agronomy

MDPI

Article

Genome-Wide Association Studies of Plant Architecture-Related Traits in the Chinese Soybean Mini Core Collection

Wentao Ding [1], Xiaoli Zhang [1,2], Dandan Liu [1], Chen Li [1], Congcong Wang [1], Ruidong Sun [1], Xiangpei Jin [1], Na Guo [1], Jinming Zhao [1,*] and Han Xing [1,*]

[1] National Center for Soybean Improvement, Key Laboratory of Biology and Genetics and Breeding for Soybean, Ministry of Agriculture, State Key Laboratory for Crop Genetics and Germplasm Enhancement, College of Agriculture, Nanjing Agricultural University, Nanjing 210095, China; wentaoding@126.com (W.D.); xiaolizhang625@163.com (X.Z.); 2019101153@njau.edu.cn (D.L.); 2021801247@stu.njau.edu.cn (C.L.); 2018201069@njau.edu.cn (C.W.); 2018201070@njau.edu.cn (R.S.); 2020801242@stu.njau.edu.cn (X.J.); guona@njau.edu.cn (N.G.)

[2] Institute of Crop and Nuclear Technology Utilization, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China

* Correspondence: jmz3000@126.com (J.Z.); hanx@njau.edu.cn (H.X.); Tel.: +86-025-8439-5795 (J.Z.); +86-025-8439-9526 (H.X.)

Abstract: Plant architecture traits are closely related to plant biomass, lodging, and photosynthetic efficiency, which in turn affect soybean yield. In this study, we investigated a Chinese soybean mini core collection consisting of 224 germplasm accessions for four plant architecture-related traits (plant height (PH), number of nodes on main stem (NN), branch number (BN), and stem diameter (DI)) under three environments and conducted a genome-wide association study (GWAS) based on 1514 single nucleotide polymorphisms (SNPs). A total of 41 SNPs were found to be significantly associated with PH, NN, BN, and DI in two or more environments. Among these SNPs, 15 were located in regions in which plant architecture-related QTLs had been reported in previous studies, and 26 were new genetic loci. In addition, 18 potential candidate genes for plant architecture-related traits were obtained by predicting the genes in the interval of four large-effect markers (BARC-017097-02199, Map-2213, BARC-014639-01604, and Map-2223). This research will help to illuminate the genetic basis of soybean plant architecture-related traits and accelerate the process of plant architecture breeding by molecular marker-assisted selection in soybean.

Keywords: soybean (*Glycine max* (L.) Merr.); plant architecture-related traits; genome-wide association study (GWAS); candidate genes

Citation: Ding, W.; Zhang, X.; Liu, D.; Li, C.; Wang, C.; Sun, R.; Jin, X.; Guo, N.; Zhao, J.; Xing, H. Genome-Wide Association Studies of Plant Architecture-Related Traits in the Chinese Soybean Mini Core Collection. *Agronomy* **2022**, *12*, 817. https://doi.org/10.3390/agronomy12040817

Academic Editor: Diego Rubiales

Received: 18 February 2022
Accepted: 25 March 2022
Published: 28 March 2022

1. Introduction

Soybean (*Glycine max* (L.) Merr.) is one of the most important oil seed crops in the world, providing high-quality vegetable oil and protein for human beings and playing a vital role in solving the problem of insufficient protein supply [1,2]. It is therefore particularly important to improve soybean yield. Ideotype breeding of soybean can increase yield by reducing the strong intragroup competition between soybean reproductive and vegetative growth and fleshy pods [3,4]. This could improve the adaptability of crops to the environment by preventing lodging, optimizing canopy architecture, determining light distribution, and enhancing photosynthesis [5,6]. The plant height (PH), number of nodes on main stem (NN), branch number (BN), and stem diameter (DI) are important plant architecture-related traits of soybean that are highly important for improving biological yield [7]. Therefore, ideotype breeding is an effective way to achieve a breakthrough in soybean yield.

Plant architecture-related traits are complex quantitative traits affected by multiple quantitative trait loci (QTLs) and environmental factors, and combining dense molecular

markers and genomic information is an effective method for genetic research [8–10]. Although some progress in traditional QTL linkage mapping has been made in soybean plant architecture, there are certain limitations in biparental segregation populations, including the limited allelic variation in each cross, time consumption for mapping population construction, and limited mapping resolution resulting in large confidence intervals for the detected loci [11,12]. In addition, well-established genome sequences will increase the understanding of the genetic basis of these traits, thereby promoting the development of soybean marker-assisted breeding [13].

Genome-wide association analysis (GWAS) overcomes the limitations of traditional biparental QTL mapping to enable more accurate and effective QTL mapping. Based on the linkage disequilibrium relationship between the loci within the population and the identification of the association between genetic markers and the target trait phenotype [14], its wide application has benefited from the development of next-generation sequencing (NGS) and high-density molecular markers.

In recent years, GWAS has been applied to the analysis of complex quantitative traits of model organisms and crops [15–18]. Currently, GWAS has been successfully applied to research on plant architecture-related traits in soybean, and a large number of high-precision loci have been identified. For example, 27 loci associated with plant height have been identified from 309 early maturing soybean germplasms, and a major effect locus on Gm 19, *PH24*, explained 10% of phenotypic variation [19]. Fang et al. (2017) used 809 soybean accessions to perform GWAS on 84 agronomic traits and identified 245 important loci, of which 95 genetically interacted with other loci, revealing the genetic network of associated loci among different soybean traits [18]. At different developmental stages of the soybean, 34 and 30 quantitative trait nucleotides (QTNs) were found to be associated with PH and NN, respectively, using the mrMLM model [20]. Five QTNs have been shown to be related to BN, and the TCP transcription factor GmBRC1 in the LD block of *qtnBR6-1* may regulate the branches' development [21]. Chen et al. (2020) detected 114 SNPs associated with NN on chromosomes 1, 3, 6, and 14 as well as 39 SNPs related to DI on chromosomes 4, 9, and 19 under well-watered and drought conditions [22]. Based on a restricted two-stage multilocus genome-wide association study (RTM-GWAS), 76 QTLs and 183 alleles of NN were found in 306 soybean materials [23]. Using RILs consisting of 427 lines, 89 NN QTLs were detected with a phenotypic contribution of 84.50%, of which 58 were new QTLs [24]. In total, 12, 12, 4, and 4 SNPs associated with PH, NN, BN, and DI, respectively, were repeatedly detected in at least two environments through 133 soybean landraces, among which seven SNPs had particularly strong effects [7]. Although a large number of soybean plant architecture-related loci have been identified, the genetic basis of plant architecture regulation has not been fully elucidated due to the complexity of its genetic mechanism. Therefore, there may be new loci and genes for plant architecture-related traits that have not been discovered in soybean. Ideotype is an important way to improve soybean yield. Therefore, elucidating the genetic mechanism of soybean plant architecture-related traits can lay the foundation for molecular design breeding of soybean ideotype and genetic improvement of varieties.

In this study, we used a Chinese soybean mini core collection of 224 germplasm accessions, which has been used in the study of soybean phytophthora root and stem rot (PRR) [25], seed hardness [26], and yield-related traits at the R6 stage of soybean [27]. This study is a supplement to the research on plant architecture-related traits in these collections and enriches the research scope of this population. Combined with 1514 high-quality SNPs, GWAS was carried out for four plant architecture-related traits (PH, NN, BN, and DI) under three environments for two consecutive years. This study aims to exploit more new loci and genes to expand the regulatory loci of architecture-related traits in soybean, and this will be used to reveal the genetic mechanism of soybean plant architecture-related traits more comprehensively.

2. Materials and Methods

2.1. Plant Materials

The plant materials in this study were 224 soybean mini core accessions selected from 23,587 soybean germplasm resources in the Chinese National Soybean GeneBank (CNSGB), which has abundant genetic diversity [25]. According to the geographical source of the materials, the 224 soybean mini core collections originated from four soybean ecological regions in China: the northeast soybean ecological region, the northern soybean ecological region, the Huanghuaihai region, and the southern soybean ecoregion. These were the origin of 39, 37, 48, and 100 materials, respectively. This set of materials was widely distributed in many provinces and autonomous regions in China.

2.2. Field Trials and Trait Measurement

The field trials were conducted at the Jiangpu Experimental Station of Nanjing Agricultural University (32°12′ N, 118°37′ E) in 2016 and 2017 and at the Anhui Dangtu Experimental Station (31°55′ N, 118°49′ E) in 2017. The three environments were identified as 2016JP, 2017JP, and 2017DT. The field trial was designed in a completely randomized block design with three blocks in each environment. Each material was planted in a row with a length of 4 m and a row spacing of 0.5 m. The seedlings were planted as 40 plants per row, with a plant spacing of 10 cm. A mixture of 2.25 L of 48% Trifluralin EC and 525 kg water was applied per hectare, sprayed evenly on the soil surface, and mixed in the soil for 1–3 cm to suppress weed growth after land consolidation. Before sowing, the special compound fertilizer was applied for soybean (the rate of N-P_2O_5-K_2O was 13:25:12), the application rate was 375 kg/hm^2, and no more topdressing was performed. There were few pests in the three environments, therefore, we added 150 mL of 20% Chlorantraniliprole SC into 450 kg water per hectare according to the pest situation in the field and sprayed only once at the pod stage. Other field management measures were consistently performed during soybean growth to minimize test error.

Five plants were randomly selected from each block under the three environments 2016JP, 2017JP, and 2017DT to examine plant height, number of nodes on main stem, branch number, and stem diameter at the maturity period. For soybean plant architecture, plant height is the length from the cotyledon node to the top of the main stem (PH, measured in cm); the nodes of the main stem indicate the number of nodes from the cotyledonary node to the top of the main stem (NN); the branch numbers indicate the effective branches of the main stem (BN) [28]; and stem diameter is measured at the third internode diameter on the main stem from the cotyledon node with a micrometer (DI, measured in mm) [10].

2.3. Phenotypic Data Analysis

IBM SPSS Statistics v20.0 was used to perform descriptive statistical analysis on the four plant architecture-related traits, and the GLM process in SAS 9.4 [29] was used to perform joint variance analysis on the phenotype data under three environment conditions. Based on the average of the four plant architecture-related traits under the three environments, the CORR process was used to calculate the Pearson correlation coefficient between plant architecture-related traits. The PROC VARCOMP method was used to calculate the variance components of the four plant architecture-related traits in genotype, environment, and genotype–environment interaction, and the estimated heritability calculation formula was as follows [30]:

$$h^2 = \frac{\sigma_g^2}{\left(\sigma_g^2 + \frac{\sigma_{ge}^2}{n} + \frac{\sigma_e^2}{rn}\right)}$$

where σ_g^2 represents the genotype variance, σ_e^2 represents the error, σ_{ge}^2 represents the genotype and environment interaction variance, r represents the number of repeats, and n represents the number of environments.

2.4. Genotyping and Linkage Disequilibrium (LD) Analysis

The SNP markers used in this study included 554 SNPs related to genetic diversity in wild and cultivated soybeans and 1536 SNPs for fatty acid composition in soybean grains [31,32]. Through the integration of the two groups, 1654 SNPs were obtained. To improve the efficiency of association analysis, TASSEL v5.2 was used to screen and filter with minor allele frequency (MAF < 0.05) and heterozygosity (H > 0.15) to obtain 1514 high quality SNP markers. These markers have been used to analyze the association between soybean resistance to *Phytophthora sojae* root rot and soybean mature grain hardness [25,26].

The LD degree r^2 between SNP markers was calculated with TASSEL v5.2. The LD decay trend line is shown in Figure S1. When r^2 is 0.04, the average LD decay distance of the population on 20 chromosomes is 544.01 kb [25].

2.5. Population Structure and Kinship

The kinship in this study was calculated based on the pairwise IBS method in TASSEL v5.2. The calculated K matrix represents the kinship between different materials (Figure S2). The K matrix has the same effect as the Q matrix in the association analysis. It is used as a covariate in association analysis to reduce false positives [33]. STRUCTURE 2.2 was used to estimate the population structure of 224 soybean mini core collection, and the population structure analysis showed that when K was equal to 2, the Δk value was the largest, allowing the population to be divided into two subgroups [25] (Figure S3).

2.6. Genome-Wide Association Analysis

This study is based on TASSEL v5.2, using population structure (Q) as a covariate and kinship (K) as a covariate. The MLM (Q + K) model was used to perform GWAS on the values of the four plant architecture traits of the soybean mini core collection under three environments and the average of the phenotypes in the three environments. To fully exploit the valuable genetic information in the soybean mini core collection, the threshold for significant association between the marker and the target trait was set to $p \leq 0.01$ ($-\log10(p) \geq 2$), which has been used in the study of soybean fatty acid content [32], mature grain hardness [26], and R6 stage yield-related traits [27] association analysis loci. Manhattan Plot and Q-Q Plot were drawn based on the results of association analysis by the R (V4.0.0) package"qqman" (V0.1.8, Stephen Turner, VA, USA, https://github.com/stephenturner/qqman) [34,35].

2.7. Prediction and Annotation of Candidate Genes

According to the LD decay distance (±554.01 kb), SNP markers that were significantly correlated in at least two environments and had the highest rate of phenotypic variation were selected for candidate gene searches. The SoyBase database (https://www.soybase.org/, accessed on 24 December 2021) and Phytozome (https://phytozome-next.jgi.doe.gov/, accessed on 24 December 2021) functional annotation of genes and different tissue expression data were combined to predict candidate genes for target traits. The soybean reference genome was Wm82.a2.v1.

3. Results

3.1. Phenotypic Variation in the Four Plant Architecture-Related Traits

A two-year survey of four plant architecture-related traits of 224 diverse soybean mini core accessions was conducted in Jiangpu, Nanjing, and Dangtu, Anhui. The results showed that PH, NN, BN, and DI had extensive phenotypic variation in the soybean mini core collection in all three environments (2016JP, 2017JP, and 2017DT) (Table 1). The variation ranges of PH, NN, BN, and DI in the three environments were 28.06–165.52 cm, 10.15–27.6, 0.67–6.28, and 2.94–9.29 mm, respectively. The distributions of PH, NN, and DI in 2016JP were significantly lower than in 2017JP and 2017DT, while the distribution of BN in the three environments was basically the same (Figure 1). The results of the analysis of variance showed that the four plant architecture-related traits had extremely

significant differences in genotype, environment, and genotype–environment interactions (Table 1), indicating that these plant architecture traits may be quantitative traits controlled by multiple genes and easily affected by environmental factors. The heritability of PH, NN, BN, and DI in these collections was 89.25%, 93.92%, 69.97%, and 73.05%, respectively. Among them, the heritability of PH and NN was higher, while the heritability of BN and DI was lower, indicating that BN and DI are greatly affected by the environment and other factors.

Table 1. Descriptive statistic, analysis of variance, and heritability (h^2) of plant architecture-related traits in soybean mini core collection.

Traits	Environments	Mean	*SD*	Min	Max	Significance			h^2 (%)
						F_G	F_E	$F_{G\times E}$	
PH	2016JP	45.43	16.58	21.64	101	113.37 **	7305.66 **	13.19 **	89.25
	2017JP	73.17	30.66	27.5	177.67				
	2017DT	100.72	40.34	24	249.33				
	Mean	69.7	28.06	28.06	165.52				
NN	2016JP	15.36	3.98	9.11	29.4	82.87 **	627.76 **	5.39 **	93.92
	2017JP	17.43	4.41	9.33	27.67				
	2017DT	18	4.23	10	28.83				
	Mean	16.75	4	10.15	27.6				
BN	2016JP	2.82	1.27	0.33	9.89	13.62 **	20.63 **	4.22 **	69.97
	2017JP	2.54	1.23	0	7.33				
	2017DT	2.68	1.18	0.17	6.67				
	Mean	2.68	0.99	0.67	6.28				
DI	2016JP	4.51	0.83	2.9	8.85	22.41 **	1797.21 **	4.87 **	73.05
	2017JP	5.72	1.24	3.36	10.66				
	2017DT	6.69	1.42	2.37	11.8				
	Mean	5.56	1	2.94	9.29				

PH: plant height; NN: number of nodes on main stem; BN: branch number; DI: stem diameter; F_G represents the genotype effect; F_E represents the environment effect; $F_{G\times E}$ represents the interaction effect of genotype and environment; ** represents significant at $p \leq 0.01$.

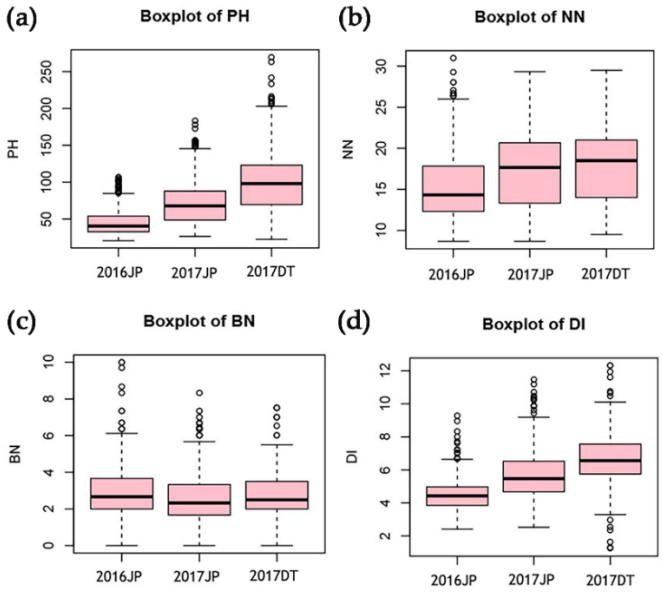

Figure 1. Boxplots of four plant architecture-related traits in soybean mini core collection in three environments. (**a**–**d**) Boxplots of PH, NN, BN, and DI.

In addition, the four plant architecture-related traits of the soybean mini core collection were continuously distributed and approximately normally distributed, consistent with

the characteristics of quantitative traits (Figure 2). The correlation analysis based on the phenotypic mean values of the four traits was computed, and the results are shown in Table 2. There were highly significant positive correlations among PH, NN, BN, and DI (r = 0.422–0.878, $p \leq 0.01$). The correlation analysis showed positive correlations between soybean plant architecture-related traits.

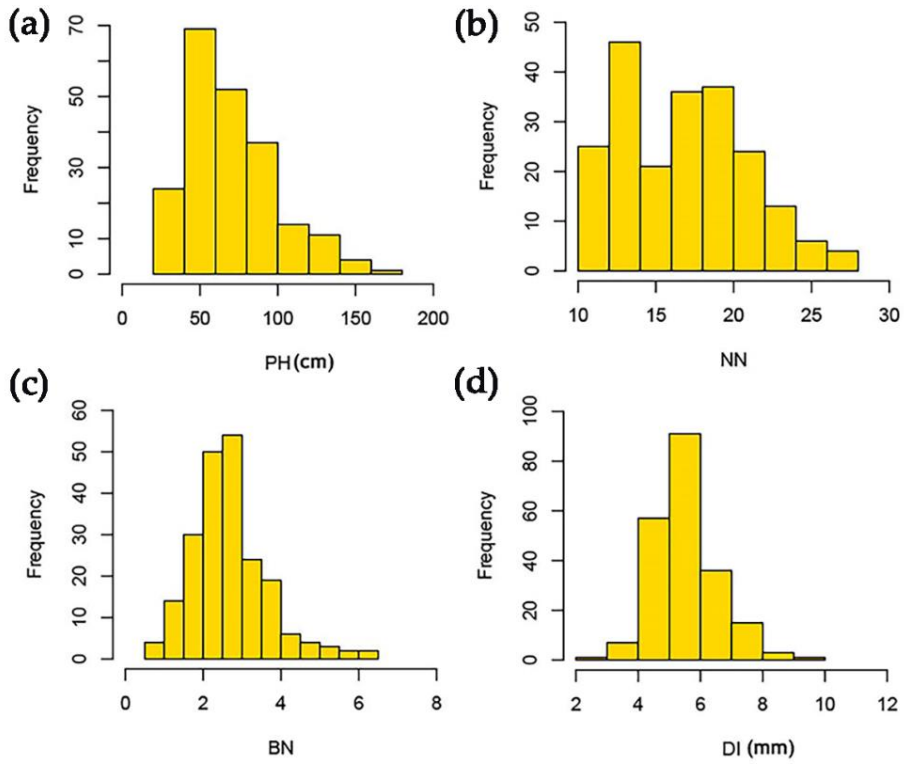

Figure 2. Frequency distribution of plant architecture-related traits in soybean mini core collection. (**a**–**d**) Frequency distribution of PH, NN, BN, and DI.

Table 2. Correlation analysis of plant architecture-related traits in soybean.

Traits	PH	NN	BN	DI
PH	1			
NN	0.878 **	1		
BN	0.487 **	0.507 **	1	
DI	0.542 **	0.522 **	0.422 **	1

** Represents significant at $p \leq 0.01$; PH: plant height; NN: number of nodes on main stem; BN: branch number; DI: stem diameter.

3.2. Genetic Diversity, Linkage Disequilibrium, and Population Structure

The analysis of SNP data, linkage disequilibrium, and population structure of 224 soybean mini core collections in this study has been reported by Huang et al. [25]. The population was genotyped using the 1645 SNPs obtained by integration, from which 1514 high-quality SNPs with missing data < 15%, minor allele frequency MAF ≥ 5%, and heterozygous alleles ≤ 15% [36] were selected for subsequent analysis. The average LD decay of the collections on 20 chromosomes was 554.01 kb at $r^2 < 0.04$, as calculated using TASSEL v5.2 [25] (Figure S1). The kinship coefficients were mostly concentrated from 0.3 to 0.4 using the pairwise IBS method, indicating that the pairwise kinship relatedness among different soybean materials was weak [25] (Figure S2). Estimation of the population structure by STRUCTURE 2.2 showed that the maximum Δk value was obtained when the K value was two, thus the Chinese soybean mini core collection could be divided into two subgroups [25] (Figure S3).

3.3. GWAS of the Four Plant Architecture-Related Traits

In this study, 1514 high-quality SNP markers were used to perform the GWAS based on the MLM (Q + K) model for plant architecture-related traits (PH, NN, BN, DI) of 224 soybean mini core collections using the values for three environments: 2016JP, 2017JP, and 2017DT, and for the phenotypic means in the three environments. A total of 41 SNP loci ($-\log10(p) \geq 2$) were significantly associated with four traits in at least two environments, of which 19, 15, 8, and 6 SNPs were significantly associated with PH, NN, BN, and DI, respectively. Among the 19 PH SNPs, five co-localized with NN and two co-located with DI at the same time. (Figures 3 and S4–S7). Among the markers that are significantly associated with PH, three markers could be repeatedly detected in all environments, namely, markers Q-02-0159217, BARC-041421-07980, and BARC-017097-02199 on chromosomes 2, 9, and 11, and the contribution of BARC-017097-02199 with larger effects on the observed phenotypic variation was 5.08–10.48% (Table 3). Three of the markers significantly associated with NN were detected in all environments, including the marker Map-2026 on chromosome 11 and the markers Map-2211 and Map-2213 on chromosome 12. Among them, Map-2213 had the highest interpretation rate of phenotypic variation under the four environments, at 7.76%, 7.41%, 6.2%, and 7.62%, respectively.

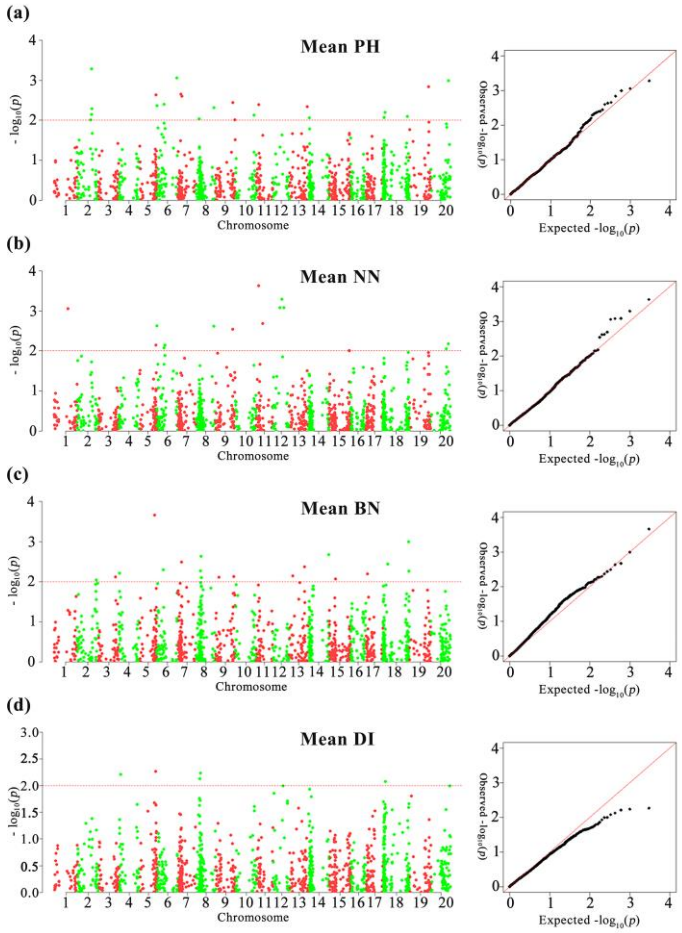

Figure 3. Manhattan and Q-Q plots of the GWAS for the four plant architecture-related traits in soybean mini core collection. The horizontal red line indicates the genome-wide significance threshold ($-\log10(p) \geq 2$). (**a**–**d**) Association mapping of PH, NN, BN, and DI based on the mean values of three environments, respectively.

Table 3. Significantly associated SNPs of plant architecture-related traits in soybean mini core collection.

Trait	Markers	Chromosome	Position (bp)	Allelic	Environments	−log10(*p*)	R^2 (%)	References
PH	Q-02-0147081	2	29,025,631	A/C	2016JP/2017JP/Mean	2.01–3.15	3.22–6.04	
	Q-02-0158174	2	31,120,457	A/C	2016JP/2017JP/Mean	2.14–3.07	3.39–5.54	[37–39]
	Q-02-0159217	2	31,236,629	A/G	2016JP/2017JP/2017DT/Mean	2.07–4.14	4.15–8.19	
	Q-02-0164155	2	31,724,101	A/G	2016JP/2017JP/Mean	2.29–3.27	3.68–5.72	
	Q-05-0207994	5	40,057,732	A/G	2016JP/2017JP/Mean	2.27–2.64	3.91–4.38	[28]
	BARC-014527-01571	6	656,098	C/T	2016JP/2017DT/Mean	2.13–2.69	5.18–5.97	[40]
	Q-06-0128380	6	17,613,829	A/C	2016JP/2017JP/Mean	2.08–2.4	4.54–5.3	[37,41–43]
	BARC-021425-04104	6	48,403,137	A/C	2016JP/2017JP/Mean	2.27–3.06	5.22–6.76	[44,45]
	Map-1255	7	7,190,281	A/T	2017JP/Mean	2.66–3.46	4.42–6.6	[46]
	BARC-041421-07980	9	44,020,453	A/T	2016JP/2017JP/2017DT/Mean	2.02–3.44	4.89–7.96	
	BARC-028861-06032	9	48,977,064	A/G	2017JP/2017DT/Mean	2.01–2.2	4.24–5.23	
	Map-1899	10	44,738,812	A/C	2016JP/2017DT	2.25–2.31	3.62–4.75	
	BARC-017097-02199	11	6,030,644	A/G	2016JP/2017JP/2017DT/Mean	2.14–4.62	5.08–10.48	[28]
	Map-2516	13	43,492,402	A/C	2016JP/Mean	2.34–2.52	3.79–4.18	[10]
	Q-14-0026827	14	3,164,755	A/C	2017JP/2017DT/Mean	2.07–2.54	3.25–5.26	
	Q-18-0016625	18	1,921,789	A/T	2017DT/Mean	2.05–2.07	4.64–5.79	[47]
	Q-18-0040883	18	4,329,135	A/G	2016JP/Mean	2.2–2.3	3.5–3.71	[38]
	Map-3990	19	45,340,653	A/G	2016JP/Mean	2.25–2.84	5.12–6.34	[48,49]
	Q-20-0260630	20	43,442,146	A/C	2017DT/Mean	2.04–2.99	4.09–5.12	
NN	Map-0076	1	33,989,570	C/G	2016JP/2017DT/Mean	2.06–5.05	5.62–11.27	
	BARC-014287-01306	5	4,279,362	C/T	2017JP/2017DT	2.77–3.37	6.5–9.44	
	BARC-014527-01571	6	656,098	C/T	2016JP/Mean	2.63–2.88	5.88–6.35	
	Q-06-0128380	6	17,613,829	A/C	2017JP/Mean	2.08–2.61	4.55–5.96	
	Q-08-0277446	8	43,136,899	A/T	2017JP/2017DT/Mean	2.11–2.62	3.89–4.38	
	BARC-041421-07980	9	44,020,453	A/T	2017DT/Mean	2.22–2.54	5.75–6.19	
	BARC-028861-06032	9	48,977,064	A/G	2017JP/2017DT	2.08–2.26	4.59–4.84	
	BARC-017097-02199	11	6,030,644	A/G	2016JP/2017JP/Mean	2.23–3.63	5.21–8.23	[28]
	Map-2026	11	25,164,685	A/T	2016JP/2017JP/2017DT/Mean	2.12–3.04	3.57–6.56	
	Map-2211	12	17,971,162	A/G	2016JP/2017JP/2017DT/Mean	2.47–3.11	6.61–6.83	
	Map-2213	12	18,455,502	A/C	2016JP/2017JP/2017DT/Mean	2.2–3.4	6.2–7.76	
	Map-2218	12	23,616,522	C/G	2017JP/Mean	2.92–3.08	5.23–5.36	
	Q-15-0364055	15	49,261,407	A/T	2017DT/Mean	2.02–2.33	3.17–4.72	
	Q-15-0364629	15	49,325,995	A/C	2017DT/Mean	2–2.33	3.14–4.73	
	BARC-030259-06840	20	38,262,189	A/G	2016JP/2017JP/Mean	2.06–2.08	3.31–3.51	
BN	BARC-014639-01604	5	37,602,081	A/T	2017JP/Mean	3.67–4.2	8.42–10.1	[9]
	Q-07-0088101	7	8,771,610	A/G	2016JP/Mean	2.25–2.49	3.85–4.34	
	BARC-013587-01169	8	10,563,212	C/G	2017JP/Mean	2.02–2.28	4.74–5.15	
	Q-08-0094591	8	12,454,445	A/G	2017DT/Mean	2.01–2.64	3.65–4.48	
	Map-2491	13	37,273,176	A/T	2016JP/Mean	2.37–2.41	4.09–4.11	
	BARC-039561-07508	14	48,880,009	C/T	2017JP/Mean	2.68–3.99	4.64–7.82	
	BARC-016029-02040	15	15,125,233	A/G	2017JP/Mean	2.07–2.19	4.7–5.23	
	BARC-018645-03217	17	4,097,240	A/G	2017JP/Mean	2.2–2.36	5.04–5.67	
DI	Q-05-0193181	5	42,009,549	A/C	2017DT/Mean	2.08–2.27	3.61–3.81	
	Q-08-0059708	8	8,402,455	A/G	2017JP/Mean	2.13–2.33	4.66–5.29	
	Map-1899	10	44,738,812	A/C	2016JP/2017JP	2.16–2.3	3.52–3.8	
	Map-2223	12	34,510,897	A/G	2017JP/2017DT	2.24–2.78	4.98–7.09	
	Q-15-0012218	15	1,858,944	A/G	2016JP/2017DT	2.26–2.34	5.28–5.73	
	Q-18-0040883	18	4,329,135	A/G	2017JP/Mean	2.08–2.27	3.24–3.74	

In addition, four markers (BARC-014527-01571, Q-06-0128380, BARC-041421-07980, and BARC-017097-02199) were significantly associated with PH and were located on chromosomes 6, 9, and 11. A total of eight SNP loci were significantly associated with BN in at least two environments, and the phenotypic variation interpretation rate ranged from 3.65% to 10.10%. Among them, BARC-014639-01604, which is located on chromosome 5, had the highest phenotypic variation explanation rate in the 2017JP and mean environments, at 10.10% and 8.42%, respectively. With a total of six SNPs significantly associated with DI in at least two environments, the phenotypic variation interpretation rate for individual SNP markers ranged from 3.52% to 7.09%, with Map-2223, located on chromosome 12, showing the highest phenotypic variation interpretation rates of 4.98% and 7.09% in 2017JP and 2017DT, respectively (Table 3). According to the SNPs that are stably and significantly associated with a high phenotypic variation interpretation rate in different environments [7], four SNP markers BARC-017097-02199, Map-2213, BARC-014639-01604, and Map-2223 were found to be remarkably related to PH, NN, BN, and DI in at least two environments with large effects on subsequent candidate genes prediction.

3.4. Functional Annotation of Candidate Genes

According to the four SNP markers BARC-017097-02199, Map-2213, BARC-014639-01604, and Map-2223, which have been determined to have great effects on phenotypic variation, and combined with the average LD decay distance of 544.01 kb, candidate genes were detected in the candidate intervals near four markers. In this study, candidate genes were predicted for PH, NN, BN, and DI, and a total of 18 candidate genes were obtained. According to homologous gene and function annotation, these candidate genes encode the pentatricopeptide repeat (PPR) superfamily protein, cytochrome P450 family proteins, SAUR auxin-responsive protein and auxin efflux carrier protein, GDSL lipase, cyclin and ubiquitin proteins, ubiquitin family protein, calcium-binding proteins, cellulose synthase, bZIP transcription factor family proteins, SAUR-like auxin-responsive protein, and NAC domain family proteins (Table 4).

Table 4. Function annotation for candidate genes of plant architecture-related traits in soybean.

Traits	Candidate Genes	Homologous Gene	Function Annotation
PH	*Glyma.11g074100*	*AT4G36220*	Cytochrome P450 (CYP84A1, FAH1)
	Glyma.11g076200	*AT4G38860*	SAUR-like auxin-responsive protein family
	Glyma.11g078800	*AT1G80550*	Pentatricopeptide repeat (PPR) superfamily protein
	Glyma.11g079400	*AT1G75900*	GDSL-like Lipase/Acylhydrolase superfamily protein
	Glyma.11g084500	*AT1G20610*	Cyclin B2;3
	Glyma.11g084700	*AT2G17200*	Ubiquitin family protein
	Glyma.11g085600	*AT1G18485*	Pentatricopeptide repeat (PPR) superfamily protein
	Glyma.11g086800	*AT2G17525*	Pentatricopeptide repeat (PPR) superfamily protein
	Glyma.11g086900	*AT2G17525*	Pentatricopeptide repeat (PPR) superfamily protein
	Glyma.11g087300	*AT5G65980*	Auxin efflux carrier family protein
NN	*Glyma.12g142900*	*AT4G27280*	Calcium-binding EF-hand family protein
BN	*Glyma.05g187300*	*AT5G64740*	Cellulose synthase 6
	Glyma.05g195000	*AT1G01720*	NAC domain transcriptional regulator superfamily protein
	Glyma.05g196300	*AT2G46690*	SAUR-like auxin-responsive protein family
DI	*Glyma.12g180200*	*AT5G22810*	GDSL-like Lipase/Acylhydrolase superfamily protein
	Glyma.12g184200	*AT3G62890*	Pentatricopeptide repeat (PPR) superfamily protein
	Glyma.12g184500	*AT1G08320*	bZIP transcription factor family protein
	Glyma.12g186200	*AT5G22380*	NAC domain containing protein 90

The expression patterns of these candidate genes were downloaded from the Phytozome database, and comparative analysis found that: the PH candidate gene *Glyma.11G74100* is specifically and highly expressed in stems, shoot apical meristems (SAM), and pods, while *Glyma.11G084500* is in SAMs, nodules, and root hairs. In addition, *Glyma.11G076200* is specifically expressed in stems. Furthermore, the candidate gene *Glyma.12G142900* for NN is expressed in the node. The BN candidate gene *Glyma.05G187300* has higher expression in the stem, and the DI candidate gene *Glyma.12G184500* has higher expression in the stem and nodules. These genes are expressed significantly differentially in specific tissues, such as stems, SAMs, nodules, and leaves during soybean plant architecture development (Figure 4). Therefore, these genes may be potential candidate genes for plant architecture-related traits and may be involved in the process of regulating soybean plant architecture morphogenesis.

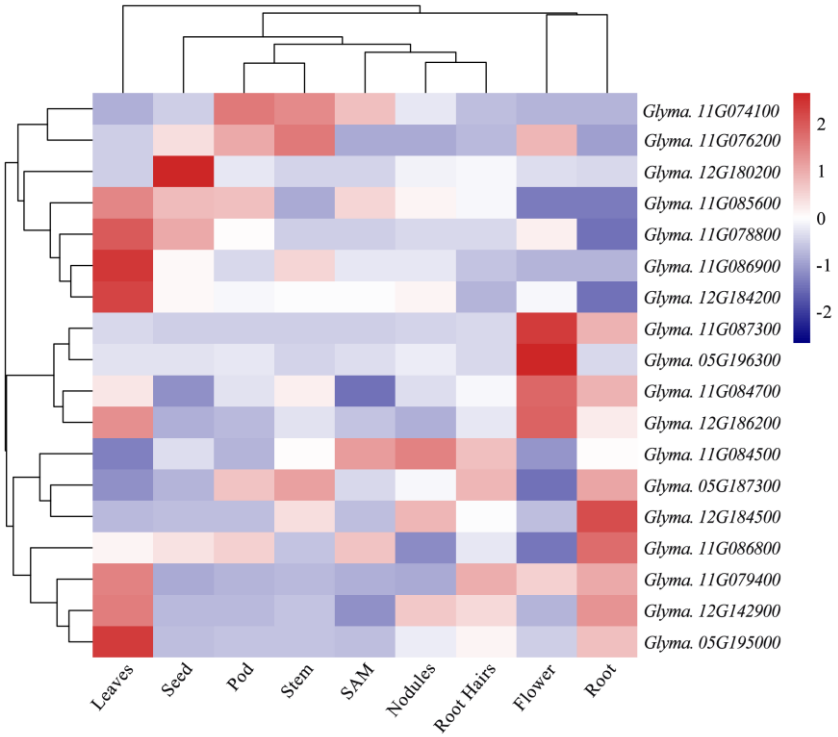

Figure 4. Heatmap profiles of the candidate genes in tissues. The color scale represents increased (red) and decreased (navy blue) fold change in the levels of candidate genes transcripts. The horizontal axis represents different soybean tissues (leaves, seed, pod, stem, SAM, nodules, root hairs, flower, and root), and the vertical axis stands for different candidate genes.

4. Discussion

Population size is an important factor in GWAS that affects the accuracy of association analysis. Increasing the number of samples in the population will broaden the genetic variation of the population and thus improve the power of the associations [50], thus generally using natural populations with wide genetic variation. Research has shown that a population of more than 300 materials may have high-precision association results [51–54]. However, in actual research, the population cannot expand indefinitely. The core collection uses the minimum set of accessions to represent the maximum genetic diversity of the entire germplasm resources [55,56], which is a favorable sampling population for GWAS mapping between target traits and markers [57,58]. There is abundant genetic variation in the Chinese soybean mini core collection utilized in this study, which is an ideal population for mining excellent loci or genes for complex traits.

In this study, field phenotype identification of four plant architecture-related traits in the soybean mini core collection was performed in three environments. There were extensive phenotypic variations in the four plant architecture traits in this collection. In addition, the heritability of PH and NN was higher, 93.92% and 89.25%, respectively. The heritability of BN and DI was lower, 69.97% and 73.05%, respectively (Table 1). This result is basically consistent with the heritability calculated by Zhang et al. [10]. Compared with PH and NN, BN and DI were easily affected by environmental factors, which is consistent with the ANOVA results. The NN of soybeans was positively correlated with PH [20], which is consistent with the results of our study (Table 2). Previous studies have confirmed that PH, NN, BN, and DI play a crucial role in soybean yield [59,60]. Therefore, the coordination and balance of soybean plant architecture-related traits could guarantee a high soybean yield.

The MLM (Q + K) model was used to analyze the association of four plant architecture traits in the soybean mini core collection. A total of 41 SNP markers that were significantly associated with the target traits identified under at least two environments, among

which 19, 15, 8, and 6 were associated with PH, NN, BN, and DI, respectively. Among the 19 PH SNPs, five markers (BARC-014527-01571, Q-06-0128380, BARC-041421-07980, BARC-028861-06032, and BARC-017097-02199) co-located with NN, and two markers (Map-1899 and Q-18-0040883) co-located with DI at the same time. Among the 41 stable and significantly associated markers identified in total, 15 were located in the QTLs of the related traits reported, and 26 were new loci identified in this study (Table 3). Among these markers, BARC-017097-02199 and the co-localization marker Q-05-0207994 with NN are located within the QTL for PH and NN identified by an $F_{2:10}$ RIL population [28]. In addition, four markers (Q-06-0128380, BARC-021425-04104, Q-18-0016625, and Q-18-0040883) overlapped with the reported plant height QTLs [37,38,41–45,47]. Among the eight SNP markers with a significant association with BN, the marker BARC-014639-01604 was located in the BN QTL detected using 236 F_2 populations derived from the cross between Jiyu 50 and Jinong 18 [9]. In addition, PH and NN were significantly correlated, and the five co-localization markers for PH and NN indicated that there was a greater possibility of QTLs for PH and NN in the region where these markers were located.

The 18 candidate genes of plant architecture-related traits were preliminarily predicted based on the candidate intervals of four SNP markers (BARC-017097-02199, Map-2213, BARC-014639-01604, and Map-2223). Among them, the four PH candidate genes *Glyma.11g078800*, *Glyma.11g085600*, *Glyma.11g086800,* and *Glyma.11g086900* and the DI candidate genes *Glyma.12g184200* encode proteins belonging to members of the pentatricopeptide repeat (PPR) protein family. PPR proteins are at the core of post-transcriptional RNA modification [61]. In maize, *ZmSMK9* encodes a P-type PPR protein, and *Zmsmk9* mutants show not only delayed endosperm and embryo development but also upright leaf architecture [62]. These candidate genes encoding PPR proteins may be involved in the embryo development process, thereby affecting PH and DI. In addition, *Glyma.11g074100* encodes a protein belonging to a member of the cytochrome P450 family. Cytochrome P450 plays important roles in biosynthesis pathways [63]. The cytochrome P450 gene, *CYP85A2*, was significantly associated with soybean plant architecture and yield traits, and the CC genotype of this locus was an excellent genotype for designing ideotypes and increasing yield [64]. In addition, tissue expression analysis showed that *Glyma.11g074100* expressed specifically in stems and SAM (Figure 4). These results suggest that *Glyma.11g074100* is a potential candidate gene for PH. The BN and PH candidate genes *Glyma.05g196300* and *Glyma.11g076200* both encode SAUR auxin response proteins. SAURs belong to the primary auxin response gene families [65]. *SAUR39* acts as a negative regulator of auxin synthesis and transportation, regulating the growth rate, and increasing leaf and tiller angles in rice [66]. The PH candidate gene *Glyma.11g084500* encodes a cyclin protein. Cyclin genes regulate the cell cycle of plants during growth and development. Ectopic overexpression of the *Arabidopsis cyclinD2;1* (*Arath; CYCD2;1*) gene resulted in leaf curling, decreased fruiting branch number, and increased plant height in cotton [67].

Among NN candidate genes, *Glyma.12g142900* encodes proteins belonging to the calcium-binding proteins family, which is an important component of the plant calcium signal transduction pathway. The overexpression of StCaM2, a potato calcium-binding protein, promotes plant growth and improves stress tolerance in tobacco [68]. For BN, the candidate gene *Glyma.05g187300* encodes cellulose synthetase, and *Glyma.05g195000* and the DI candidate gene *Glyma.12g186200* both encode NAC transcription regulators. NAC transcription factors are involved in mediating auxin signaling to promote the formation of SAM and for regulating secondary wall synthesis in fibers [69,70]. The overexpression of OsmiR164b-resistant *OsNAC2*, which showed better plant architecture, produced more grains and significantly upregulated the grain number and plant architecture-related genes *IPA1* and *DEP*1 [71]. The DI candidate gene *Glyma.12g184500* encodes a bZIP transcription factor family protein. The rice transcription factor OsbZIP49 determines plant architecture by influencing the local homeostasis of the auxin [72]. Plant architecture-related traits are complex in the growth and development of soybean plants. The above candidate genes may be involved in growth and development, auxin synthesis and transportation, glycosylation

reactions, cellulose synthesis, and other processes. These candidate genes will provide information for cloning ideotype genes of soybean.

At present, the trend of molecular breeding research all over the world is clarifying the genetic mechanism of important traits and mine-related genes for breeding and utilization. These new loci and candidate genes found in the Chinese soybean mini core collection could both enrich the gene network's regulation of soybean plant architecture and be used to develop molecular markers to provide new gene resources for plant architecture molecular design breeding. The results of this study lay a foundation for the cloning and functional verification of ideotype genes, as well as provide a reference for the study of plant architecture of other crops.

5. Conclusions

In this study, a total of 41 stable and significant SNPs for plant architecture-related traits were found using 224 Chinese soybean mini core germplasm via GWAS, of which 26 were novel loci. In addition, four large-effect markers BARC-017097-02199, Map-2213, BARC-014639-01604, and Map-2223 were identified for PH, NN, BN, and DI, respectively. Based on gene annotation and tissue expression analysis, 18 promising candidate genes related to soybean plant architecture were predicted in the candidate interval where the large-effect markers were located. This study provides a reference for the genetic mechanism of soybean plant architecture-related traits, which has been the basis of molecular design breeding of the soybean ideotype.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy12040817/s1, Figure S1: Scatter plots and LD decay against physical distance among co-chromosomes. Figure S2: Distributions of pairwise relative kinship estimates between mini core collections. Figure S3: Population structure of soybean mini core collection. Figures S4–S7: Manhattan plots and Q-Q plots of GWAS for PH, NN, BN, and DI in three environments.

Author Contributions: Conceptualization, W.D., X.Z. and J.Z.; investigation, X.Z. and C.W.; formal analysis, W.D., X.Z. and D.L.; data curation, C.L. and R.S.; methodology, N.G.; Software, W.D. and X.J.; funding acquisition, J.Z. and H.X.; writing—original draft, W.D.; writing—review and editing, J.Z. and H.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by China Agriculture Research System of MOF and MARA (CARS-04), National Natural Science Foundation of China (32072082), Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT_17R55), Jiangsu Collaborative Innovation Center for Modern Crop Production (JCIC-MCP202101), Cyrus Tang Innovation Center for Seed Industry (CTIC-SI201902), and the Key R&D project of Jiangsu Province (BE2019376).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets supporting the results of this article are included within the main text and the supplementary materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Qin, J.; Zhang, J.A.; Liu, D.; Yin, C.C.; Wang, F.M.; Chen, P.Y.; Chen, H.; Ma, J.B.; Zhang, B.; Xu, J.; et al. iTRAQ-based analysis of developmental dynamics in the soybean leaf proteome reveals pathways associated with leaf photosynthetic rate. *Mol. Genet. Genomics* **2016**, *291*, 1595–1605. [CrossRef]
2. Kim, S.W.; Less, J.F.; Wang, L.; Yan, T.H.; Kiron, V.; Kaushik, S.J.; Lei, X.G. Meeting Global Feed Protein Demand: Challenge, Opportunity, and Strategy. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 221–243. [CrossRef]
3. Donald, C.M. The breeding of crop ideotypes. *Euphytica* **1968**, *17*, 385–403. [CrossRef]
4. Huyghe, C. Genetics and genetic modifications of plant architecture in grain legumes: A review. *Agronomie* **1998**, *18*, 383–411. [CrossRef]

5. Sarlikioti, V.; de Visser, P.H.B.; Buck-Sorlin, G.H.; Marcelis, L.F.M. How plant architecture affects light absorption and photosynthesis in tomato: Towards an ideotype for plant architecture using a functional–structural plant model. *Ann. Bot.* **2011**, *108*, 1065–1073. [CrossRef]
6. Pan, Q.C.; Xu, Y.C.; Li, K.; Peng, Y.; Zhan, W.; Li, W.Q.; Li, L.; Yan, J.B. The Genetic Basis of Plant Architecture in 10 Maize Recombinant Inbred Line Populations. *Plant Physiol.* **2017**, *175*, 858–873. [CrossRef]
7. Zhang, X.L.; Ding, W.T.; Xue, D.; Li, X.N.; Zhou, Y.; Shen, J.C.; Feng, J.Y.; Guo, N.; Qiu, L.J.; Xing, H.; et al. Genome-wide association studies of plant architecture-related traits and 100-seed weight in soybean landraces. *BMC Genom. Data* **2021**, *22*, 1–14. [CrossRef]
8. Hill, W.G. Understanding and using quantitative genetic variation. *Philos. Trans. R. Soc. B* **2010**, *365*, 73–85. [CrossRef]
9. Yao, D.; Liu, Z.Z.; Zhang, J.; Liu, S.Y.; Qu, J.; Guan, S.Y.; Pan, L.D.; Wang, D.; Liu, J.W.; Wang, P.W. Analysis of quantitative trait loci for main plant traits in soybean. *GMR Genet. Mol. Res.* **2015**, *14*, 6101–6109. [CrossRef]
10. Zhang, H.R.; Hao, D.R.; Sitoe, H.M.; Yin, Z.T.; Hu, Z.B.; Zhang, G.Z.; Yu, D.Y. Genetic dissection of the relationship between plant architecture and yield component traits in soybean (*Glycine max*) by association analysis across multiple environments. *Plant Breed.* **2015**, *134*, 564–572. [CrossRef]
11. Gupta, P.K.; Rustgi, S.; Kulwal, P.L. Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol. Biol.* **2005**, *57*, 461–485. [CrossRef]
12. Sehgal, D.; Singh, R.; Rajpal, V.R. Quantitative trait loci mapping in plants: Concepts and approaches. In *Molecular Breeding for Sustainable Crop Improvement*; Rajpal, V.R., Rao, S.R., Raina, S.N., Eds.; Springer International Publishing: Cham, Switzerland, 2016; Volume 2, pp. 31–59.
13. Wang, Z.; Tian, Z.X. Genomics progress will facilitate molecular breeding in soybean. *Sci. China Life Sci.* **2015**, *58*, 813–815. [CrossRef]
14. Flint-Garcia, S.A.; Thuillet, A.C.; Yu, J.M.; Pressoir, G.; Romero, S.M.; Mitchell, S.E.; Doebley, J.; Kresovich, S.; Goodman, M.M.; Buckler, E.S. Maize association population: A high-resolution platform for quantitative trait locus dissection. *Plant J.* **2005**, *44*, 1054–1064. [CrossRef]
15. Atwell, S.; Huang, Y.S.; Vilhjalmsson, B.J.; Willems, G.; Horton, M.; Li, Y.; Meng, D.Z.; Platt, A.; Tarone, A.M.; Hu, T.T.; et al. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature* **2010**, *465*, 627–631. [CrossRef]
16. Chen, W.; Gao, Y.Q.; Xie, W.B.; Gong, L.; Lu, K.; Wang, W.S.; Li, Y.; Liu, X.Q.; Zhang, H.Y.; Dong, H.X.; et al. Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* **2014**, *46*, 714–721. [CrossRef]
17. Guo, Z.F.; Chen, D.J.; Alqudah, A.M.; Roder, M.S.; Ganal, M.W.; Schnurbusch, T. Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. *New Phytol.* **2017**, *214*, 257–270. [CrossRef]
18. Fang, C.; Ma, Y.M.; Wu, S.W.; Liu, Z.; Wang, Z.; Yang, R.; Hu, G.H.; Zhou, Z.K.; Yu, H.; Zhang, M.; et al. Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol.* **2017**, *18*, 161. [CrossRef]
19. Zhang, J.P.; Song, Q.J.; Cregan, P.B.; Nelson, R.L.; Wang, X.Z.; Wu, J.X.; Jiang, G.L. Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genom.* **2015**, *16*, 1–11. [CrossRef]
20. Chang, F.G.; Guo, C.Y.; Sun, F.L.; Zhang, J.S.; Wang, Z.L.; Kong, J.J.; He, Q.Y.; Sharmin, R.A.; Zhao, T.J. Genome-Wide Association Studies for Dynamic Plant Height and Number of Nodes on the Main Stem in Summer Sowing Soybeans. *Front. Recent Dev. Plant Sci.* **2018**, *9*, 1184. [CrossRef]
21. Shim, S.; Ha, J.; Kim, M.Y.; Choi, M.S.; Kang, S.T.; Jeong, S.C.; Moon, J.K.; Lee, S.H. *GmBRC1* is a Candidate Gene for Branching in Soybean (*Glycine max* (L.) Merrill). *Int. J. Mol. Sci.* **2019**, *20*, 135. [CrossRef]
22. Chen, L.M.; Fang, Y.H.; Li, X.Y.; Zeng, K.; Chen, H.F.; Zhang, H.B.; Yang, H.L.; Cao, D.; Hao, Q.N.; Yuan, S.L.; 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] [PubMed]
23. Fu, M.M.; Wang, Y.P.; Ren, H.X.; Du, W.G.; Yang, X.Y.; Wang, D.L.; Cheng, Y.X.; Zhao, J.M.; Gai, J.Y. Exploring the QTL–allele constitution of main stem node number and its differentiation among maturity groups in a Northeast China soybean population. *Crop Sci.* **2020**, *60*, 1223–1238. [CrossRef]
24. Fahim, A.M.; Pan, L.Y.; Li, C.Y.; He, J.B.; Xing, G.N.; Wang, W.B.; Zhang, F.K.; Li, N.; Gai, J.Y. QTL-allele system of main stem node number in recombinant inbred lines of soybean (*Glycine max*) using association versus linkage mapping. *Plant Breed.* **2021**, *140*, 870–883. [CrossRef]
25. Huang, J.; Guo, N.; Li, Y.H.; Sun, J.T.; Hu, G.J.; Zhang, H.P.; Li, Y.F.; Zhang, X.; Zhao, J.M.; Xing, H.; et al. Phenotypic evaluation and genetic dissection of resistance to *Phytophthora sojae* in the Chinese soybean mini core collection. *BMC Genet.* **2016**, *17*, 85. [CrossRef] [PubMed]
26. Zhang, X.; Zhao, J.M.; Bu, Y.P.; Xue, D.; Liu, Z.X.; Li, X.N.; Huang, J.; Guo, N.; Wang, H.T.; Xing, H.; et al. Genome-Wide Association Studies of Soybean Seed Hardness in the Chinese Mini Core Collection. *Plant Mol. Biol. Rep.* **2018**, *36*, 605–617. [CrossRef]
27. Li, X.N.; Zhou, Y.; Bu, Y.P.; Wang, X.F.; Zhang, Y.M.; Guo, N.; Zhao, J.M.; 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]
28. Chen, Q.S.; Zhang, Z.C.; Liu, C.Y.; Xin, D.W.; Qiu, H.M.; Shan, D.P.; Shan, C.Y.; Hu, G.H. QTL Analysis of Major Agronomic Traits in Soybean. *Agric. Sci. China* **2007**, *6*, 399–405. [CrossRef]

29. SAS Institute. *Base SAS 9.4 Procedures Guide: Statistical Procedures*; SAS Institute: Cary, NC, USA, 2013.
30. Nyquist, W.E.; Baker, R.J. Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.* **1991**, *10*, 235–322. [CrossRef]
31. Li, Y.H.; Li, W.; Zhang, C.; Yang, L.A.; Chang, R.Z.; Gaut, B.S.; Qiu, L.J. Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. *New Phytol.* **2010**, *188*, 242–253. [CrossRef]
32. Li, Y.H.; Reif, J.C.; Ma, Y.S.; Hong, H.L.; Liu, Z.X.; Chang, R.Z.; Qiu, L.J. Targeted association mapping demonstrating the complex molecular genetics of fatty acid formation in soybean. *BMC Genom.* **2015**, *16*, 841. [CrossRef]
33. Zhang, D.; Kan, G.Z.; Hu, Z.B.; Cheng, H.; Zhang, Y.; Wang, Q.; Wang, H.; Yang, Y.M.; Li, H.Y.; Hao, D.R.; et al. Use of single nucleotide polymorphisms and haplotypes to identify genomic regions associated with protein content and water-soluble protein content in soybean. *Theor. Appl. Genet.* **2014**, *127*, 1905–1915. [CrossRef] [PubMed]
34. Turner, S.D. qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv* **2014**. *in preprint*.
35. Cichy, K.A.; Wiesinger, J.A.; Mendoza, F.A. Genetic diversity and genome-wide association analysis of cooking time in dry bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* **2015**, *128*, 1555–1567. [CrossRef]
36. Liu, K.J.; Muse, S.V. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* **2005**, *21*, 2128–2129. [CrossRef]
37. Rossi, M.E.; Orf, J.H.; Liu, L.J.; Dong, Z.M.; Rajcan, I. Genetic basis of soybean adaptation to North American vs. Asian mega-environments in two independent populations from Canadian × Chinese crosses. *Theor. Appl. Genet.* **2013**, *126*, 1809–1823. [CrossRef] [PubMed]
38. Sun, D.S.; Li, W.B.; Zhang, Z.C.; Chen, Q.S.; Ning, H.L.; Qiu, L.J.; Sun, G.L. Quantitative trait loci analysis for the developmental behavior of Soybean (*Glycine max* L. Merr.). *Theor. Appl. Genet.* **2006**, *112*, 665–673. [CrossRef]
39. Kim, K.S.; Diers, B.W.; Hyten, D.L.; Mian, M.A.R.; Shannon, J.G.; Nelson, R.L. Identification of positive yield QTL alleles from exotic soybean germplasm in two backcross populations. *Theor. Appl. Genet.* **2012**, *125*, 1353–1369. [CrossRef]
40. Guzman, P.S.; Diers, B.W.; Neece, D.J.; Martin, S.K.S.; Leroy, A.R.; Grau, C.R.; Hughes, T.J.; Nelson, R.L. QTL Associated with Yield in Three Backcross-Derived Populations of Soybean. *Crop Sci.* **2007**, *47*, 111–122. [CrossRef]
41. Orf, J.H.; Chase, K.; Jarvik, T.; Mansur, L.M.; Cregan, P.B.; Adler, F.R.; Lark, K.G. Genetics of Soybean Agronomic Traits: I. Comparison of Three Related Recombinant Inbred Populations. *Crop Sci.* **1999**, *39*, 1642–1651. [CrossRef]
42. Kabelka, E.A.; Diers, B.W.; Fehr, W.R.; LeRoy, A.R.; Baianu, I.C.; You, T.; Neece, D.J.; Nelson, R.L. Putative Alleles for Increased Yield from Soybean Plant Introductions. *Crop Sci.* **2004**, *44*, 784–791. [CrossRef]
43. Palomeque, L.; Liu, L.J.; Li, W.B.; Hedges, B.; Cober, E.R.; Rajcan, I. QTL in mega-environments: II. Agronomic trait QTL co-localized with seed yield QTL detected in a population derived from a cross of high-yielding adapted × high-yielding exotic soybean lines. *Theor. Appl. Genet.* **2009**, *119*, 429–436. [CrossRef] [PubMed]
44. Gai, J.Y.; Wang, Y.J.; Wu, X.L.; Chen, S.Y. A comparative study on segregation analysis and QTL mapping of quantitative traits in plants—with a case in soybean. *Front. Agric. China* **2007**, *1*, 1–7. [CrossRef]
45. Liu, W.X.; Kim, M.Y.; Van, K.; Lee, Y.-H.; Li, H.; Liu, X.H.; Lee, S.H. QTL Identification of Yield-Related Traits and their Association with Flowering and Maturity in Soybean. *J. Crop Sci. Biotechnol.* **2011**, *14*, 65–70. [CrossRef]
46. Wang, D.; Graef, G.L.; Procopiuk, A.M.; Diers, B.W. Identification of putative QTL that underlie yield in interspecific soybean backcross populations. *Theor. Appl. Genet.* **2004**, *108*, 458–467. [CrossRef]
47. Reinprecht, Y.; Poysa, V.W.; Yu, K.F.; Rajcan, I.; Ablett, G.R.; Pauls, K.P. Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (*Glycine max* (L.) Merrill) germplasm. *Genome* **2006**, *49*, 1510–1527. [CrossRef] [PubMed]
48. Mansur, L.M.; Lark, K.G.; Kross, H.; Oliveira, A. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). *Theor. Appl. Genet.* **1993**, *86*, 907–913. [CrossRef]
49. Lee, S.H.; Bailey, M.A.; Mian, M.A.R.; Shipe, E.R.; Ashley, D.A.; Parrott, W.A.; Hussey, R.S.; Boerma, H.R. Identification of quantitative trait loci for plant height, lodging, and maturity in a soybean population segregating for growth habit. *Theor. Appl. Genet.* **1996**, *92*, 516–523. [CrossRef]
50. Korte, A.; Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **2013**, *9*, 29. [CrossRef]
51. Wang, X.L.; Wang, H.W.; Liu, S.X.; Ferjani, A.; Li, J.S.; Yan, J.B.; Yang, X.H.; Qin, F. Genetic variation in *ZmVPP1* contributes to drought tolerance in maize seedlings. *Nat. Genet.* **2016**, *48*, 1233–1241. [CrossRef]
52. Wang, M.; Li, W.Z.; Fang, C.; Xu, F.; Liu, Y.C.; Wang, Z.; Yang, R.; Zhang, M.; Liu, S.L.; Lu, S.J.; et al. Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nat. Genet.* **2018**, *50*, 1435–1441. [CrossRef]
53. Li, N.; Lin, B.; Wang, H.; Li, X.M.; Yang, F.F.; Ding, X.H.; Yan, J.B.; Chu, Z.H. Natural variation in *ZmFBL41* confers banded leaf and sheath blight resistance in maize. *Nat. Genet.* **2019**, *51*, 1540–1548. [CrossRef] [PubMed]
54. Lu, S.J.; Dong, L.D.; Fang, C.; Liu, S.L.; Kong, L.P.; Cheng, Q.; Chen, L.Y.; Su, T.; Nan, H.Y.; Zhang, D.; et al. Stepwise selection on homeologous PRR genes controlling flowering and maturity during soybean domestication. *Nat. Genet.* **2020**, *52*, 428–436. [CrossRef] [PubMed]
55. Frankel, O.H.; Brown, A.H.D. Current plant genetic resources—A critical appraisal. In *Genetics: New Frontiers*; Oxford & IBH Publishing: New Delhi, India, 1984; Volume IV, pp. 1–11.
56. Brown, A.H.D. Core collections: A practical approach to genetic resources management. *Genome* **1989**, *31*, 818–824. [CrossRef]

57. Zhang, P.; Zhong, K.Z.; Shahid, M.Q.; Tong, H.H. Association Analysis in Rice: From Application to Utilization. *Front. Recent Dev. Plant Sci.* **2016**, *7*, 1202. [CrossRef]
58. Schlappi, M.R.; Jackson, A.K.; Eizenga, G.C.; Wang, A.J.; Chu, C.C.; Shi, Y.; Shimoyama, N.; Boykin, D.L. Assessment of Five Chilling Tolerance Traits and GWAS Mapping in Rice Using the USDA Mini-Core Collection. *Front. Plant Sci.* **2017**, *8*, 957. [CrossRef]
59. Oz, M.; Karasu, A.; Goksoy, A.T.; Turan, Z.M. Interrelationships of Agronomical Characteristics in Soybean (*Glycine max*) Grown in Different Environments. *Int. J. Agric. Biol.* **2009**, *11*, 85–88.
60. Asanome, N.; Ikeda, T. Effect of Branch Direction's Arrangement on Soybean Yield and Yield Components. *J. Agron. Crop Sci.* **1998**, *181*, 95–102. [CrossRef]
61. Guillaumot, D.; Lopez-Obando, M.; Baudry, K.; Avon, A.; Rigaill, G.; de Longevialle, A.F.; Broche, B.; Takenaka, M.; Berthome, R.; De Jaeger, G.; et al. Two interacting PPR proteins are major Arabidopsis editing factors in plastid and mitochondria. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8877–8882. [CrossRef]
62. Pan, Z.Y.; Liu, M.; Xiao, Z.Y.; Ren, X.M.; Zhao, H.L.; Gong, D.M.; Liang, K.; Tan, Z.D.; Shao, Y.Q.; Qiu, F.Z. ZmSMK9, a pentatricopeptide repeat protein, is involved in the cis-splicing of *nad5*, kernel development and plant architecture in maize. *Plant Sci.* **2019**, *288*, 110205. [CrossRef]
63. Schuler, M.A.; Werck-Reichhart, D. Functional Genomics of P450S. *Annu. Rev. Plant Biol.* **2003**, *54*, 629–667. [CrossRef]
64. Yang, C.M.; Yan, J.; Jiang, S.Q.; Li, X.; Min, H.W.; Wang, X.F.; Hao, D.Y. Resequencing 250 soybean accessions: New insights into genes associated with agronomic traits and genetic networks. *bioRxiv* **2021**, *in press*. [CrossRef] [PubMed]
65. Hagen, G.; Guilfoyle, T. Auxin-responsive gene expression: Genes, promoters and regulatory factors. *Plant Mol. Biol.* **2002**, *49*, 373–385. [CrossRef] [PubMed]
66. Kant, S.; Bi, Y.M.; Zhu, T.; Rothstein, S.J. *SAUR39*, a Small Auxin-Up RNA Gene, Acts as a Negative Regulator of Auxin Synthesis and Transport in Rice. *Plant Physiol.* **2009**, *151*, 691–701. [CrossRef] [PubMed]
67. Liu, G.Z.; Jin, S.X.; Liu, X.Y.; Tan, J.F.; Yang, X.Y.; Zhang, X.L. Overexpression of *Arabidopsis cyclin D2;1* in cotton results in leaf curling and other plant architectural modifications. *Plant Cell, Tissue Organ Cult.* **2012**, *110*, 261–273. [CrossRef]
68. Raina, M.; Kumar, A.; Yadav, N.; Kumari, S.; Yusuf, M.A.; Mustafiz, A.; Kumar, D. StCaM2, a calcium binding protein, alleviates negative effects of salinity and drought stress in tobacco. *Plant Mol. Biol.* **2021**, *106*, 85–108. [CrossRef] [PubMed]
69. Zimmermann, R.; Werr, W. Pattern Formation in the Monocot Embryo as Revealed by *NAM* and *CUC3* Orthologues from *Zea mays* L. *Plant Mol. Biol.* **2005**, *58*, 669–685. [CrossRef]
70. Zhong, R.Q.; Richardson, E.A.; Ye, Z.H. Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*. *Planta* **2007**, *225*, 1603–1611. [CrossRef]
71. Jiang, D.G.; Chen, W.T.; Dong, J.F.; Li, J.; Yang, F.; Wu, Z.C.; Zhou, H.; Wang, W.S.; Zhuang, C.X. Overexpression of miR164b-resistant *OsNAC2* improves plant architecture and grain yield in rice. *J. Exp. Bot.* **2018**, *69*, 1533–1543. [CrossRef]
72. Ding, C.H.; Lin, X.H.; Zuo, Y.; Yu, Z.L.; Baerson, S.R.; Pan, Z.Q.; Zeng, R.S.; Song, Y.Y. Transcription factor OsbZIP49 controls tiller angle and plant architecture through the induction of indole-3-acetic acid-amido synthetases in rice. *Plant J.* **2021**, *108*, 1346–1364. [CrossRef]