.

Liyanage, D. K., Chathuranga, I., Mori, B. A., and Thilakarathna, M. S. (2022). A simple , semi-automated , gravimetric method to simulate drought stress on plants. *Agronomy* 12, 349.

Lü, H., Yang, Y., Li, H., Liu, Q., Zhang, J., Yin, J., et al. (2018). Genome-wide association studies of photosynthetic traits related to phosphorus efficiency in soybean. *Front. Plant Sci.* 9. doi:10.3389/fpls.2018.01226.

Ma, J., Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., et al. (2022). Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4 .0 , r10285). *Geosci. Model Dev.* 15, 815–839.

Mak, M., Babla, M., Xu, S. C., O'Carrigan, A., Liu, X. H., Gong, Y. M., et al. (2014). Leaf mesophyll K^+, H+ and Ca^{2+} fluxes are involved in drought-induced decrease in photosynthesis and stomatal closure in soybean. *Environ. Exp. Bot.* 98, 1–12. doi:10.1016/j.envexpbot.2013.10.003.

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

Marchin, R. M., Ossola, A., Leishman, M. R., and Ellsworth, D. S. (2020). A simple method for simulating drought effects on plants. *Front. Plant Sci.* 10, 1715. doi:10.3389/fpls.2019.01715.

Marcolino-Gomes, J., Rodrigues, F. A., Fuganti-Pagliarini, R., Bendix, C., Nakayama, T. J., Celaya, B., et al. (2014). Diurnal oscillations of soybean circadian clock and drought responsive genes. *PLoS One* 9. doi:10.1371/journal.pone.0086402.

Mateos-Aparicio, I., Cuenca, A. R., Villanueva-Suárez M.J., and Zapata-Revilla, M. A. (2008). Soybean , a promising health source. *Nutr. Hosp.* 23, 305–312.

McCauley, D., Levin, A., and Nackley, L. (2021). Reviewing mini-lysimeter controlled irrigation in container crop systems. *Hort Technol.*, 1–8. doi:10.21273/horttech04826-21.

Medina, V., and Gilbert, M. E. (2016). Physiological trade-offs of stomatal closure under high evaporative gradients in field grown soybean. *Funct. Plant Biol.* 43, 40–51. doi:10.1071/FP15304.

Meena, H. M., R.K.Singh, and P.Santra (2015). Design and development of a load-cell based cost effective mini- lysimeter. *J. Agric. Phys.* 15, 1–6.

Minamisawa, K., Isawa, T., Nakatsuka, Y., and Ichikawa, N. (1998). New *Bradyrhizobium japonicum* strains that possess high copy numbers of the repeated sequence RSα. *Appl. Environ. Microbiol.* 64, 1845–1851. doi:10.1128/aem.64.5.1845-1851.1998.

Misra, R. K., Padhi, J., and Payero, J. O. (2011). A calibration procedure for load cells to improve accuracy of mini-lysimeters in monitoring evapotranspiration. *J. Hydrol.* 406, 113–118. doi:10.1016/j.jhydrol.2011.06.009.

Mustofa, A., Zubaidah, S., and Kuswantoro, H. (2021). Correlation and path analysis on yield and yield components in segregating populations. *Genetika* 53, 157–166. doi:10.2298/GENSR2101157R.

Mutava, R. N., Prince, S. J. K., Syed, N. H., Song, L., Valliyodan, B., Chen, W., et al. (2015). Understanding abiotic stress tolerance mechanisms in soybean: a comparative evaluation of soybean response to drought and flooding stress. *Plant Physiol. Biochem.* 86, 109–120.

Mwenye, O. J., van Rensburg, L., van Biljon, A., and van der Merwe, R. (2016). The role of proline and root traits on selection for drought-stress tolerance in soybeans: a review. *South African J. Plant Soil* 33, 245–256. doi:10.1080/02571862.2016.1148786.

Naderi, A., Naseri, R., Fathi, A., Bahamin, S., and Maleki, R. (2013). Physiological performance of soybean cultivars under drought stress. *Bull. Environ. Pharmacol. Life Sci.* 2, 38–44.

Naya, L., Ladrera, R., Ramos, J., González, E. M., Arrese-Igor, C., Minchin, F. R., et al. (2007). The response of carbon metabolism and antioxidant defenses of alfalfa nodules to drought stress and to the subsequent recovery of plants. *Plant Physiol.* 144, 1104–1114.

doi:10.1104/pp.107.099648.

Neo, H. H., and Layzell, D. B. (1997a). Phloem glutamine and the regulation of O_2 diffusion in legume nodules. *Plant Physiol.* 113, 259–267. doi:10.1104/pp.113.1.259.

Neo, H. H., and Layzell, D. B. (1997b). Phloem glutamine and the regulation of oxygen , diffusion in legume nodules '. *Plant Physiol.* 113, 259–267.

Nouri, M. Z., Moumeni, A., and Komatsu, S. (2015). Abiotic stresses: Insight into gene regulation and protein expression in photosynthetic pathways of plants. *Int. J. Mol. Sci.* 16, 20392–20416. doi:10.3390/ijms160920392.

Ohashi, Y., Nakayama, N., Saneoka, H., and Fujita, K. (2006). Effects of drought stress on photosynthetic gas exchange, chlorophyll fluorescence and stem diameter of soybean plants. *Biol. Plant.* 50, 138–141. doi:10.1007/s10535-005-0089-3.

Ortiz, D., Litvin, A. G., and Salas Fernandez, M. G. (2018). A cost-effective and customizable automated irrigation system for precise high-throughput phenotyping in drought stress studies. *PLoS One* 13, e0198546. doi:10.1371/journal.pone.0198546.

Oya, T., Nepomuceno, A. L., Neumaier, N., Farias, J. R. B., Tobita, S., and Ito, O. (2004). Drought tolerance characteristics of Brazilian soybean cultivars - evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field. *Plant Prod. Sci.* 7, 129–137. doi:10.1626/pps.7.129.

Pantalião, G. F., Narciso, M., Guimarães, C., Castro, A., Colombari, J. M., Breseghello, F., et al. (2016). Genome wide association study (GWAS) for grain yield in rice cultivated under water deficit. *Genetica* 144, 651–664. doi:10.1007/s10709-016-9932-z.

Patinni, I. R. G., de Andrade, C. A., Campos, C. N. S., Teodoro, L. P. R., Andrade, S. M., Roque, C. G., et al. (2020). Agronomic performance and water-use efficiency of F3 soybean populations grown under contrasting base saturation. *J. Agron. Crop Sci.* 206, 806–814. doi:10.1111/jac.12413.

Pejić, B., Maksimović, L., Cimpeanu, S., Bucur, D., Milić, S., and Ćupina, B. (2011). Response of soybean to water stress at specific growth stages. *J. Food, Agric. Environ.* 9, 280–284.

Phang, T.-H., Li, M.-W., Cheng, C.-C., Wong, F.-L., Chan, C., and Lam, H.-M. (2011). Molecular Responses to Osmotic Stresses in Soybean. *Soybean - Mol. Asp. Breed.* doi:10.5772/15155.

Pitumpe Arachchige, P. S., Rosso, L. H. M., Hansel, F. D., Ramundo, B., Torres, A. R., Asebedo, R., et al. (2020). Temporal biological nitrogen fixation pattern in soybean inoculated with *Bradyrhizobium*. *Agrosystems, Geosci. Environ.* 3, 1–11. doi:10.1002/agg2.20079.

Priyanatha, C., Torkamaneh, D., and Rajcan, I. (2022). Genome-Wide Association Study of soybean germplasm derived from Canadian × Chinese crosses to mine for novel alleles to improve seed yield and seed quality traits. *Front. Plant Sci.* 13. doi:10.3389/fpls.2022.866300.

Purcell, L. C., Salmeron, M., and Ashlock, L. (2014). "Soybean growth and development," in *Arkansas soybean production handbook* (University of Arkansas), 1–8.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795.

Quero, G., Simondi, S., Ceretta, S., Otero, Á., Garaycochea, S., Fernández, S., et al. (2021). An integrative analysis of yield stability for a GWAS in a small soybean breeding population. *Crop Sci.* 61, 1903–1914. doi:10.1002/csc2.20490.

R Core Team (2021). R: a language and environment for statistical computing. Available at: https://www.r-project.org/.

Raj, A., Stephens, M., and Pritchard, J. K. (2014). FastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. doi:10.1534/genetics.114.164350.

Ravelombola, W., Qin, J., Shi, A., Song, Q., Yuan, J., Wang, F., et al. (2021). Genome-wide association study and genomic selection for yield and related traits in soybean. *PLoS One* 16, e0255761. doi:10.1371/journal.pone.0255761.

Rebolledo, M. C., Dingkuhn, M., Courtois, B., Gibon, Y., Clément-Vidal, A., Cruz, D. F., et al. (2015). Phenotypic and genetic dissection of component traits for early vigour in rice using plant growth modelling, sugar content analyses and association mapping. *J. Exp. Bot.* 66, 5555–5566. doi:10.1093/jxb/erv258.

Robinson, D. A., Jones, S. B., Wraith, J. M., Or, D., and Friedman, S. P. (2003). A Review of advances in dielectric and electrical conductivity measurement in soils using time domain reflectometry. *Vadose Zo. J.* 2, 444–475. doi:10.2136/vzj2003.4440.

Rodríguez-Navarro, D. N., Margaret Oliver, I., Albareda Contreras, M., and Ruiz-Sainz, J. E. (2011). Soybean interactions with soil microbes, agronomical and molecular aspects. *Agron. Sustain. Dev.* 31, 173–190. doi:10.1051/agro/2010023.

Rutten, P. J., and Poole, P. S. (2019). *Oxygen regulatory mechanisms of nitrogen fixation in rhizobia*. Elsevier Ltd doi:10.1016/bs.ampbs.2019.08.001.

Saeidi, M., and Abdoli, M. (2015). Effect of drought stress during grain filling on yield and its components, gas exchange variables, and some physiological traits of wheat cultivars. *J. Agric. Sci. Technol.* 17, 885–898.

Santachiara, G., Salvagiotti, F., and Rotundo, J. L. (2019). Nutritional and environmental effects on biological nitrogen fixation in soybean: A meta-analysis. *F. Crop. Res.* 240, 106–115. doi:10.1016/j.fcr.2019.05.006.

Sebastián, D. I., Fernando, F. D., Raúl, D. G., and Gabriela, G. M. (2020). Overexpression of *Arabidopsis* aspartic protease APA1 gene confers drought tolerance. *Plant Sci.* 292, 110406. doi:10.1016/j.plantsci.2020.110406.

Seck, W., Torkamaneh, D., and Belzile, F. (2020). Comprehensive genome-wide association analysis reveals the genetic basis of root system srchitecture in soybean. *Front. Plant Sci.* 11, 590740. doi:10.3389/fpls.2020.590740.

Serraj, R. (2003). Effects of drought stress on legume symbiotic nitrogen fixation: physiological mechanisms. *Indian J. Exp. Biol.* 41, 1136–1141.

Serraj, R. S., Adez, V. V, and Inclair, T. R. S. (2001). Feedback regulation of symbiotic N_2 fixation under drought stress. *Agronomie* 21, 621–626.

Serraj, R., Sinclair, T. R., and Purcell, L. C. (1999a). Symbiotic N_2 fixation response to drought. *J. Exp. Bot.* 50, 143–155. doi:10.1093/jxb/50.331.143.

Serraj, R., Sinclair, T. R., and Purcell, L. C. (1999b). Symbiotic N_2 fixation response to drought. *J. Exp. Bot.* 50, 143–155. doi:10.1093/jxb/50.331.143.

Sheffield, J., and Wood, E. F. (2008). Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Clim. Dyn.* 31, 79–105. doi:10.1007/s00382-007-0340-z.

Sheffield, J., Wood, E. F., and Roderick, M. L. (2012). Little change in global drought over the past 60 years. *Nature* 491, 435–438. doi:10.1038/nature11575.

Shi, W. Y., Du, Y. T., Ma, J., Min, D. H., Jin, L. G., Chen, J., et al. (2018). The WRKY transcription factor *GmWRKY12* confers drought and salt tolerance in soybean. *Int. J. Mol. Sci.* 19, 4087. doi:10.3390/ijms19124087.

Shibata, M., Breuer, C., Kawamura, A., Clark, N. M., Rymen, B., Braidwood, L., et al. (2018). GTL1 and DF1 regulate root hair growth through transcriptional repression of ROOT HAIR DEFECTIVE 6-LIKE 4 in Arabidopsis. *Dev.* 145, 1–11. doi:10.1242/dev.159707.

Shirinbayan, S., Khosravi, H., and Malakouti, M. J. (2019). Alleviation of drought stress in maize (*Zea mays*) by inoculation with *Azotobacter* strains isolated from semi-arid regions. *Appl. Soil Ecol.* 133, 138–145. doi:10.1016/j.apsoil.2018.09.015.

Shurtleff, W., and Aoyagi, A. (2010). *History of soybeans and soyfood in Canada (1831-2010): Extensively annotated bibliography and sourcebook*. Soyinfo Center,Lafayette, CA.

Siddiqui, M. H., Al-Khaishany, M. Y., Al-Qutami, M. A., Al-Whaibi, M. H., Grover, A., Ali, H. M., et al. (2015a). Response of different genotypes of faba bean plant to drought stress. *Int. J. Mol. Sci.* 16, 10214–10227. doi:10.3390/ijms160510214.

Siddiqui, Z. S., Cho, J. Il, Kwon, T. R., Ahn, B. O., Lee, K. S., Jeong, M. J., et al. (2014). Physiological mechanism of drought tolerance in transgenic rice plants expressing *Capsicum annuum* methionine sulfoxide reductase B2 (CaMsrB2) gene. *Acta Physiol. Plant.* 36, 1143–1153. doi:10.1007/s11738-014-1489-9.

Siddiqui, Z. S., Cho, J. Il, Park, D. B., Lee, G. S., Ryu, T. H., Shahid, H., et al. (2015b). Field assessment of CaMsrB2 transgenic lines in a drought stress environment. *Turk. J. Botany* 39, 973–981. doi:10.3906/bot-1502-18.

Silva, A. F., Sediyama, T., Silva, F. C. S., Bezerra, A. R. G., and Ferreira, L. V. (2015). Correlation and path analysis of yield components in soybean varieties. *Turkish J. F. Crop.*

10, 177–179.

Silvente, S., Sobolev, A. P., and Lara, M. (2012). Metabolite adjustments in drought tolerant and sensitive soybean genotypes in response to water stress. *PLoS One* 7, e38554. doi:10.1371/journal.pone.0038554.

Singh, G. (2010). *The soybean: botany, production and uses.* CABI doi:10.1016/s0016-0032(32)91432-x.

Singh, J., and Verma, P. K. (2021). NSP1 allies with GSK3 to inhibit nodule symbiosis. *Trends Plant Sci.* 26, 999–1001. doi:10.1016/j.tplants.2021.07.001.

Son, S., Oh, C. J., Bae, J. H., Lee, H., and An, C. S. (2015). GmREM1.1 and GmREM2.1, which encode the remorin proteins in soybean, have distinct roles during root nodule development. *J. Plant Biol.* 58, 17–25. doi:10.1007/s12374-014-0273-0.

Sonah, H., O'Donoughue, L., Cober, E., Rajcan, I., and Belzile, F. (2015). Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. *Plant Biotechnol. J.* 13, 211–221. doi:10.1111/pbi.12249.

Specht, J. E., and Williams (2022). Breeding for drought and heat resistance: Prerequisites and examples. in *World Soybean Research Conference III: Proceedings* (CRC Press), 468–475.

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

Steketee, C. J., Sinclair, T. R., Riar, M. K., Schapaugh, W. T., and Li, Z. (2019). Unraveling the genetic architecture for carbon and nitrogen related traits and leaf hydraulic conductance in soybean using genome-wide association analyses. *BMC Genomics* 20, 811. doi:10.1186/s12864-019-6170-7.

Sulieman, S., and Tran, L. S. P. (2016). *Legume nitrogen fixation in a changing environment: achievements and challenges*. doi:10.1007/978-3-319-06212-9.

Susha Lekshmi, S. U., Singh, D. N., and Shojaei Baghini, M. (2014). A critical review of soil moisture measurement. *J. Int. Meas. Confed.* 54, 92–105. doi:10.1016/j.measurement.2014.04.007.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., et al. (2002). Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J.* 29, 417–426. doi:10.1046/j.0960-7412.2001.01227.x.

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484. doi:10.1038/s41576-019-0127-1.

Tarawneh, R. A., Alqudah, A. M., Nagel, M., and Börner, A. (2020). Genome-wide association mapping reveals putative candidate genes for drought tolerance in barley. *Environ. Exp. Bot.* 180, 104237. doi:10.1016/j.envexpbot.2020.104237.

Taylor, H. M., Burnett, E., and Booth, G. D. (1978). Taproot elongation rates of soybeans.

Zeitschrift fuer Acker und Pflanzenbau 146, 33–39.

Thilakarathna, M. S., McElroy, M. S., Chapagain, T., Papadopoulos, Y. A., and Raizada, M. N. (2016a). Belowground nitrogen transfer from legumes to non-legumes under managed herbaceous cropping systems. A review. *Agron. Sustain. Dev.* 36. doi:10.1007/s13593-016-0396-4.

Thilakarathna, M. S., Moroz, N., and Raizada, M. N. (2017). A biosensor-based leaf punch assay for glutamine correlates to symbiotic nitrogen fixation measurements in legumes to permit rapid screening of *rhizobia* inoculants under controlled conditions. *Front. Plant Sci.* 8, 1714. doi:10.3389/fpls.2017.01714.

Thilakarathna, M. S., Papadopoulos, Y. A., Rodd, A. V., Grimmett, M., Fillmore, S. A. E., Crouse, M., et al. (2016b). Nitrogen fixation and transfer of red clover genotypes under legume–grass forage based production systems. *Nutr. Cycl. Agroecosystems* 106, 233–247. doi:10.1007/s10705-016-9802-1.

Thilakarathna, M. S., and Raizada, M. N. (2017). A meta-analysis of the effectiveness of diverse rhizobia inoculants on soybean traits under field conditions. *Soil Biol. Biochem.* 105, 177–196. doi:10.1016/j.soilbio.2016.11.022.

Thilakarathna, M. S., Torkamaneh, D., Bruce, R. W., Rajcan, I., Chu, G., Grainger, C. M., et al. (2021). Testing whether pre-pod-fill symbiotic nitrogen fixation in soybean is subject to drift or selection over 100 years of soybean breeding. *Front. Agron.* 3, 725813. doi:10.3389/fagro.2021.725813.

Tombesi, S., Nardini, A., Frioni, T., Soccolini, M., Zadra, C., Farinelli, D., et al. (2015). Stomatal closure is induced by hydraulic signals and maintained by ABA in drought-stressed grapevine. *Nature* 5, 12449. doi:10.1038/srep12449.

Torkamaneh, D., Chalifour, F. P., Beauchamp, C. J., Agrama, H., Boahen, S., Maaroufi, H., et al. (2020). Genome-wide association analyses reveal the genetic basis of biomass accumulation under symbiotic nitrogen fixation in African soybean. *Theor. Appl. Genet.* 133, 665–676. doi:10.1007/s00122-019-03499-7.

Torkamaneh, D., Laroche, J., Tardivel, A., O'Donoughue, L., Cober, E., Rajcan, I., et al. (2018). Comprehensive description of genomewide nucleotide and structural variation in short-season soya bean. *Plant Biotechnol. J.* 16, 749–759. doi:10.1111/pbi.12825.

Turner, N. C. (2019). Imposing and maintaining soil water deficits in drought studies in pots. *Plant Soil* 439, 45–55. doi:10.1007/s11104-018-3893-1.

Turner, N. C., Wright, G. C., and Siddique, K. H. M. (2001). *Adaptation of grain legumes (pulses) to water-limited environments.*

Turuspekov, Y., Baibulatova, A., Yermekbayev, K., Tokhetova, L., Chudinov, V., Sereda, G., et al. (2017). GWAS for plant growth stages and yield components in spring wheat (*Triticum aestivum* L.) harvested in three regions of Kazakhstan. *BMC Plant Biol.* 17. doi:10.1186/s12870-017-1131-2.

Tuteja, N., Gill, S. S., and Tuteja, R. (2011). Plant responses to abiotic stresses : shedding light on salt , drought , cold and heavy metals. *Omi. Plant Abiotic Stress Toler.* 1, 39–64.

United Nations (2019). Department of economics and social affairs. https://www.un.org/development/desa/en/news/population/world-population-prospects-2019.html Malinda Thilakarathna.

Vadez, V., Sinclair, T. R., and Serraj, R. (2000). Asparagine and ureide accumulation in nodules and shoots as feedback inhibitors of N_2 fixation in soybean. *Physiol. Plant.* 110, 215–223. doi:10.1034/j.1399-3054.2000.110211.x.

Valentine, A. J., A., B. V., and Yun, K. (2011). "Legume nitrogen fixation and soil abiotic stress : from physiology to genomics and beyond," in *Nitrogen Metabolism in Plants in the Post-genomic Era* (Blackwell Publishing Ltd), 207–248.

Vance, C. P., and Gantt, S. J. (1992). Control of nitrogen and carbon metabolism in root nodules. *Physiol. Plant.* 85, 266–274.

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., et al. (2017). ePlant: Visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *Plant Cell* 29, 1806–1821. doi:10.1105/tpc.17.00073.

Walden-Coleman, A. E., Rajcan, I., and Earl, H. J. (2013). Dark-adapted leaf conductance, but not minimum leaf conductance, predicts water use efficiency of soybean (*Glycine max* L. Merr.). *Can. J. Plant Sci.* 93, 13–22. doi:10.4141/CJPS2012-178.

Wang, J., Chu, S., Zhang, H., Zhu, Y., Cheng, H., and Yu, D. (2016a). Development and application of a novel genome-wide SNP array reveals domestication history in soybean. *Nature* 6, 20728. doi:10.1038/srep20728.

Wang, L., Yang, Y., Zhang, S., Che, Z., Yuan, W., and Yu, D. (2020). GWAS reveals two novel loci for photosynthesis-related traits in soybean. *Mol. Genet. Genomics* 295, 705–716. doi:10.1007/s00438-020-01661-1.

Wang, W., Wang, C., Pan, D., Zhang, Y., Luo, B., and Ji, J. (2018). Effects of drought stress on photosynthesis and chlorophyll fluorescence images of soybean (*Glycine max*) seedlings. *Int. J. Agric. Biol. Eng.* 11, 196–201. doi:10.25165/j.ijabe.20181102.3390.

Wang, X., Oh, M. W., and Komatsu, S. (2016b). Characterization of S-adenosylmethionine synthetases in soybean under flooding and drought stresses. *Biol. Plant.* 60, 269–278. doi:10.1007/s10535-016-0586-6.

Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS One* 14, 0208217. doi:10.1371/journal.pone.0208217.

Wasaya, A., Zhang, X., Fang, Q., and Yan, Z. (2018). Root phenotyping for drought tolerance: A review. *Agronomy* 8, 241. doi:10.3390/agronomy8110241.

Wei, Y., Jin, J., Jiang, S., Ning, S., and Liu, L. (2018). Quantitative response of soybean development and yield to drought stress during different growth stages in the Huaibei plain, China. *Agronomy* 8, 97. doi:10.3390/agronomy8070097.

Wilhite, D. A. (2000). Drought as a natural hazard: concepts and definitions. *Drought Mitig. Cent. Fac. Publ.*, 1–21.

Wilker, J., Navabi, A., Rajcan, I., Marsolais, F., Hill, B., Torkamaneh, D., et al. (2019). Agronomic performance and nitrogen fixation of heirloom and conventional dry bean varieties under low-nitrogen field conditions. *Front. Plant Sci.* 10, 952. doi:10.3389/fpls.2019.00952.

Wilkinson, S., and Davies, W. J. (2002). ABA-based chemical signalling : the co-ordination of responses to stress in plants. *Plant. Cell Environ.* 25, 195–210.

Wright, D. E. J., Cline, J. A., and Earl, H. J. (2019). Physiological responses of four apple (*Malus* × *domestica* borkh.) rootstock genotypes to soil water deficits. *Can. J. Plant Sci.* 99, 510–524. doi:10.1139/cjps-2018-0276.

Wu, C., Ma, Q., Yam, K. M., Cheung, M. Y., Xu, Y., Han, T., et al. (2006). In situ expression of the GmNMH7 gene is photoperiod-dependent in a unique soybean (*Glycine max* [L.] Merr.) flowering reversion system. *Planta* 223, 725–735. doi:10.1007/s00425-005-0130-y.

Wu, T. ting, Li, J. yu, Wu, C. xiang, Sun, S., Mao, T. ting, Jiang, B. jun, et al. (2015). Analysis of the independent- and interactive-photo-thermal effects on soybean flowering. *J. Integr. Agric.* 14, 622–632. doi:10.1016/S2095-3119(14)60856-X.

Yan, L., Hofmann, N., Li, S., Ferreira, M. E., Song, B., Jiang, G., et al. (2017). Identification of QTL with large effect on seed weight in a selective population of soybean with genome-wide association and fixation index analyses. *BMC Genomics* 18, 529. doi:10.1186/s12864-017-3922-0.

Yang, W., Wu, T., Zhang, X., Song, W., Xu, C., Sun, S., et al. (2019). Critical photoperiod measurement of soybean genotypes in different maturity groups. *Crop Sci.* 59, 2055–2061. doi:10.2135/cropsci2019.03.0170.

Yang, Y., Wang, L., Zhang, D., Cheng, H., Wang, Q., Yang, H., et al. (2020). GWAS identifies two novel loci for photosynthetic traits related to phosphorus efficiency in soybean. *Mol. Breed.* 40. doi:10.1007/s11032-020-01112-0.

Yang, Y., Zhu, X., Cui, R., Wang, R., Li, H., Wang, J., et al. (2021). Identification of soybean phosphorous efficiency QTLs and genes using chlorophyll fluorescence parameters through GWAS and RNA-seq. *Planta* 254, 1–13. doi:10.1007/s00425-021-03760-8.

Yanti, E., and Dafriani, P. (2021). The effect of soybean power on blood sugar and cholesterol levels in type 2 diabetes mellitus patients in the work area of the Puskesmas Andalas Padang. *Adv. Heal. Sci. Res.* 39, 340–344.

Ye, H., Song, L., Schapaugh, W. T., Ali, M. L., Sinclair, T. R., Riar, M. K., et al. (2020). The importance of slow canopy wilting in drought tolerance in soybean. *J. Exp. Bot.* 71, 642–652. doi:10.1093/jxb/erz150.

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, Proteomics Bioinforma.* 19, 619–628. doi:10.1016/j.gpb.2020.10.007.

Yuan, S., Li, X., Li, R., Wang, L., Zhang, C., Chen, L., et al. (2018). Genome-wide identification and classification of soybean C2H2 Zinc Finger proteins and their expression analysis in legume-rhizobium symbiosis. *Front. Microbiol.* 9, 126. doi:10.3389/fmicb.2018.00126.

Zapata, F., Danso, F., Hardarson, G., and Fried, M. (1987). Nitrogen fixation and translocation in field-grown fababean 1. *Agron. J.* 79, 505–509. doi:10.2134/agronj1987.00021962007900030020x.

Zargar, S. M., Gupta, N., Nazir, M., Mahajan, R., Malik, F. A., Sofi, N. R., et al. (2017). Impact of drought on photosynthesis: molecular perspective. *Plant Gene* 11, 154–159. doi:10.1016/j.plgene.2017.04.003.

Zhang, J., Song, Q., Cregan, P. B., and Jiang, G. L. (2016). Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). *Theor. Appl. Genet.* 129, 117–130. doi:10.1007/s00122-015-2614-x.

Zhang, J., Song, Q., Cregan, P. B., Nelson, R. L., Wang, X., Wu, J., et al. (2015). Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics* 16, 1–11. doi:10.1186/s12864-015-1441-4.

Zhang, J., Subramanian, S., Stacey, G., and Yu, O. (2009). Flavones and flavonols play distinct critical roles during nodulation of *Medicago truncatula* by *Sinorhizobium meliloti*. *Plant J.* 57, 171–183. doi:10.1111/j.1365-313X.2008.03676.x.

Zhang, L. X., Kyei-Boahen, S., Zhang, J., Zhang, M. H., Freeland, T. B., Watson, C. E., et al. (2007). Modifications of optimum adaptation zones for soybean maturity groups in the USA. *Crop Manag.* 6, 1–11. doi:10.1094/cm-2007-0927-01-rs.

Zhang, T., Wu, T., Wang, L., Jiang, B., Zhen, C., Yuan, S., et al. (2019). A combined linkage and GWAS analysis identifies QTLs linked to soybean seed protein and oil content. *Int. J. Mol. Sci.* 20, 5915. doi:10.3390/ijms20235915.

Zhang, W., Xu, W., Zhang, H., Liu, X., Cui, X., Li, S., et al. (2021). Comparative selective signature analysis and high-resolution GWAS reveal a new candidate gene controlling seed weight in soybean. *Theor. Appl. Genet.* 134, 1329–1341. doi:10.1007/s00122-021-03774-6.

Zhang, Z., Ober, U., Erbe, M., Zhang, H., Gao, N., He, J., et al. (2014). Improving the accuracy of whole genome prediction for complex traits using the results of genome wide association studies. *PLoS One* 9, e93017. doi:10.1371/journal.pone.0093017.

Zhao, J. Y., Lu, Z. W., Sun, Y., Fang, Z. W., Chen, J., Zhou, Y. Bin, et al. (2020). The ankyrin-repeat gene *GmANK114* confers drought and salt tolerance in *Arabidopsis* and soybean. *Front. Plant Sci.* 11, 584167. doi:10.3389/fpls.2020.584167.

Zhao, Z., and Assmann, S. M. (2011). The glycolytic enzyme, phosphoglycerate mutase, has critical roles in stomatal movement, vegetative growth, and pollen production in *Arabidopsis thaliana*. *J. Exp. Bot.* 62, 5179–5189. doi:10.1093/jxb/err223.

Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and prospects of association mapping in plants. *Plant Genome* 1, 5–20. doi:10.3835/plantgenome2008.02.0089.

Appendix

Appendix - A

File A.1: Original source of different components of the irrigation system.

- Breadboard: 400 tie Points Interlocking Solderless Breadboard Mini Universal Test Protoboard DIY Bread Board Bus Test Circuit Board|breadboard point solderless|breadboard solderlessbreadboard 400 AliExpress.
- Compact Wire wiring connector: 222 412 222 413 222 415 Compact Wire Wiring Connector Conductor Terminal Block With Lever 0.08 2.5mm2 214 218 SPL 2 3|terminal blocks wire connector|block terminal connectorblock connector - AliExpress.
- Flexible Hose: 8mm ID x 10mm OD Food Grade Flexible Hose Silicone Tubing Tube 1Meter|Cable Sleeves| - AliExpress.
- Jumper wire (2020). Cable Dupont,Jumper Wire Dupont,30CM Male to Male + Female to Male + Female to Female Jumper Copper Wire Dupont Cable DIY KIT|Wires & Cables| - AliExpress.
- Power Adapter: AC 110 240V DC 3V 5V 6V 9V 12V 15V 24V 0.5A 1A 2A 3A 5A 6A 8A Universal Power Adapter Supply Charger adapter Eu Us for LED light|AC/DC Adapters| - AliExpress.
- Pressure sensor and HX711 module: 5KG 10KG 20KG pressure sensor + HX711AD module weighing electronics accessories feeder lines and DuPont|pressure sensor|sensor | sensorweighing sensor - AliExpress.
- Relay Module: 1 Channel 5V Relay Module for arduino 1 Channel relay KY 019 For PIC AVR DSP ARM for Arduino|Integrated Circuits| - AliExpress.
- Submersible Water Pump: Ultra quiet Mini DC12V 4.2W Micro Brushless Water Oil Pump Waterproof Submersible Aquarium Circulating 240L/H Lift 300cm|pump waterproof submersible|water oil pumpbrushless water oil pump - AliExpress.
- USB Data Cable: 1m/1.8m/3m/5m High Speed Black/transparent 480Mpbs USB 2.0 Printer Data Cable Scanner Cord For Xerox Canon HP Brother Arduino|Data Cables| - AliExpress.

File A.2: C++ Code for Arduino

```
#include <HX711_ADC.h>
#include <EEPROM.h>
//pins:
const int HX711_dout_1 = 4; //mcu > HX711 no 1 dout pin
const int HX711_sck_1 = 5; //mcu > HX711 no 1 sck pin
const int HX711_dout_2 = 2; //mcu > HX711 no 2 dout pin
const int HX711_sck_2 = 3; //mcu > HX711 no 2 sck pin
//HX711 constructor (dout pin, sck pin)
HX711_ADC LoadCell_1(HX711_dout_1, HX711_sck_1); //HX711 1
HX711_ADC LoadCell_2(HX711_dout_2, HX711_sck_2); //HX711 2
const int calVal_eepromAdress_1 = 0; // eeprom adress for calibration value load cell 1 (4 bytes)
const int calVal_eepromAdress_2 = 4; // eeprom adress for calibration value load cell 2 (4 bytes)
float Wlimit = 5363.36; // Water limit for Well watered pots
float Dlimit = 4025.96; // Water limit for Drought watered pots
float potw ;
long t;
float k=0;
float u=0;
String pot;
int m =0;
float p=0;
float w=0;
float e=20;
float tdiff=0;
float tendd;
float tini;
int rrr = 0;
float timet1 ;
```

```
float timet2;
float timettdif;
float pp1 = 0;
float ppdif = 0;
int nb = 0;
float kk;
int tankt1;
int tankt0;
int tanktd=0;
int lengthh;
String type;
int len;
float TOTW = 0;
void setup() {
pinMode(7,OUTPUT);
Serial.begin(57600); delay(10);
Serial.println();
Serial.println("Starting...");
delay(5000);
float calibrationValue_1; // calibration value load cell 1
float calibrationValue_2; // calibration value load cell 2
calibrationValue_1 = -107; // Calibration value for loadcell 1
calibrationValue_2 = -107; // Calibration value for loadcell 2
LoadCell_1.begin();
LoadCell_2.begin();
long stabilizingtime = 2000; // tare preciscion can be improved by adding a few seconds of
stabilizing
time
boolean _tare = true; //set this to false if you don't want tare to be performed in the next step
byte loadcell_1_rdy = 0;
byte loadcell_2_rdy = 0;
```

```
while ((loadcell_1_rdy + loadcell_2_rdy) < 2) { //run startup, stabilization and tare, both
modules
simultaniously
if (!loadcell_1_rdy) loadcell_1_rdy = LoadCell_1.startMultiple(stabilizingtime, _tare);
if (!loadcell_2_rdy) loadcell_2_rdy = LoadCell_2.startMultiple(stabilizingtime, _tare);
}
if (LoadCell_1.getTareTimeoutFlag()) {
Serial.println("Timeout, check MCU>HX711 no.1 wiring and pin designations");
}
if (LoadCell_2.getTareTimeoutFlag()) {
Serial.println("Timeout, check MCU>HX711 no.2 wiring and pin designations");
}
LoadCell_1.setCalFactor(calibrationValue_1); // user set calibration value (float)
LoadCell_2.setCalFactor(calibrationValue_2); // user set calibration value (float)
Serial.println("Startup is complete");
int kkk = 1;
// Below code will tare the loadcell until it reaches zero.
while (kkk != 0){
// if (Serial.available() > 0) {
float i;
//char inByte = Serial.read();
//if (inByte == 't') {
LoadCell_1.tareNoDelay();
LoadCell_2.tareNoDelay();
//check if last tare operation is complete
if (LoadCell_1.getTareStatus() == true) {
Serial.println("Tare load cell 1 complete");
}
if (LoadCell_2.getTareStatus() == true) {
Serial.println("Tare load cell 2 complete");
}
```

```
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float a = LoadCell_1.getData();
float b = LoadCell_2.getData();
kkk = a+b;
newDataReady = 0;
t = millis();
//Serial.println(kk);
}
}
}
// below will print the well watered and drought watered water limits, so the user can double
check
whether the figures are correct.
Serial.print("Well watered pot weight default value = ");
Serial.println(Wlimit);
Serial.print("Drought watered pot weight default value = ");
Serial.println(Dlimit);
Serial.print("residual water value = ");
Serial.println(e);
Serial.println("Keep the 1st pot on the scale to initiate");
}
void loop() {
pp1 = 0;
ppdif = 0;
```

```
nb = 0;
u = 0;
kk = 1;
while ( kk < 3000){ // This will wait until the pot is kept on the loadcell, Also this will prevent algorithm identifying impact readings or falls readings as the initial weight of the pot.
u = kk;
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float a = LoadCell_1.getData();
float b = LoadCell_2.getData();
kk = a+b;
newDataReady = 0;
t = millis();
}
} }
u = -3;
kk = -2;
while (kk > u){ // This will wait until the load cell has reached its maximum value
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
```

```
if (millis() > t + serialPrintInterval) {
float a = LoadCell_1.getData();
float b = LoadCell_2.getData();
u = kk;
kk = a+b;
newDataReady = 0;
t = millis();
}
} }
for(m; m<1 ; m++){
Serial.print("please enter the pot number: , ");
while(Serial.available() == 0){
}
delay(4000);
pot = Serial.readString();
//////////////// Below code identifies the pot type ////////////
lengthh = pot.length();
while ( type != "W" && type != "w" && type != "d" && type != "D"){
type = pot.charAt(len);
len = len+1;
if (len > lengthh){
Serial.println("");
Serial.println("");
Serial.println("***Error: Pot number didn't include W or D please try again , ");
Serial.println("");
Serial.print("please enter the pot number: , ");
while(Serial.available() == 0){
}
delay(1000);
pot = Serial.readString();
len = 0;
```

```
}
}
if (type == "W" || type == "w"){
Serial.print(", A Well watered Pot , ");
potw = Wlimit;
type = "A";
len = 0;
}
if (type == "D" || type == "d"){
Serial.print(", A Drought watered Pot , ");
potw = Dlimit;
type = "A";
len = 0;
}
/////////////// pot identification finish /////////////
Serial.print(" ,");
Serial.print(" Initial weight is: ");
Serial.print(" , ");
Serial.print(u);
Serial.print(" ,");
}
if(k<potw){
/////// Below code checks whether the load cell is working properly //////
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
tini = millis();
while (newDataReady == 0) {
if (tdiff > 4000){
Serial.println(" ");
tdiff = 0;
Serial.println("there was a problem with the Scale, please remove the pot and restart the
```

```
program");
while (tdiff > 10000){
} }
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
tendd = millis();
tdiff = tendd - tini;
}
//////// Checking the load cell (end) ///////
digitalWrite(7,HIGH); // Turn ON the water pump
timet2 = millis();
tankt0 = millis();
while(p < potw-e) { // Keep the pump ON until the weight scale reading hits the set values.
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update() && LoadCell_2.update()) {
newDataReady = true;
}
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float aa = LoadCell_1.getData();
float bb = LoadCell_2.getData();
p = aa+bb;
newDataReady = 0;
t = millis();
}
}
timet1 = millis();
timettdif = timet1 - timet2;
```

```
if (timettdif > 3000){
Serial.print(" ");
Serial.print(p);
Serial.print(" ");
timet2 = timet1;
ppdif = p - pp1;
if (nb >1){
if (ppdif < 20) {
Serial.println(" ");
Serial.println(" ");
Serial.println("*********** An Error detected due to below reasons **********************
");
Serial.println(" 1) Water Tank is empty");
Serial.println(" 2) Pot is way off from the center ");
Serial.println(" ");
Serial.println(" Please remove the pot and restart the Putty program ");
Serial.println(" ");
Serial.println("************************************************************************
*");
digitalWrite(7,LOW);
while (1){
}
}
}
nb = nb+1;
pp1 = p;
} }
//delay(1200);
digitalWrite(7,LOW);
}
else{
```

```
p=u;
}
if(p > potw-e){
m = 0;
Serial.print(" , ");
Serial.print("Watering completed");
Serial.print(" ,");
Serial.print(p);
Serial.print(" , ");
Serial.print("remove the pot and wait......");
TOTW = TOTW + (p - u);
tankt1 = millis();
tanktd = tanktd + (tankt1 - tankt0);
/*Serial.print("testing: ");
Serial.print("k= ");
Serial.print(k);
Serial.print("u= ");
Serial.print(u);
Serial.print("p= ");
Serial.print(p);*/
w=p;
k=0;
u=0;
p=0;
while(w > 100) {
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
```

```
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float aaa = LoadCell_1.getData();
float bbb = LoadCell_2.getData();
w = aaa+bbb;
newDataReady = 0;
t = millis();
}
}
}
Serial.print(" , ");
LoadCell_1.tareNoDelay();
LoadCell_2.tareNoDelay();
if (LoadCell_1.getTareStatus() == true) {
Serial.print(" T1 ");
}
if (LoadCell_2.getTareStatus() == true) {
Serial.print(" T2 ");
}
Serial.print(" totW: ");
Serial.print(TOTW);
Serial.println(" keep the next pot");
Serial.println(" ");
if (TOTW > 12500){ // This detects when the tank is near empty
Serial.println(" ");
Serial.println(" ");
Serial.println("********************************************************************
******");
Serial.println("Please refill the tank and press Y ");
Serial.println("Don't keep pot on the basket while filling the tank");
Serial.println("********************************************************************
```

```
******");
Serial.println(" ");
while(Serial.available() == 0){
}
Serial.println(" ");
delay(2000);
Serial.println("Thanks");
Serial.println(" ");
//////////// check whether there's a pot on the machine///////////////
float ww = 200;
int nw = 1;
while(ww > 100) {
if (nw == 1){
Serial.println("************ Attention : Please remove the pot and wait ***********");
Serial.println(" ");
nw = nw +1;
}
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float aaaw = LoadCell_1.getData();
float bbbw = LoadCell_2.getData();
ww = aaaw+bbbw;
newDataReady = 0;
t = millis();
}
```

```
}
}
//////////// check whether there's a pot on the machine (end) //////////////////
Serial.println(" ");
TOTW = 0;
delay(2000);
Serial.println("Keep the Next Pot on the Machine to Continue Watering");
Serial.println(" ");
}
}
static boolean newDataReady = 0;
t = millis();
}
```

File A.3: C++ Code for Arduino with dynamic watering conditions

```
#include <HX711_ADC.h>
#include <EEPROM.h>
//pins:
const int HX711_dout_1 = 4; //mcu > HX711 no 1 dout pin
const int HX711_sck_1 = 5; //mcu > HX711 no 1 sck pin
const int HX711_dout_2 = 2; //mcu > HX711 no 2 dout pin
const int HX711_sck_2 = 3; //mcu > HX711 no 2 sck pin
HX711_ADC LoadCell_1(HX711_dout_1, HX711_sck_1); //HX711 1
HX711_ADC LoadCell_2(HX711_dout_2, HX711_sck_2); //HX711 2
const int calVal_eepromAdress_1 = 0; // eeprom adress for calibration value load cell 1 (4 bytes)
const int calVal_eepromAdress_2 = 4; // eeprom adress for calibration value load cell 2 (4 bytes)
float potw ;
long t;
float k=0;
float u=0;
```

```
String pot;
int m =0;
float p=0;
float w=0;
float e=20;
float tdiff=0;
float tendd;
float tini;
int rrr = 0;
float timet1 ;
float timet2;
float timettdif;
float pp1 = 0;
float ppdif = 0;
int nb = 0;
float kk;
int tankt1;
int tankt0;
int tanktd=0;
int lengthh;
int len;
float TOTW = 0;
char ea;
int condN;
int condN_array;
int xcondN;
char condition;
float condition_w;
int ar_size;
String ready_level;
char result;
```

```
char type;
String condition_name[50];
char condition_letter[50];
float condition_weight[50];
void setup() {
pinMode(7,OUTPUT);
Serial.begin(57600); delay(10);
Serial.println();
Serial.println("Starting...");
delay(5000);
float calibrationValue_1;
float calibrationValue_2;
calibrationValue_1 = -107;
calibrationValue_2 = -107;
LoadCell_1.begin();
LoadCell_2.begin();
long stabilizingtime = 2000; // tare preciscion can be improved by adding a few seconds of stabilizing time
boolean _tare = true; //set this to false if you don't want tare to be performed in the next step
byte loadcell_1_rdy = 0;
byte loadcell_2_rdy = 0;
while ((loadcell_1_rdy + loadcell_2_rdy) < 2) { //run startup, stabilization and tare, both modules simultaniously
if (!loadcell_1_rdy) loadcell_1_rdy = LoadCell_1.startMultiple(stabilizingtime, _tare);
if (!loadcell_2_rdy) loadcell_2_rdy = LoadCell_2.startMultiple(stabilizingtime, _tare);
}
if (LoadCell_1.getTareTimeoutFlag()) {
Serial.println(F("Timeout, check MCU>HX711 no.1 wiring and pin designations"));
}
if (LoadCell_2.getTareTimeoutFlag()) {
Serial.println(F("Timeout, check MCU>HX711 no.2 wiring and pin designations"));
```

```
}
LoadCell_1.setCalFactor(calibrationValue_1); // user set calibration value (float)
LoadCell_2.setCalFactor(calibrationValue_2); // user set calibration value (float)
Serial.println(F("Startup is complete"));
// receive command from serial terminal, send 't' to initiate tare operation:
int kkk = 1;
while (kkk != 0){
// if (Serial.available() > 0) {
float i;
//char inByte = Serial.read();
//if (inByte == 't') {
LoadCell_1.tareNoDelay();
LoadCell_2.tareNoDelay();
//check if last tare operation is complete
if (LoadCell_1.getTareStatus() == true) {
Serial.println(F("Tare load cell 1 complete"));
}
if (LoadCell_2.getTareStatus() == true) {
Serial.println(F("Tare load cell 2 complete"));
}
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float a = LoadCell_1.getData();
float b = LoadCell_2.getData();
kkk = a+b;
```

```
newDataReady = 0;
t = millis();
//Serial.println(kk);
}
}
}
//////////////// This section can be used to define the watering conditions /////////////////
while(ready_level != "Y" && ready_level != "y"){
Serial.print(F("How many watering conditions: "));
while(Serial.available() == 0){
}
condN = Serial.parseInt();
//String condition_name[condN];
String condition_n;
Serial.print(condN);
Serial.println(F(" "));
Serial.println(F(" "));
condN_array = condN-1;
for(xcondN=1; xcondN < condN+1 ; ){
///////// name of the condition start //////////
Serial.print(F("Enter the name of the condition: "));
Serial.print(xcondN);
delay(4000);
while(Serial.available() == 0){
}
//delay(1000);
condition_n = Serial.readString();
while (Serial.available() >0 ) {
//Serial.read();
}
Serial.println(F(""));
```

```
condition_name[xcondN-1] = condition_n;
///////// name of the condition end /////////
Serial.print(F("Enter the Identifying Letter for the condition: "));
Serial.print(xcondN);
delay(4000);
while(Serial.available() == 0){
}
//delay(1000);
condition = Serial.read();
if(isLowerCase(condition)){
condition = toUpperCase(condition);
}
while (Serial.available() >0 ) {
}
Serial.println(F(""));
condition_letter[xcondN-1] = condition;
//////////////////////////// weights input/////////
Serial.print(F("Enter the weight for the condition "));
Serial.println(xcondN);
delay(4000);
while(Serial.available() == 0){
}
//delay(1000);
condition_w = Serial.parseFloat();
while (Serial.available() >0 ) {
Serial.read();
}
Serial.println(F(""));
condition_weight[xcondN-1] = condition_w;
xcondN = xcondN+1 ;
/////////////////////////
```

```
}
Serial.println(F(""));
Serial.println(F("data in the array"));
ar_size = sizeof(condition_weight);
for(int cond_i = 0; cond_i < condN; cond_i++)
{
Serial.print(F("Weight of the Condition "));
Serial.print(condition_name[cond_i]);
Serial.print(F(" : "));
Serial.print(condition_letter[cond_i]);
Serial.print(F(" is: "));
Serial.print(condition_weight[cond_i]);
Serial.println(F(" grams"));
}
Serial.println(F(""));
Serial.println(F("Press Y to save above configuration"));
Serial.println(F("Press N to redefine the configuration values"));
while(Serial.available() == 0){
}
ready_level = Serial.readString();
while ( ready_level != "Y" && ready_level != "y" && ready_level != "N" && ready_level !=
"n"){
Serial.println(F(""));
Serial.println(F("Input was incorrect"));
Serial.println(F("Press Y to save above configuration"));
Serial.println(F("Press N to redefine the configuration values"));
while(Serial.available() == 0){
}
ready_level = Serial.read();
}
Serial.println(F(""));
```

```
}
Serial.println(F("Configurations saved"));
Serial.println(F(" "));
////////////////////// end of the code ////////////////////////////////////////
/////////////// Define e value ///////////////////////////////////////////////
Serial.print(F("default e value = "));
Serial.println(e);
Serial.println(F("Press Y if you need to keep the default value"));
Serial.println(F("Press N if you need to change the value"));
while(Serial.available() == 0){
}
ea = Serial.read();
if(ea == 'Y'){
Serial.println(F(""));
Serial.print(F("Enter the new value in grams: "));
while(Serial.available() == 0){
}
e = Serial.parseFloat();
Serial.println(F(""));
Serial.println(F(""));
Serial.println(F("New e value is: "));
Serial.println(e);
Serial.println(F("Saved the new e value"));
}
else{
Serial.println(F("Default e value will be used"));
}
//////////////////////////////////////////////////////////////////////
Serial.println(F(""));
Serial.println(F("Keep the 1st pot on the scale to initiate"));
}
```

```
void loop() {
pp1 = 0;
ppdif = 0;
nb = 0;
u = 0;
kk = 1;
while ( kk < 3000){ // This will wait until the load cell has reached its maximu value
u = kk; // Allways U will end up with the maximum load cell value
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float a = LoadCell_1.getData();
float b = LoadCell_2.getData();
kk = a+b;
newDataReady = 0;
t = millis();
//Serial.println(kk);
}
} }
u = -3;
kk = -2;
while (kk > u){ // This will wait until the load cell has reached its maximu value
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
```

```
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float a = LoadCell_1.getData();
float b = LoadCell_2.getData();
u = kk;
kk = a+b;
newDataReady = 0;
t = millis();
}
} }
for(m; m<1 ; m++){
Serial.print(F("please enter the pot number: , "));
while(Serial.available() == 0){
}
delay(4000);
pot = Serial.readString();
/////////////////// (This will identify the waterning condition by the name of the POT, following the
above saved configurations) /////////////
lengthh = pot.length();
result = 'N';
while(result == 'N'){
for(int pot_code_i=0; pot_code_i < lengthh; pot_code_i++){
type = pot.charAt(pot_code_i);
if(pot_code_i == lengthh-1){
if(pot_code_i == lengthh-1 && result == 'T'){
}
else{
result = 'F';
}
```

```
}
if(isLowerCase(type)){
type = toUpperCase(type);
}
for(int array_i=0; array_i < condN; array_i++){
if (type == condition_letter[array_i]){
Serial.print(F(", This is a "));
Serial.print(condition_name[array_i]);
Serial.print(F(" Condition maintained Pot"));
potw = condition_weight[array_i];
result = 'T';
}
}
if(result == 'F'){
result = 'F';
Serial.println(F(""));
Serial.println(F(""));
Serial.println(F("***Error: Pot number didn't include any letter defined for the conditions please
try
again , "));
Serial.println(F(""));
Serial.print(F("please enter the pot number: , "));
while(Serial.available() == 0){
}
delay(1000);
pot = Serial.readString();
lengthh = pot.length();
}
}
}
////////////////// end of pot identification ////////////////////////////////////////
```

```
Serial.print(F(" Initial weight is: "));
//Serial.print(pot);
Serial.print(F(" , "));
Serial.print(u);
Serial.print(F(" ,"));
Serial.print(F(" Watering the Plant.... "));
}
if(k<potw){
digitalWrite(7,HIGH);
timet2 = millis();
tankt0 = millis();
while(p < potw-e) {
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update() && LoadCell_2.update()) {
newDataReady = true;
}
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float aa = LoadCell_1.getData();
float bb = LoadCell_2.getData();
p = aa+bb;
newDataReady = 0;
t = millis();
}
}
timet1 = millis();
timettdif = timet1 - timet2;
}
```

```
//delay(1200);
digitalWrite(7,LOW);
}
else{
p=u;
}
if(p > potw-e){
m = 0;
Serial.print(F(" , "));
Serial.print(F(" Watering completed at"));
Serial.print(F(" ,"));
Serial.print(p);
Serial.print(F(" , "));
Serial.print(F("remove the pot and wait......"));
TOTW = TOTW + (p - u);
tankt1 = millis();
tanktd = tanktd + (tankt1 - tankt0);
w=p;
k=0;
u=0;
p=0;
while(w > 100) {
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float aaa = LoadCell_1.getData();
```

```
float bbb = LoadCell_2.getData();
w = aaa+bbb;
newDataReady = 0;
t = millis();
}
}
}
LoadCell_1.tareNoDelay();
LoadCell_2.tareNoDelay();
if (LoadCell_1.getTareStatus() == true) {
}
if (LoadCell_2.getTareStatus() == true) {
// Serial.print(" T2 ");
}
//Serial.print(" totW: ");
//Serial.print(TOTW);
Serial.println(F(" keep the next pot"));
Serial.println(F(" "));
if (TOTW > 12500){
Serial.println(F(" "));
Serial.println(F(" "));
Serial.println(F("****************************************************************"));
Serial.println(F("Please refill the tank and press Y "));
Serial.println(F("Don't keep pot on the basket while filling the tank"));
Serial.println(F("****************************************************************"));
Serial.println(F(" "));
while(Serial.available() == 0){
}
Serial.println(F(" "));
delay(2000);
Serial.println(F("Please Wait....."));
```

```
Serial.println(F(" "));
//////////// check whether there's a pot on the machine, start ///////////////
float ww = 200;
int nw = 1;
while(ww > 100) {
if (nw == 1){
Serial.println(F(" ****** Attention : Please remove the pot and wait ***************"));
Serial.println(F(" "));
nw = nw +1;
}
static boolean newDataReady = 0;
const int serialPrintInterval = 0; //increase value to slow down serial print activity
// check for new data/start next conversion:
if (LoadCell_1.update()) newDataReady = true;
LoadCell_2.update();
//get smoothed value from data set
if ((newDataReady)) {
if (millis() > t + serialPrintInterval) {
float aaaw = LoadCell_1.getData();
float bbbw = LoadCell_2.getData();
ww = aaaw+bbbw;
newDataReady = 0;
t = millis();
}
}
}
//////////// check whethrr there's a pot on the machine end ///////////////////
Serial.println(F(" "));
TOTW = 0;
delay(2000);
Serial.println(F("Keep the Next Pot on the Machine to Continue Watering"));
```

```
Serial.println(F(" "));
}
}
static boolean newDataReady = 0;
t = millis();
}
```

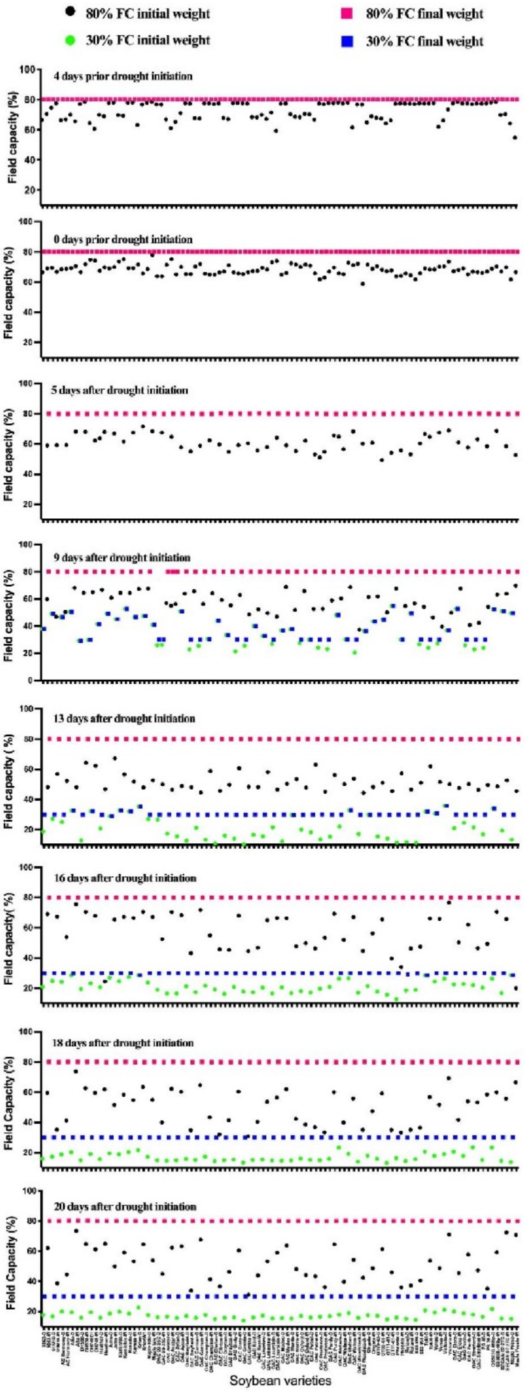


Figure A.1: Filed capacity adjustment in pots growing soybean plants using the semi-automated Arduino-based lysimeter. Field capacity of the pots were adjusted to 80% and 30%. Soybean variety names followed by D represent the drought treatment (30% FC) pots and by W represent well-watered treatment (80% FC) pots. Black circles represent pot weight of the 80% FC treatment pots before adding water; pink circles represent pot weight of the 80% FC treatment pots after adding water; Green circles represent pot weight of the 30% FC treatment pots before adding water; blue circles represent pot weight of the 30% FC treatment pots after adding water.

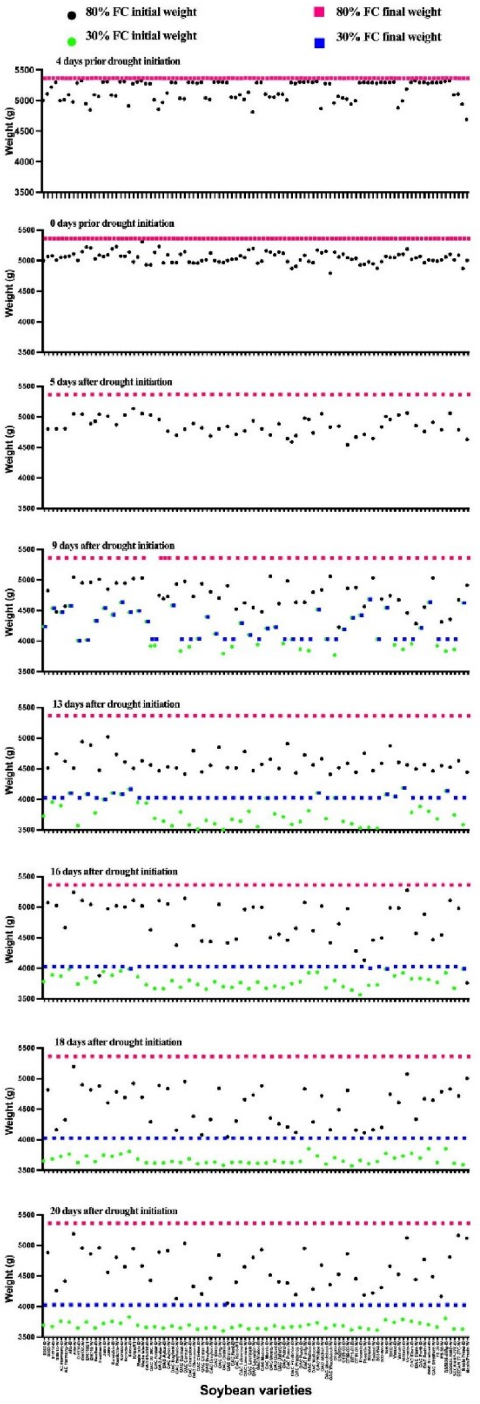


Figure A.2: Gravimetric moisture adjustment in pots growing soybean plants using the semi automated Arduino-based lysimeter. Final weights of the pots were adjusted to 5363.4 g (80% field capacity-FC) and 4025.9 g (30% FC). Soybean variety names followed by D represent the drought treatment (30% FC) pots and by W represent well-watered treatment (80% FC) pots. Black circles represent FC of 80% treatment pots before adding water; pink circles represent FC of 80% treatment pots after adding water; green circles represent FC of 30% treatment pots before adding water; blue circles represent FC of 30% treatment pots after adding water.

Video A.1: A video of the Arduino-based lysimeter system in operation.

Link to the Video : https://ualbertaca-my.sharepoint.com/:v:/g/personal/kombalal_ualberta_ca/EZx00SsTcUFDkBKcDQ6Leh8Bautr2_f854-7Ug8ANGBk1Q?e=jVHeFF

Appendix - B

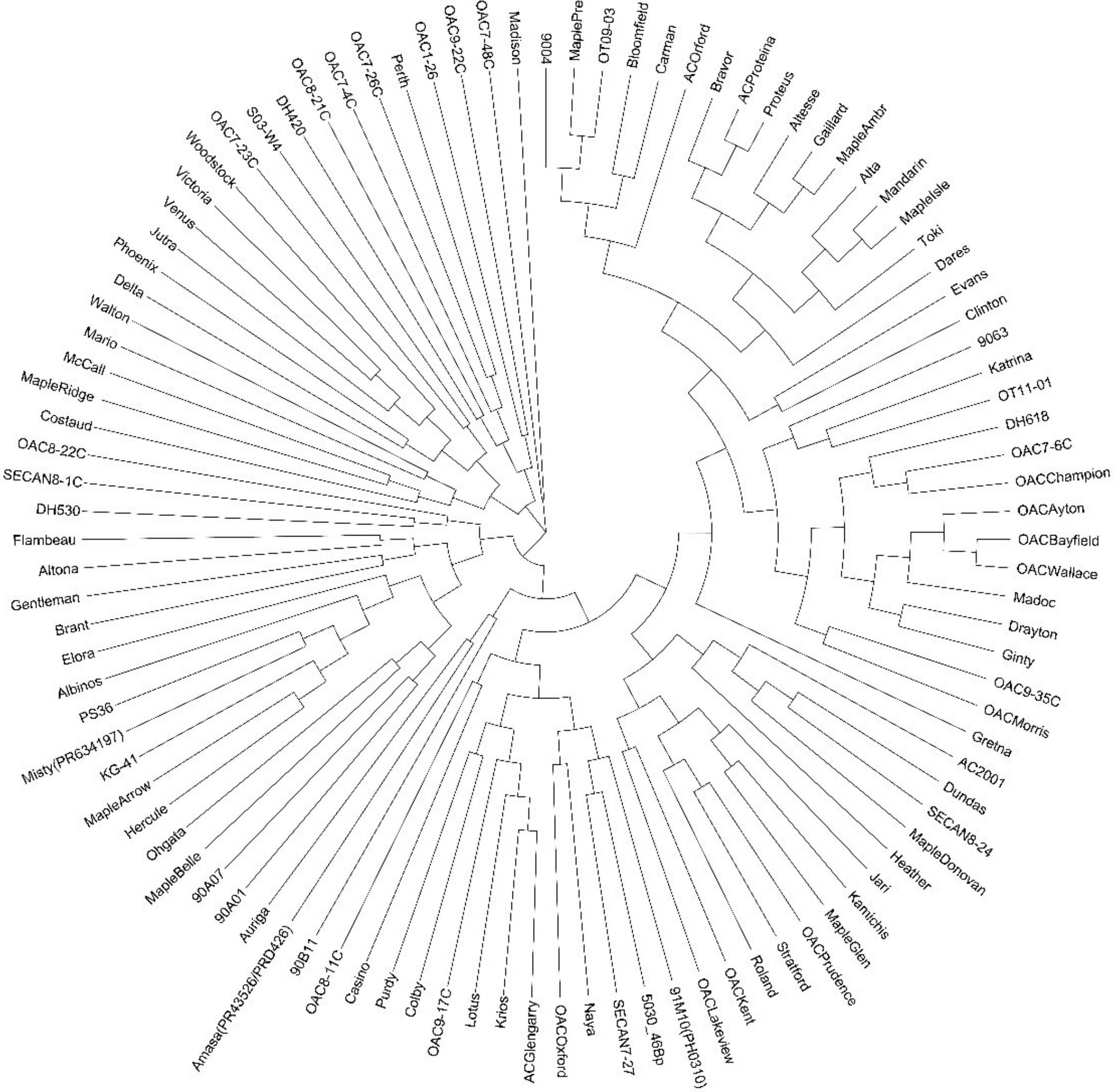


Figure B.1: Phylogenetic tree constructed using a core set of 103 Canadian soybean lines in TASSEL 5.0.

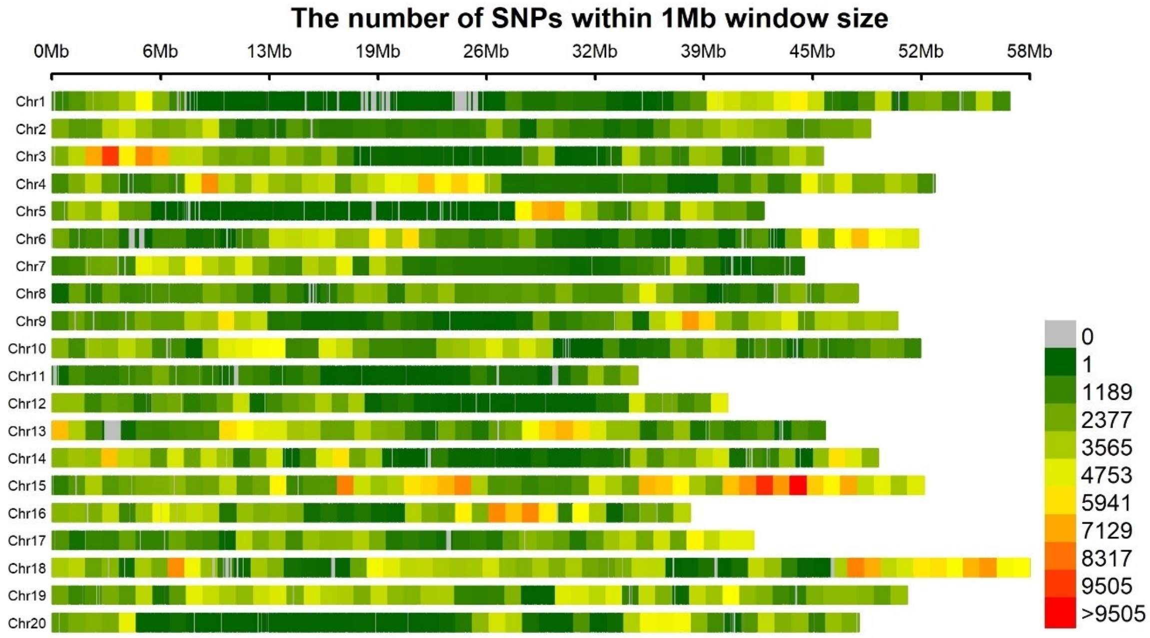


Figure B.2: Distribution of SNP markers across the soybean genome.

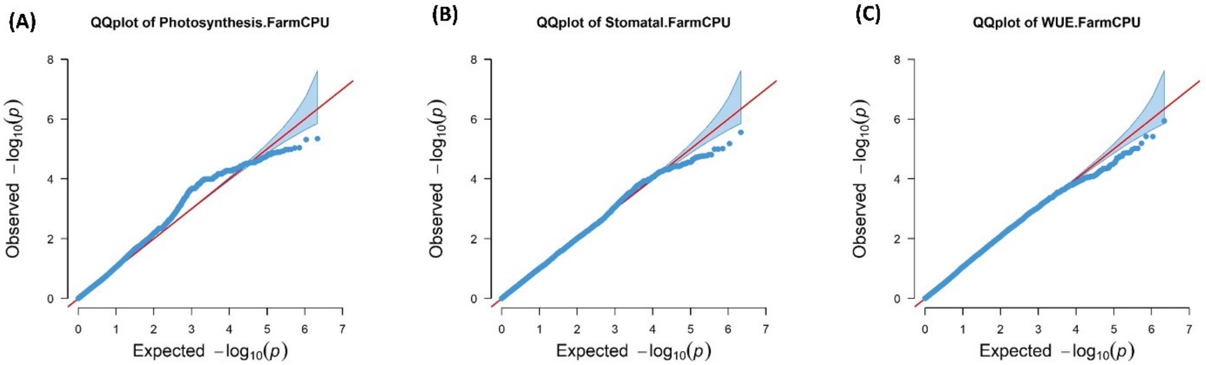


Figure B.3: Quantile-quantile (Q-Q) plot of p-values for the association between SNP markers (blue dots) and photosynthesis (A), stomatal conductance (B) and water use efficiency (C) for 30% field capacity using fixed and random model circulating probability unification (FarmCPU) (Liu et al. 2016) in a panel of 103 soybean lines. The y-axis is the observed negative base 10 logarithm of the p-values, and the x-axis is the expected observed negative base 10 logarithm of the p-values under the assumption that the p-values follow a uniform (0,1) distribution. The blue line shows the 95% confidence interval for the Q-Q plot under the null hypothesis of no association between the SNP and the trait.

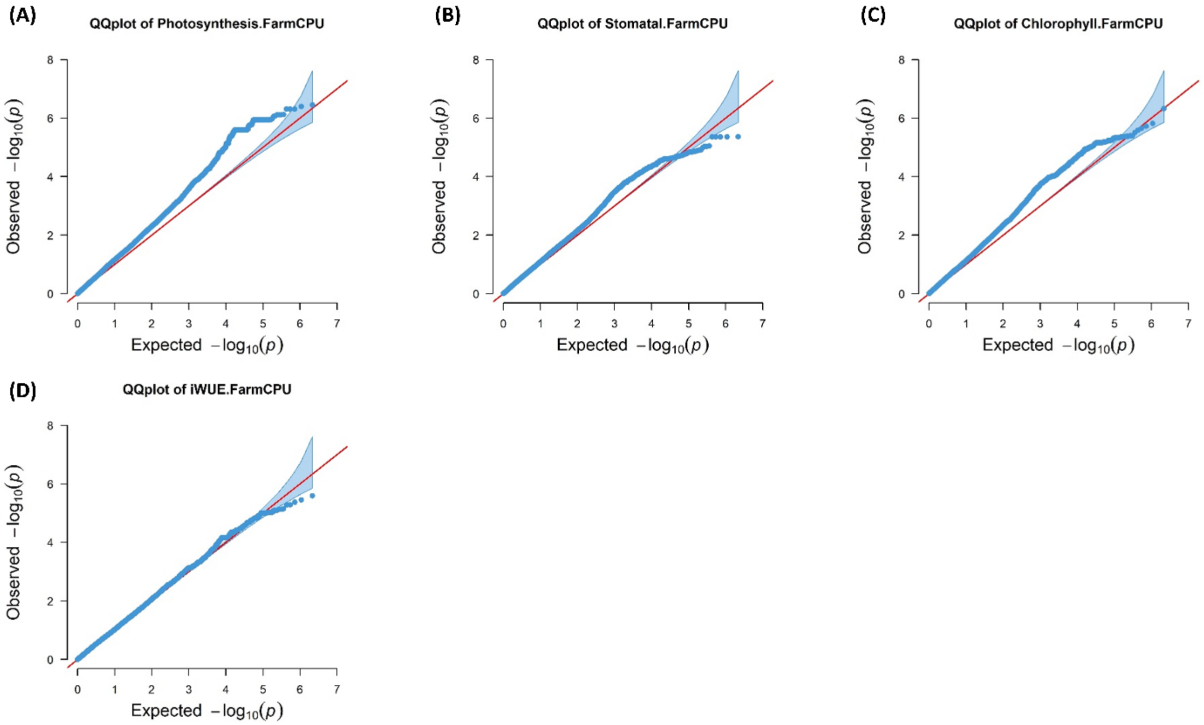


Figure B.4: Quantile-quantile (Q-Q) plot of p-values for the association between SNP markers (blue dots) and photosynthesis (A), stomatal conductance (B), leaf chlorophyll content (C), and intrinsic water use efficiency (D) for 80% field capacity (FC) using fixed and random model circulating probability unification (FarmCPU) (Liu et al. 2016) in a panel of 103 soybean lines. The y-axis is the observed negative base 10 logarithm of the p-values, and the x-axis is the expected observed negative base 10 logarithm of the p-values under the assumption that the p-values follow a uniform (0,1) distribution. The blue line shows the 95% confidence interval for the Q-Q plot under the null hypothesis of no association between the SNP and the trait.

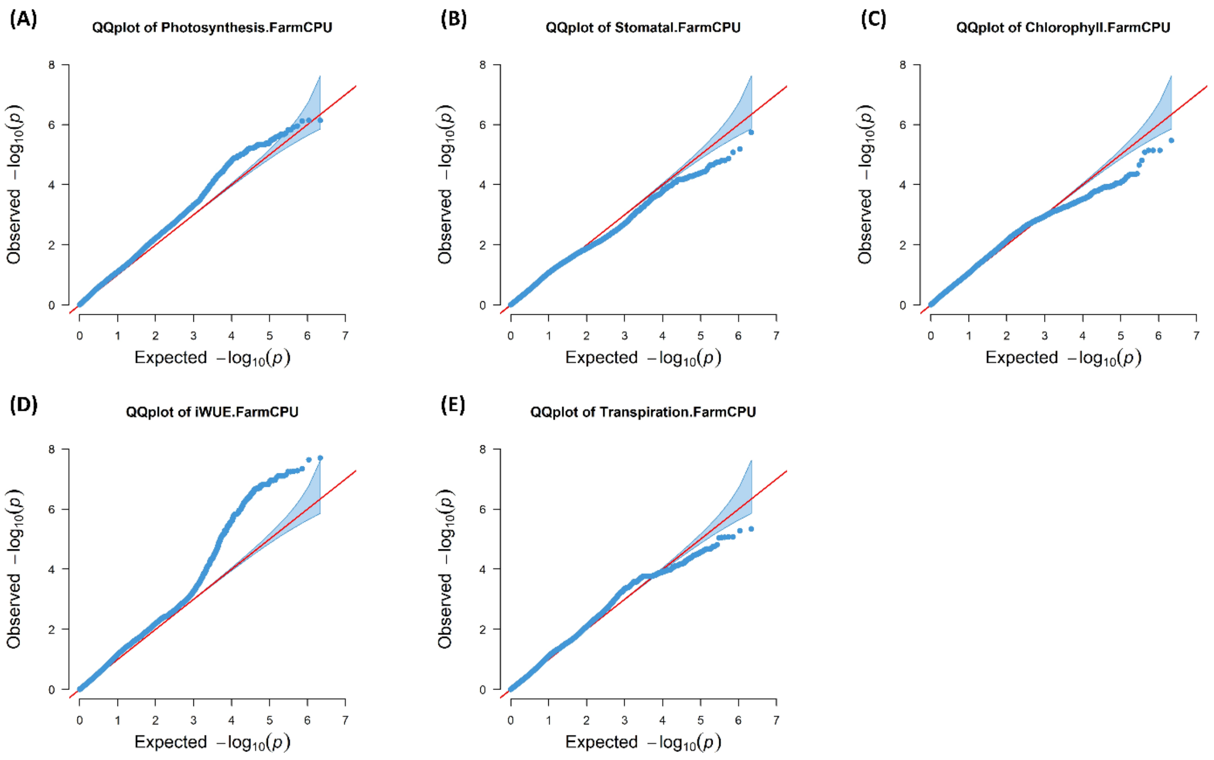


Figure B.5: Quantile-quantile (Q-Q) plot of p-values for the association between SNP markers (blue dots) and photosynthesis (A), stomatal conductance (B), leaf chlorophyll content (C), intrinsic water use efficiency (D), and transpiration (E) for relative performance using fixed and random model circulating probability unification (FarmCPU) (Liu et al. 2016) in a panel of 103 soybean lines. The y-axis is the observed negative base 10 logarithm of the p-values, and the x-axis is the expected observed negative base 10 logarithm of the p-values under the assumption that the p-values follow a uniform (0,1) distribution. The blue line show the 95% confidence interval for the Q-Q plot under the null hypothesis of no association between the SNP and the trait.

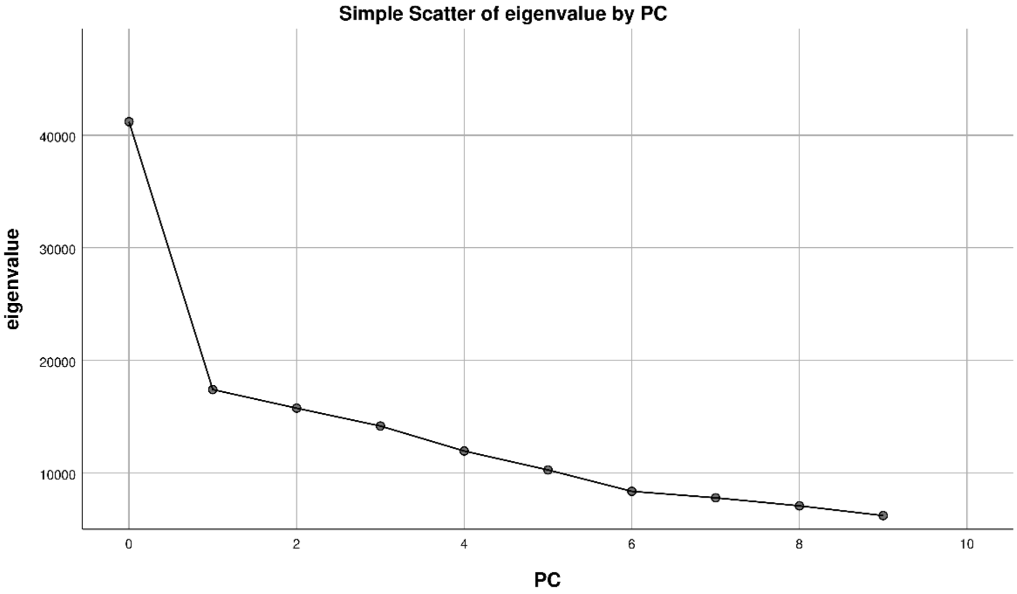


Figure B.6: Scree plot for 10 components

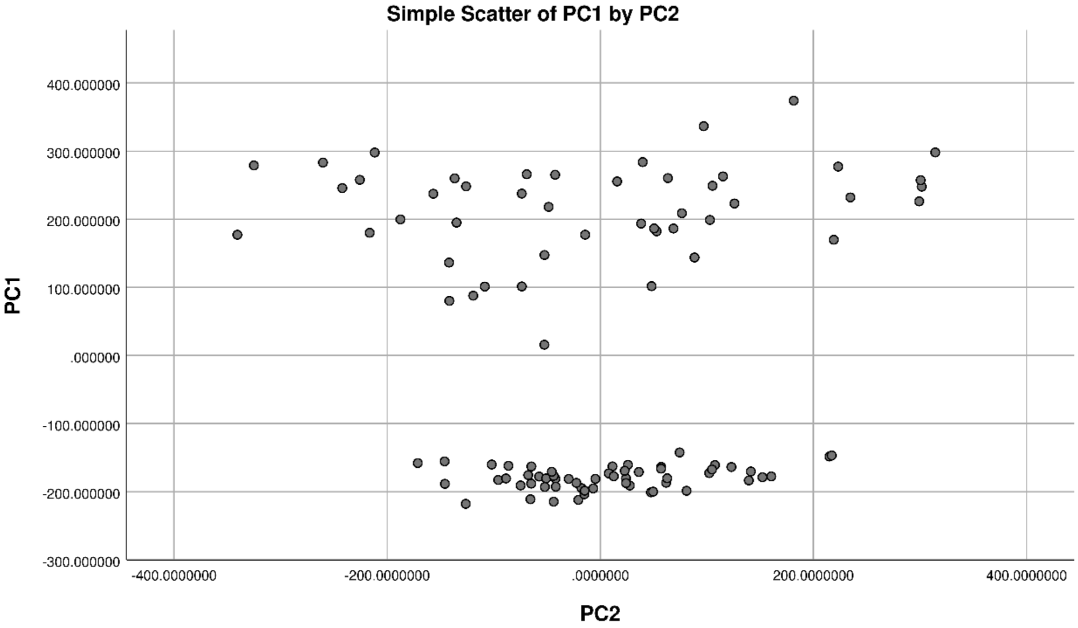


Figure B.7: PCA plot of first two components (PC1 and PC2)

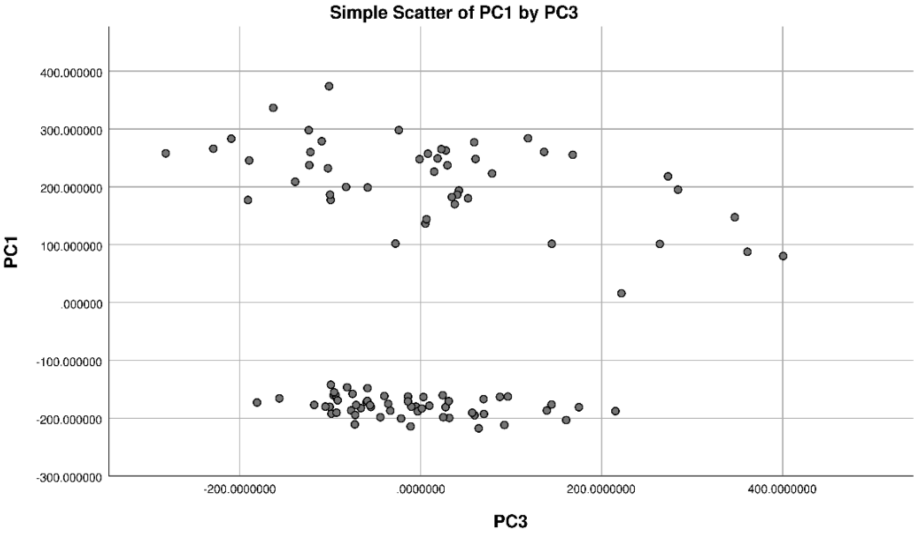


Figure B.8: PCA plot of first and third components (PC1 and PC3)

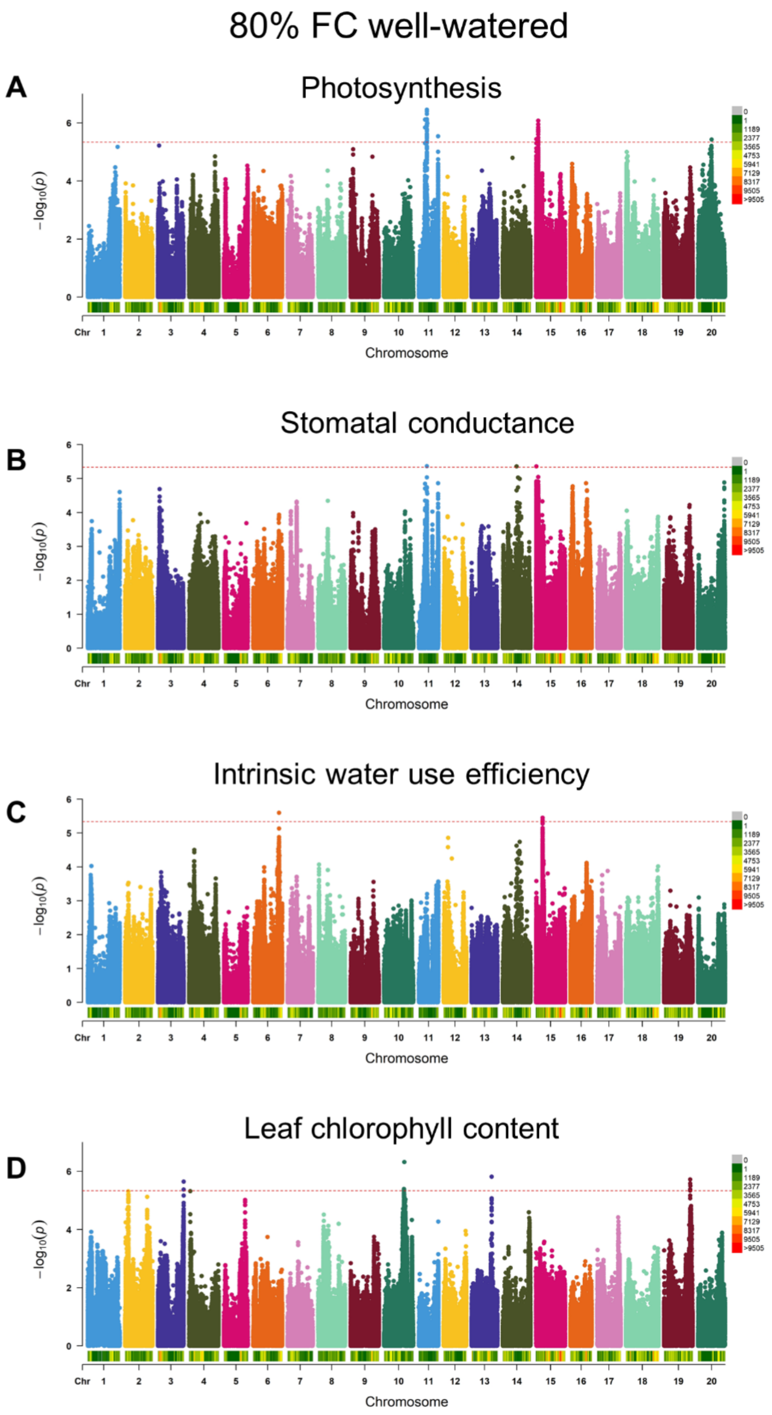


Figure B.9: Manhattan plots of the genome-wide association results for different plant physiological traits measured at the flowering stage under well-watered (80% FC) conditions. (A) photosynthesis, (B) stomatal conductance, (C) intrinsic water-use efficiency, (D) leaf chlorophyll content. Negative log10 (P-values) (y-axis) describing the strength of the association between each marker and trait are plotted against the physical position of each marker (x-axis). The pink dashed line indicates the significance threshold (FDR = 5%) and beyond that are considered as significant association

Table B.1: The list of short season soybean genotypes used in the study

9004
9063
90A01
90A07
90B11
91M10
AC 2001
AC Brant
AC Bravor
AC Glengarry
AC Harmony
AC Hercule
AC Orford
AC Proteina
AC Proteus
Albinos
Alta
Altesse
Altona
Amasa
Auriga
Bloomfield
Casino
Colby
Costaud
Dares
Delta
DH420
DH530
DH618
DH748
Dundas
Evans
Flambeau
Gaillard
Gentleman
Heather
Jari
Jutra
Kamichis
Katrina
KG 41
Korada
Krios
Lotus
Madison
Mandarin
Maple Amber
Maple Arrow
Maple Belle
Maple Donovan
Maple Glen
Maple Isle
Maple Presto
Maple Ridge
Mario
McCall
Misty
Naya
OAC 01-26
OAC 07-04C
OAC 07-06C
OAC 07-26C
OAC 08-11C
OAC 08-21C
OAC 08-22C
OAC 09-01C
OAC 09-17C
OAC 09-22C
OAC 09-35C
OAC Avatar
OAC Ayton
OAC Bayfield
OAC Carman
OAC Champion
OAC Clinton
OAC Drayton
OAC Ginty
OAC Gretna
OAC Kent
OAC Lakeview
OAC Lauralain
OAC Madoc
OAC Morris
OAC Oxford
OAC Petrel
OAC Prudence
OAC Purdy
OAC Wallace
OAC Walton
OAC Woodstock
Ohgata
OT09-03
OT11-01
Phoenix
Roland
S03-W4
Toki
Venus
Victoria
OAC Elora
OAC Perth
OAC Stratford
PS 36
QS5030.46Bp
SECAN 07-27C

Table B.2: Candidate genes associated under drought stress (30% Field capacity)

Trait	Chr	MSS position	REF/ALT	QTL ID	Candidate genes	Orthologous genes in *Arabidopsis*	Annotations
Stomatal conductance	15	4877172	C/T	qSC1	Glyma.15g061300		Soluble inorganic pyrophosphatase 1% 2C Chloroplastic like - BT096213
					Glyma.15g06581		proline/glycine/tyrosine-rich protein- HM131450
					Glyma.15g061100	Gene Model: AT4G21860.1	Methionine sulfoxide reductase B2
					Glyma.15g064300	Gene Model: AT1G11430.1	Plastid developmental protein DAG, putative
					Glyma.15g060800		Phosphoglycerate mutase family protein

Table B.3: Candidate genes associated with relative performance

Trait	Chr	MSS position	REF/ALT	QTL ID	Candidate genes	Orthologous genes in Arabidopsis	Annotations
IWUE	10	48220950	A/G	qIWUE3	Glyma.10g256000	Gene Model: AT1G49975.1	Photosystem I reaction centre subunit N (PSAN or PSI-N)
					Glyma.10g257500	Gene Model: AT4G33880.1	ROOT HAIR DEFECTIVE 6-LIKE 2
	16	28085463	C/A	qIWUE4	Glyma.16g130200	Gene Model: AT2G18060.1	Vascular related NAC-domain protein 1
	17	8878535	C/T		Glyma.17g113000	Gene Model: AT5G53870.1	Early nodulin-like protein 1
	17	8978069	C/T	qIWUE5	Glyma.17g117200	Gene Model: AT5G25940.1	Early nodulin 93 ENOD93 protein
					Glyma.17g117100	Gene Model: AT5G25940.1	Early nodulin-93-like-BT096343
Photosynthesis	16	3072259	T/A	qPH3	Glyma.16g030400	Gene Model: AT4G26500.1	Chloroplast sulfur E

Appendix – C

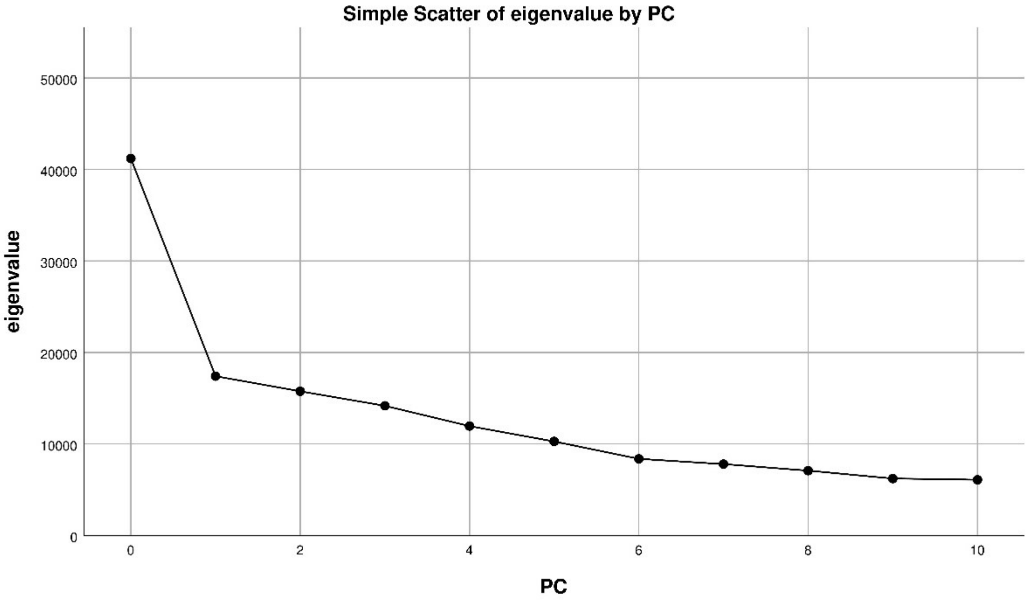


Figure C.1: Scree plot for 10 components

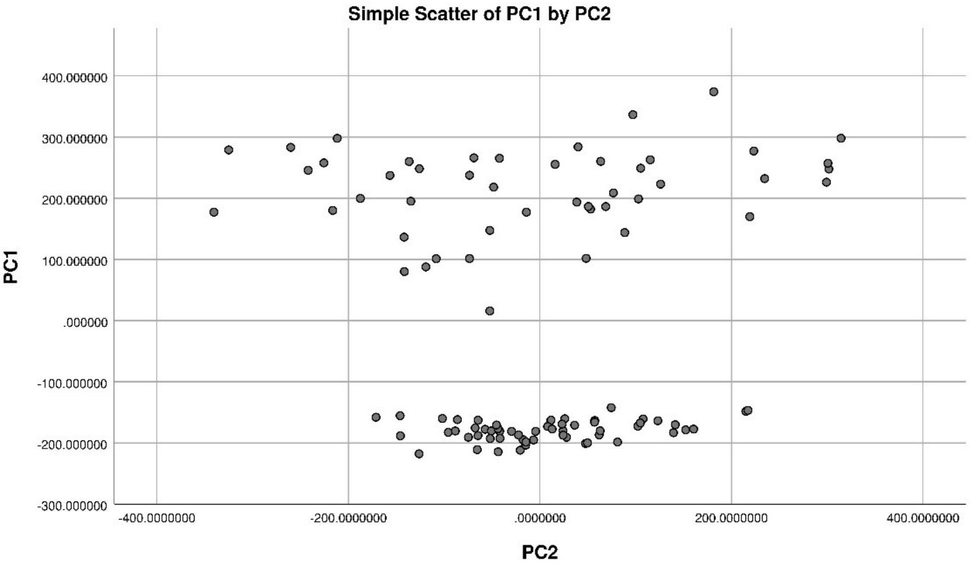


Figure C.2: PCA plot of first two components (PC1 and PC2)

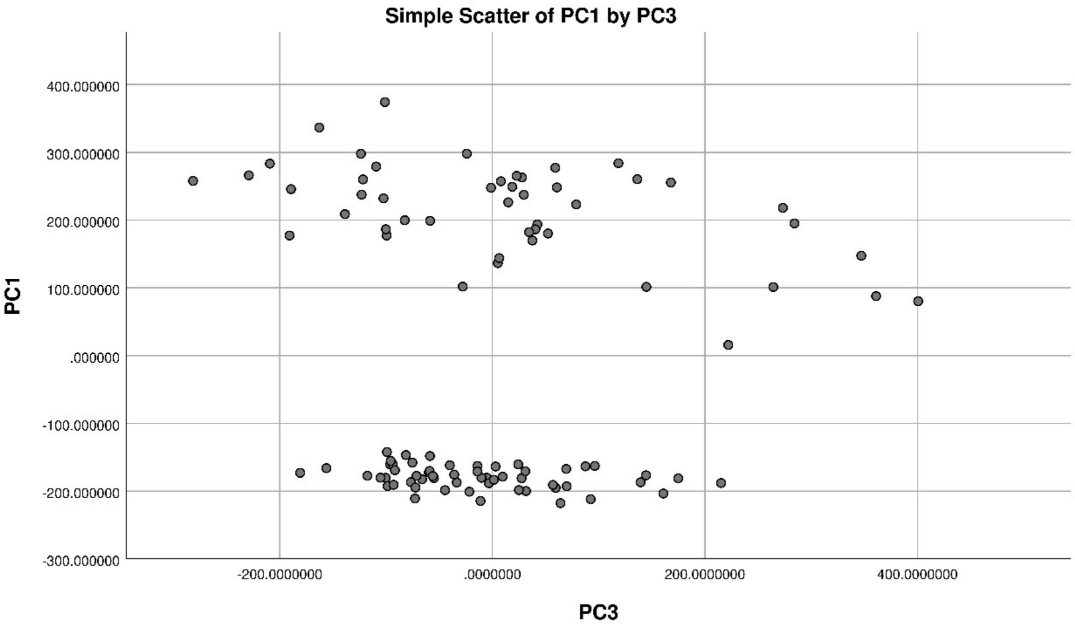


Figure C.3: PCA plot of first and third components (PC1 and PC3)

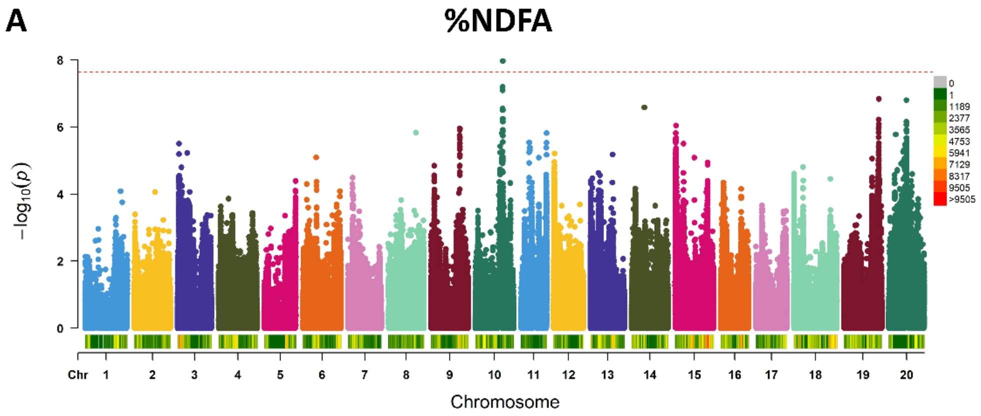


Figure C.4: Manhattan plots of the genome-wide association results for % nitrogen derived from the atmosphere (Ndfa) at 80% FC conditions. Negative log10 (P-values) (y-axis) describing the strength of the association between each marker and trait are plotted against the physical position of each marker (x-axis). The pink dashed line indicates the significance threshold (FDR = 5%) and beyond that are considered as significant association.

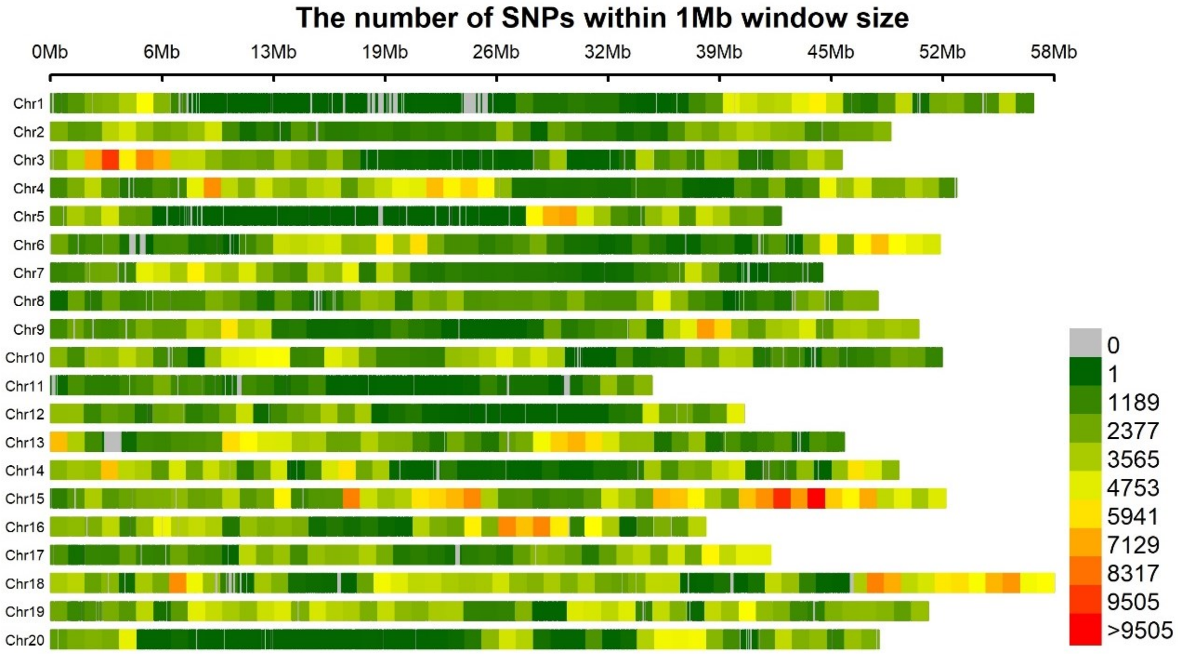


Figure C.5: Distribution of SNP markers across the soybean genome.

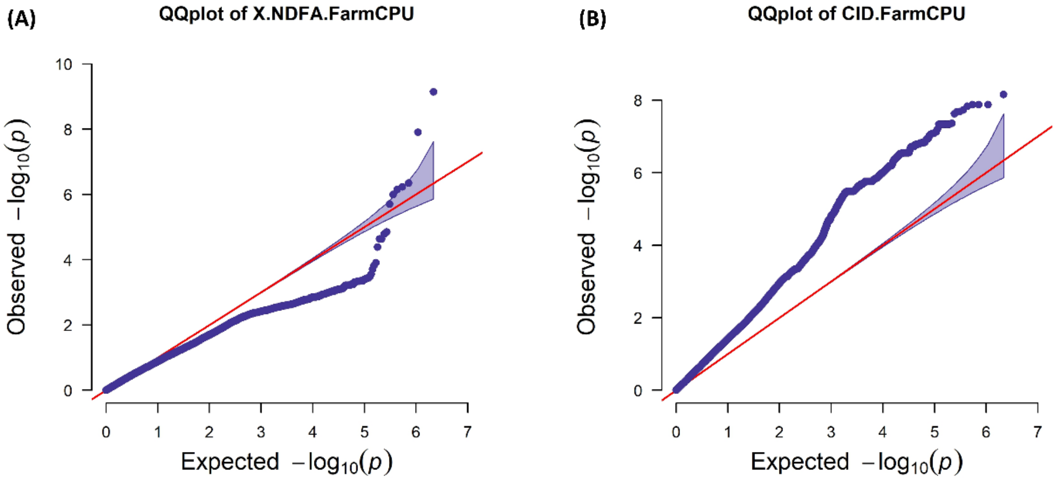


Figure C.6: Quantile-quantile (Q-Q) plot of p-values for the association between SNP markers (blue dots) and %Ndfa (A), CID (B) for 30% field capacity using fixed and random model circulating probability unification (FarmCPU) (Liu et al. 2016) in a panel of 103 soybean lines. The y-axis is the observed negative base 10 logarithm of the p-values, and the x-axis is the expected observed negative base 10 logarithm of the p-values under the assumption that the p-values follow a uniform (0,1) distribution. The blue line shows the 95% confidence interval for the Q-Q plot under the null hypothesis of no association between the SNP and the trait.

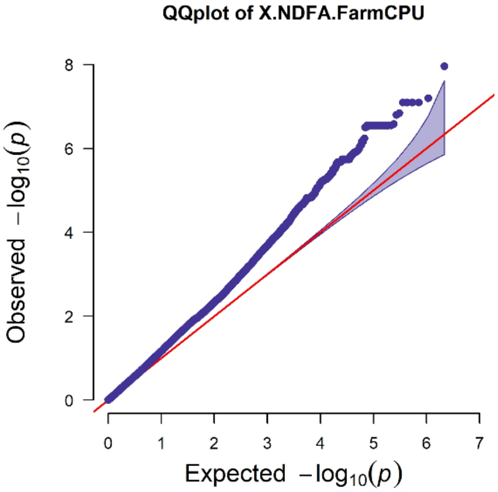


Figure C.7: Quantile-quantile (Q-Q) plot of p-values for the association between SNP markers (blue dots) and %Ndfa for 80% field capacity (FC) using fixed and random model circulating probability unification (FarmCPU) (Liu et al. 2016) in a panel of 103 soybean lines. The y-axis is the observed negative base 10 logarithm of the p-values, and the x-axis is the expected observed negative base 10 logarithm of the p-values under the assumption that the p-values follow a uniform (0,1) distribution. The blue line shows the 95% confidence interval for the Q-Q plot under the null hypothesis of no association between the SNP and the trait.

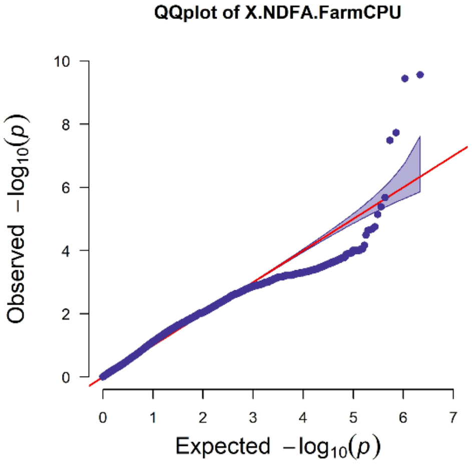


Figure C.8: Quantile-quantile (Q-Q) plot of p-values for the association between SNP markers (blue dots) and %Ndfa for relative performance using fixed and random model circulating probability unification (FarmCPU) (Liu et al. 2016) in a panel of 103 soybean lines. The y-axis is the observed negative base 10 logarithm of the p-values, and the x-axis is the expected observed negative base 10 logarithm of the p-values under the assumption that the p-values follow a uniform (0,1) distribution. The blue line shows the 95% confidence interval for the Q-Q plot under the null hypothesis of no association between the SNP and the trait.