

Article

The Genotypic Variability among Short-Season Soybean Cultivars for Nitrogen Fixation under Drought Stress

Dilrukshi Kombala Liyanage ¹, Davoud Torkamaneh ^{2,3} , François Belzile ^{2,3}, Parthiba Balasubramanian ⁴, Brett Hill ⁴ and Malinda S. Thilakarathna ^{1,*}

¹ Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB T6G 2P5, Canada

² Département de Phytologie, Université Laval, Québec City, QC G1V 0A6, Canada

³ Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec City, QC G1V 0A6, Canada

⁴ Agriculture and Agri-Food Canada, Lethbridge Research and Development Centre, Lethbridge, AB T1J 4B1, Canada

* Correspondence: malinda.thilakarathna@ualberta.ca

Abstract: Soybean fixes atmospheric nitrogen through the symbiotic rhizobia bacteria that inhabit root nodules. Drought stress negatively affect symbiotic nitrogen fixation (SNF) in soybean. The main objective of this study was to identify allelic variations associated with SNF in short-season Canadian soybean varieties under drought stress. A diversity panel of 103 early-maturity Canadian soybean varieties was evaluated under greenhouse conditions to determine SNF-related traits under drought stress. 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, seed nitrogen content, % nitrogen derived from the atmosphere (%Ndfa), and total seed nitrogen fixed compared to those under well-watered conditions. Significant genotypic variability among soybean varieties was found for yield, yield parameters, and nitrogen fixation traits. A genome-wide association study (GWAS) was conducted using 2.16 M single nucleotide polymorphisms (SNPs) for different yield and nitrogen fixation related parameters for 30% FC and their relative performance (30% FC/80% FC). In total, five quantitative trait locus (QTL) regions, including candidate genes, were detected as significantly associated with %Ndfa 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 genes; quantitative trait locus; genome wide association study



Citation: Liyanage, D.K.; Torkamaneh, D.; Belzile, F.; Balasubramanian, P.; Hill, B.; Thilakarathna, M.S. The Genotypic Variability among Short-Season Soybean Cultivars for Nitrogen Fixation under Drought Stress. *Plants* **2023**, *12*, 1004. <https://doi.org/10.3390/plants12051004>

Academic Editors: Roberto Barbato and Veronica De Micco

Received: 26 January 2023

Revised: 16 February 2023

Accepted: 20 February 2023

Published: 22 February 2023



Copyright: © 2023 His Majesty the King in Right of Canada. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soybean (*Glycine max* (L) Merr) is a main grain legume crop grown around the world, and is primarily used for oil, food, and feed production [1]. Soybean originated and was domesticated in China 6000–9000 years ago from *Glycine soja* (Sieb. & Zucc.) [2,3]. Soybean production has increased significantly over the last five decades as a result of the expansion of the growing areas [4]. In Canada, the highest amounts of soybean production were recorded in Ontario, followed by Quebec, Manitoba, the Maritimes, and Saskatchewan, and the total production in Canada accounts for 6.54 million metric tons (MMT) of soybean in 2022 [5].

It is predicted that drought stress will be one of the world's costliest climatic problems in the near future [6]. The agricultural regions of the Canadian Prairies are particularly vulnerable to frequent drought conditions [7]. Drought stress limits soybean productivity and affects yield stability [8]. Soybean yield can decrease by more than half due to drought stress, resulting in significant losses for farmers [9]. Drought stress reduces grain legume productivity in all growth stages; however, drought stress during the reproductive and

grain filling stages causes significant yield loss [10]. This yield loss is represented 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 [11]. The plant growth stage and the duration of drought stress are important in determining the level of impact in soybean [9,12]. Drought stress induced on early maturity soybean varieties at vegetative stages results in reduced plant height, decline in seed number at the early reproductive stages, and reduced seed weight at late reproductive stages [13]. The drought conditions between flowering and early seed filling stages can influence the vegetative growth of branches and lead to a decrease in seed yield in branches [14]. For example, research in China found that drought stress at the flowering and seed filling stages reduced yield by 73 and 82%, respectively [9]. Long-term drought stress during the reproductive stages reduces biomass allocation to reproductive organs and results in lower seed yield in soybean [15].

One of the most intriguing characteristics of soybean is its ability to form a symbiotic relationship with rhizobia bacteria in root nodules. These bacteria can convert atmospheric nitrogen into ammonia inside the root nodules and, in return, the host plant provides photosynthesis products for rhizobia metabolism [16]. Symbiotic nitrogen fixation (SNF) is sensitive to different biotic and abiotic factors [16,17] 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 [1,16,18,19]. Among the different abiotic factors, drought stress is a major factor that limits SNF in soybean [16,20,21]. Drought stress affects different stages of legume-rhizobia symbiosis, such as root hair infection, nodule growth and development, and nodule function [22,23]. Furthermore, drought stress inhibits nitrogenase activity which is the key enzyme in catalyzing the reduction of dinitrogen (N_2) to ammonia (NH_3) [24]. Reduced SNF ultimately leads to reductions in grain yield, seed nitrogen, and grain protein production. Drought stress reduces SNF in nodules due to oxygen limitation in nodules, carbon scarcity, and nitrogen fixation feedback inhibition [25–27].

In the last decade, genome-wide association studies (GWAS) have become a feasible approach for discovering beneficial alleles from genetic diversity panels. GWAS have 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 [28], including 100-seed weight [29–33], seed yield [29–32,34], number of pods [29,31], pod weight [33–36], yield stability [8], number of seeds per plant [31,37], number of seeds per pod [37], seed moisture content [33], and %Ndfa [38]. However, there is a research gap in identifying alleles involved in SNF in soybean under drought conditions.

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-tolerant SNF in short-season soybean varieties. A diversity 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 yield, 100-seed weight, %Ndfa, seed nitrogen, and total seed nitrogen fixed under drought conditions to perform a GWAS analysis.

2. Materials and Methods

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 (Table S1) [39]. First, seeds were surface sterilized using 70% ethanol for two minutes, and then washed five times with autoclaved double-distilled water [40]. The professional growing mix (Sun Gro Horticultural Canada Ltd., Seba Beach, AB, Canada) and sand (Target Products Ltd., Morinville, AB, Canada) were mixed in a 3:1 ratio in 6.52 L pots (H.J.S. Wholesale Ltd., Winnipeg, MB, Canada). 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 $OD_{600} = 0.1$) [40]. 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, Smithfield, UT, USA). Plants were labeled with 25 mL of 0.5 mM $K^{15}NO_3$ solution (10 atom% ^{15}N ; 348481-25G; Sigma Aldrich, Oakville, ON, Canada) two and three weeks after planting to measure SNF.

2.2. Field Capacity of Growth Media

The determination of field capacity of growth media was completed 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 leakage of potting mixture. Then, pots were filled with a mixture of sand (Target Products Ltd., Morinville, AB, Canada) and growing mix (Sun Gro Horticultural Canada Ltd., Seba Beach, AB, 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 (D_w) was measured after drying in an oven at 80 °C until a constant weight was obtained [41]. The pots were watered until the growth media was saturated and water drained out from the bottom. The top of the pots was covered using aluminum foil to avoid evaporation. The pots were undisturbed for 24 h until no further water drainage was observed, and then the final saturated weight was recorded (S_w). 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 [42].

2.3. Drought Treatment and Yield Data Collection

After three weeks of germination, the soil moisture content of the pots was maintained at 80% FC (well-watered) and 30% FC (drought). 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 [42]. The main treatments were genotypes and moisture levels, which were allocated according to a randomized complete block design with four replicates per treatment ($n = 4$). 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.

2.4. Determination of Nitrogen Fixation-Related Parameters

The %Ndfa was measured using the isotope dilution method [43]. Seeds were oven-dried at 60°C for three days and ground to a coarse powder by using a coffee grinder. A subsample 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., Calgary, AB, Canada) using a microbalance. Samples were enveloped and compressed into a tiny pellet to make sure no air remained. The tin capsules were arranged in a 96-well plate and sent to the Stable Isotope Facility, Agriculture and Agri-Food Canada, Lethbridge Research and Development Centre to analyze ^{15}N and total N% [44]. 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, The 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 ^{15}N (USGA41A) (United States Geological Survey). The %Ndfa of the soybean was calculated using the following formula according to the isotope dilution technique:

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

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

2.5. 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 among different response variables [45].

2.6. Genotyping Data

The whole population was genotyped previously by using genotyping-by-sequencing (GBS) and whole-genome-sequencing (WGS) approaches [39,46,47] with 56 samples in common. Briefly, the SNPs were called using the Fast-GBS and Fast-WGS pipelines for GBS and WGS data, respectively [47,48], using soybean Williams 82 reference genome (Gmax_275_Wm82.a2.v1) [49]. These analyses resulted in two datasets of 56 K (GBS) and 4.3 M (WGS) SNPs. A missing genotype imputation using WGS dataset as a reference panel was previously performed using BEAGLE v4.1 [50] as described by Torkamaneh and Belzile 2021 [39,51]. The genotyping quality control and filtering measures were extensively described by Malle et al., 2020 [39]. Here, we extracted genotypic data for 103 samples from this larger collection (137 samples) and used VCFtools [52] to retain SNPs with minor allele frequency (MAF) ≥ 0.05 and heterozygosity ≤ 0.1 that resulted in a panel of 2.16 M SNPs.

2.7. Population Structure

In the panel of 2.18 M SNPs, LD-based pruning ($r^2 > 0.5$) was done with PLINK [53] to get a reduced and uniformly distributed set of 14 K markers. The fastSTRUCTURE algorithm [54] was used to characterize the population structure using tested subpopulations (K) from 1 to 13 with 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 (Supplementary Figure S1), and a scree plot (Supplementary Figure S2) was used to evaluate the most informative Principal Components (PCs) (Supplementary Figures S3 and S4). Furthermore, a Kinship matrix was calculated using the efficient mixed-model association (EMMA) method.

2.8. Genome-Wide Association Study

A GWAS was carried out with 2.16 M SNPs utilizing the fixed and random model circulation probability unification (FarmCPU) model [55] implemented in rMVP package on Microsoft R Open [56]. To reduce false positives, 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 significant associations using the false discovery rate (FDR) test [57].

2.9. Candidate Gene Identification

The soybean public database SoyBase 2020 [58] and soybean reference genome annotation were used to identify candidate genes for the yield- and nitrogen-fixation-related parameters. The quantitative trait loci (QTL) flanking areas were set to 100 kb 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 [59].

3. Results

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, and total seed nitrogen fixed ($p < 0.0001$) (Figure 1). The soil moisture content also had a significant effect on the yield parameters, grain yield, and SNF-related traits (Figure 1). In comparison to the well-watered treatment, drought stress significantly reduced the number of pods per plant (67.2 vs. 39.6) (Figure 1A), the number of seeds per plant (150.0 vs. 87.9) (Figure 1B), and grain yield (24.7 vs. 16.1 g) (Figure 1C). However, 100-seed weight was higher under 30% FC (18.5 g) compared to the well-watered 80% FC treatment (17.5 g) (Figure 1D). Regarding the 80% FC well-watered treatment, drought treatment reduced the seed total nitrogen content (1.8 vs. 1.2 g N plant⁻¹) (Figure 1E), %Ndfa (84.4 vs. 73.1%) (Figure 1F), and total seed nitrogen fixed (1.6 vs. 0.9 g N plant⁻¹) (Figure 1G).

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 2). The %Ndfa and total nitrogen fixation (g N plant⁻¹) were positively correlated with the number of pods ($r = 0.79, 0.83$), number of seeds ($r = 0.79, 0.83$), seed yield ($r = 0.85, 0.84$), and seed nitrogen content ($r = 0.86, 0.85$) under both 80% FC and 30% FC treatments. Seed yield was positively correlated with the number of pods ($r = 0.95, 0.94$), number of seeds ($r = 0.96, 0.96$), %Ndfa ($r = 0.85, 0.84$), seed nitrogen content ($r = 0.99, 0.99$), and total nitrogen fixed ($r = 0.98, 0.98$) under 80% FC and 30% FC treatments. Interestingly, the number of pods, number of seeds, seed yield, %Ndfa, seed nitrogen, and total seed nitrogen fixed were not correlated with 100-seed weight under the 30% FC treatment. The number of seeds was negatively correlated with 100-seed weight ($r = -0.65$) under the 80% FC treatment.

3.3. Genome-Wide Association of Yield- and Nitrogen Fixation-Related Traits

As reported in Seck et al., 2020, both GBS and WGS genotyping techniques were utilized in this study to cover the entire soybean genome. A subset of 14 K trimmed SNPs was used to characterize population structure. The optimal number of subpopulations (K) was six to nine, and confirmed with both principal component analysis (PCA) and phylogeny analysis. An extensive genomic SNP coverage for the panel of 103 soybean genotypes was obtained with almost 1 SNP every 455 bp (Supplementary Figure S5), perfectly suitable for GWAS analysis. GWAS analyses were performed for yield- and nitrogen fixation-related traits, and yield parameters under 30% FC and their relative performance (30% FC/80% FC) using 2.16 M SNPs and the FarmCPU statistical model. We found two SNPs were associated with %Ndfa-qNDFA-30-1, qNDFA-30-2 in 30% FC (Figure 3A), and another three SNPs were associated with %Ndfa—qNDFA-RP-1, qNDFA-RP-2, qNDFA-RP-3 in relative performance (Figure 3B) (Table 1).

3.4. Yield- and Nitrogen Fixation-Trait-Related Candidate Genes

All the genes that are residing in whole or in part within the five QTLs of interest were extracted from SoyBase. Table 2 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. We found four strong candidate genes for qNDFA-30-1, three for qNDFA-RP-1, and two for qNDFA-RP-3 based on the literature review and previous studies.

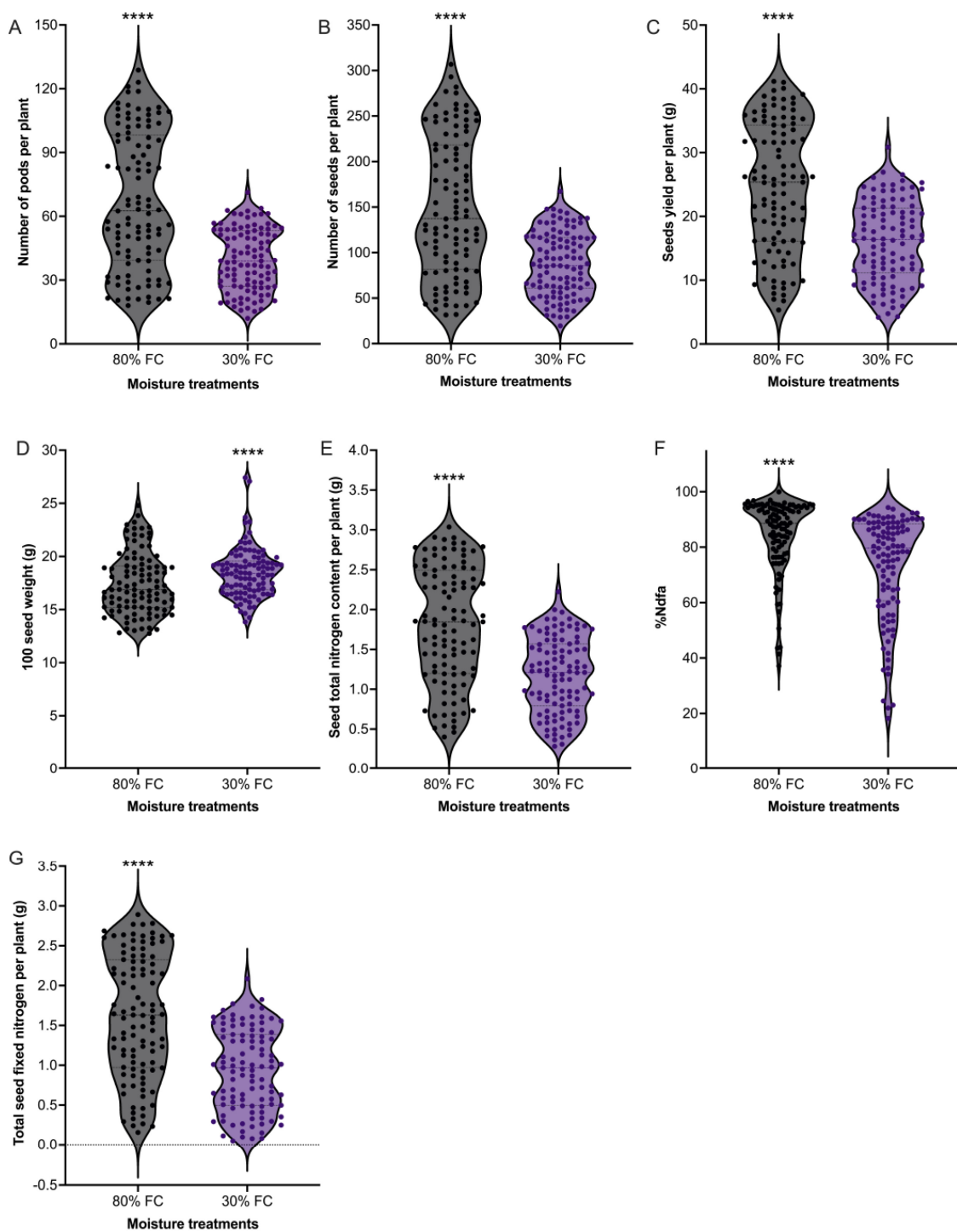


Figure 1. Distribution of yield- and SNF-related traits under well-watered and drought conditions in different soybean cultivars. (A) Number of pods per plant. (B) Number of seeds per plant. (C) Seed yield per plant. (D) 100-seed weight. (E) Seed total nitrogen content per plant. (F) Percentage nitrogen derived from the atmosphere (%Ndfa). (G) Total seed nitrogen fixed per plant. Each data point represents the mean value of a single cultivar. The asterisks (****) indicate significant differences between the 80% field capacity (FC) and 30% FC at $p < 0.0001$.

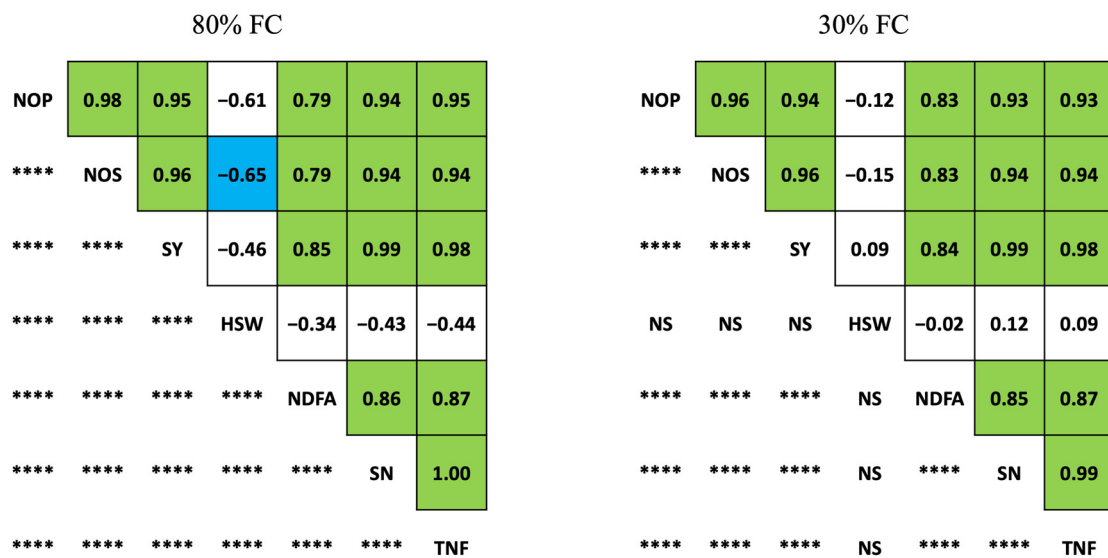


Figure 2. Correlations among yield- and nitrogen fixation-related traits in soybean under drought and well-watered conditions. Numbers above the diagonal correspond to Pearson’s correlation coefficients (r). Green boxes highlight the positive values exceeding 0.65 and blue boxes highlight the negative values exceeding or equal to 0.65 [60]. Below the diagonal show the degree of significance of the corresponding correlations between traits (**** $p < 0.0001$, 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.

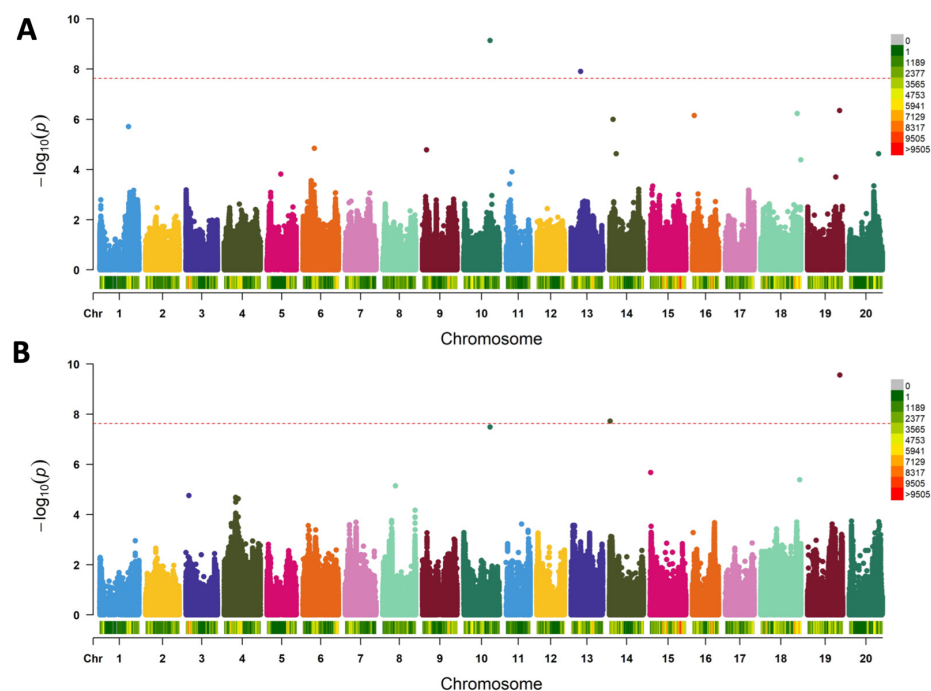


Figure 3. Manhattan plots showing significantly associated SNPs detected in genome-wide association results for percentage nitrogen derived from the atmosphere (%Ndfa) under (A) %Ndfa—30% field capacity (FC); (B) %Ndfa—relative performance. Relative performance was calculated as the ratio of %Ndfa under 30% FC vs. 80% FC. Negative $\log_{10}(p)$ -values (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. Each colored dot represents a SNP.

Table 1. List of quantitative trait loci (QTL) associated with percent nitrogen derived from the atmosphere (%Ndfa) under 30% field capacity (FC) and relative performance.

Moisture Effect	Chr Number	MSS Position	QTL ID	Minor Allele Frequency	p Value	Effect
30% FC	10	37,995,110	qNDFA-30-1	0.19	<0.0001	−6.87
	13	13,866,995	qNDFA-30-2	0.29	<0.0001	−4.79
Relative Performance	6	18,244,365	qNDFA-RP-1	0.06	<0.0001	−0.12
	14	994,141	qNDFA-RP-2	0.12	<0.0001	−0.05
	19	46,575,733	qNDFA-RP-3	0.11	<0.0001	−0.09

FC, field capacity; Chr, Chromosome; MSS, Most Significant SNP. Relative performance was calculated as the ratio of %Ndfa under 30% FC vs. 80% FC.

Table 2. Candidate genes associated with drought stress (30% field capacity) and relative performance.

Treatment	Chr Number	MSS Position	REF/ALT	QTL ID	Candidate Genes	Orthologous Genes in <i>Arabidopsis</i>	Annotations
30% FC	10	37,995,110	G/A	qNDFA-30-1	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
Relative Performance	6	18,244,365	G/A	qNDFA-RP-1	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	46,575,733	C/T	qNDFA-RP-3	Glyma.19g212800	Gene Model: AT4G02280.1	Sucrose synthase 3
					Glyma.19g213900	Gene Model: AT4G02200.1	Drought-responsive family protein

FC, field capacity; Chr number, Chromosome number; MSS, Most significant SNP. Relative performance was calculated as the ratio of %Ndfa under 30% FC vs. 80% FC.

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 (qNDFA-30-1) (Table 2). The gene *Glyma.10g145300* that encodes for galactinol synthase was also found as a candidate gene for qNDFA-30-1. The gene *Glyma.10g144300* annotated for S-adenosylmethionine synthetase 2 was also found in this region (qNDFA-30-1).

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 (qNDFA-RP-1) (Table 2). The gene *Glyma.06g198600* encodes for Ankyrin repeat family protein identified in this region. Moreover, the *Glyma.06g199700* gene encoding for Remorin family protein was recognized as an important candidate gene (qNDFA-RP-1). The gene *Glyma.19g212800* annotated as sucrose synthase 3 is another strong candidate gene for qNDFA-RP-3 (Table 2). The gene *Glyma.19g213900* encodes for drought-responsive family protein was also identified for this trait.

4. Discussion

4.1. Significant Phenotypic Variation of Yield- and Nitrogen Fixation-Related Traits in Soybean

As previously stated, we discovered significant phenotypic variability for several 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 [29], number of seeds [37], seed weight [34,61], 100-seed weight [29,30,62–64], and %Ndfa [38]. The presence of phenotypic variation within a germplasm pool for various yield- and nitrogen fixation-related traits is critical for plant breeders to make breeding selections.

Under drought stress, varieties such as DH 748, OAC Lakeview, Mario, OAC Madoc, and OAC Champion produced the most pods per plant, whereas Albions, 9004, Auriga, OAC-07-04C, and AC 2001 produced the least. 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 highest 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, and the lower in Maple presto, 90B11, Maple Donovan, 90A01, and OAC Madoc.

All the soybean cultivars were ranked based on their %Ndfa and grain yield under 30% FC (Table S2). 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 (Table S2). The soybean cultivars, which ranked higher for %Ndfa, also ranked for higher grain yield and vice versa under drought conditions (Table S2). The nitrogen content of seed was highest in DH 748, OAC Ginty, OAC Avatar, OAC Oxford, and 91M10 under drought stress, whereas Maple Donovan, 9004, Albinos, Auriga, and AC 2001 had the lowest. Total seed nitrogen fixation was highest in DH 748, OAC Ginty, OAC Avatar, OAC Oxford, and OAC Stratford, while Maple Donovan, Costaud, Albinos, 9004, and Gaillard were lowest under drought stress.

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. 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) [22]. Oxaloacetate is reduced to malate by malate dehydrogenase (MDH) regenerating NAD⁺ [65]. During drought stress, nodule sucrose synthase activity sharply declines [66], hence limiting the carbon flux required for bacteroid respiration. Drought stress directly affects nodule activity due to increased oxygen diffusion resistance and decreased nitrogenase enzyme activities, thereby affecting the metabolic activities of nitrogen-fixing rhizobia [21]. Furthermore, feedback inhibition of nitrogen fixation can also take place due to increased ureides and free amino acids in soybean plant tissues [22,67].

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 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 [29,68–70]. We found a negative correlation between the number of seeds and 100-seed weight. The negative relationship could be a result of the seed size decreasing when the number of seeds per plant increases [69,70]. Soybean seeds contain a significant amount of plant nitrogen, (~71%) [71] and it mainly derives from SNF. The strong positive correlation between the %Ndfa and seed nitrogen content under drought conditions indicates that symbiotic nitrogen fixation is crucial for seed nitrogen accumulation and protein production in drought-stressed soybean.

4.4. Genome-Wide Association Using Whole-Genome Data Revealed Five QTLs Controlling %Ndfa

In this study, we discovered five genomic regions, or QTLs, that contribute to %Ndfa under drought stress and relative performance. In the same manner, some previous GWAS in soybean have also revealed loci for %Ndfa [38]. We discovered that the majority of the QTL regions identified in our study are novel. Furthermore, the majority of GWAS in soybean were based on seed yield because this is a key trait for crop improvement. A limited number of GWAS have been conducted for the %Ndfa in soybean, specifically under drought conditions.

4.5. Candidate Genes for %Ndfa-Associated QTLs

Candidate genes under drought stress:

The gene *Glyma.10g144600* associated with %Ndfa (qNDFA-30-1) is a strong candidate gene annotated for glycogen synthase kinase-3 and was highly expressed in roots (Table 2). Under salt stress, the glycogen synthase kinase 3 (GSK3)-like kinase plays an important role in inhibiting symbiotic signaling and nodule formation in soybean [72–74]. The GSK3-like kinases directly regulate the activities of *G. max* Nodulation Signaling Pathway 1 (*GmNSP1*) proteins, facilitating legume-rhizobium symbiosis under salt stress [72]. Galactinol and raffinose act as osmoprotectants for drought-stress tolerance in plants [75] and the gene *Glyma.10g145300*, annotated for galactinol synthase, was highly expressed in roots and identified for qNDFA-30-1 (Table 2). Overexpression of galactinol synthase—which catalyzes the first step in the biosynthesis of raffinose family oligosaccharides—results in increased galactinol and raffinose accumulation and improved drought tolerance in soybean [75,76]. 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 (qNDFA-30-1) [77] (Table 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 qNDFA-RP-1 (Table 2). It is found that in soybean, C2H2 zinc finger proteins are involved in nodule development, nodule function, and nodule signal transduction [78]. The gene *Glyma.06g198600* encodes for Ankyrin repeat family protein (ANK) and was found in the same region (qNDFA-RP-1) (Table 2). It has been found that the overexpression of *GmANK114* improved the survival rate of transgenic soybean hairy roots under drought and salt stresses [79]. *GmANK114* overexpression in soybean hairy root showed higher proline, lower malondialdehyde contents, and lower H₂O₂ and O²⁻ contents in response to drought or salt stress [79]. The gene *Glyma.06g199700* encodes for Remorin family protein and is highly expressed in soybean roots (qNDFA-RP-1) (Table 2). Remorin participates in a wide range of biotic and abiotic stresses, and root nodule development [80]. *GmREM1.1*, for example, was found to be highly expressed in the nodule primordia and the inner cortex region of root nodules. Furthermore, *GmREM2.1* transcription was mostly found in rhizobia-infected cells [80]. The gene *Glyma.19g212800*, annotated as sucrose synthase 3, was found in qNDFA-RP-3 (Table 2). Sucrose synthase plays a key role in the regulation of nodule carbon metabolism [23,66,81]. Therefore, this gene will directly affect SNF in soybean, as photosynthesis and carbon supply are hampered under the drought stress. The gene *Glyma.19g213900*, which encodes for drought-responsive family protein, was found in the same QTL region (Table 2). These drought-responsive family genes are differentially expressed upon exposure to drought stress [82,83]. Importantly, drought-responsive candidate genes contribute to the development of drought-tolerant soybean cultivars [83].

4.6. Limitations

There are some limitations associated with the current study. This study was conducted under controlled environmental conditions, where plants were grown in pots using peat-

sand-based media with two moisture levels (30% FC and 80% FC). These conditions may not precisely mimic the drought conditions that occur under field conditions.

5. Conclusions

This study found significant genotypic variability among soybean genotypes in terms of the number of pods per plant, the number of seeds per plant, seed weight, 100-seed weight, %Ndfa, seed nitrogen, and total nitrogen fixed. The GWAS conducted for this research revealed five QTLs for %Ndfa 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.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants12051004/s1>, Figure S1: Phylogenetic tree constructed using a core set of 103 Canadian short-season soybean lines in TASSEL 5.0; Figure S2: Scree plot created for genotypic data using 10 principal components. The *y*-axis is the eigenvalue, and the *x*-axis is the ten principal components; Figure S3: Principal Component Analysis (PCA) scatter plot for the genotypic data using the principal component 1 (PC1) and principal component 2 (PC2). The *y*-axis is the PC1 and the *x*-axis is the PC2; Figure S4: Principal Component Analysis (PCA) scatter plot for the genotypic data using the principal component 1 (PC1) and principal component 3 (PC3). The *y*-axis is the PC1 and the *x*-axis is the PC3; Figure S5: Distribution of SNP markers across the soybean genome. Chromosomes appear horizontally with the density of SNPs depicted in the scale shown to the right; Table S1: The list of 103 short season soybean genotypes used in the study; Table S2: Ranking of soybean cultivars based on the percentage nitrogen derived from the atmosphere (%Ndfa) and grain yield under 30% field capacity.

Author Contributions: Conceptualization, D.K.L., D.T. and M.S.T.; methodology, D.K.L., D.T. and M.S.T.; investigation, D.K.L.; resources, M.S.T., D.T. and F.B.; data curation, D.K.L.; writing—original draft preparation, D.K.L.; review and editing, D.K.L., M.S.T., D.T., F.B. and P.B.; isotope analysis, P.B. and B.H.; supervision, M.S.T.; project administration, M.S.T.; funding acquisition, M.S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2020-04425) to MST.

Data Availability Statement: The data presented in this study (mean values) are available in Supplementary File S1.

Acknowledgments: We thank Ishan Chathuranga, Kelley Dunfield, and Shay Missiaen (University of Alberta) for their help with the greenhouse trial and Nina Mori for editorial assistance.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Herridge, D.F.; Peoples, M.B.; Boddey, R.M. Global Inputs of Biological Nitrogen Fixation in Agricultural Systems. *Plant Soil* **2008**, *311*, 1–18. [[CrossRef](#)]
2. Guo, J.; Wang, Y.; Song, C.; Zhou, J.; Qiu, L.; Huang, H.; Wang, Y. A Single Origin and Moderate Bottleneck during Domestication of Soybean (*Glycine Max*): Implications from Microsatellites and Nucleotide Sequences. *Ann. Bot.* **2010**, *106*, 505–514. [[CrossRef](#)]
3. Wang, J.; Chu, S.; Zhang, H.; Zhu, Y.; Cheng, H.; Yu, D. Development and Application of a Novel Genome-Wide SNP Array Reveals Domestication History in Soybean. *Nature* **2016**, *6*, 20728. [[CrossRef](#)]
4. Liu, S.; Zhang, M.; Feng, F.; Tian, Z. Toward a “Green Revolution” for Soybean. *Mol. Plant* **2020**, *13*, 688–697. [[CrossRef](#)]
5. SOY Canada 2022. Available online: <https://soycanada.ca/> (accessed on 5 January 2023).
6. Wilhite, D.A. Drought as a Natural Hazard: Concepts and Definitions. In *Drought Mitigation Center Faculty Publications*; Routledge: London, UK, 2000; pp. 1–21.
7. Bonsor, B.; Regier, M. Historical Comparison of the 2001/2002 Drought in the Canadian Prairies. *Clim. Res.* **2007**, *33*, 229–242. [[CrossRef](#)]
8. Quero, G.; Simondi, S.; Ceretta, S.; Otero, Á.; Garaycochea, S.; Fernández, S.; Borsani, O.; Bonnacarrère, V. An Integrative Analysis of Yield Stability for a GWAS in a Small Soybean Breeding Population. *Crop Sci.* **2021**, *61*, 1903–1914. [[CrossRef](#)]

9. Wei, Y.; Jin, J.; Jiang, S.; Ning, S.; Liu, L. Quantitative Response of Soybean Development and Yield to Drought Stress during Different Growth Stages in the Huaibei Plain, China. *Agronomy* **2018**, *8*, 97. [[CrossRef](#)]
10. Farooq, M.; Gogoi, N.; Barthakur, S.; Baroowa, B.; Bharadwaj, N.; Alghamdi, S.S.; Siddique, K.H.M. Drought Stress in Grain Legumes during Reproduction and Grain Filling. *J. Agron. Crop Sci.* **2017**, *203*, 81–102. [[CrossRef](#)]
11. Arya, H.; Singh, M.B.; Bhalla, P.L. Towards Developing Drought-Smart Soybeans. *Front. Plant Sci.* **2021**, *12*, 750664. [[CrossRef](#)] [[PubMed](#)]
12. Dong, S.; Jiang, Y.; Dong, Y.; Wang, L.; Wang, W.; Ma, Z.; Yan, C.; Ma, C.; Liu, L. A Study on Soybean Responses to Drought Stress and Rehydration. *Saudi J. Biol. Sci.* **2019**, *26*, 2006–2017. [[CrossRef](#)]
13. Desclaux, D.; Huynh, T.T.; Roumet, P. Identification of Soybean Plant Characteristics That Indicate the Timing of Drought Stress. *Crop Sci.* **2000**, *40*, 716–722. [[CrossRef](#)]
14. Frederick, J.R.; Camp, C.R.; Bauer, P.J. Crop Ecology, Production & Management: Drought-Stress Effects on Branch and Mainstem Seed Yield and Yield Components of Determinate Soybean. *Crop Sci.* **2001**, *41*, 759–763. [[CrossRef](#)]
15. Du, Y.; Zhao, Q.; Chen, L.; Yao, X.; Xie, F. Effect of Drought Stress at Reproductive Stages on Growth and Nitrogen Metabolism in Soybean. *Agronomy* **2020**, *10*, 302. [[CrossRef](#)]
16. Thilakarathna, M.S.; Raizada, M.N. A Meta-Analysis of the Effectiveness of Diverse Rhizobia Inoculants on Soybean Traits under Field Conditions. *Soil Biol. Biochem.* **2017**, *105*, 177–196. [[CrossRef](#)]
17. Hungria, M.; Vargas, M.A.T. Environmental Factors Affecting N₂ Fixation in Grain Legumes in the Tropics, with an Emphasis on Brazil. *Field Crops Res.* **2000**, *65*, 151–164. [[CrossRef](#)]
18. Hungria, M.; Bohrer, T.R.J. Variability of Nodulation and Dinitrogen Fixation Capacity among Soybean Cultivars. *Biol. Fertil. Soils* **2000**, *31*, 45–52. [[CrossRef](#)]
19. Santachiara, G.; Salvagiotti, F.; Rotundo, J.L. Nutritional and Environmental Effects on Biological Nitrogen Fixation in Soybean: A Meta-Analysis. *Field Crops Res.* **2019**, *240*, 106–115. [[CrossRef](#)]
20. Serraj, R.; Sinclair, T.R.; Purcell, L.C. Symbiotic N₂ Fixation Response to Drought. *J. Exp. Bot.* **1999**, *50*, 143–155. [[CrossRef](#)]
21. Serraj, R. Effects of Drought Stress on Legume Symbiotic Nitrogen Fixation: Physiological Mechanisms. *Indian J. Exp. Biol.* **2003**, *41*, 1136–1141.
22. Sulieman, S.; Tran, L.S.P. *Legume Nitrogen Fixation in a Changing Environment: Achievements and Challenges*; Springer International Publishing: Cham, Switzerland, 2016; ISBN 9783319062129.
23. Arrese-Igor, C.; Gonzalez, E.M.; Anthony, J.; Minchin, F.R.; Galvez, L.; Cabrerizol, P.M.; Aricio-tejolo, P.M.A.P. Sucrose Synthase and Nodule Nitrogen Fixation under Drought and Other Environmental Stresses. *Symbiosis* **1999**, *27*, 189–212.
24. Einsle, O.; Rees, D.C. Structural Enzymology of Nitrogenase Enzymes. *Chem. Rev.* **2020**, *120*, 4969–5004. [[CrossRef](#)] [[PubMed](#)]
25. Arrese-Igor, C.; Esther González, B.M.; Daniel Marino, B.; Rubén Ladrera, B.; Estibaliz Larrainzar, B.; Erena Gil-Quintana, B. Physiological Responses of Legume Nodules to Drought. *Plant Stress* **2011**, *5*, 24–31.
26. Collier, R.; Tegeder, M. Soybean Ureide Transporters Play a Critical Role in Nodule Development, Function and Nitrogen Export. *Plant J.* **2012**, *72*, 355–367. [[CrossRef](#)] [[PubMed](#)]
27. Kunert, K.J.; Vorster, B.J.; Fenta, B.A.; Kibido, T.; Giuseppe, D.; Foyer, C.H. Drought Stress Responses in Soybean Roots and Nodules. *Front. Plant Sci.* **2016**, *7*, 1015. [[CrossRef](#)]
28. Torkamaneh, D.; Chalifour, F.P.; Beauchamp, C.J.; Agrama, H.; Boahen, S.; Maaroufi, H.; Rajcan, I.; Belzile, F. Genome-Wide Association Analyses Reveal the Genetic Basis of Biomass Accumulation under Symbiotic Nitrogen Fixation in African Soybean. *Theor. Appl. Genet.* **2020**, *133*, 665–676. [[CrossRef](#)]
29. Hu, D.; Zhang, H.; Du, Q.; Hu, Z.; Yang, Z.; Li, X.; Wang, J.; Huang, F.; Yu, D.; Wang, H.; et al. Genetic Dissection of Yield-Related Traits via Genome-Wide Association Analysis across Multiple Environments in Wild Soybean (*Glycine soja* Sieb. and Zucc.). *Planta* **2020**, *251*, 39. [[CrossRef](#)]
30. Contreras-Soto, R.I.; Mora, F.; De Oliveira, M.A.R.; Higashi, W.; Scapim, C.A.; Schuster, I. A Genome-Wide Association Study for Agronomic Traits in Soybean Using SNP Markers and SNP-Based Haplotype Analysis. *PLoS ONE* **2017**, *12*, 0171105. [[CrossRef](#)]
31. Adeboye, K.A.; Akhter Bhat, J.; Ganie, S.A.; Varshney, R.K.; Yu, D. Genome-Wide Association Analysis of Yield-Related Traits of Soybean Using Haplotype-Based Framework. *Biol. Life Sci. Forum* **2021**, *11*, 49.
32. Priyanatha, C.; Torkamaneh, D.; Rajcan, I. 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.* **2022**, *13*, 823. [[CrossRef](#)]
33. Li, X.; Zhou, Y.; Bu, Y.; Wang, X.; Zhang, Y.; Guo, N.; Zhao, J.; Xing, H. Genome-Wide Association Analysis for Yield-Related Traits at the R6 Stage in a Chinese Soybean Mini Core Collection. *Genes Genom.* **2021**, *43*, 897–912. [[CrossRef](#)]
34. Ravelombola, W.; Qin, J.; Shi, A.; Song, Q.; Yuan, J.; Wang, F.; Chen, P.; Yan, L.; Feng, Y.; Zhao, T.; et al. Genome-Wide Association Study and Genomic Selection for Yield and Related Traits in Soybean. *PLoS ONE* **2021**, *16*, e0255761. [[CrossRef](#)] [[PubMed](#)]
35. Zhang, W.; Xu, W.; Zhang, H.; Liu, X.; Cui, X.; Li, S.; Song, L.; Zhu, Y.; Chen, X.; Chen, H. Comparative Selective Signature Analysis and High-Resolution GWAS Reveal a New Candidate Gene Controlling Seed Weight in Soybean. *Theor. Appl. Genet.* **2021**, *134*, 1329–1341. [[CrossRef](#)] [[PubMed](#)]
36. Jo, H.; Lee, J.Y.; Lee, J.D. Genome-Wide Association Mapping for Seed Weight in Soybean with Black Seed Coats and Green Cotyledons. *Agronomy* **2022**, *12*, 250. [[CrossRef](#)]

37. Chen, L.; Fang, Y.; Li, X.; Zeng, K.; Chen, H.; Zhang, H.; Yang, H.; Cao, D.; Hao, Q.; Yuan, S.; et al. Identification of Soybean Drought-Tolerant Genotypes and Loci Correlated with Agronomic Traits Contributes New Candidate Genes for Breeding. *Plant Mol. Biol.* **2020**, *102*, 109–122. [[CrossRef](#)]
38. Dhanapal, A.P.; Ray, J.D.; Singh, S.K.; Hoyos-Villegas, V.; Smith, J.R.; Purcell, L.C.; King, C.A.; Fritsch, F.B. Genome-Wide Association Analysis of Diverse Soybean Genotypes Reveals Novel Markers for Nitrogen Traits. *Plant Genome* **2015**, *8*, 1–15. [[CrossRef](#)]
39. Malle, S.; Eskandari, M.; Morrison, M.; Belzile, F. Genome-Wide Association Identifies Several QTLs Controlling Cysteine and Methionine Content in Soybean Seed Including Some Promising Candidate Genes. *Sci. Rep.* **2020**, *10*, 21812. [[CrossRef](#)]
40. Thilakarathna, M.S.; Moroz, N.; Raizada, M.N. 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.* **2017**, *8*, 1714. [[CrossRef](#)]
41. Earl, H.J. A Precise Gravimetric Method for Simulating Drought Stress in Pot Experiments. *Crop Sci.* **2003**, *43*, 1868–1873. [[CrossRef](#)]
42. Liyanage, D.K.; Chathuranga, I.; Mori, B.A.; Thilakarathna, M.S. A Simple, Semi-Automated, Gravimetric Method to Simulate Drought Stress on Plants. *Agronomy* **2022**, *12*, 349. [[CrossRef](#)]
43. Thilakarathna, M.S.; Papadopoulos, Y.A.; Rodd, A.V.; Grimmett, M.; Fillmore, S.A.E.; Crouse, M.; Prithiviraj, B. Nitrogen Fixation and Transfer of Red Clover Genotypes under Legume–Grass Forage Based Production Systems. *Nutr. Cycl. Agroecosyst.* **2016**, *106*, 233–247. [[CrossRef](#)]
44. Thilakarathna, M.S.; Torkamaneh, D.; Bruce, R.W.; Rajcan, I.; Chu, G.; Grainger, C.M.; Szczyglowski, K.; Hill, B.; Raizada, M.N. Testing Whether Pre-Pod-Fill Symbiotic Nitrogen Fixation in Soybean Is Subject to Drift or Selection over 100 Years of Soybean Breeding. *Front. Agron.* **2021**, *3*, 725813. [[CrossRef](#)]
45. R Core Team. *R: A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2021.
46. Sonah, H.; Bastien, M.; Iquira, E.; Tardivel, A.; Légaré, G.; Boyle, B.; Normandeau, É.; Laroche, J.; Larose, S.; Jean, M.; et al. An Improved Genotyping by Sequencing (GBS) Approach Offering Increased Versatility and Efficiency of SNP Discovery and Genotyping. *PLoS ONE* **2013**, *8*, e54603. [[CrossRef](#)] [[PubMed](#)]
47. Torkamaneh, D.; Laroche, J.; Tardivel, A.; O'Donoghue, L.; Cober, E.; Rajcan, I.; Belzile, F. Comprehensive Description of Genomewide Nucleotide and Structural Variation in Short-Season Soya Bean. *Plant Biotechnol. J.* **2018**, *16*, 749–759. [[CrossRef](#)] [[PubMed](#)]
48. Torkamaneh, D.; Laroche, J.; Bastien, M.; Abed, A.; Belzile, F. Fast-GBS: A New Pipeline for the Efficient and Highly Accurate Calling of SNPs from Genotyping-by-Sequencing Data. *BMC Bioinform.* **2017**, *18*, 5. [[CrossRef](#)] [[PubMed](#)]
49. Schmutz, J.; Cannon, S.B.; Schlueter, J.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D.L.; Song, Q.; Thelen, J.J.; Cheng, J.; et al. Genome Sequence of the Palaeopolyploid Soybean. *Nature* **2010**, *463*, 178–183. [[CrossRef](#)]
50. Browning, B.L.; Browning, S.R. Genotype Imputation with Millions of Reference Samples. *Am. J. Hum. Genet.* **2016**, *98*, 116–126. [[CrossRef](#)]
51. Torkamaneh, D.; Belzile, F. Accurate Imputation of Untyped Variants from Deep Sequencing Data. *Deep. Seq. Data Anal.* **2021**, *2243*, 271–281.
52. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The Variant Call Format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)]
53. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
54. Raj, A.; Stephens, M.; Pritchard, J.K. FastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics* **2014**, *197*, 573–589. [[CrossRef](#)]
55. Liu, X.; Huang, M.; Fan, B.; Buckler, E.S.; Zhang, Z. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *Public Libr. Sci.* **2016**, *12*, 1–24. [[CrossRef](#)] [[PubMed](#)]
56. Yin, L.; Zhang, H.; Tang, Z.; Xu, J.; Yin, D.; Zhang, Z.; Yuan, X.; Zhu, M.; Zhao, S.; Li, X.; et al. RMVP: A Memory-Efficient, Visualization-Enhanced, and Parallel-Accelerated Tool for Genome-Wide Association Study. *Genom. Proteom. Bioinform.* **2021**, *19*, 619–628. [[CrossRef](#)] [[PubMed](#)]
57. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc.* **1995**, *57*, 289–300. [[CrossRef](#)]
58. Grant, D.; Nelson, R.T.; Cannon, S.B.; Shoemaker, R.C. SoyBase, the USDA-ARS Soybean Genetics and Genomics Database. *Nucleic Acids Res.* **2009**, *38*, D843–D846. [[CrossRef](#)]
59. Waese, J.; Fan, J.; Pasha, A.; Yu, H.; Fucile, G.; Shi, R.; Cumming, M.; Kelley, L.A.; Sternberg, M.J.; Krishnakumar, V.; et al. ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* **2017**, *29*, 1806–1821. [[CrossRef](#)]
60. Seck, W.; Torkamaneh, D.; Belzile, F. Comprehensive Genome-Wide Association Analysis Reveals the Genetic Basis of Root System Architecture in Soybean. *Front. Plant Sci.* **2020**, *11*, 590740. [[CrossRef](#)]

61. Yan, L.; Hofmann, N.; Li, S.; Ferreira, M.E.; Song, B.; Jiang, G.; Ren, S.; Quigley, C.; Fickus, E.; Cregan, P.; et al. Identification of QTL with Large Effect on Seed Weight in a Selective Population of Soybean with Genome-Wide Association and Fixation Index Analyses. *BMC Genom.* **2017**, *18*, 529. [[CrossRef](#)]
62. Zhang, J.; Song, Q.; Cregan, P.B.; Jiang, G.L. Genome-Wide Association Study, Genomic Prediction and Marker-Assisted Selection for Seed Weight in Soybean (*Glycine Max*). *Theor. Appl. Genet.* **2016**, *129*, 117–130. [[CrossRef](#)]
63. Sonah, H.; O'Donoghue, L.; Cober, E.; Rajcan, I.; Belzile, F. Identification of Loci Governing Eight Agronomic Traits Using a GBS-GWAS Approach and Validation by QTL Mapping in Soya Bean. *Plant Biotechnol. J.* **2015**, *13*, 211–221. [[CrossRef](#)]
64. Copley, T.R.; Duceppe, M.O.; O'Donoghue, L.S. Identification of Novel Loci Associated with Maturity and Yield Traits in Early Maturity Soybean Plant Introduction Lines. *BMC Genom.* **2018**, *19*, 167. [[CrossRef](#)]
65. Vance, C.P.; Gantt, S.J. Control of Nitrogen and Carbon Metabolism in Root Nodules. *Physiol. Plant* **1992**, *85*, 266–274. [[CrossRef](#)]
66. Gonzalez, E.M.; Gordon, A.J.; James, C.L. The Role of Sucrose Synthase in the Response of Soybean Nodules to Drought. *J. Exp. Bot.* **1995**, *46*, 1515–1523. [[CrossRef](#)]
67. King, C.A.; Purcell, L.C. Inhibition of N₂ Fixation in Soybean Is Associated with Elevated Ureides and Amino Acids. *Plant Physiol.* **2005**, *137*, 1389–1396. [[CrossRef](#)]
68. Mustofa, A.; Zubaidah, S.; Kuswanto, H. Correlation and Path Analysis on Yield and Yield Components in Segregating Populations. *Genetika* **2021**, *53*, 157–166. [[CrossRef](#)]
69. Hao, D.; Cheng, H.; Yin, Z.; Cui, S.; Zhang, D.; Wang, H.; Yu, D. Identification of Single Nucleotide Polymorphisms and Haplotypes Associated with Yield and Yield Components in Soybean (*Glycine Max*) Landraces across Multiple Environments. *Theor. Appl. Genet.* **2012**, *124*, 447–458. [[CrossRef](#)]
70. Silva, A.F.; Sediya, T.; Silva, F.C.S.; Bezerra, A.R.G.; Ferreira, L.V. Correlation and Path Analysis of Yield Components in Soybean Varieties. *Turk. J. Field Crops* **2015**, *10*, 177–179.
71. Ciampitti, I.A.; Salvaggiotti, F. New Insights into Soybean Biological Nitrogen Fixation. *Agron. J.* **2018**, *110*, 1185–1196. [[CrossRef](#)]
72. He, C.; Gao, H.; Wang, H.; Guo, Y.; He, M.; Peng, Y.; Wang, X. GSK3-Mediated Stress Signaling Inhibits Legume–Rhizobium Symbiosis by Phosphorylating *GmNSP1* in Soybean. *Mol. Plant* **2021**, *14*, 488–502. [[CrossRef](#)]
73. Singh, J.; Verma, P.K. NSP1 Allies with GSK3 to Inhibit Nodule Symbiosis. *Trends Plant Sci.* **2021**, *26*, 999–1001. [[CrossRef](#)]
74. Phang, T.-H.; Li, M.-W.; Cheng, C.-C.; Wong, F.-L.; Chan, C.; Lam, H.-M. Molecular Responses to Osmotic Stresses in Soybean. In *Soybean Molecular Aspects of Breeding*; InTech: Rijeka, Croatia, 2011; pp. 215–240. [[CrossRef](#)]
75. Taji, T.; Ohsumi, C.; Iuchi, S.; Seki, M.; Kasuga, M.; Kobayashi, M.; Yamaguchi-Shinozaki, K.; Shinozaki, K. Important Roles of Drought- and Cold-Inducible Genes for Galactinol Synthase in Stress Tolerance in *Arabidopsis thaliana*. *Plant J.* **2002**, *29*, 417–426. [[CrossRef](#)]
76. Bellaloui, N.; Hu, Y.; Mengistu, A.; Kassem, M.A.; Abel, C.A. Effects of Foliar Boron Application on Seed Composition, Cell Wall Boron, and Seed $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ Isotopes in Water-Stressed Soybean Plants. *Front. Plant Sci.* **2013**, *4*, 270. [[CrossRef](#)] [[PubMed](#)]
77. Wang, X.; Oh, M.W.; Komatsu, S. Characterization of S-Adenosylmethionine Synthetases in Soybean under Flooding and Drought Stresses. *Biol. Plant* **2016**, *60*, 269–278. [[CrossRef](#)]
78. Yuan, S.; Li, X.; Li, R.; Wang, L.; Zhang, C.; Chen, L.; Hao, Q.; Zhang, X.; Chen, H.; Shan, Z.; et al. Genome-Wide Identification and Classification of Soybean C2H2 Zinc Finger Proteins and Their Expression Analysis in Legume–Rhizobium Symbiosis. *Front. Microbiol.* **2018**, *9*, 126. [[CrossRef](#)] [[PubMed](#)]
79. Zhao, J.Y.; Lu, Z.W.; Sun, Y.; Fang, Z.W.; Chen, J.; Zhou, Y.B.; Chen, M.; Ma, Y.Z.; Xu, Z.S.; Min, D.H. The Ankyrin-Repeat Gene *GmANK114* Confers Drought and Salt Tolerance in *Arabidopsis* and Soybean. *Front. Plant Sci.* **2020**, *11*, 584167. [[CrossRef](#)]
80. Son, S.; Oh, C.J.; Bae, J.H.; Lee, H.; An, C.S. GmREM1.1 and GmREM2.1, Which Encode the Remorin Proteins in Soybean, Have Distinct Roles during Root Nodule Development. *J. Plant Biol.* **2015**, *58*, 17–25. [[CrossRef](#)]
81. Gordon, A.J.; Minchin, F.R.; Skøt, L.; James, C.L. Stress-Induced Declines in Soybean N₂ Fixation Are Related to Nodule Sucrose Synthase Activity. *Plant Physiol.* **1997**, *114*, 937–946. [[CrossRef](#)]
82. Marcolino-Gomes, J.; Rodrigues, F.A.; Fuganti-Pagliarini, R.; Bendix, C.; Nakayama, T.J.; Celaya, B.; Molinari, H.B.C.; De Oliveira, M.C.N.; Harmon, F.G.; Nepomuceno, A. Diurnal Oscillations of Soybean Circadian Clock and Drought Responsive Genes. *PLoS ONE* **2014**, *9*, e86402. [[CrossRef](#)]
83. Le, D.T.; Nishiyama, R.; Watanabe, Y.; Tanaka, M.; Seki, M.; Ham, L.H.; Yamaguchi-Shinozaki, K.; Shinozaki, K.; Tran, L.S.P. Differential Gene Expression in Soybean Leaf Tissues at Late Developmental Stages under Drought Stress Revealed by Genome-Wide Transcriptome Analysis. *PLoS ONE* **2012**, *7*, e495. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.