BMC Genomics

RESEARCH

Open Access

Identification of quantitative trait loci for lodging and related agronomic traits in soybean (*Glycine max* [L.] Merr.)

Bo Chen[1,2,3†], Cheng Chai[1,2,3†], Mingming Duan[1,2,3†], Ximeng Yang[1,2,3], Zhandong Cai[1,2,3], Jia Jia[1,2,3], Qiuju Xia[4], Shilin Luo[1,2,3], Lu Yin[1,2,3], Yunxia Li[1,2,3], Nianen Huang[1,2,3], Qibin Ma[1,2,3], Hai Nian[1,2,3*] and Yanbo Cheng[1,2,3*]

Abstract

Background Lodging, a crucial agronomic trait linked to soybean yield, poses a significant challenge in soybean production. Nevertheless, there has been less research on soybean lodging compared to other important agronomic traits, hindering progress in breeding high-yield soybeans. Our goals were to investigate lodging, pinpoint quantitative trait loci (QTL) linked to lodging, and forecast potential candidate genes linked to this trait. To achieve this, we employed a recombinant inbred line (RIL) population derived from a cross between Guizao 1 and B13 (GB) across various environments.

Results The lodging score of the RIL population was found to be significantly positively correlated with flowering time, maturity time, plant height, number of main stem nodes, stem diameter, and internode length, with correlation coefficients ranging from 0.457 to 0.783. A total of 84 QTLs associated with soybean lodging and related traits were identified using the GB population. The contribution of phenotypic variance ranged from 1.26 to 66.87%, with LOD scores ranging from 2.52 to 69.22. Additionally, within these QTLs, a stable major QTL associated with lodging was newly discovered in the GB population. Out of the ten major QTLs associated with other related traits, nine of them were situated within the *qLD-4-1* interval of the major lodging score locus, displaying phenotypic variations ranging from 12.10 to 66.87%. Specific alterations in gene expression were revealed through the analysis of resequencing data from the two parental lines, potentially indicating their significant roles in lodging. Subsequently, it was determined through qRT-PCR that four genes are likely to be the major genes controlling soybean lodging.

Conclusions This study's findings offer valuable insights into the genetic underpinnings of soybean lodging resistance traits. By comprehending the potential genetic factors associated with lodging, this research lays the groundwork for breeding high-yield soybeans with improved lodging resistance.

Keywords Soybean, Lodging, Lodging related traits, Quantitative trait loci, Candidate genes

†Bo Chen, Cheng Chai and Mingming Duan are contributed equally to this work.

*Correspondence:
Hai Nian
hnian@scau.edu.cn
Yanbo Cheng
ybcheng@scau.edu.cn

Full list of author information is available at the end of the article

Background

Soybean (*Glycine max* [L.] Merr.) is a significant oil and protein crop with origins dating back approximately 5,000 years to ancient China [1, 2]. Given the rising global population and improving living standards, there is an expectation that soybean production will need to meet consumer demands by 2050 [3]. Consequently, the aim of crop breeding is to develop agricultural plants with desirable traits [4].

In recent decades, numerous studies have employed molecular markers and QTL analysis to investigate and identify desired agronomic traits in crops [5–8]. The rapid advancement of molecular biology technology has facilitated the widespread adoption of DNA molecular marker techniques in various domains, including gene mapping and identification [9–11]. Conventional molecular markers include restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and simple sequence repeat (SSR) markers [12–16]. With the ongoing advancement of molecular biology technology, a novel molecular marker known as SNP has emerged, demonstrating rich and stable genetic variations within the genome [17, 18]. The creation of high-density genetic linkage maps through high-throughput SNP genotyping platforms is of paramount importance for precise QTL mapping and the exploration of essential agricultural traits [19, 20]. Subsequently, many scholars have identified a total of 27 QTLs using RFLP and SSR molecular markers.

Crop lodging represents an agronomic trait influenced by multiple factors, including genetics, cultivation practices, climate, and ecology. It is closely linked to plant height, stem thickness, stem strength, root weight, root length, and root volume [21–24]. Numerous QTLs associated with soybean lodging have been identified in prior research. Mansur et al. [25] detected a QTL on chromosome 19 closely linked to RFLP markers using an $F_{2:5}$ population resulting from a cross between PI27890 × PI27890. Subsequently, Lee et al. [26] successfully identified 18 lodging QTLs employing F_2 populations and RFLP molecular markers. Orf et al. [27] discovered two QTLs on chromosome 6 tightly linked to RFLP markers using a recombinant inbred line (RIL) population derived from a cross between Minsoy × Archer. Specht et al. [28] found five QTLs on chromosomes 6, 12, and 19, using SSR and RFLP markers in an $F_{7:11}$ population from a Minsoy × Noir1 cross. Chapman et al. [29] uncovered a QTL on chromosome 11 closely linked to SSR markers in an $F_{4:6}$ population from an Essex × Essex cross. Wang et al. [30] identified a QTL on chromosome 9 closely linked to SSR markers in a BC2F4 population from an IA2008 × PI468916 cross. Zhang et al. [31] identified eight QTLs on chromosomes 6 and 13, using SSR and RFLP markers in an $F_{2:7:11}$ population from a Kefeng No.1 × Nannong 1138-2 cross. Subsequently, SSR molecular markers became widely adopted by scholars for detecting 15 lodging QTLs on various chromosomes [32–36]. Additionally, some scholars successfully detected lodging QTLs using SNP molecular markers [37]. Lee et al. [22] identified five QTLs on chromosomes 4, 6, and 19 closely linked to SNPs in an F_7 population from crosses between Wyandot × PI567301B. In summary, previous research has identified 20 chromosomes containing QTLs associated with soybean lodging, utilizing various molecular markers.

This study aims to utilize the GB populations and high-density genetic linkage maps to pinpoint the locus associated with lodging across multiple environments and to identify stable major-effect QTLs related to soybean lodging. Furthermore, we aim to identify candidate genes responsible for regulating soybean lodging. The lodging QTLs and genes identified in this study may contribute to a deeper understanding of the genetic mechanisms underlying lodging and provide valuable guidance for future efforts to improve soybean yields.

Materials and methods

Plant materials

The RIL populations were created through a cross between two distinct soybean accessions, Guizao 1 × B13, and comprised 248 $F_{8:11}$ lines [38]. Guizao 1 belongs to the early-maturing soybean varieties of southern China, selected by the Guangxi Zhuang Autonomous Region Academy of Agricultural Sciences. B13 is an imported variety from Brazil. Both the parental lines and the 248 RILs were cultivated in three replicates in July of 2020, 2021, and 2022 and in March of 2021 at the Teaching and Research Base of Zengcheng (23°14'N, 113°38'E), and also in July 2020 at the Experimental Teaching Base of Guangzhou (23°10'N, 113°22'E), denoted as 20ZC, 21ZC-2, 22ZC, 21ZC-1, and 20GZ, respectively. The experimental planting followed a completely randomized block design, with one row allocated per line, a row length of 1.5 m, a row spacing of 0.5 m, and a plant spacing of 0.1 m. Field management adhered to conventional practices, and no occurrences of pests or diseases were observed during the growth period.

Phenotypic statistics and analysis

The improvement in lodging score is categorized into five levels, and during the soybean R8 phase, it is required that a minimum of 80% of the plants display a tilted angle [22, 25]. These levels are defined as follows: Level 1: Nearly all plants are upright. Level 2: The angle of plant inclination does not exceed 15°. Level 3: The tilt angle ranges from 15° to 45°. Level 4: The tilt angle ranges from 45° to 85°. Level 5: The tilt angle exceeds 85°. Other key traits measured include: Flowering time: The presence of

an open flower on any section of the main stem. Maturity time: When 95% of the pods achieve maturity based on pod color. Plant height: The distance from the cotyledon to the top of the plant stem, excluding the inflorescence. Number of main stem nodes: The count of nodes starting above the cotyledon node, excluding the apical inflorescence and cotyledon nodes. Stem diameter: The diameter of the first internode of the stem. Internode length: The ratio of plant height to the number of main stem nodes. Grain weight per plant: The weight of seeds produced per plant. Among these traits, plant height, stem diameter, and the number of main stem nodes were assessed using two plants for each repetition [32, 33, 39, 40].

The mean of the phenotypic data collected from the GB RIL population across five natural environments was computed to represent the phenotypic data of the combined environment. These data encompassed lodging score, flowering time, maturity time, plant height, number of main stem nodes, stem diameter, internode length, and grain weight per plant across various natural environments. Statistical parameters for lodging and related traits were computed between the parental lines and their respective RIL individuals in each environment and for the combined environment using IBM SPSS Statistics 26.0 (SPSS, Inc., Chicago, IL, USA). Descriptive statistics of phenotype data were obtained through analysis of variance (ANOVA), including measures such as maximum, minimum, mean, range of variation, standard deviation, coefficient of variation (CV), kurtosis, and skewness for the two RIL populations.

Analysis of variance (ANOVA) and the estimation of broad-sense heritability (h^2) were performed on the phenotypic data collected from the five natural environments utilizing Genstat 12th Edition software. The formulae [41] were as follows:

$h^2=\sigma^2_g=(\sigma^2_g+\sigma^2_{ge}/n+\sigma^2_e/nr)$.

where σ^2_g denotes genotype variance, σ^2_{ge} denotes genotype × environment interaction variance, and σ^2_e denotes variance error. "n" denotes the number of environments, and "r" denotes the number of replicates per environment.

Linkage map construction and QTL mapping

The construction of genetic maps for the RIL populations in this study followed previously described methods [38]. These maps encompassed 3,748 bin markers identified in the 248 individuals, covering a total of 3,031.93 cM. The average genetic distance between adjacent bin markers on each chromosome within the GB RIL population ranged from 0.60 to 1.03 cM. These high-density linkage maps provided enhanced resolution compared to traditional maps. To identify the positions of QTLs, we employed the Composite Interval Mapping (CIM) method using WinQTLCart 2.5 software. The threshold for LOD (logarithm of odds) scores for different traits was set at 2.5, corresponding to the 5% significance level. LOD scores exceeding 2.5 were considered indicative of the presence of QTLs. The distribution of all identified QTLs on the linkage map was visualized using MapChart software. Each detected QTL was denoted by a combination of one or more letters representing trait abbreviations and chromosome numbers [42]. Specifically, "q" represented QTL, LD represented lodging score, FT represented flowering time, MT represented maturity time, PH represented plant height, SD represented stem diameter, NMSN represented the number of main stem nodes, IL represented internode length, and GWPP represented grain weight per plant. The number immediately following indicated the chromosome number, followed by the sequence number [43].

Prediction of candidate genes in the major stable QTL interval

For this analysis, information regarding the Glyma.Wm82.a2.v1 gene model was sourced from SoyBase (http://www.soybase.org), focusing on the physical intervals of both stable and novel QTLs. Functional gene annotation information for candidate genes within these QTLs was assigned based on the Glyma.Wm82.a2.v1 reference genome sequence available at (https://phytozome.jgi.doe.gov) and SoyBase. To further elucidate the roles of these candidate genes, an online data analysis platform and cloud-based tools (https://www.omicstudio.cn/tool/22) were utilized to conduct Gene Ontology (GO) annotation, categorizing the functions associated with all genes found within the stable and novel QTLs. In addition to functional annotation, reference gene expression data were obtained from the ePlant database (https://bar.utoronto.ca/eplant_soybean/), which includes RNA-seq data from various soybean samples, specifically from Shoot Apex, Hypocotyl, and Stem, all located within the major and stable QTL interval. To provide insights into tissue-specific expression patterns, an online resource (https://www.omicstudio.cn/tool/59) was employed to generate heat maps depicting the expression patterns of candidate genes [44]. Moreover, an assessment of genetic variation among candidate genes between the parental lines was conducted based on resequencing data of the parental lines. Whole-genome sequencing of Guizao 1 and B13 was performed using the Illumina HiSeq X Ten platform, with an average sequencing depth of 8× [45]. High-quality sequencing data from these parental lines were analyzed to predict structural variations in the genes.

Quantitative real-time PCR (qRT-PCR) analysis

Root and stem tissues from the two parental lines (Guizao 1 and B13) were preserved at -80 °C. Total RNA

extraction from the roots and stems was performed using the Plant Total RNA Purification Kit (Promega (Beijing) Biotech Co., Ltd). Subsequently, one microgram of RNA underwent genomic DNA removal and reverse transcription using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix kit (Novoprotein). Gene expression levels were assessed through qRT-PCR analysis conducted with the CFX96 Real-Time System (Bio-Rad). The PCR cycling conditions were as follows: an initial denaturation step at 95 °C for 1 min, followed by 40 cycles of denaturation at 95 °C for 10 s, annealing at 55.0–60.0 °C (depending on the gene) for 10 s, and extension at 72 °C for 30 s [46]. The soybean ACT3 gene [47] was amplified as a reference control. The resulting expression data were calculated using the comparative cycle threshold method ($2^{-\triangle\triangle ct}$), and the experiments were independently repeated three times [48]. Specific primer sequences for each gene used in this analysis can be found in Table S1.

Results

Correlation analysis of lodging score and related traits

The box plots displaying phenotypic data for various environmental traits (Fig. 1) indicate minimal variation in the data across the years. Consequently, the mean values of the data for each trait were employed in the correlation analysis (Table 1). In the GB population, the lodging score displays highly significant positive correlations with several traits, including flowering time, maturity time, plant height, number of main stem nodes, stem diameter, and internode length. The correlation coefficients for these relationships range from 0.457 to 0.783. However, there is no significant correlation observed between lodging score and grain weight per plant. The correlation coefficients between lodging score and each of the correlated traits are as follows: Flowering time: 0.698 Maturity time: 0.483 Plant height: Highest correlation at 0.783 Number of main stem nodes: 0.749 Stem diameter: 0.457 Internode length: 0.564 In contrast, grain weight per plant exhibits no significant correlation, with a correlation coefficient of 0.050. These findings indicate that the correlations between soybean lodging and other agronomic

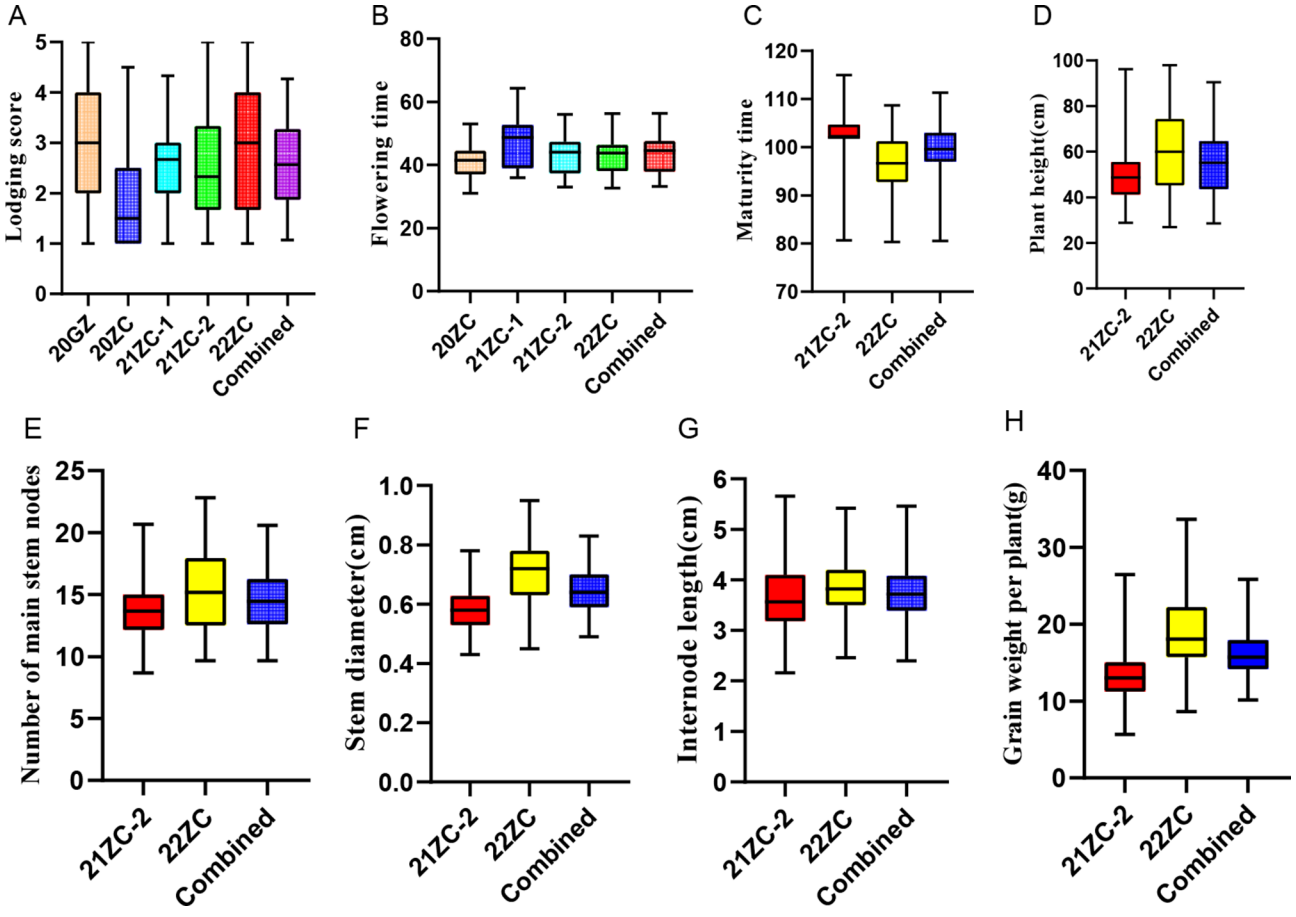

Fig. 1 Box Plot of Lodging and Related Traits. (**A**) Lodging Score, (**B**) Flowering Time, (**C**) Maturity Time, (**D**) Plant Height, (**E**) Number of Main Stem Nodes, (**F**) Stem Diameter, (**G**) Internode Length, (**H**) Grain Weight per Plant. Environments: Guangzhou 2020 (20GZ), Zengcheng 2020 (20ZC), Zengcheng Spring 2021 (21ZC-1), Zengcheng Summer 2021 (21ZC-2), Zengcheng 2022 (22ZC), Combined Environment (CE)

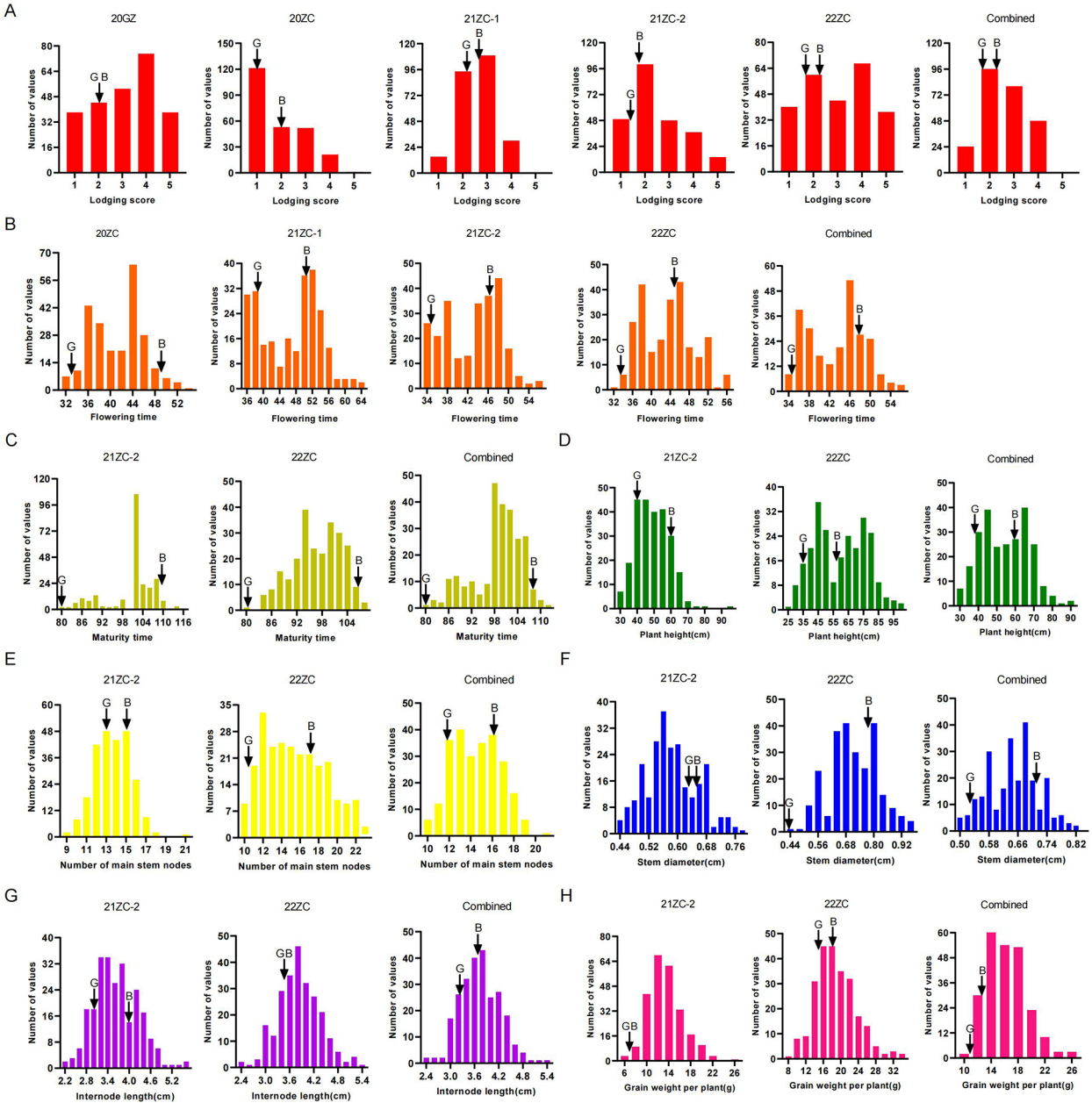

Fig. 2 Frequency Distribution of Lodging-Related Traits in the GB Population Across Multiple Environments. Environments: Guangzhou 2020 (20GZ), Zengcheng 2020 (20ZC), Zengcheng Spring 2021 (21ZC-1), Zengcheng Summer 2021 (21ZC-2), Zengcheng 2022 (22ZC), Combined Environment (CE). "G" represents the female parent Guizao1, and "B" represents the male parent B13. Figures **A** to **H** indicate the lodging score, flowering time, maturity time, plant height, number of main stem nodes, stem diameter, internode length and grain weight per plant, respectively

traits can serve as valuable references for selecting high-yield varieties in field-based breeding programs.

Descriptive statistical analysis of soybean lodging related traits and estimates of broad-sense heritability

Descriptive statistical analysis was conducted on the phenotypic data (Table 2). The results revealed that both parental lines exhibited mild lodging scores. However, the RIL populations displayed a wide range of family variation, with coefficients of variation ranging from 28.44 to 51.94%. This wide variation encompasses extreme traits, providing a solid foundation for QTL mapping of lodging. Similar patterns were observed in other lodging-related traits within the RIL population, indicating parental segregation and laying the groundwork for QTL mapping. The coefficients of variation for various traits are as follows: 11.34–13.00% for flowering time, 5.99–7.00% for maturity time, 20.20–28.51% for

Table 1 Correlation Analysis between Lodging Score and Related Traits in the GB Population

Trait [#]	LD	PH	NMSN	FT	IL	MT	SD	GWPP
LD	1							
PH	0.783**	1						
NMSN	0.749**	0.849**	1					
FT	0.698**	0.818**	0.819**	1				
IL	0.564**	0.796**	0.387**	0.540**	1			
MT	0.483**	0.696**	0.648**	0.801**	0.503**	1		
SD	0.457**	0.625**	0.677**	0.698**	0.353**	0.614**	1	
GWPP	0.050	0.077	0.057	0.025	0.081	0.041	0.191**	1

[#] Lodging score (LD), Plant height (PH), Number of main stem nodes (NMSN), Flowering time (FT), Internode length (IL), Maturity time (MT), Stem diameter (SD), Grain weight per plant (GWPP)

** Significant at the 0.01 probability level

Table 2 Descriptive statistical table for lodging and related traits of parents and the GB Population in various environments

Trait [#]	Env. [a]	Parents [b]		RILs				
		G [c]	B [d]	Range	Mean ± SD [e]	CV(%) [f]	Skewness	Kurtosis
LD	20ZC	2.00	2.00	1.00–5.00	3.13 ± 1.30	41.68	-0.234	-1.079
	20GZ	1.00	2.00	1.00–5.00	1.79 ± 0.93	51.94	0.802	-0.640
	21ZC-1	2.33	2.67	1.00–5.00	2.60 ± 0.74	28.44	-0.115	-0.609
	21ZC-2	1.60	1.75	1.00–5.00	2.46 ± 1.13	46.00	0.577	-0.482
	22ZC	1.75	2.17	1.00–5.00	2.98 ± 1.27	42.92	-0.020	-1.331
	CE	1.74	2.12	1.00–5.00	2.60 ± 0.85	32.68	0.115	-1.056
FT	20ZC	33.00	49.00	31.00–53.00	41.04 ± 4.65	11.34	0.039	-0.811
	21ZC-1	38.43	50.82	36.00–64.33	46.51 ± 7.24	15.57	-0.042	-1.167
	21ZC-2	34.40	46.15	33.00–56.00	42.73 ± 5.58	13.00	-0.097	-1.040
	22ZC	33.50	45.08	32.67–56.33	43.14 ± 5.40	12.51	0.284	-0.666
	CE	34.83	47.76	33.17–56.42	43.36 ± 5.44	12.55	-0.030	-1.088
MT	21ZC-2	79.45	109.85	80.67–115.00	100.79 ± 7.19	7.00	-1.090	0.571
	22ZC	80.58	106.83	80.33–108.67	96.57 ± 5.79	5.99	-0.358	-0.538
	CE	80.02	108.34	80.50–111.33	98.71 ± 6.31	6.39	-0.787	0.110
PH	21ZC-2	40.07	60.27	28.83–96.17	49.08 ± 9.92	20.20	0.611	1.254
	22ZC	35.80	58.70	26.92–98.00	59.56 ± 16.98	28.51	0.073	-1.038
	CE	37.93	59.48	28.54–90.42	54.27 ± 12.74	23.48	0.125	-0.729
NMSN	21ZC-2	13.07	15.20	8.67–20.67	13.65 ± 1.84	13.51	0.145	0.290
	22ZC	10.60	17.20	9.67–22.83	15.33 ± 3.27	21.35	0.298	-0.904
	CE	11.83	16.20	9.67–20.58	14.45 ± 2.22	15.33	0.073	-0.743
SD	21ZC-2	0.63	0.65	0.43–0.78	0.58 ± 0.07	12.00	0.291	-0.309
	22ZC	0.43	0.77	0.45–0.95	0.71 ± 0.10	14.45	0.016	-0.474
	CE	0.53	0.71	0.49–0.83	0.65 ± 0.07	10.80	-0.019	-0.459
IL	21ZC-2	3.07	3.96	2.16–5.66	3.62 ± 0.62	17.17	0.383	0.148
	22ZC	3.46	3.44	2.46–5.42	3.85 ± 0.52	13.48	0.174	0.246
	CE	3.27	3.70	2.40–5.46	3.74 ± 0.50	13.31	0.229	0.216
GWPP	21ZC-2	7.08	7.37	5.68–26.47	13.38 ± 3.13	23.40	0.608	0.948
	22ZC	15.40	18.08	8.63–33.66	18.96 ± 4.63	24.42	0.597	0.423
	CE	11.24	12.73	10.15–25.86	16.17 ± 2.95	18.25	0.722	0.524

[#] Lodging score (LD), Flowering time (FT), Maturity time (MT), Plant height (PH), Number of main stem nodes (NMSN), Stem diameter (SD), Internode length (IL), Grain weight per plant (GWPP)

[a] Environment; Guangzhou in 2020 (20GZ), Zengcheng in 2020 (20ZC), Zengcheng Spring in 2021 (21ZC-1), Zengcheng Summer in 2021 (21ZC-2), Zengcheng in 2022 (22ZC), Combined Environment (CE)

[b] Parents of GB RIL population

[c] Female parent of GB RIL population (Guizao 1)

[d] Male parent of GB RIL population (B13)

[e] Standard deviation

[f] Coefficient of variation

plant height, 13.51–21.35% for the number of main stem nodes, 10.80–14.45% for stem diameter, 13.31–17.17% for internode length, and 18.25–24.42% for grain weight per plant.

Several traits, including lodging score, flowering time, maturity time, and plant height, exhibit kurtosis and skewness values greater than 1. These values indicate the presence of numerous influential factors influencing these traits. The segregation of these traits in the RILs is governed by multiple genes, aligning with the characteristics of quantitative genetic traits. Conversely, the absolute values of kurtosis and skewness for the number of main stem nodes, stem diameter, internode length, and grain weight per plant are all less than 1, signifying that these traits follow a normal or approximately normal distribution and are consistent with quantitative genetic traits. Additionally, the frequency distribution map (Fig. 2) vividly illustrates the continuous variations in the phenotypic data of lodging score and its related traits. These findings collectively suggest that the lodging score and associated traits within the GB population conform to a normal or partially normal distribution, aligning with the characteristics of the RIL population and categorizing as quantitative genetic traits. In summary, the results indicate that lodging and related traits in the GB population adhere to a normal distribution pattern, consistent with the characteristics of the RIL population and indicating their classification as quantitative genetic traits.

ANOVA results for the lodging score of the GB RIL population across five natural environments demonstrate significant effects of genotype, environment, and the interaction between genotype and environment on lodging and related traits of the GB population (Table 3). The lodging trait in the GB population exhibited a substantially high heritability estimate (h^2) of 93.18%, indicating that the lodging phenotype in soybean is primarily influenced by genotype.

Identification of QTLs for lodging score and related traits

In total, 84 QTL loci were identified, accounting for phenotypic variation ranging from 1.26 to 66.87% across six environments. These QTLs were distributed across various traits, with 20, 11, 11, 12, 9, 10, 6, and 5 QTLs detected for lodging score, flowering time, maturity time, plant height, number of main stem nodes, stem diameter, internode length, and grain weight per plant, respectively (Fig. 3; Table S2). All of these QTLs displayed LOD values exceeding 2.5. The major and stable QTL locus, named *qLD-4-1*, associated with lodging score was identified. It is positioned within the physical interval of 3,513,907–5,769,624 bp on chromosome 4, spanning between bin15 to bin39 markers. This QTL was consistently detected in all six environments and exhibited phenotypic variation ranging from 15.38 to 38.68%, with LOD values ranging from 10.36 to 34.70. Additionally, nine out of the ten primary QTLs for other related traits (*qFT-4*, *qMT-4-1*, *qMT-4-2*, *qPH-4*, *qPH-19-2*, *qNMSN-4*, *qSD-4-1*, *qSD-4-2*, *qIL-4-1*, *qIL-4-2*) were found within the physical region of the primary QTL, *qLD-4-1*, for lodging score (Table 4). These QTLs exhibited phenotypic variation ranging from 55.93 to 66.87% for flowering time, 31.58–48.70% for maturity time, 41.41–51.32% for plant height, 20.46–48.39% for the number of main stem nodes, 12.10–29.12% for stem diameter, and 13.60–30.07% for internode length. The stable QTLs mentioned above provide valuable insights for the exploration of genes that regulate soybean lodging and related agronomic traits.

Candidate gene prediction within stable and major QTL interval

To identify potential genes associated with lodging within the QTL region (*qLD-4-1*), a search was conducted for 271 gene models located within this interval. Subsequently, 225 gene functions linked to *qLD-4-1* underwent GO annotation and were categorized through GO annotation analysis, providing functional annotations

Table 3 Analysis of Variance and Broad-Sense Heritability for the GB Population across Five Natural Environments

Trait [a]	Sources [b]	Df [c]	SS [d]	MS [e]	*P* [f]	VC [g]	h^2(%) [h]
LD	Genotype	247	2134.67	8.64	< 0.0001	0.53	93.18%
	Environment	4	217.36	54.34	< 0.0001	0.07	
	Interaction	988	1028.06	1.04	< 0.0001	0.13	
	Error	2480	1598.67	0.64			
	Total variation	3719	4978.76	1.34			

[a] Lodging score (LD)

[b] Sources of variation

[c] Degree of freedom

[d] Sum of deviation squares

[e] mean square

[f] The P value of the F-test (join t hypotheses test)

[g] Variance components for different sources of variation

[h] Broad-sense heritability

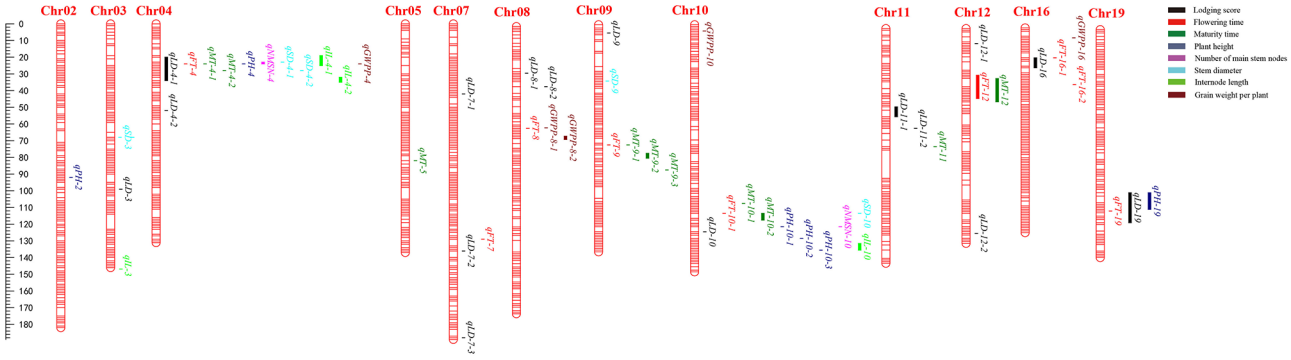

Fig. 3 Distribution of QTLs for Lodging and Related Traits on Soybean Chromosomes The left ruler indicates the physical distance between markers in cM. The right side of the chromosome shows the position of QTLs for lodging-related traits. The eight colors in the upper right corner of all chromosomes represent different traits

spanning cellular composition, biological processes, and molecular function (Fig. 4). The majority of genes within *qLD-4-1* were found to be involved in processes such as the regulation of DNA-templated transcription, plasma membrane, chloroplast, membrane, and ATP binding. In order to narrow down the list of candidate genes associated with lodging, a comparison was made of differentially expressed genes across various soybean tissues (Fig. 5) during three specific periods: Shoot Apex, Hypocotyl, and Stem. Through gene GO annotation analysis, followed by gene expression screening and functional annotation, a total of 13 candidate genes were identified (Table 5), indicating their potential roles in critical processes governing soybean lodging. In the high-quality resequencing data, seven out of the 13 genes exhibited structural variation between the parental lines of the RIL population (Guizao 1 and B13). These genes are *Glyma.04g050200*, *Glyma.04g050800*, *Glyma.04g051300*, *Glyma.04g052100*, *Glyma.04g053600*, *Glyma.04g056200*, and *Glyma.04g063800* (Table 6).

Expression for the identification of candidate genes

This study conducted a comprehensive analysis of the expression levels of candidate genes in the root and stem tissues of the two parental lines. The genes exhibited differential expression in the stems and leaves of the two parents, as determined by qRT-PCR analysis (Fig. 6). Among these genes, *Glyma.04g051300*, *Glyma.04g053600*, *Glyma.04g056200*, and *Glyma.04g063800* demonstrated significant differences in expression between Guizao 1 and B13 in both root and stem tissues, and these differences were highly significant. These findings strongly suggest that these four genes are the primary candidates responsible for regulating soybean lodging.

Discussion

Improvement of evaluation method for lodging score

The grain weight measurement conducted by our research institute specifically applies to seeds with smooth surfaces. However, due to the prolonged maturity of the populations studied, lodging plants experienced seed mold and rot issues. This led to a significant decrease in the data related to grain weight per plant, resulting in data errors that rendered the results less meaningful. The recombinant inbred lines exhibited diverse phenotypic variations, and relying solely on a single indicator has certain limitations. This approach may overlook the influence of other traits on lodging. To comprehensively evaluate the lodging phenomenon, it is essential to detect specific morphological and physiological indicators in plants. These indicators often involve assessing traits like stem bending resistance and stem composition [49, 50]. Currently, the most commonly employed method for lodging assessment is intuitive observation [22, 25]. This method assesses the degree of lodging based on actual conditions, capturing the combined effects of various factors. It is a straightforward approach, providing authentic and easily interpretable results, all while avoiding damage to the plants. To address the potential differences resulting from various degrees of edge lodging, adjustments can be made to enhance the suitability of this method for the studied population. Therefore, in this study, lodging scores were determined based on the intuitive method and the inclination of the main stem. Additionally, the proportion of lodging plants was calculated in relation to the total number of plants. These modifications aimed to facilitate data analysis and improve the visualization of lodging distribution within the population.

Comparisons of QTLs localization with previous results

The primary objective of crop breeding is to develop agricultural plants with desirable traits, such as increased grain yield, improved nutritional quality, and enhanced

Table 4 Key QTLs Associated with Lodging and related traits in the GB Population across different environmental conditions

QTL [#]	Chr. [a]	Interval	CI (bp) [b]	Position (cM)	LOD [c]	Add. [d]	PVE (%) [e]	Env. [f]
qLD-4-1	4	Bin15 ~ Bin39	3,513,907–5,769,624	22.4–26.5	10.36	-0.53	15.38	20GZ
					13.16	-0.45	17.69	20ZC
					21.95	-0.43	29.28	21ZC-1
					17.30	-0.54	22.20	21ZC-2
					33.43	-0.82	37.28	22ZC
					34.70	-0.56	38.68	CE
qFT-4-1	4	Bin19	4,069,992–4,097,410	24.2	59.90	-3.80	59.89	20ZC
					50.37	-5.66	55.93	21ZC-1
					69.22	-4.81	66.87	21ZC-2
					55.88	-4.29	56.60	22ZC
					68.07	-4.68	66.74	CE
qMT-4-1	4	Bin18 ~ Bin19	4,033,995–4,097,410	23.8–24.2	31.76	-4.53	35.90	21ZC-2
					46.88	-4.24	48.70	22ZC
					41.13	-4.38	43.51	CE
qMT-4-2	4	Bin31	4,820,821–4,871,065	28.1	26.80	-4.21	31.58	21ZC-2
qPH-4	4	Bin18 ~ Bin19	4,033,995–4,097,410	23.8–24.2	32.38	-6.71	41.41	21ZC-2
					52.10	-12.81	51.32	22ZC
					48.80	-9.51	50.14	CE
qPH-19-2	19	Bin139	45,022,952–45,248,245	107.6	5.11	2.20	4.89	21ZC-2
					17.25	5.86	11.79	22ZC
					15.11	4.18	10.63	CE
qNMSN-4	4	Bin17 ~ Bin18	3,979,305–4,069,991	23.3–23.8	13.21	-0.86	20.46	21ZC-2
					46.48	-2.37	47.23	22ZC
					44.19	-1.63	48.39	CE
qSD-4-1	4	Bin16	3,904,813–4,097,410	22.9–24.2	7.63	-0.03	12.10	21ZC-2
					22.48	-0.06	29.11	22ZC
					22.72	-0.04	29.12	CE
qSD-4-2	4	Bin31	4,820,821–4,871,065	28.1	20.82	-0.06	27.35	22ZC
qIL-4-1	4	Bin14 ~ Bin21	3,371,562–4,180,494	19.4–24.9	9.02	-0.24	13.60	21ZC-2
					22.38	-0.31	30.07	22ZC
					22.32	-0.29	30.07	CE
qIL-4-2	4	Bin36 ~ Bin40	5,461,549–5,817,504	33.3–34.7	11.53	-0.22	17.54	CE
					4.07	-0.16	6.42	21ZC-2

[#] The name of each QTL

[a] Chromosome

[b] Physical position

[c] Logarithm of odds

[d] Additive effect

[e] Phenotypic variance

[f] Environment; Guangzhou in 2020 (20GZ), Zengcheng in 2020 (20ZC), Zengcheng Spring in 2021 (21ZC-1), Zengcheng Summer in 2021 (21ZC-2), Zengcheng in 2022 (22ZC), Combined Environment (CE)

adaptability to changing environmental conditions, with grain yield being a predominant factor [4, 51]. Lodging significantly impacts crop yield, making it essential to study lodging characteristics in soybeans to reduce production costs and enhance yield. In this study, a total of 20 QTLs were identified for lodging score, including a previously unreported lodging QTL (*qLD-4-1*). Comparing the results of this study with previous research, it was discovered that the *Glyma.04G050200* gene within the major QTL (*qLD-4-1*) for lodging score has been previously associated with controlling the growth period. Its physical location corresponds to a range of 4,075,901 to 4,081,260 bp, suggesting a potential link between genes regulating lodging and those governing the growth period [52–54]. Furthermore, QTLs for three lodging scores (*qLD-4-2*, *qLD-7-2*, and *qLD-19-2*) overlap or coincide with loci related to growth stage traits from previous studies [39, 55, 56]. Additionally, *qLD-7-2*, *qLD-11-1*, and *qLD-19-2* are associated with QTLs for grain weight per plant, number of pods per plant, stem diameter, and plant height, as reported in previous research [21, 57, 58]. Comparing the 64 QTLs related to other traits with previous studies (Table 3), it was observed that 25 of them overlap or fall within the range of previously

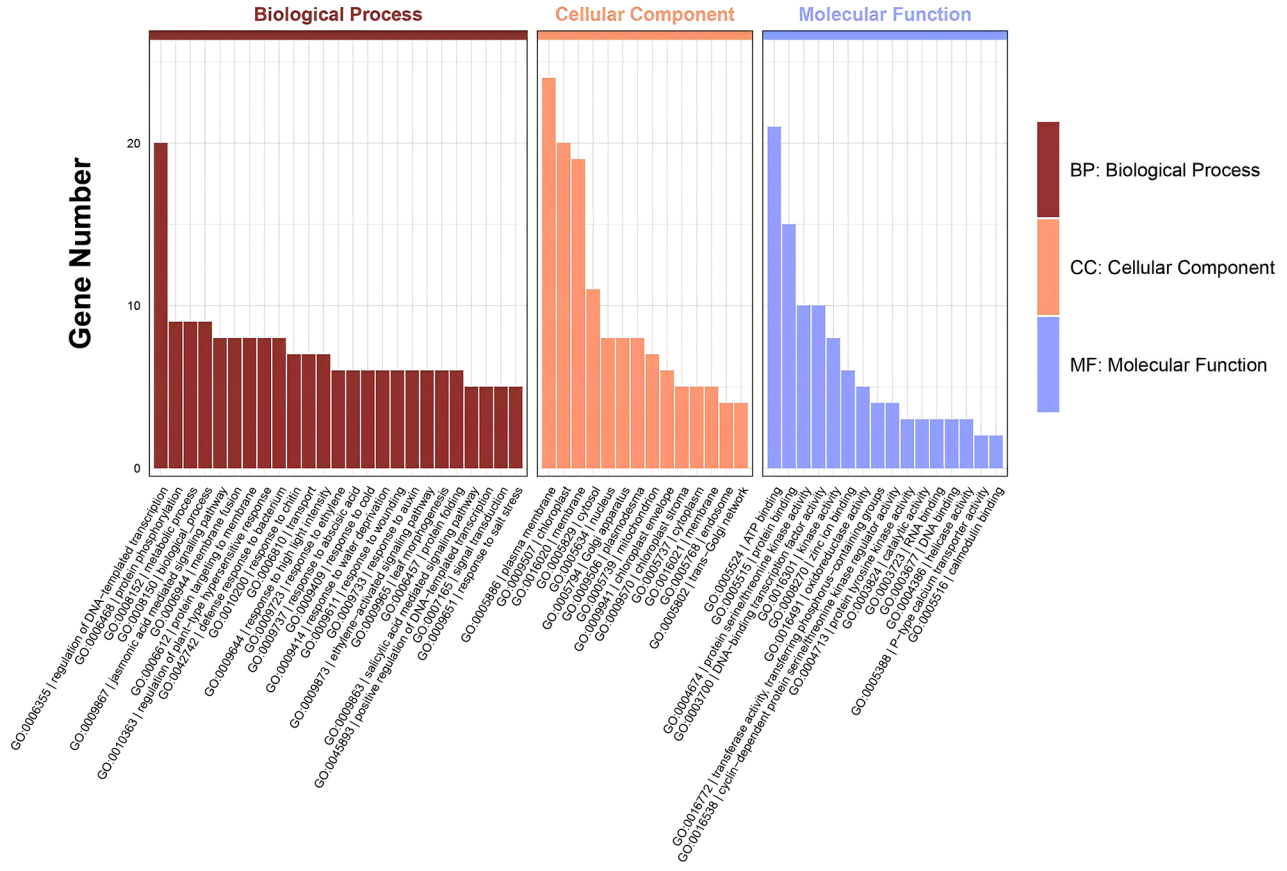

Fig. 4 Gene Ontology Annotation Analysis of Main and Stable QTL (*qLD-4-1*) Interval Genes

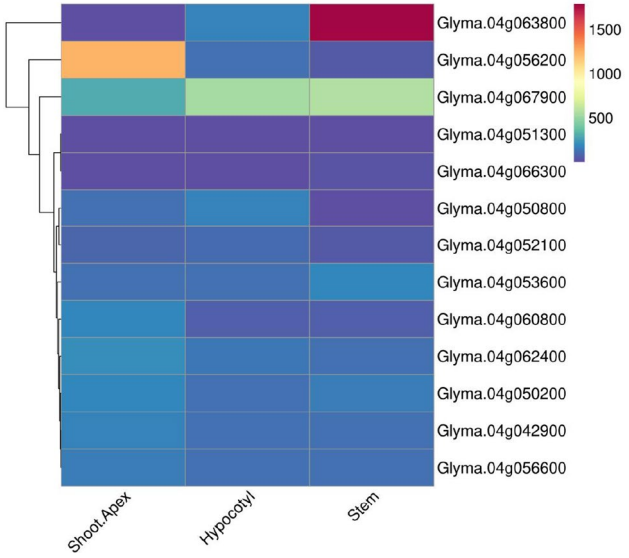

Fig. 5 Expression of Candidate Genes in the *qLD-4-1* Region in Different Soybean Tissues The heat map displays the expression levels of 13 candidate genes in the *qLD-4-1* region across various soybean tissues, as analyzed using ePlant

identified QTL locations. Notably, the primary QTLs for these traits consistently align with the main QTL for lodging score (*qLD-4-1*). In summary, many QTLs associated with lodging identified in this study show correlations with prior research results [21, 22, 30, 35, 37, 52–56, 59–70]. The presence of shared QTL loci for various traits, such as growth period, seed type, grain weight per plant, number of pods per plant, stem diameter, and plant height, suggests the reliability of the QTL mapping in this study. It is plausible that soybean lodging may be associated with genes controlling the growth period, a trait influenced by multiple genes.

Putative genes for the lodging resistant trait in soybean

Lodging resistance in soybean is influenced by various factors. One key factor is wheat lignin production, which plays a critical role in countering lodging [71, 72]. Wheat modifies the structure of its cell walls through the regulation of pathways related to hormones, reactive oxygen species, and nitrogen assimilation. This process limits cell wall loosening, restricts cell elongation, and enhances lodging resistance [73]. Wheat that is sensitive to gibberellin exhibits improved lodging resistance due to its semi-dwarf stature [74]. The mechanical strength of sorghum

Table 5 Candidate genes in the *qLD-4-1* region of the Population

Gene IDs [a]	Start	Stop	Homolog in Arabidopsis[b]	Gene descriptions	PFAM descriptions [c]	Panther descriptions [d]
Glyma.04g042900	3,452,611	3,454,681	*AT4G24210.1*	F-box family protein	NA	NA
Glyma.04g050200	4,075,901	4,081,260	*AT2G25930.1*	Hydroxyproline-rich glycoprotein family protein	NA	NA
Glyma.04g050800	4,128,984	4,133,372	*AT4G32980.1*	Homeobox gene 1	Homeobox domain; Associated with HOX	Homeobox protein transcription factors
Glyma.04g051300	4,162,013	4,163,673	*AT4G26150.1*	Cytokinin-responsive GATA factor 1	GATA zinc finger	Transcription factor GATA (GATA binding factor)
Glyma.04g052100	4,228,426	4,236,655	*AT2G26170.1*	Cytochrome P450, family 711, subfamily A, polypeptide 1	Cytochrome P450	NA
Glyma.04g053600	4,354,294	4,355,060	*AT4G01800.2*	Albino or Glassy Yellow 1	SecA Wing and Scaffold domain	NA
Glyma.04g056200	4,548,094	4,559,223	*AT2G26330.1*	Leucine-rich receptor-like protein kinase family protein	Protein kinase domain; Leucine Rich Repeat; Leucine rich repeat N-terminal domain	Leucine-rich repeat receptor-like protein kinase
Glyma.04g056600	4,593,407	4,599,841	*AT2G26300.1*	G protein alpha subunit 1	G-protein alpha subunit	GTP-binding protein alpha subunit
Glyma.04g060800	4,961,252	4,966,129	*AT4G31820.1*	Phototropic-responsive NPH3 family protein	BTB/POZ domain; NPH3 family	NA
Glyma.04g062400	5,134,682	5,156,284	*AT2G25170.1*	Chromatin remodeling factor CHD3	SNF2 family N-terminal domain; Helicase conserved C-terminal domain; Chromo (CHRromatin Organisation MOdifier) domain; PHD-finger	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily-related; chromatin remodeling 4 protein
Glyma.04g063800	5,293,426	5,299,847	*AT4G18780.1*	Cellulose synthase family protein	Cellulose synthase	X-box transcription factor-related
Glyma.04g066300	5,525,939	5,528,040	*AT4G32280.1*	Indole-3-acetic acid inducible 29	AUX/IAA family	NA
Glyma.04g067900	5,686,681	5,692,676	*AT4G32410.1*	Cellulose synthase 1	Cellulose synthase	X-box transcription factor-related

[a] Genes IDs from the Williams 82 soybean reference genome Wm82.a2.v1

[b] Genes IDs of *Arabidopsis* orthologs were obtained from the Phytozome 12

[c] PFAM (Protein families database of alignments and hidden Markov models) descriptions were obtained from the Soybase

[d] Panther (Protein analysis through evolutionary relationships) descriptions were obtained from the Soybase

stems may be linked to the production of secondary cell wall cellulose [75]. In corn, an increase in ethylene content leads to reduced plant height, thereby strengthening lodging resistance [76, 77]. Additionally, various plant hormones, including auxin, abscisic acid, jasmonic acid, and salicylic acid, may play roles in maize lodging resistance [78]. Rice employs dwarfing breeding techniques to reduce plant height and enhance lodging resistance [79]. Mutations at the CESA4 site in rice impact cell wall characteristics, particularly cellulose structure, resulting in improved biomass digestion and lodging resistance [80]. In *Arabidopsis*, gibberellin (GA) biosynthesis may also be involved in lodging control [81]. Furthermore, the accumulation of lignin and cellulose in soybeans can inhibit lodging [50].

We identified gene models within the *qLD-4-1* interval using publicly available data from Soybase. Within the physical interval of *qLD-4-1*, we found a total of 271 genes. Using high-quality resequencing data, we identified structural variations in 7 out of the 13 candidate genes among the parents in the RIL population. These seven candidate genes are as follows: (1) *Glyma.04g050200*, which is homologous to *Arabidopsis ELF3* (*AT2G25930.1*), and control of soybean flowering genes and gibberellin biosynthesis [82, 83]; (2) *Glyma.04g050800*, which is homologous to *Arabidopsis ATH1* (*AT4G32980.1*), and inhibiting stem growth and affecting lodging [84]; (3) *Glyma.04g051300*, which is homologous to *Arabidopsis CGA1* (*AT4G26150.1*), and involved in influencing internal plant hormones, including cytokinins and gibberellin (GA) [85]; (4) *Glyma.04g052100*, which is homologous to *Arabidopsis CYP711A1* (*AT2G26170.1*), and inhibition of axillary bud growth may be achieved by regulating flavonoid dependent auxin retention in buds and stems [86]; (5) *Glyma.04g053600*, which is homologous to *Arabidopsis AGY1* (*AT4G01800.2*), and AtcpSecA plays a crucial role in the biogenesis of chloroplasts, as its deletion triggers retrograde signaling, ultimately leading to chloroplast weight programming and mitochondrial gene expression

Table 6 Information on candidate genes in *qLD-4-1* in the Population

Number	Gene ID	Loci	Ref [a]	Alt [b]	Guizao 1	B13	Region
1	*Glyma.04g050200*	4,074,911	TC	T	T	TC	downstream
		4,077,171	AT	A	AT	A	exonic frameshift deletion
		4,083,007	G	GA	GA	G	Intergenic
		4,083,045	AT	A	A	AT	Intergenic
		4,084,083	G	GTT, GT	GT	G	Intergenic
		4,084,277	T	TAA	TAA	T	Intergenic
		4,086,985	T	TTC	TTC	T	Intergenic
		4,087,396	AT	A	A	AT	Intergenic
2	*Glyma.04g050800*	4,132,841	GATTC	G	G	GATTC	UTR5
		4,133,337	CA	C	C	CA	UTR5
		4,134,036	TTTC	T	T	TTTC	upstream
		4,135,015	T	TA	TA	T	Intergenic
3	*Glyma.04g051300*	4,162,474	G	GC	GC	G	intronic
		4,165,298	A	AAC	AAC	A	Intergenic
4	*Glyma.04g052100*	4,231,206	GA	G	G	GA	intronic
		4,233,151	A	ATG	ATG	A	intronic
		4,234,057	TTATA	T	T	TTATA	intronic
		4,234,943	C	CTAG, CTAGTAG	CTAGTAG	C	exonic nonframeshift insertion
5	*Glyma.04g053600*	4,355,043	CG	C	C	CG	UTR3
		4,355,492	A	ATC	ATC	A	downstream
6	*Glyma.04g056200*	4,548,436	T	TCTCTCC	TCTCTCC	T	UTR5
		4,548,811	T	TG	TG	T	intronic
		4,550,545	GT	G	G	GT	intronic
7	*Glyma.04g063800*	5,293,286	TTA	T	T	TTA	downstream
		5,293,287	TA	T	T	TA	downstream
		5,293,380	TA	T	T	TA	downstream
		5,295,813	CT	C	C	CT	intronic
		5,296,757	TTAGCTTTATA	T	T	TTAGCTTTATA	intronic
		5,298,299	G	GTT, GT	GTT	G	intronic
		5,302,171	C	CAA	CAA	C	Intergenic
		5,304,243	T	TATG	TATG	T	Intergenic
		5,304,264	C	CT	CT	C	Intergenic
		5,304,351	C	CGATTGCATAGT	CGATTGCATAGT	C	Intergenic
		5,304,501	CG	C	C	CG	Intergenic
		5,304,614	C	CG	CG	C	Intergenic
		5,304,796	TC	T	T	TC	Intergenic
		5,305,881	T	TA	TA	T	Intergenic

[a] Reference indicates the genotype of from the soybean reference genome Wm82.a2.v1

[b] Alt indicates the genotype of alter

[c] Female parent of GB RIL population

[d] Male parent of GB RIL population

[e] Region of variation in genes

[87]; (6) *Glyma.04g056200*, which is homologous to *Arabidopsis ER* (*AT2G26330.1*), and the important role of soybean native genes (*GmER* and *GmERL*) in soybean growth and stress response, and the truncation of *Arabidopsis ERECTA* gene can be used to regulate the growth and stress response of different crop varieties [88, 89]; (7) *Glyma.04g063800*, which is homologous to *Arabidopsis IRX1* (*AT4G18780.1*), and *Arabidopsis* xylem is the *CesA* gene synthesized by cellulose in the secondary wall of cotton fibers, which is used in different ways to construct specific specialized cell walls [90]. Further analysis is necessary to validate the specific functions of these 7 genes. Subsequently, it was determined through qRT-PCR that four genes are likely to be the major genes controlling soybean lodging. The findings of this study will provide valuable insights for future research. To determine the true controlling gene of lodging, this study consulted the expression information of different genes in public databases and conducted qRT-PCR experiments on the presumed candidate genes. The relative expression level of

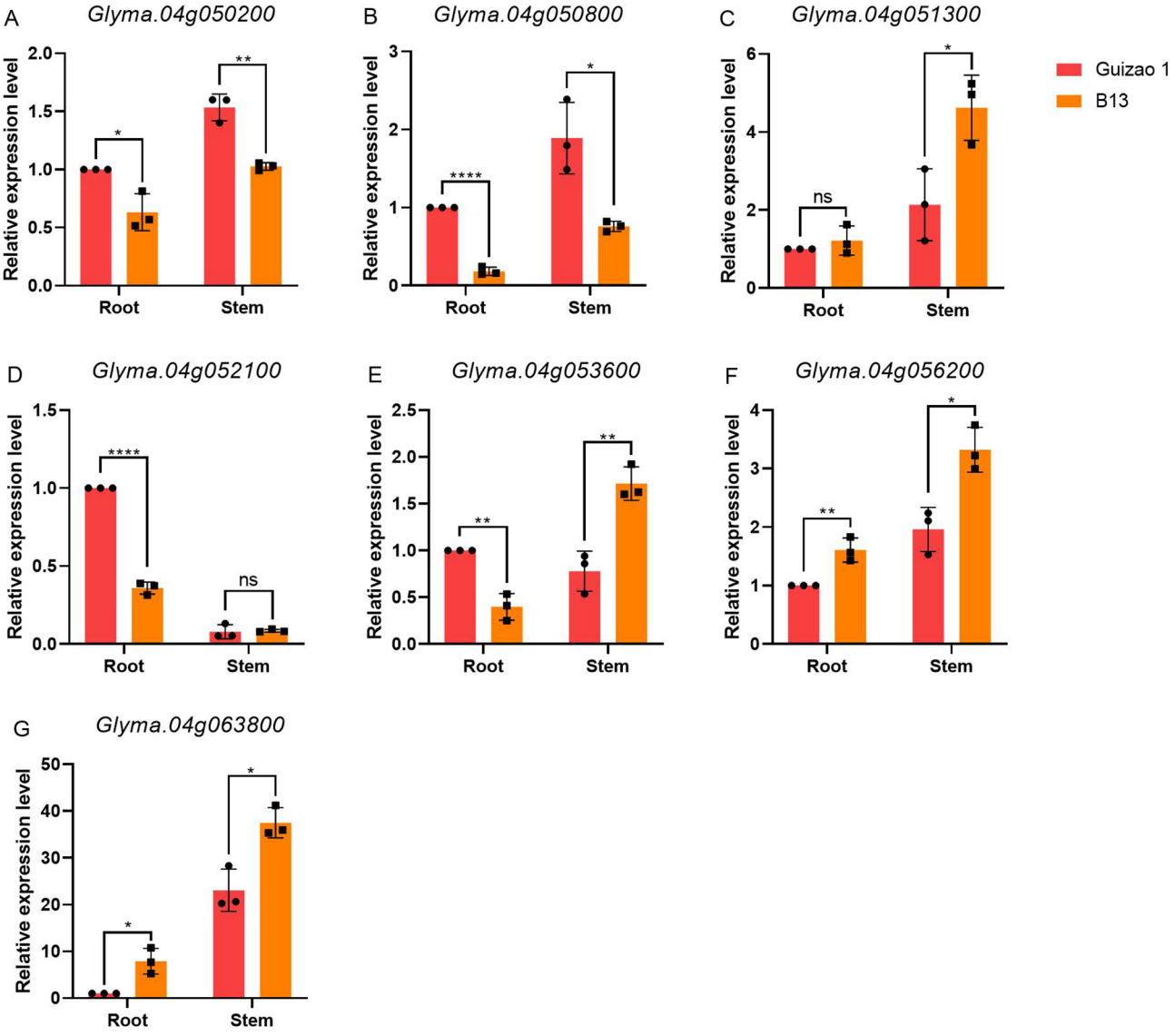

Fig. 6 Expression of Seven Candidate Genes in Root and Stem Tissues of the Two Parent Plants Detected by qRT-PCR. The y-axis represents the relative expression levels of candidate genes compared to the expression in the roots of Guizao 1. A to G denote the relative expression of different candidate genes, respectively. The female parent of the GB RIL population is Guizao 1, and the male parent is B13. Error bars indicate standard deviation ($n=3$). Asterisks indicate significant differences determined by Student's t test (*, $P<0.05$; **, $P<0.01$; ***, $P<0.001$; ****, $P<0.0001$)

Glyma.04g063800 is the highest. Therefore, we believe that this gene is most likely the gene controlling lodging.

Conclusions

In conclusion, this study employed high-density genetic linkage maps for QTL mapping and conducted correlation analysis between lodging traits and other traits across six different environments. The lodging score displayed a strong and significant correlation with various traits, except for single plant grain weight. A total of 84 QTLs were identified, comprising 20 QTLs associated with lodging score, 11 with flowering time, 11 with maturity time, 12 with plant height, 9 with the number of main stem nodes, 10 with stem diameter, 6 with internode length, and 5 with grain weight per plant. These QTLs contributed to a phenotypic variation ranging from 1.26 to 66.87%, with LOD scores ranging from 2.52 to 69.22. Furthermore, this study placed particular emphasis on the consistent detection of the QTL *qLD-4-1* across all six environments and identified seven new candidate genes related to lodging. Subsequent qRT-PCR analysis revealed that four of these genes are likely to play a major role in controlling soybean lodging. The findings from this research provide valuable insights for a better understanding of lodging in soybeans and hold promise for future investigations aimed at enhancing soybean yield and lodging resistance through breeding efforts.

Abbreviations

QTL	Quantitative trait loci
RIL	Recombinant inbred line
LOD	Logarithm of odds
GO	Gene Ontology
SNP	Single nucleotide polymorphism

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10794-1.

Supplementary Material 1

Author contributions

B. C., C. C., M. D., X. Y., J. J., S. L., L. Y., Y. L., and N. H. collected the plant materials used in this study. Q. X. performed QTL mapping. B. C. and Y. C. prepared the first draft of the manuscript. B. C. and S. L. contributed to data analysis. Y. C., Q. M., Z. C., and H. N. planned, supervised, and financed this work, as well as edited the manuscript. All authors have read and approved the final version of the manuscript.

Funding

This work was supported by the China Agriculture Research System (CARS-04-PS12), the Key-Areas Research and Development Program of Guangdong Province (2020B02022008), and Guangdong Agricultural Research System (2023KJ136-03).

Data availability

The data that support the findings of this study are available from the Genome Sequence Archive database at the National Genomics Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, with accession number CRA004753 (https://bigd.big.ac.cn/gsa/browse/CRA004753) and CRA004754 (https://bigd.big.ac.cn/gsa/browse/CRA004754). The phenotype dataset used during the current study is provided in the Supplementary table: Table S3.

Declarations

Ethics approval and consent to participate

This study complies with relevant institutional, national, and international guidelines.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]The Key Laboratory of Plant Molecular Breeding of Guangdong Province, College of Agriculture, South China Agricultural University, Guangzhou, Guangdong 510642, People's Republic of China
[2]The State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, South China Agricultural University, Guangzhou, Guangdong 510642, People's Republic of China
[3]Guangdong Provincial Laboratory of Lingnan Modern Agricultural Science and Technology, South China Agricultural University, Guangzhou, Guangdong 510642, People's Republic of China
[4]Rice Molecular Breeding Institute, Granlux Associated Grains, Shenzhen, Guangdong 518023, People's Republic of China

Received: 23 February 2024 / Accepted: 11 September 2024
Published online: 30 September 2024

References

1. Dong L, Fang C, Cheng Q, Su T, Kou K, Kong L, Zhang C, Li H, Hou Z, Zhang Y, et al. Genetic basis and adaptation trajectory of soybean from its temperate origin to tropics. NAT COMMUN. 2021;12(1):5445.
2. Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, Li W, Yu Y, Shu L, Zhao Y, Ma Y, et al. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. NAT BIOTECHNOL. 2015;33(4):408–14.
3. Ray DK, Mueller ND, West PC, Foley JA. Yield trends are insufficient to double global crop production by 2050. PLoS ONE. 2013;8(6):e66428.
4. Liu S, Zhang M, Feng F, Tian Z. Toward a Green Revolution for soybean. MOL PLANT. 2020;13(5):688–97.
5. *Arachis hypogaea* L.). *THEOR APPL GENET* 2023, 136(5):105.
6. Kumar J, Gupta DS, Gupta S, Dubey S, Gupta P, Kumar S. Quantitative trait loci from identification to exploitation for crop improvement. PLANT CELL REP. 2017;36(8):1187–213.
7. Des Marais DL, Razzaque S, Hernandez KM, Garvin DF, Juenger TE. Quantitative trait loci associated with natural diversity in water-use efficiency and response to soil drying in Brachypodium distachyon. PLANT SCI. 2016;251:2–11.
8. *THEOR APPL GENET* 1990, 79(4):465–469.
9. Kushanov FN, Turaev OS, Ernazarova DK, Gapparov BM, Oripova BB, Kudratova MK, Rafieva FU, Khalikov KK, Erjigitov DS, Khidirov MT, et al. Genetic diversity, QTL mapping, and marker-assisted Selection Technology in Cotton (*Gossypium* spp). FRONT PLANT SCI. 2021;12:779386.
10. Sahu BB, Sumit R, Srivastava SK, Bhattacharyya MK. Sequence based polymorphic (SBP) marker technology for targeted genomic regions: its application in generating a molecular map of the *Arabidopsis thaliana* genome. BMC Genomics. 2012;13:20.
11. Kumar LS. DNA markers in plant improvement: an overview. BIOTECHNOL ADV. 1999;17(2–3):143–82.
12. Kantety RV, La Rota M, Matthews DE, Sorrells ME. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. PLANT MOL BIOL. 2002;48(5–6):501–10.
13. Karp A, Edwards KJ, Bruford M, Funk S, Vosman B, Morgante M, Seberg O, Kremer A, Boursot P, Arctander P, et al. Molecular technologies for biodiversity evaluation: opportunities and challenges. NAT BIOTECHNOL. 1997;15(7):625–8.
14. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. AFLP: a new technique for DNA fingerprinting. NUCLEIC ACIDS RES. 1995;23(21):4407–14.
15. *THEOR APPL GENET* 1989, 77(3):353–359.
16. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. AM J HUM GENET. 1980;32(3):314–31.
17. Stolting KN, Nipper R, Lindtke D, Caseys C, Waeber S, Castiglione S, Lexer C. Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. MOL ECOL. 2013;22(3):842–55.
18. Ganal MW, Altmann T, Roder MS. SNP identification in crop plants. CURR OPIN PLANT BIOL. 2009;12(2):211–7.
19. Berraies S, Cuthbert R, Knox R, Singh A, DePauw R, Ruan Y, Bokore F, Henriquez MA, Kumar S, Burt A, et al. High-density genetic mapping of Fusarium head blight resistance and agronomic traits in spring wheat. FRONT PLANT SCI. 2023;14:1134132.
20. Gao F, Wen W, Liu J, Rasheed A, Yin G, Xia X, Wu X, He Z. Genome-wide linkage mapping of QTL for Yield Components, Plant Height and Yield-related physiological traits in the Chinese Wheat Cross Zhou 8425B/Chinese Spring. FRONT PLANT SCI. 2015;6:1099.
21. Sun CY, Yang YM, Jia L, Liu XQ, Xu HQ, Lv HY, Huang ZW, Zhang D. QTL mapping of the genetic basis of stem diameter in soybean. Planta. 2021;253(5):109.
22. Lee S, Jun TH, Michel AP, Rouf Mian MA. SNP markers linked to QTL conditioning plant height, lodging, and maturity in soybean. Euphytica. 2015;203(3):521–32.
23. *Lolium multiflorum* Lam.). *THEOR APPL GENET* 2004, 109(8):1576–1585.
24. Tar'An B, Warkentin T, Somers DJ, Miranda D, Vandenberg A, Blade S, Woods S, Bing D, Xue A, DeKoeyer D et al: Quantitative trait loci for lodging resistance, plant height and partial resistance to mycosphaerella blight in field pea (Pisum sativum L.). THEOR APPL GENET 2003, 107(8):1482-1491.
25. Glycine max L. Merr. THEOR APPL GENET. 1993;86(8):914–8.
26. *THEOR APPL GENET* 1996, 92(5):516–523.

27. Orf JH, Chase K, Adler FR, Mansur LM, Lark KG. Genetics of soybean agronomic traits: II. Interactions between Yield Quantitative Trait Loci in soybean. CROP SCI. 1999;39(6):1652–7.
28. Specht JE, Chase K, Macrander M, Graef GL, Chung J, Markwell JP, Germann M, Orf JH, Lark KG. Soybean response to Water: a QTL analysis of Drought Tolerance. CROP SCI. 2001;41(2):493–509.
29. Chapman A, Pantalone VR, Ustun A, Allen FL, Landau-Ellis D, Trigiano RN, Gresshoff PM. Quantitative trait loci for agronomic and seed quality traits in an F_2 and $F_{4:6}$ soybean population. Euphytica. 2003;129(3):387–93.
30. *THEOR APPL GENET* 2004, 108(3):458–467.
31. *Glycine max* L. Merr.) genetic map and their association with EST markers. *THEOR APPL GENET* 2004, 108(6):1131–1139.
32. *THEOR APPL GENET* 2013, 126(7):1809–1823.
33. × high-yielding exotic soybean lines. *THEOR APPL GENET* 2009, 119(3):429–436.
34. Guzman PS, Diers BW, Neece DJ, St. Martin SK, LeRoy AR, Grau CR, Hughes TJ, Nelson RL. QTL Associated with yield in three backcross-derived populations of soybean. CROP SCI. 2007;47(1):111–22.
35. Reinprecht Y, Poysa VW, Yu K, Rajcan I, Ablett GR, Pauls KP. Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (Glycine max (L.) Merrill) germplasm. GENOME. 2006;49(12):1510–27.
36. *THEOR APPL GENET* 2010, 120(5):997–1003.
37. *THEOR APPL GENET* 2012, 125(6):1353–1369.
38. *THEOR APPL GENET* 2019, 132(6):1833–1845.
39. Kong L, Lu S, Wang Y, Fang C, Wang F, Nan H, Su T, Li S, Zhang F, Li X, et al. Quantitative trait locus mapping of flowering time and Maturity in soybean using next-generation sequencing-based analysis. FRONT PLANT SCI. 2018;9:995.
40. Chen H, Shan Z, Sha A, Wu B, Yang Z, Chen S, Zhou R, Zhou X. Quantitative trait loci analysis of stem strength and related traits in soybean. Euphytica. 2011;179(3):485–97.
41. Luo S, Jia J, Liu R, Wei R, Guo Z, Cai Z, Chen B, Liang F, Xia Q, Nian H, et al. Identification of major QTLs for soybean seed size and seed weight traits using a RIL population in different environments. FRONT PLANT SCI. 2022;13:1094112.
42. Asekova S, Kulkarni KP, Kim M, Kim JH, Song JT, Shannon JG, Lee JD. Novel quantitative trait loci for Forage Quality traits in a Cross between PI 483463 and 'Hutcheson' in soybean. CROP SCI. 2016;56(5):2600–11.
43. *THEOR APPL GENET* 2014, 127(3):659–675.
44. JIA J, WANG H, CAI Z, WEI R, HUANG J, XIA Q, XIAO X, NIAN MAQ, CHENG H. Identification and validation of stable and novel quantitative trait loci for pod shattering in soybean [Glycinemax (L.) Merr]. J INTEGR AGR. 2022;21(11):3169–84.
45. Xian P, Cai Z, Jiang B, Xia Q, Cheng Y, Yang Y, Zhou Q, Lian T, Ma Q, Wang Y, et al. GmRmd1 encodes a TIR-NBS-BSP protein and confers resistance to powdery mildew in soybean. PLANT COMMUN. 2022;3(6):100418.
46. *THEOR APPL GENET* 2018, 131(8):1715–1728.
47. Wang Y, Yu K, Poysa V, Shi C, Zhou Y. Selection of reference genes for normalization of qRT-PCR analysis of differentially expressed genes in soybean exposed to cadmium. MOL BIOL REP. 2012;39(2):1585–94.
48. Xian P, Cai Z, Cheng Y, Lin R, Lian T, Ma Q, Nian H. Wild Soybean Oxalyl-CoA Synthetase Degrades Oxalate and Affects the Tolerance to Cadmium and Aluminum Stresses. *INT J MOL SCI* 2020, 21(22).
49. Wang H, Lu H, Yang Z, Zhang Z, Li M, Zhang Z, Dai W, Song X, Olsen KM, Qiang S. Characterization of lodging variation of weedy rice. J EXP BOT. 2023;74(5):1403–19.
50. Hussain S, Shuxian L, Mumtaz M, Shafiq I, Iqbal N, Brestic M, Shoaib M, Sisi Q, Li W, Mei X, et al. Foliar application of silicon improves stem strength under low light stress by regulating lignin biosynthesis genes in soybean (*Glycine max* (L.) Merr). J HAZARD MATER. 2021;401:123256.
51. Mehrabi Z, Ramankutty N. Synchronized failure of global crop production. NAT ECOL EVOL. 2019;3(5):780–6.
52. *THEOR APPL GENET* 2021, 134(4):1007–1014.
53. Fang T, Bai Y, Huang W, Wu Y, Yuan Z, Luan X, Liu X, Sun L. Identification of potential Gene Regulatory pathways affecting the ratio of four-seed pod in soybean. FRONT GENET 2021, 12.
54. Yue Y, Liu N, Jiang B, Li M, Wang H, Jiang Z, Pan H, Xia Q, Ma Q, Han T, et al. A single nucleotide deletion in J encoding GmELF3 confers long juvenility and is Associated with adaption of tropic soybean. MOL PLANT. 2017;10(4):656–8.
55. Zimmer G, Miller MJ, Steketee CJ, Jackson SA, de Tunes L, Li Z. Genetic control and allele variation among soybean maturity groups 000 through IX. PLANT GENOME-US. 2021;14(3):e20146.
56. Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, Xia Z, Nagamatsu A, Arai M, Yamada T, Kitamura K, et al. The soybean stem growth habit gene Dt1 is an ortholog of *Arabidopsis TERMINAL FLOWER1*. PLANT PHYSIOL. 2010;153(1):198–210.
57. Yao D, Liu ZZ, Zhang J, Liu SY, Qu J, Guan SY, Pan LD, Wang D, Liu JW, Wang PW. Analysis of quantitative trait loci for main plant traits in soybean. GENET MOL RES. 2015;14(2):6101–9.
58. Glycine max L. Merr. THEOR APPL GENET. 2006;112(4):665–73.
59. Lu Y, Zhang J, Guo X, Chen J, Chang R, Guan R, Qiu L. Identification of Genomic Regions Associated with Vine Growth and Plant Height of soybean. INT J MOL SCI 2022, 23(10).
60. Hu B, Li Y, Wu H, Zhai H, Xu K, Gao Y, Zhu J, Li Y, Xia Z. Identification of quantitative trait loci underlying five major agronomic traits of soybean in three biparental populations by specific length amplified fragment sequencing (SLAF-seq). PEERJ. 2021;9:e12416.
61. *THEOR APPL GENET* 2020, 133(3):1039–1054.
62. Fang Y, Liu S, Dong Q, Zhang K, Tian Z, Li X, Li W, Qi Z, Wang Y, Tian X, et al. Linkage Analysis and Multi-locus genome-wide Association studies identify QTNs Controlling soybean plant height. FRONT PLANT SCI. 2020;11:9.
63. Ogiso-Tanaka E, Shimizu T, Hajika M, Kaga A, Ishimoto M. Highly multiplexed AmpliSeq technology identifies novel variation of flowering time-related genes in soybean (*Glycine max*). DNA RES. 2019;26(3):243–60.
64. Assefa T, Otyama PI, Brown AV, Kalberer SR, Kulkarni RS, Cannon SB. Genome-wide associations and epistatic interactions for internode number, plant height, seed weight and seed yield in soybean. BMC Genomics. 2019;20(1):527.
65. Li R, Jiang H, Zhang Z, Zhao Y, Xie J, Wang Q, Zheng H, Hou L, Xiong X, Xin D et al. Combined linkage mapping and BSA to identify QTL and candidate genes for Plant Height and the number of nodes on the main stem in soybean. INT J MOL SCI 2019, 21(1).
66. Cao Y, Li S, Chen G, Wang Y, Bhat JA, Karikari B, Kong J, Gai J, Zhao T. Deciphering the Genetic Architecture of Plant Height in soybean using two RIL populations sharing a common M8206 parent. PLANTS-BASEL 2019, 8(10).
67. Oki N, Sayama T, Ishimoto M, Yokota I, Kaga A, Takahashi M, Takahashi M. Quantitative trait loci associated with short inter-node length in soybean. Breed SCI. 2018;68(5):554–60.
68. Zhang X, Wang W, Guo N, Zhang Y, Bu Y, Zhao J, Xing H. Combining QTL-seq and linkage mapping to fine map a wild soybean allele characteristic of greater plant height. BMC Genomics. 2018;19(1):226.
69. *THEOR APPL GENET* 2017, 130(11):2297–2314.
70. Watanabe S, Xia Z, Hideshima R, Tsubokura Y, Sato S, Yamanaka N, Takahashi R, Anai T, Tabata S, Kitamura K, et al. A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. Genetics. 2011;188(2):395–407.
71. Muszynska A, Guendel A, Melzer M, Tandron Moya YA, Röder MS, Rolletschek H, Rutten T, Munz E, Melz G, Ortleb S, et al. A mechanistic view on lodging resistance in rye and wheat: a multiscale comparative study. PLANT BIOTECHNOL J. 2021;19(12):2646–61.
72. Ma QH. Functional analysis of a cinnamyl alcohol dehydrogenase involved in lignin biosynthesis in wheat. J EXP BOT. 2010;61(10):2735–44.
73. Chai S, Yao Q, Zhang X, Xiao X, Fan X, Zeng J, Sha L, Kang H, Zhang H, Li J et al. The semi-dwarfing gene rht-dp from dwarf Polish wheat (*Triticum polonicum* L.) is the Green Revolution gene Rht-B1b. BMC Genomics 2021, 22(1).
74. Ford BA, Foo E, Sharwood R, Karafiatova M, Vrana J, MacMillan C, Nichols DS, Steuernagel B, Uauy C, Doležel J, et al. Rht18 semidwarfism in wheat is due to increased GA 2-oxidaseA9 expression and reduced GA Content. PLANT PHYSIOL. 2018;177(1):168–80.
75. Li P, Liu Y, Tan W, Chen J, Zhu M, Lv Y, Liu Y, Yu S, Zhang W, Cai H. Brittle culm 1 encodes a COBRA-Like protein involved in secondary cell Wall Cellulose Biosynthesis in Sorghum. PLANT CELL PHYSIOL. 2019;60(4):788–801.
76. Li H, Wang L, Liu M, Dong Z, Li Q, Fei S, Xiang H, Liu B, Jin W. Maize Plant Architecture is regulated by the Ethylene Biosynthetic Gene ZmACS7. PLANT PHYSIOL. 2020;183(3):1184–99.
77. Shi J, Drummond BJ, Habben JE, Brugire N, Weers BP, Hakimi SM, Lafitte HR, Schussler JR, Mo H, Beatty M, et al. Ectopic expression of ARGOS8 reveals a role for ethylene in root-lodging resistance in maize. Plant J. 2019;97(2):378–90.
78. Liu L, Liu S, Lu H, Tian Z, Zhao H, Wei D, Wang S, Huang Z. Integration of transcriptome and metabolome analyses reveals Ke lodging-resistance-related genes and metabolic pathways in maize. FRONT GENET. 2022;13:1001195.

79. Ke S, Luan X, Liang J, Hung Y, Hsieh T, Zhang X. Rice OsPEX1, an extensin-like protein, affects lignin biosynthesis and plant growth. PLANT MOL BIOL. 2019;100(1–2):151–61.
80. Li F, Liu S, Xu H, Xu Q. A novel FC17/CESA4 mutation causes increased biomass saccharification and lodging resistance by remodeling cell wall in rice. BIOTECHNOL BIOFUELS 2018, 11(1).
81. Guo S, Zhang X, Bai Q, Zhao W, Fang Y, Zhou S, Zhao B, He L, Chen J. Cloning and Functional Analysis of Dwarf Gene *Mini Plant 1* (*MNP1*) in *Medicago truncatula*. *INT J MOL SCI* 2020, 21(14).
82. Liew LC, Singh MB, Bhalla PL. A novel role of the soybean clock gene *LUX ARRHYTHMO* in male reproductive development. SCI REP-UK. 2017;7(1):10605.
83. Lu S, Zhao X, Hu Y, Liu S, Nan H, Li X, Fang C, Cao D, Shi X, Kong L, et al. Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. NAT GENET. 2017;49(5):773–9.
84. Gomez-Mena C, Sablowski R. *ARABIDOPSIS THALIANA HOMEOBOX GENE1* establishes the basal boundaries of shoot organs and controls stem growth. PLANT CELL. 2008;20(8):2059–72.
85. Hudson D, Guevara D, Yaish MW, Hannam C, Long N, Clarke JD, Bi YM, Rothstein SJ. GNC and CGA1 modulate chlorophyll biosynthesis and glutamate synthase (GLU1/Fd-GOGAT) expression in Arabidopsis. PLoS ONE. 2011;6(11):e26765.
86. Lazar G, Goodman HM. MAX1, a regulator of the flavonoid pathway, controls vegetative axillary bud outgrowth in Arabidopsis. P NATL ACAD SCI USA. 2006;103(2):472–6.
87. Liu D, Gong Q, Ma Y, Li P, Li J, Yang S, Yuan L, Yu Y, Pan D, Xu F, et al. cpSecA, a thylakoid protein translocase subunit, is essential for photosynthetic development in Arabidopsis. J EXP BOT. 2010;61(6):1655–69.
88. Shanmugam S, Zhao S, Nandy S, Srivastava V, Khodakovskaya M. Modification of soybean growth and abiotic stress tolerance by expression of truncated ERECTA protein from *Arabidopsis thaliana*. PLoS ONE. 2020;15(5):e233383.
89. Villagarcia H, Morin AC, Shpak ED, Khodakovskaya MV. Modification of tomato growth by expression of truncated ERECTA protein from *Arabidopsis thaliana*. J EXP BOT. 2012;63(18):6493–504.
90. Betancur L, Singh B, Rapp RA, Wendel JF, Marks MD, Roberts AW, Haigler CH: Phylogenetically distinct cellulose synthase genes support secondary wall thickening in arabidopsis shoot trichomes and cotton fiber. J Integr Plant Biol 2010, 52(2):205–220.

Publisher's note