

RESEARCH ARTICLE

# Phenotypic Characterization and Genetic Dissection of Growth Period Traits in Soybean (*Glycine max*) Using Association Mapping

Zhangxiong Liu<sup>1</sup>✉, Huihui Li<sup>1</sup>✉, Xuhong Fan<sup>2</sup>✉, Wen Huang<sup>3</sup>, Jiyu Yang<sup>4</sup>, Candong Li<sup>5</sup>, Zixiang Wen<sup>6</sup>, Yinghui Li<sup>1</sup>, Rongxia Guan<sup>1</sup>, Yong Guo<sup>1</sup>, Ruzhen Chang<sup>1</sup>, Dechun Wang<sup>6</sup>, Shuming Wang<sup>2\*</sup>, Li-Juan Qiu<sup>1\*</sup>

**1** National Key Facility for Gene Resources and Genetic Improvement, Key Laboratory of Crop Germplasm Utilization, Ministry of Agriculture, Institute of Crop Sciences, Chinese Academy of Agricultural Science, Beijing, China, **2** Institute of Soybean Research, Jilin Academy of Agricultural Sciences, Changchun, China, **3** Tonghua Academy of Agricultural Sciences, Meihekou, China, **4** Jilin City Academy of Agricultural Sciences, Jilin, China, **5** Jiamusi Branch of Heilongjiang Agricultural Sciences, Jiamusi, China, **6** Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, Michigan, United States of America

✉ These authors contributed equally to this work.

\* [qjulijuan@caas.cn](mailto:qjulijuan@caas.cn) (LQ); [shumingw@263.net](mailto:shumingw@263.net) (SW)



## OPEN ACCESS

**Citation:** Liu Z, Li H, Fan X, Huang W, Yang J, Li C, et al. (2016) Phenotypic Characterization and Genetic Dissection of Growth Period Traits in Soybean (*Glycine max*) Using Association Mapping. PLoS ONE 11(7): e0158602. doi:10.1371/journal.pone.0158602

**Editor:** Xianlong Zhang, National Key Laboratory of Crop Genetic Improvement, CHINA

**Received:** April 7, 2016

**Accepted:** June 18, 2016

**Published:** July 1, 2016

**Copyright:** © 2016 Liu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the Crop Germplasm Resources Protection (2014NWB030, 2015NWB030-05), Platform of National Crop Germplasm Resources of China (2014-004, 2015-004), The National Key Technology R&D Program (2011BAD35B06-2-9), and The Agricultural Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS).

## Abstract

The growth period traits are important traits that affect soybean yield. The insights into the genetic basis of growth period traits can provide theoretical basis for cultivated area division, rational distribution, and molecular breeding for soybean varieties. In this study, genome-wide association analysis (GWAS) was exploited to detect the quantitative trait loci (QTL) for number of days to flowering (ETF), number of days from flowering to maturity (FTM), and number of days to maturity (ETM) using 4032 single nucleotide polymorphism (SNP) markers with 146 cultivars mainly from Northeast China. Results showed that abundant phenotypic variation was presented in the population, and variation explained by genotype, environment, and genotype by environment interaction were all significant for each trait. The whole accessions could be clearly clustered into two subpopulations based on their genetic relatedness, and accessions in the same group were almost from the same province. GWAS based on the unified mixed model identified 19 significant SNPs distributed on 11 soybean chromosomes, 12 of which can be consistently detected in both planting densities, and 5 of which were pleiotropic QTL. Of 19 SNPs, 7 SNPs located in or close to the previously reported QTL or genes controlling growth period traits. The QTL identified with high resolution in this study will enrich our genomic understanding of growth period traits and could then be explored as genetic markers to be used in genomic applications in soybean breeding.

## Introduction

Soybean is a typical short-day plant [1]. Soybean cultivars adapted to specific geographical regions often differ in day-length perception that affect the length of time required to reach a

**Competing Interests:** The authors have declared that no competing interests exist.

given growth stage [2]. Soybean growth stages show high correlations with grain yield, plant height, pod number, and seed weight [3]. Therefore, understanding the genetic architecture of growth period traits will be useful for cultivated area division, rational distribution, and molecular breeding for soybean varieties.

Soybean was grown at a wide range of latitudes, from at least 50°N to 35°S [4], however, each soybean cultivar generally has a specific geographic or latitudinal distribution to reach its high yield. Thirteen soybean maturity groups (MG 000, MG 00, MG 0, and MG I—X) have been well developed to estimate the range of adaptability to latitudinal or geographic zones of soybean cultivars in the USA and Canada [5–6]. The growth period of soybean can be divided into vegetative growth period, from seedling to flowering, and reproductive growth period, from flowering to maturity. Hence, flowering is the transition from the two stages, which is a critical event in the life cycle of plants. Since the genes controlling soybean maturity were first reported by Woodworth (1923) [7], a series of loci (*E1* to *E9*, and *J*) that control flowering time and maturity of soybean were identified, and characterized at the phenotypic and genetic levels by classic methods [4, 8–15]. To date, the molecular mechanism for *E1* to *E4* loci has been uncovered [16–19], while other loci such as *E5* to *E8* remain unknown. It was reported that most of the *E* series genes had pleiotropic effects on vegetative growth period, reproductive growth period, and growth period. For example, the dominant alleles of *E2* to *E4* prolonged the reproductive period from stages R1 (days to flowering of the first flower) to R8 (full maturity) to some extent [20–21]. Relationship between maturity groups and genotypes of the *E* loci was inferred by Harosoy or Clark near isogenic lines (NILs) [22]. Besides the cloned maturity genes, two homologs of *Flowering Locus T* (*FT*), *GmFT2a*, and *GmFT5a*, were found to encode the components of ‘florigen’, which is the mobile flowering promotion signal involved in the transition to flowering, and to coordinately control flowering in soybean [23]. Recently, the *GmFT2a* was validated to be gene *E9* [24].

For the last decades, QTL mapping based on bi-parental populations has been extensively used to dissect the genetic architecture of complex traits for soybean. From SoyBase (<http://www.soybase.org/>) 106 QTL related to R1, 2 QTL related to reproductive stage length (the number of days between R1 and R8), and 156 QTL related to R8, identified by bi-parental populations, could be retrieved. These QTLs were mainly distributed on the 6<sup>th</sup>, 7<sup>th</sup>, and 19<sup>th</sup> chromosomes, but the physical positions and molecular functions for most of these QTL were not confirmed yet. Genome-wide association analysis (GWAS), which exploits historical recombination events, has become a powerful complementary strategy to linkage mapping for complex trait dissection at the sequence level. It has been successfully used in *Arabidopsis thaliana* [25], rice [26], maize [27], wheat [28], etc. In soybean, GWAS has been conducted for series of complex quantitative traits, such as disease resistance [29], stress tolerance [30], yield and quality related traits [31–32], and photosynthesis [33].

The aim of our study was to investigate the genetic architecture of three growth period traits, the number of days to flowering (ETF), the number of days from flowering to maturity (FTM), and the number of days to maturity (ETM) by high-throughput genetic markers in soybean. The uncovered genetic architecture will enrich our genomic understanding for growth period traits, and enhance the genetic gain to breed high yield soybean cultivars by genomic assistant breeding.

## Materials and Methods

### Plant materials and phenotypic data collection

To construct a diversity panel of phenotypes, 146 soybean accessions mainly from Northeast China were selected. The phenotypes were evaluated in two planting densities conditions, low

density with one seed per 10 cm, and high density with two seeds per 10 cm. All the accessions were grown in three-row plots with 3 m in length and 0.60 m spacing between rows in a randomized blocks design with 2 replications at 4 locations, Fanjiatun experiment station of Jilin Academy of Agricultural Sciences in Jilin province from 2011 to 2015, Jilin experiment station of Jilin City Academy of Agricultural Sciences in Jilin province from 2011 to 2015, Tonghua experiment station of Tonghua Academy of Agricultural Sciences in Jilin province from 2012 to 2015, and Jiamusi experiment station of Heilongjiang Academy of Agricultural Sciences in Heilong province in 2015. The three growth period traits, the number of days to flowering (ETF), the number of days from flowering to maturity (FTM), and the number of days to maturity (ETM) were measured by the number of days from stages of plant emergence (VE) to R1, from stages of R1 to R8, and from stages of VE to R8, respectively. The criteria for the growth stages were determined following the method of Fehr and Caviness (1977) [34].

### Phenotypic data analysis

Analysis of variance (ANOVA), implemented by procedure GLM in SAS software (Release 9.1.3; SAS Institute, Cary, NC, USA), was conducted for each of the three traits by combining two planting densities, 5 years, and 4 locations. Best linear unbiased predictors (BLUPs), excluding the variations of years, locations, and experimental error, were used in the following genetic studies. The BLUPs were calculated by procedure MIXED in SAS software (Release 9.1.3; SAS Institute, Cary, NC, USA).

### DNA extraction and SNP genotyping

Genomic DNA sample were extracted from the leaf of soybean seedlings following method described by Kisha et al. (1997) [35]. All the accessions were genotyped via the Illumina SoySNP6k iSelect BeadChip (Illumina, San Diego, Calif. USA), which consisted of 5361 SNPs [36]. The chromosomal distributions, coding, and quality controlling of the SNPs were previously documented in Wen et al. (2014) [29]. By excluding the SNPs with missing rate higher than 0.25, and MAF (minor allele frequency) lower than 0.05, 4032 SNP markers were retained for the following up analysis.

### Population genetics analysis

Genetic diversity characteristics, including MAF, polymorphic information content (PIC), heterozygosity, and gene diversity were evaluated using software Powermarker V3.25 [37]. An admixture model-based clustering method was used to infer population structure and to assign 146 genotypes to subpopulations by 4032 SNP markers using the software STRUCTURE 2.3 [38]. The hypothetical number of subpopulations ( $k$ ) was ranged from 1 to 10, and each  $k$  was run 10 times with a burn-in period of 100,000, and 100,000 Markov Chain Monte Carlo (MCMC) replications. The ad hoc statistics delta  $k$  ( $\Delta k$ ) was used to determine the number of clusters [39].

To investigate the population differentiations, analysis of molecular variance (AMOVA) [40], and  $F$ -statistics ( $F_{ST}$ ) for the inferred subpopulations were performed using Arlequin V3.11 [41]. A neighbor-joining (NJ) phylogenetic tree was constructed based on the Nei's genetic distances matrix [37, 42] using Powermarker V3.25 [37]. Genotypic similarity among 146 accessions was evaluated by Flapjack (downloaded from <https://ics.hutton.ac.uk/flapjack/>), which was previously described by Milne et al. (2010) [43]. The kinship matrix was calculated using TASSEL 4.0 [44] to determine the genetic relatedness among individuals based on the sets of SNPs. The linkage disequilibrium parameter  $r^2$  to estimate the degree of LD was

**Table 1. Descriptive statistics of best linear unbiased predictors (BLUPs) for three traits in two planting densities.**

Traits <sup>a</sup>	Low-density					High-density				
	Min	Max	Mean	SD <sup>b</sup>	CV <sup>c</sup>	Min	Max	Mean	SD <sup>b</sup>	CV <sup>c</sup>
ETF	31.0	55.8	37.4	4.8	13.0	31.1	56.7	37.5	4.8	12.9
FTM	67.9	95.0	83.7	6.6	7.9	67.2	95.2	83.7	6.8	8.2
ETM	97.6	135.1	119.9	9.6	8.0	97.0	136.5	120.1	9.8	8.2

<sup>a</sup> ETF, number of days to flowering; FTM, number of days from flowering to maturity; ETM, number of days to maturity

<sup>b</sup> SD, standard deviation

<sup>c</sup> CV, coefficient of variation.

doi:10.1371/journal.pone.0158602.t001

calculated using TASSEL 4.0 [44]. The decay distance of LD at  $r^2 = 0.1$  was assigned as the length of LD block.

### The whole—genome association analysis

GWAS was conducted for three traits related to soybean growth period by unified mixed linear model (MLM) using 4032 SNPs. Variations of population structure and kinship between accessions were both fitted into the MLM to account for multiple levels of relatedness [45]. Significant markers associated with traits were declared by Bonferroni *P*-value cut-off 0.001. Significant markers in the same LD block were viewed as one QTL region. The most significant marker in one QTL region was reported.

## Results

### Phenotypic variations, correlations, and ANOVA for growth period traits

Large phenotypic variation was observed for all the three traits of 146 accessions (Table 1). As expected, FTM was much longer than ETF, and ETM was longer than both of ETF and FTM. Thus the standard deviation (SD) of ETM was the largest, followed by FTM and ETF, while the coefficient of variation (CV) of ETF was the highest, and those of FTM and ETM were similar. Comparing the mean values of each trait under two densities, the basic statistics of each trait under two densities were fairly the same (Table 1). ETF was ranged from 31.0 d to 55.8 d with an average of 37.4 d in low planting density, and from 31.1 d to 56.7 d with an average of 37.5 d in high planting density.

There were significant positive correlations among three growth period traits (Table 2). In low planting density, the correlation coefficients were 0.7527 for ETF and ETM, 0.8876 for FTM and ETM, and 0.3665 for ETF and FTM. In the high planting density, the correlation coefficients were 0.7645 for ETF and ETM, 0.8880 for FTM and ETM, and 0.3834 for ETF and

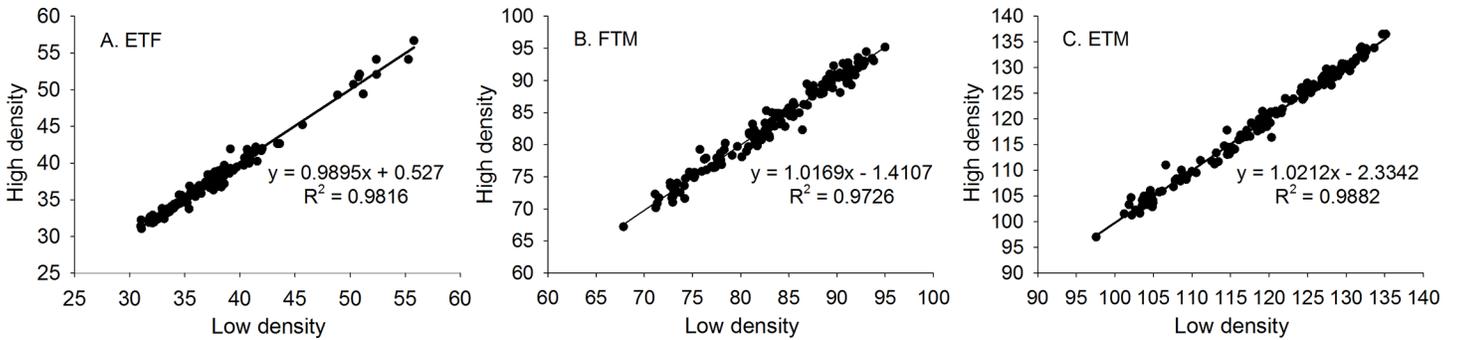
**Table 2. Correlation coefficients between three growth period traits across the four locations.**

	ETF	FTM	ETM
ETF	1	0.3665**	0.7527**
FTM	0.3834**	1	0.8876**
ETM	0.7645**	0.8880**	1

Correlation coefficients in the upper triangular are for trait performance under low density, and in the lower triangular are for trait performance under high density

\*\* denotes the significant level under 0.01 for Pearson correlation test; ETF, number of days to flowering; FTM, number of days from flowering to maturity; and ETM, number of days to maturity.

doi:10.1371/journal.pone.0158602.t002



**Fig 1. The scatter plot of the phenotype of three traits under low and high planting densities.** A is for ETF, number of days to flowering; B is for FTM, number of days from flowering to maturity; and C is for ETM, number of days to maturity.

doi:10.1371/journal.pone.0158602.g001

FTM. Since ETM was derived from ETF and FTM, the correlations between ETM and ETF, and between ETM and FTM were significantly high. The low correlation between ETF and FTM indicated that ETF and FTM were independent to some extent. It was observed that the phenotypic measurements for each of the three traits under two planting densities were highly correlated, with  $R^2$  ranging from 0.97 to 0.99 (Fig 1), which explained why the values in upper triangle and low triangle of Table 2 were almost the same.

ANOVA was conducted for each trait by combining the phenotypic data of four locations during five years for both two densities (Table 3) and for each of the density (S1 and S2 Tables). Results showed that the variation of density was significant for ETF and ETM, but not for FTM. For all the three traits, significant variations were observed among environments, replications (or blocks) within environments, genotypes, and genotype by environment (GE)

**Table 3. Analysis of variance (ANOVA) of three traits across 15 environments and two planting densities.**

Trait <sup>a</sup>	Source <sup>b</sup>	DF <sup>c</sup>	Sum of Square	Mean Square	F Value	Pr > F	H <sup>2</sup> <sup>d</sup>
ETF	Density	1	67.62	67.62	14.28	0.0002	67.72
	Env	14	106893.75	7635.27	1612.07	< .0001	
	Block(Env)	15	312.88	20.86	4.40	< .0001	
	Geno	145	207550.35	1431.38	302.21	< .0001	
	Geno*Env	2030	62063.50	30.57	6.46	< .0001	
FTM	Density	1	0.93	0.93	0.08	0.7752	64.18
	Env	14	192354.13	13739.58	1205.09	< .0001	
	Block(Env)	15	1131.55	75.44	6.62	< .0001	
	Geno	145	398352.03	2747.26	240.96	< .0001	
	Geno*Env	2025	135822.74	67.07	5.88	< .0001	
ETM	Density	1	58.71	58.71	6.75	0.0094	82.37
	Env	14	111548.24	7967.73	915.99	< .0001	
	Block(Env)	15	407.50	27.17	3.12	< .0001	
	Geno	145	810759.92	5591.45	642.81	< .0001	
	Geno*Env	2025	110074.98	54.36	6.25	< .0001	

<sup>a</sup> ETF, number of days to flowering; FTM, number of days from flowering to maturity; ETM, number of days to maturity

<sup>b</sup> Env means environment; Block(Env) means the block nested within environments; Geno means genotype; and Geno\*Env means genotype by environment interaction

<sup>c</sup> degree of freedom; and

<sup>d</sup> broad-sense heritability.

doi:10.1371/journal.pone.0158602.t003

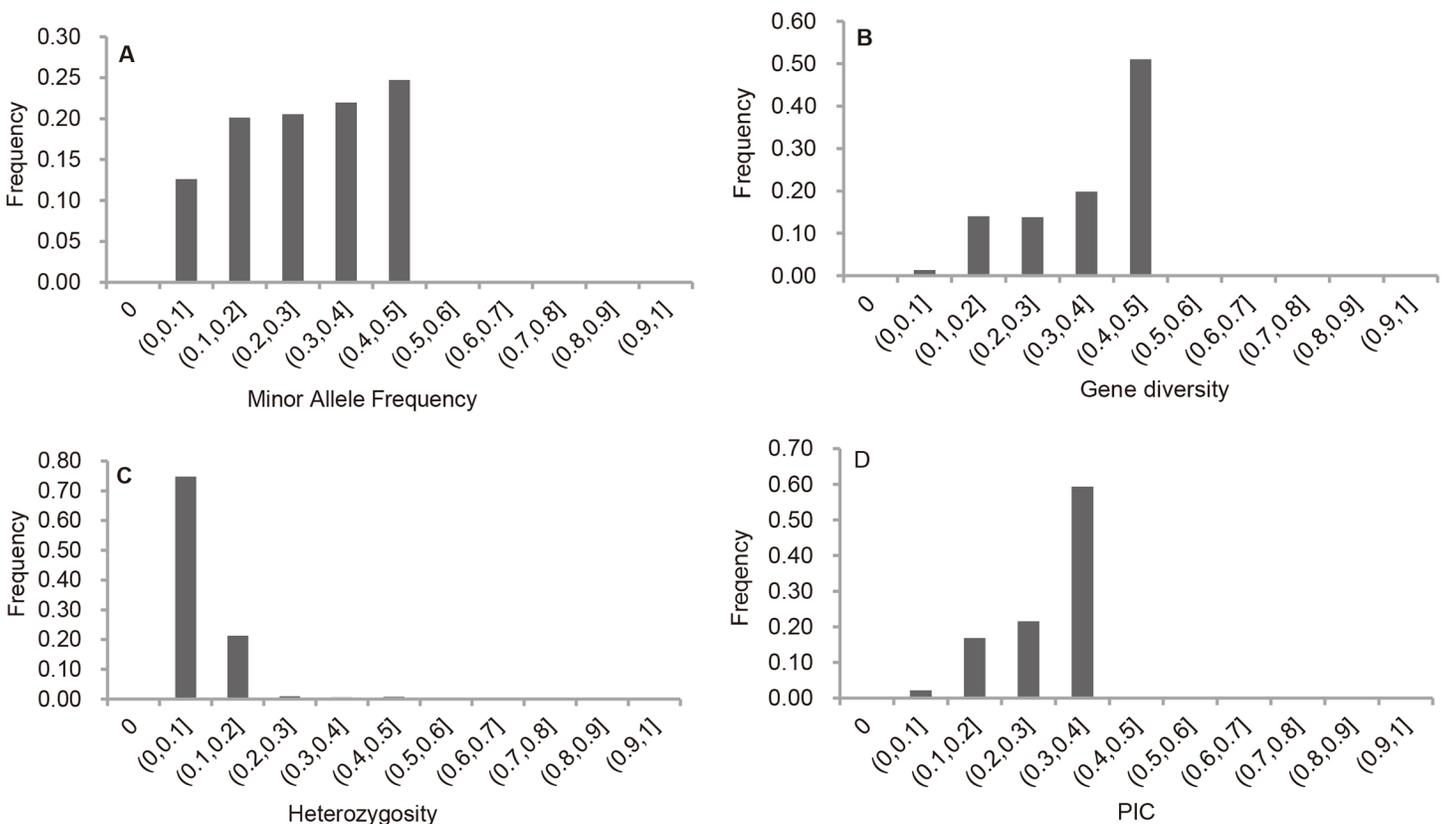
interactions (Table 3). For each trait in interest, the proportions of genotypic variation were much higher than the variation of GE interaction, indicating that the genotypic variation were the major part of the phenotypic variation, which resulted in the relatively high broad sense heritabilities for three traits, i.e., 67.72% for ETF, 64.18% for FTM, and 82.37% for ETM (Table 3).

### Genetic diversity

Among the 5361 SNPs, 4032 SNPs with MAF greater than 5% were selected to estimate the genetic diversity. All the SNPs were mapped in silico and/or genetically in soybean chromosomes, and were well distributed on the 20 chromosomes [29]. The average MAF value of 4032 SNPs was 0.28, ranging from 0.05 to 0.50. About 46.7% of the SNPs had MAF greater than 0.30. The gene diversity, heterozygosity, and PIC for 4032 SNPs were 0.37, 0.09, and 0.29 on average, with ranges of 0.10–0.50, 0.00–1.00, and 0.09–0.38, respectively (Fig 2).

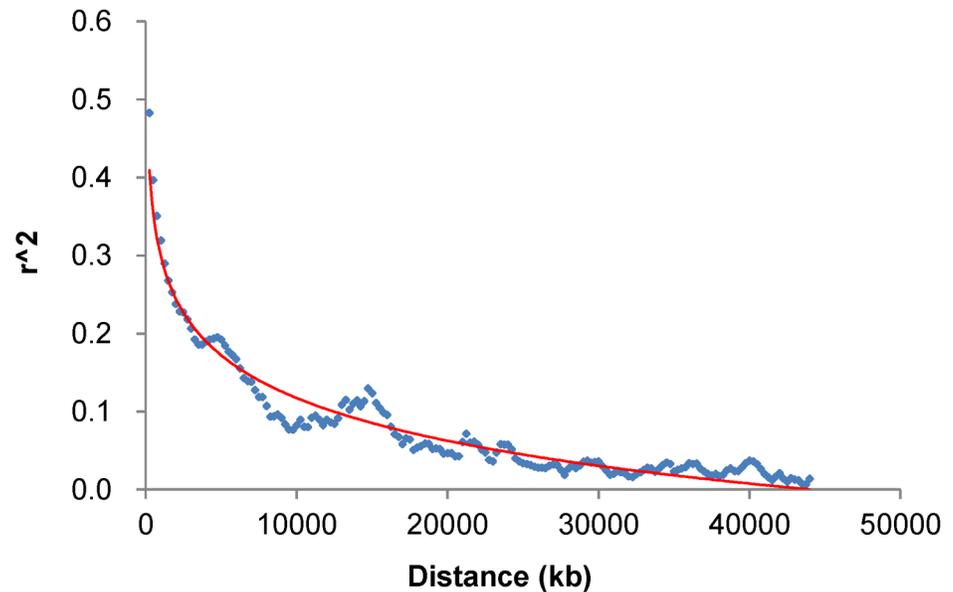
### Linkage disequilibrium (LD) pattern

To characterize the mapping resolution of GWAS in this study, the average extent of genome-wide LD decay distance in 146 accessions with 4032 SNPs was estimated and shown in Fig 3. On average, the  $r^2$  of the whole-genome for 146 accessions was 0.23. When the physical distance was around 1,800 kb, the  $r^2$  reached to half of its maximum value 0.48. When the  $r^2$  was at 0.1, the decay distance was approximately 8,000 kb, indicating a strong LD existed in the



**Fig 2. Distribution of the genetic diversity of 4,032 SNPs across 146 accessions.** A is for minor allele frequency; B is for gene diversity; C is for heterozygosity; and D is for polymorphic information content (PIC).

doi:10.1371/journal.pone.0158602.g002



**Fig 3. Linkage disequilibrium (LD) decay across soybean genome.**

doi:10.1371/journal.pone.0158602.g003

population. Given that the average marker density was 233.53 kb in the population, the SNP markers used in this study were expected to have reasonable power to identify major QTL with large effect for three traits in 146 soybean accessions.

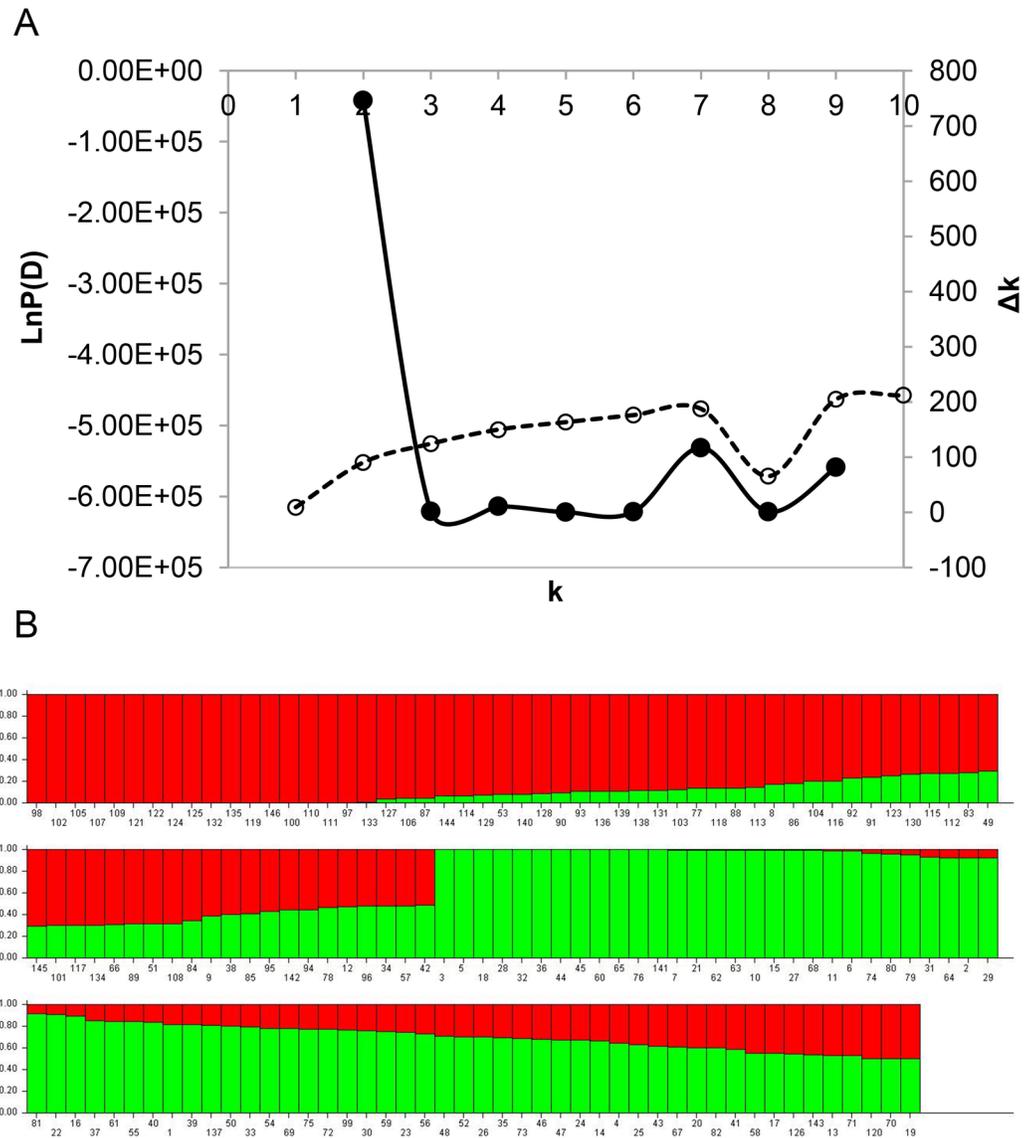
### Population structure

To avoid false-positive associations due to population stratification, three statistical methods, including STRUCTURE, NJ tree-based, and PCA (Figs 4 and 5) were exploited to estimate the relatedness among 146 accessions using 4032 SNPs. The distribution of  $\text{LnP}(D)$  value for each given  $k$  did not show a clear trend (Fig 4A). The  $\Delta k$  reached to the highest value when  $k$  was at 2 (Fig 4A), which indicated that the 146 accessions could be divided into two subpopulations (Fig 4B). The measurement of population differentiation,  $F_{ST}$ , was estimated at 0.18 ( $P < 0.001$ ) between the two subpopulations, suggesting high level of genetic difference (S3 Table). The result of AMOVA showed that 17.64% of the total genetic variation was among subpopulations, whereas 82.36% was within subpopulations (S3 Table).

The results of NJ phylogenetic tree and PCA (Fig 5) were consistent with the results from STRUCTURE software (Fig 4). For subpopulation 1 (71 accessions), 53 of which originated from Heilongjiang province with relatively early maturity. The accessions in subpopulation 2 (75 accessions) were mainly from Jilin province with relatively late maturity. These results suggested that accessions from the same subgroup were closely related and generally from the same province. Such co-existence or overlapping of genetic and special differentiation is much likely due to the selection of fitness by nature and breeders. It has previously been suggested that the photoperiod response between different maturity groups may be the primary factor driving differentiation of cultivated soybean [46]. The Q matrix outputted from STRUCTURE software for two sub-populations were used to control the variation of population structure in the subsequent genetic analysis.

### QTL identified for the three traits

From GWAS analysis based on MLM model involved with Q matrix and Kinship matrix (S1 Fig), 19 QTL regions can be clearly identified from the Manhattan plots (Figs 6–8; Table 4).

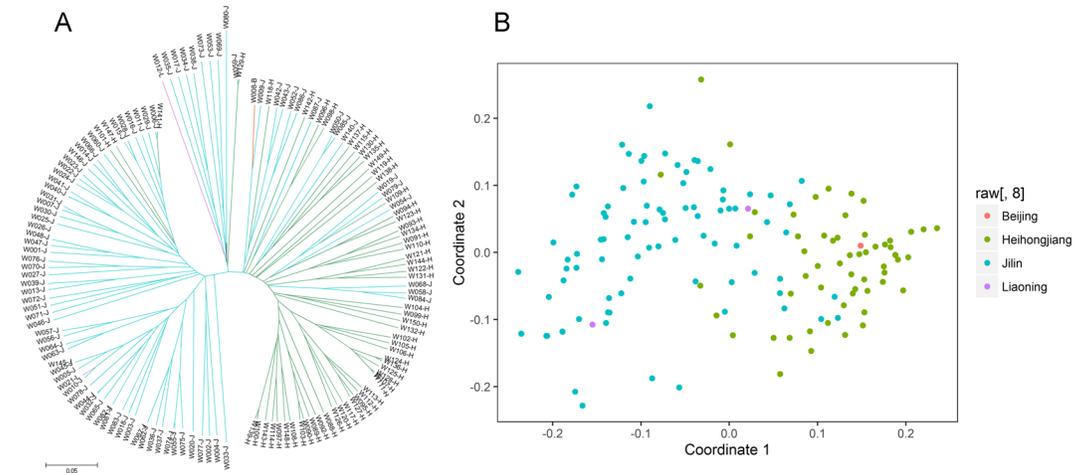


**Fig 4. Analysis of the population structure of 146 soybean accessions.** A is for the estimated  $\Delta k$  over 10 repeats of STRUCTURE analysis; and B is for the population structure estimated by STRUCTURE. Each individual is represented by a vertical bar, partitioned into colored segments with the length of each segment representing the proportion of the individual's genome when  $k = 2$ .

doi:10.1371/journal.pone.0158602.g004

They were distributed on 11 chromosomes of soybean genome, three of each on chromosomes 3, 4, and 13, 2 on each of chromosomes 11 and 15, and one on each of chromosomes 3, 5, 9, 16, 18, and 19. The largest QTL, explaining 14.62% of the phenotypic variance, was ss245775380 on chromosome 5 associated with ETF (Table 4). Of 19 QTLs, 5 QTLs could explain more than 10% of the phenotypic variance, and the other 14 QTL could explain more than 5% of the phenotypic variance. On average, 9.53% of the phenotypic variance could be explained by each QTL.

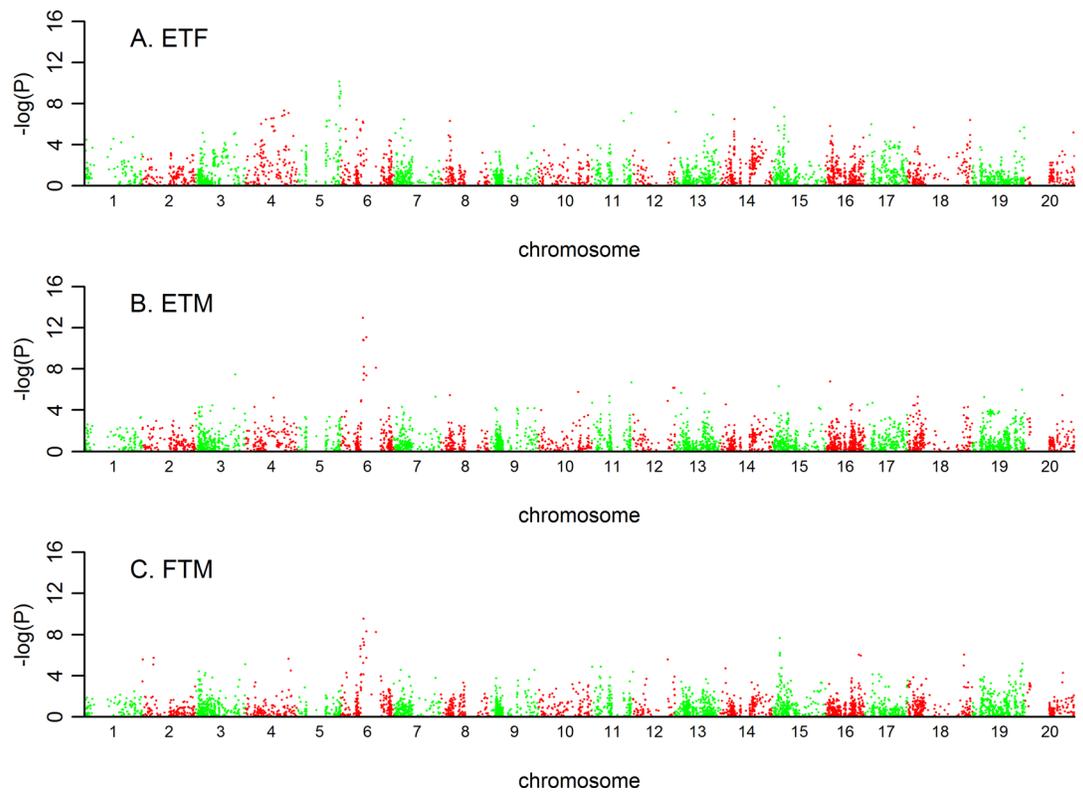
Regarding the two planting conditions, 13 QTL were identified under low density, and 23 QTL were identified under high density. Twelve of them could be consistently identified under both densities, 7 for ETF, 2 for FTM, and 3 for ETM (Table 4; Fig 8). These results were consistent with the high correlations of phenotypic values under two densities (Fig 1).



**Fig 5. Genetic relatedness based on 4032 SNPs.** A is for neighbor-joining tree; and B is for principle coordinate analysis (PCA). "W+Number" is the accession number.

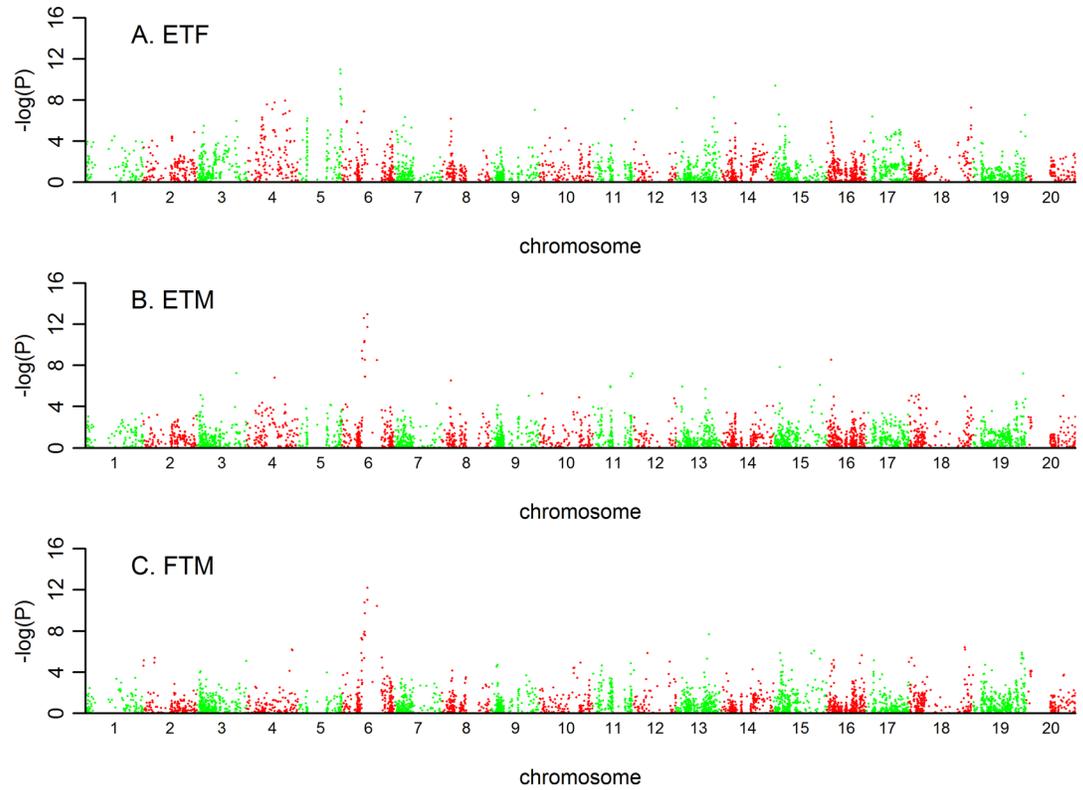
doi:10.1371/journal.pone.0158602.g005

Five QTL, three on chromosome 6 and one on each of chromosomes 11 and 15, had pleiotropic effects on more than one trait. Four of them controlled ETM and FTM, and the other one controlled ETM and ETF. None of the QTL had pleiotropic effects on ETF and FTM. These results were concordant with the high phenotypic correlations between ETM and FTM, and



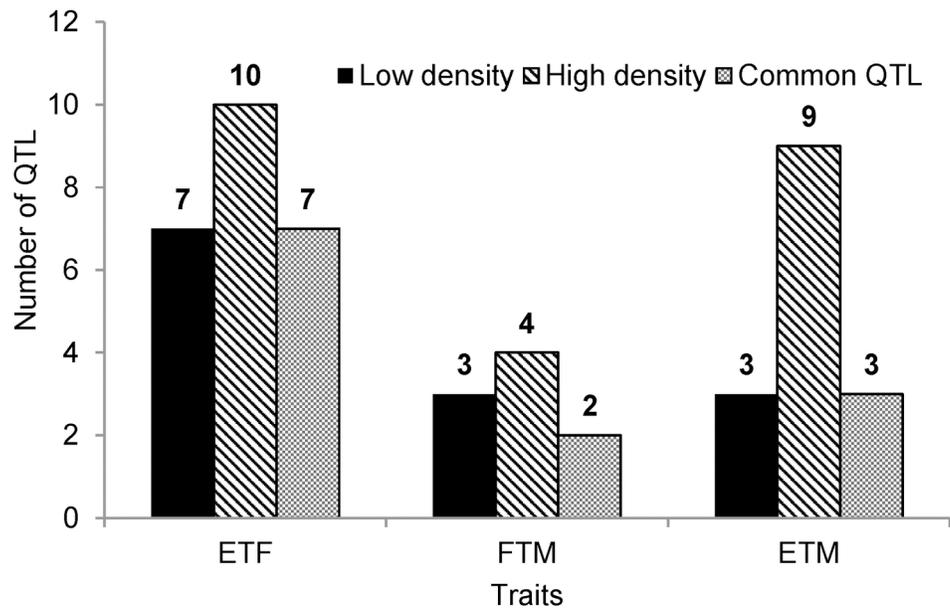
**Fig 6. Manhattan plot from Q+K model across three traits in low density.** ETF is for trait of number of days to flowering; FTM is for trait of number of days from flowering to maturity; and ETM is for trait of number of days to maturity.

doi:10.1371/journal.pone.0158602.g006



**Fig 7. Manhattan plot from Q+K model across three traits in high density.** ETF is for trait of number of days to flowering; FTM is for trait of number of days from flowering to maturity; and ETM is for trait of number of days to maturity.

doi:10.1371/journal.pone.0158602.g007



**Fig 8. Number of QTL detected under each of low and high densities and both two densities.** ETF is for trait of number of days to flowering; FTM is for trait of number of days from flowering to maturity; and ETM is for trait of number of days to maturity.

doi:10.1371/journal.pone.0158602.g008

**Table 4. SNPs significantly associated with three traits.**

Marker	Chr	Position	Associated traits ( $R^2$ ) <sup>a</sup>	Reported QTLs/genes <sup>b</sup>	Gene annotation <sup>c</sup>
ss245178465	3	37637700	ETM (1.37), <u>ETM (6.93)</u>		C2H2-type zinc finger family protein
ss245412963	4	24492629	ETF (7.82)	<i>E8</i> gene [47]	PIF1-like helicase
ss245456710	4	36658492	ETF (8.77), <u>ETF (9.56)</u>		Unknown
ss245494974	4	40933102	ETF (7.79), <u>ETF (7.60)</u>	Seed weight [48]; Seed yield [49]; Pod number [50]; Seed weight per plant [51]	Unknown
ss245775380	5	40409971	ETF (12.99), <u>ETF (14.62)</u>	Seed weight [48, 52–53]; Seed yield [54–55]; Plant height [56]	NAD (P)-binding Rossmann-fold superfamily protein
ss245937498	6	19656740	ETM (8.69), <u>FTM (7.66)</u>	Pod maturity [52, 57]; Photoperiod insensitivity [58]; First flower [59]; Seed weight [48]; Seed yield [57, 60]; Plant height [57]	Unknown
ss245950346	6	21009630	ETM (14.27), FTM (11.43), <u>ETM (14.12)</u> , <u>FTM (12.83)</u>	<i>E1</i> gene [17]; Seed yield [61]	A protein contains a putative bipartite nuclear localization signal and a region distantly related to B3 domain
ss245977002	6	33563355	ETM (8.67), FTM (9.00), <u>ETM (9.27)</u> , <u>FTM (12.04)</u>		Plant protein of unknown function
ss247025033	9	42166200	<u>ETF (8.24)</u>		Prefoldin subunit; RNI-like superfamily protein; P-loop containing nucleoside triphosphate hydrolases superfamily protein
ss247557298	11	36213236	<u>ETM (6.70)</u>	Pod maturity [62]; Seed weight [48, 63]	Tetratricopeptide repeat (TPR)-like superfamily protein; NAD (P)-binding Rossmann-fold superfamily protein
ss247571761	11	37497232	ETF (7.89), <u>ETF (7.95)</u> , <u>ETM (6.89)</u>		Uroporphyrinogen decarboxylase; Haloacid dehalogenase-like hydrolase (HAD) superfamily protein; Weak chloroplast movement under blue light
ss247852752	13	768019	ETF (7.90), <u>ETF (7.91)</u>		Yip1 domain; Voltage dependent anion channel; Arogenate dehydrogenase
ss248095215	13	31305831	<u>FTM (8.25)</u>	Pod maturity [54]; Seed fill [54]; Days to flower [64]; Days to maturity [64]; Seed weight [65–67]; Seed yield [67]; Plant height [68]; Node number [68]	Mo25 family protein; F-box/RNI-like superfamily protein
ss248139411	13	36268141	ETF (7.53), <u>ETF (9.40)</u>	Seed yield [69–70]	Sodium/Calcium exchanger family protein/Calcium-binding EF hand family protein; PHD finger family protein/Bromo-adjacent homology (BAH) domain-containing protein
ss248535946	15	1798512	ETF (8.75), <u>ETF (11.43)</u>		Protein phosphatase 2A regulatory subunit PR55; Pyridoxal phosphate phosphatase-related protein
ss248571908	15	6771127	FTM (10.14), <u>ETM (7.73)</u>		Disease resistance family protein/LRR family protein
ss248968508	16	4339640	<u>ETM (8.68)</u>	Gene <i>GmFT5a</i> [23, 71]; Reproductive stage length [72]	K+ potassium transporter
ss249946467	18	59403616	<u>ETF (8.69)</u>		Glycosyl hydrolase family 81 protein; Receptor-like protein kinase
ss250268653	19	47089771	<u>ETM (7.26)</u>	<i>E3</i> gene [18, 73]; First flower [74–76]; Flower from [75]; Seed weight [74, 77]; Seed number [74]	Pectin lyase-like superfamily protein; NAD(P)-binding Rossmann-fold superfamily protein

<sup>a</sup> Traits underlined indicated that the marker-trait associations were detected under high density, otherwise under low density; ETF, number of days to flowering; FTM, number of days from flowering to maturity; and ETM, number of days to maturity

<sup>b</sup> Reported QTLs/genes with italics were associated with growth period traits; and

<sup>c</sup> the putative biological candidate gene in the locus or the nearest annotated gene (*Glycine max* Wm82.a1.v1) to the significant SNP.

doi:10.1371/journal.pone.0158602.t004

between ETM and ETF, but low correlations between ETF and FTM (Table 2). ss245950346 on chromosome 6 was consistently associated with ETM and FTM under two planting densities, explaining an average of 13.16% of the phenotypic variance. It was worth noting that ss245950346 was also in the same region with well characterized gene *E1* (Table 4). ss245937498 on chromosome 6 was identified for traits of ETM and FTM, and also reported by other literatures controlling series of traits in interest. Three pleiotropic QTL, ss245977002 on chromosome 6, ss247571761 on chromosome 11, and ss248571908 on chromosome 15, had not been reported yet, but were adjacent to candidate genes with functions such as haloacid dehalogenase-like hydrolase superfamily protein, disease resistance family protein, etc.

We compared the positions of the significant SNPs identified in this study with the positions of the QTL previously reported in bi-parental mapping studies and found considerable overlap between these SNPs and the reported genes or QTL for growth period traits. Of the 19 loci, 7 were overlapped with previously reported QTL or genes related to growth period traits, and 8 loci were located in or close to the regions where QTL controlling yield or yield-related traits were reported. Sixteen QTL were closed to the candidate gene regions, and found that various types of genes were probably involving in natural variation for three growth period traits in soybean (Table 4).

## Discussions

### Large phenotypic variations and BLUP estimation

Tremendous phenotypic variations for three growth period traits can be observed in 146 soybean accessions in this study (Table 1 and S2–S4 Figs). For both densities, the deviations from the minimum values to the maximum values of ETF, FTM, and ETM were 25 days, 27 days, and 37 days, respectively (Table 1). When we dissected the phenotypic variations by ANOVA, the genotypic by environment variations were significant for all the three traits, but the genotypic variations were still the major and significant source (Table 3; S1 and S2 Tables). The broad sense heritability of ETM was the highest (82.37%), indicating that ETM performed relatively stable compared with the other two traits. To exclude the environmental variation, BLUPs per accession across environments were estimated for following up GWAS.

It is generally agreed that soybean growth period is positively correlated with yield in normal maturity conditions, and yield is significantly and positively correlated with the ratio of the reproductive period to growth period, but not with the ratio of the vegetative period to growth period. The appropriate ratio of the reproductive period to the vegetative period was supposed to be a secondary trait for indirect selection of yield and to facilitate the increasing of seed number and yield in soybean breeding [78]. In this study, ETF and ETM, FTM and ETM were significantly and positively correlated. Therefore, the affection of the ratio of the reproductive period to the vegetative period need be further investigated. The three growth period traits in this study were measured in the cultivation situation similar to that in the field, so the QTL detected in this study could be directly used in the real breeding scheme.

### SNP markers repeatedly detected or located in or near previously reported QTL

Quantitative traits were largely affected by environments. Mapping accurate and stable QTL across multiple environments was critical for molecular marker-assisted selection breeding and QTL cloning. In this study, 12 QTL were consistently identified across two planting densities (Table 4, Fig 8), which agreed with the strong correlation of trait measurements under two densities (Fig 1). In addition, 7 QTLs were located in or close to previously reported QTL or genes related to growth period. SNP ss245937498 at 19656740 bp on chromosome 6, was detected to

be significantly associated with ETM and FTM, and also associated with pod maturity [52, 57], photoperiod insensitivity [58], and first flower [59]. *E1* gene was cloned by Xia et al. (2012) [17] at 20,207,279–20,207,803 bps, which is 1,827 bp away from the significant marker ss245950346 associated with ETM and FTM under two densities in this study. Significant marker ss245977002 on chromosome 6, associated with ETM and FTM, were in the downstream of *E1*. While *E7* was supposed to be either in the downstream of *E1*, or a different allele of *E1* (*Private communication*). The relationship between marker ss245977002 and *E7* need to be further investigated. Marker ss250268653 at 47,089,771 bp on chromosome 19 was associated with ETM, and was 77,821 bp away from marker satt229, which was closely linked with *E3* gene [18, 73] and controlled first flower [74–76]. On chromosome 16, ss248968508 at 4,339,640 bp was associated with ETM, and 202,142 bp far away from gene *GmFT5a* [23, 71]. In the QTL region of ss248968508, there was a QTL controlling reproductive stage length [72]. Cober et al. (2010) [47] reported that the locus *E8* was in the interval flanked by Sat\_404 and Satt136 on chromosome 4. This interval is approximately 1 cM in length. Marker Satt361 located downstream of Satt136, and was approximately 0.5 cM from Satt136. SNP marker ss245412963 at 24,492,629 bp on chromosome 4 was detected to be associated with ETF in high planting density in this study, and was 378,973 bp away from Satt361. Pod maturity was similar to ETM. SNP marker ss247557298 on chromosome 11 controlling ETM was near to the marker Sat\_123, which is associated with Pod maturity [62]. Marker Satt355 on chromosome 13 was reported to be associated with Pod maturity [54, 64], seed fill [54], and days to flower [64]. Marker ss248095215 used in this study controlling FTM was physically near Satt355. Therefore, some reported QTL could be validated in different environment, population, and statistical method, indicating the results in present study is rather reliable.

### SNP markers with pleiotropic effect

Of the 19 QTL detected, 5 QTL were co-association with at least two growth period traits, which coincided with significant phenotypic correlations among the traits in interest [31, 33]. In addition, it is reported that some growth period loci have pleiotropic effects on other important agronomic traits [13], such as branching [79], yield related traits [21, 80], and cleistogamy [75]. Similarly, QTL associated with yield component traits, such as seed weight, seed yield, seed weight per plant, seed number, pod number, plant height, and node number, were located in the regions where 8 QTL were identified in our study. In QTL region of ss245937498 related to ETM and FTM on chromosome 6, there were QTL not only controlling growth period traits, but also controlling seed weight [48], seed yield [57, 60], and plant height [57]. In QTL region of ss250268653 related to ETM on chromosome 19, there were QTL controlling seed weight [74, 77], seed number [74], and cleistogamous [75]. The genomic regions where multi-traits were co-associated indicated pleiotropy of single causal gene or tight linkage of multiple causal genes. In soybean molecular breeding schemes, MAS of a co-associated genetic locus could simultaneously improve multi-associated target traits, including yield and yield related traits.

### Novel SNP markers to be verified further

For the 19 QTL reported in this study, 9 of them were not reported before. One reason could be because the accessions used in this study were nearly from the maturity groups MG 000, MG 00, MG 0, and MG I, and QTL identified in accessions from other maturity groups could not present in the accessions used in this study. Soybean cultivars adapted to high latitudes have weak or no photoperiod sensitivity. Zhai et al. (2014) [81] found that e1-nf genetic groups are approximately corresponding to cultivars of MG 000, MG 00, and MG 0 groups. The second reason may be that since most of loci reported before were detected by QTL mapping

based on bi-parental populations, where only alleles polymorphic between two parents were considered, thus the number of detected QTL were restricted by the allele distribution of two parents. The third possibility was that many growth period QTL have been identified in a number of different populations, however, if a QTL is controlled by a rare allele present only in a specific accession used in creating a QTL mapping population, it could not be detected in a GWAS such as reported here. The inability of GWAS to detect rare alleles occurring in one or a few members of a population under study is well documented [82–83].

Of 9 novel loci, we found that genomic regions around eight of them harbor candidate genes (Table 4). Therefore, further studies would be conducted using a large population size with more diverse genetic background, and a large number of SNPs to verify the associated markers identified in this study.

## Supporting Information

**S1 Fig. Kinship value between accessions among population.** Each pixel in the square indicates the kinship value of corresponding individual pair as shown on the above bar.  
(PDF)

**S2 Fig. Phenotypic distribution of ETF.** Fan is for Fanjiatun experiment station; Ji is for Jilin experiment station; Tong is for Tonghua experiment station; and Jia is for Jiamushi experiment station.  
(PDF)

**S3 Fig. Phenotypic distribution of FTM.** Fan is for Fanjiatun experiment station; Ji is for Jilin experiment station; Tong is for Tonghua experiment station; and Jia is for Jiamushi experiment station.  
(PDF)

**S4 Fig. Phenotypic distribution of ETM.** Fan is for Fanjiatun experiment station; Ji is for Jilin experiment station; Tong is for Tonghua experiment station; and Jia is for Jiamushi experiment station.  
(PDF)

**S1 Table. Analysis of variance (ANOVA) of three traits across 15 low-dense plant environments.**  
(PDF)

**S2 Table. Analysis of variance (ANOVA) of three traits across 15 high-dense plant environments.**  
(PDF)

**S3 Table. Analysis of molecular variance (AMOVA) and  $F_{ST}$  for two subpopulations of soybean accessions inferred from STRUCTURE.**  
(PDF)

## Author Contributions

Conceived and designed the experiments: LQ SW DW. Performed the experiments: ZL XF WH JY ZW CL. Analyzed the data: ZL HL YL RG YG RC. Wrote the paper: ZL HL.

## References

1. Garner WW, Allard HA. Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *J Agr Res.* 1920; 18: 553–606.

2. Cober ER, Tanner JW, Voldeng HD. Genetic control of photoperiod response in early-maturing, near isogenic soybean lines. *Crop Sci.* 1996; 36: 601–605.
3. Chen XZ, Xie H, Li X. Studies on correlationship of development stages and agronomic traits of summer sowing soybean. *Mol Plant Breed.* 2004; 2: 247–252 (in Chinese with English abstract).
4. McBlain B, Bernard R. A new gene affecting the time of flowering and maturity in soybeans. *J Hered.* 1987; 78: 160–162.
5. Hartwig EE. Varietal development. In: Caldwell BE, Editor. *Soybeans: Improvement, Production and Uses.* Madison: ASA, CSSA, and SSSA; 1973. pp. 187–210.
6. Norman AG. *Soybean physiology, agronomy and utilization.* NY: Academic Press; 1978.
7. Woodworth CM. Inheritance of growth habit, pod color, and flowering color in soybean. *J Am Soc Agron.* 1923; 15: 485–495.
8. Bernard R. Two major genes for time of flowering and maturity in soybeans. *Crop Sci.* 1971; 11: 242–244.
9. Buzzell R. Inheritance of a soybean flowering response to fluorescent-day length conditions. *Can J Genet Cytol.* 1971; 13: 703–707.
10. Buzzell R, Voldeng H. Inheritance of insensitivity to long daylength. *Soybean Genet Newsl.* 1980; 7: 26–29.
11. Ray JD, Hinson K, Mankono J, Malo MF. Genetic control of a long-juvenile trait in soybean. *Crop Sci.* 1995; 35: 1001–1006.
12. Bonato ER, Vello NA. *E6*, a dominant gene conditioning early flowering and maturity in soybeans. *Genet Mol Biol.* 1999; 22: 229–232.
13. Cober ER, Morrison MJ. Regulation of seed yield and agronomic characters by photoperiod sensitivity and growth habit genes in soybean. *Theor Appl Genet.* 2010; 120: 1005–1012. doi: [10.1007/s00122-009-1228-6](https://doi.org/10.1007/s00122-009-1228-6) PMID: [20012856](https://pubmed.ncbi.nlm.nih.gov/20012856/)
14. Cober ER, Voldeng HD. A new soybean maturity and photoperiod-sensitivity locus linked to *E1* and *T*. *Crop Sci.* 2001; 41: 698–701.
15. Kong FJ, Nan HY, Cao D, Li Y, Wu FF, Wang JL, et al. A new dominant gene *E9* conditions early flowering and maturity in soybean. *Crop Sci.* 2014; 54: 1–7.
16. Liu B, Kanazawa A, Matsumura H, Takahashi R, Harada K, Abe J, et al. Genetic redundancy in soybean photoreponses associated with duplication of the phytochrome A gene. *Genetics.* 2008; 180: 995–1007. doi: [10.1534/genetics.108.092742](https://doi.org/10.1534/genetics.108.092742) PMID: [18780733](https://pubmed.ncbi.nlm.nih.gov/18780733/)
17. Xia ZJ, Watanabe S, Yamada T, Tsubokura S, Nakashima H, Zhai H, et al. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1*, which regulates photoperiodic flowering. *Proc Natl Acad Sci USA.* 2012; 109: E2155–2164. doi: [10.1073/pnas.1117982109](https://doi.org/10.1073/pnas.1117982109) PMID: [22619331](https://pubmed.ncbi.nlm.nih.gov/22619331/)
18. Watanabe S, Hideshima R, Xia Z, Tsubokura Y, Sato S, Nakamoto Y, et al. Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics.* 2009; 182: 1251–1262. doi: [10.1534/genetics.108.098772](https://doi.org/10.1534/genetics.108.098772) PMID: [19474204](https://pubmed.ncbi.nlm.nih.gov/19474204/)
19. Watanabe S, Xia Z, Hideshima R, Tsubokura Y, Sato S, Yamanaka N, 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: 395–407. doi: [10.1534/genetics.110.125062](https://doi.org/10.1534/genetics.110.125062) PMID: [21406680](https://pubmed.ncbi.nlm.nih.gov/21406680/)
20. McBlain BA, Hesketh JD, Bernard RL. Genetic effects on reproductive phenology in soybean isolines differing in maturity genes. *Can J Plant Sci.* 1987; 67: 105–116.
21. Saindon G, Beversdorf WD, Voldeng HD. Adjustment of the soybean phenology using the *E4* locus. *Crop Sci.* 1989; 29: 1361–1365.
22. Zhang LX, Kyei-Boahen S, Zhang J, Zhang MH, Freeland TB, Watson CE, et al. Modifications of optimum adaptation zones for soybean maturity groups in the USA. *Crop Management.* 2007; 6: 1. doi: [10.1094/CM-2007-0927-01-RS](https://doi.org/10.1094/CM-2007-0927-01-RS)
23. Kong FJ, Liu BH, Xia ZJ, Sato S, Kim BM, Watanabe S, et al. Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. *Plant Physiol.* 2010; 154: 1220–1231. doi: [10.1104/pp.110.160796](https://doi.org/10.1104/pp.110.160796) PMID: [20864544](https://pubmed.ncbi.nlm.nih.gov/20864544/)
24. Zhao C, Takeshima R, Zhu J, Xu M, Sato M, Watanabe S, et al. A recessive allele for delayed flowering at the soybean maturity locus *E9* is a leaky allele of *FT2a*, a *FLOWERING LOCUS T* ortholog. *BMC Plant Biol.* 2016; 16: 20. doi: [10.1186/s12870-016-0704-9](https://doi.org/10.1186/s12870-016-0704-9) PMID: [26786479](https://pubmed.ncbi.nlm.nih.gov/26786479/)
25. Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, et al. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature.* 2010; 465: 627–631. doi: [10.1038/nature08800](https://doi.org/10.1038/nature08800) PMID: [20336072](https://pubmed.ncbi.nlm.nih.gov/20336072/)

26. Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet.* 2012; 44: 32–39. doi: [10.1038/ng.1018](https://doi.org/10.1038/ng.1018)
27. Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, et al. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nature Genet.* 2013; 45: 43–50. doi: [10.1038/ng.2484](https://doi.org/10.1038/ng.2484) PMID: [23242369](https://pubmed.ncbi.nlm.nih.gov/23242369/)
28. Raman H, Stodart B, Ryan PR, Delhaize E, Emebiri L, Raman R, et al. Genome-wide association analyses of common wheat (*Triticum aestivum* L.) germplasm identifies multiple loci for aluminium resistance. *Genome.* 2010; 53: 957–966. doi: [10.1139/G10-058](https://doi.org/10.1139/G10-058) PMID: [21076511](https://pubmed.ncbi.nlm.nih.gov/21076511/)
29. Wen Z, Tan R, Yuan J, Bales C, Du W, Zhang S, et al. Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. *BMC Genomics.* 2014; 15: 809. doi: [10.1186/1471-2164-15-809](https://doi.org/10.1186/1471-2164-15-809) PMID: [25249039](https://pubmed.ncbi.nlm.nih.gov/25249039/)
30. Korir PC, Zhang J, Wu KJ, Zhao TJ, Gai JY. Association mapping combined with linkage analysis for aluminum tolerance among soybean cultivars released in Yellow and Changjiang River Valleys in China. *Theor Appl Genet.* 2013; 126: 1659–1675. doi: [10.1007/s00122-013-2082-0](https://doi.org/10.1007/s00122-013-2082-0) PMID: [23515677](https://pubmed.ncbi.nlm.nih.gov/23515677/)
31. Hao DR, Cheng H, Yin ZT, Cui SY, Zhang D, Wang H, et al. Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. *Theor Appl Genet.* 2012; 124: 447–458. doi: [10.1007/s00122-011-1719-0](https://doi.org/10.1007/s00122-011-1719-0) PMID: [21997761](https://pubmed.ncbi.nlm.nih.gov/21997761/)
32. Hwang EY, Song QJ, Jia GF, Specht JE, Hyten DL, Costa J, et al. A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics.* 2014; 15: 1. doi: [10.1186/1471-2164-15-1](https://doi.org/10.1186/1471-2164-15-1) PMID: [24382143](https://pubmed.ncbi.nlm.nih.gov/24382143/)
33. Hao DR, Chao MN, Yin ZT, Yu DY. Genome-wide association analysis detecting significant single nucleotide polymorphisms for chlorophyll and chlorophyll fluorescence parameters in soybean (*Glycine max*) landraces. *Euphytica.* 2012; 186: 919–931.
34. Fehr WR, Caviness CE. Stages of soybean development. Ames: Iowa State University of Science and Technology; 1977.
35. Kisha TJ, Sneller CH, Diers BW. Relationship between genetic distance among parents and genetic variance in populations of soybean. *Crop Sci.* 1997; 37: 1317–1325.
36. Akond M, Liu S, Schoener L, Anderson JA, Kantartzi SK, Meksem K, et al. SNP-based genetic linkage map of soybean using the SoySNP6K Illumina Infinium BeadChip genotyping array. *J Plant Genom Sci.* 2013; 1: 80–89.
37. Liu KJ, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics.* 2005; 21: 2128–2129. PMID: [15705655](https://pubmed.ncbi.nlm.nih.gov/15705655/)
38. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000; 155: 945–959. PMID: [10835412](https://pubmed.ncbi.nlm.nih.gov/10835412/)
39. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* 2005; 14: 2611–2620. PMID: [15969739](https://pubmed.ncbi.nlm.nih.gov/15969739/)
40. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 1992; 131: 479–491. PMID: [1644282](https://pubmed.ncbi.nlm.nih.gov/1644282/)
41. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online.* 2005; 1: 47–50.
42. Nei M, Feldman MW. Identity of genes by descent within and between populations under mutation and migration pressures. *Theor Popul Biol.* 1972; 3: 460–465. PMID: [4667099](https://pubmed.ncbi.nlm.nih.gov/4667099/)
43. Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WTB, et al. Flapjack-graphical genotype visualization. *Bioinformatics.* 2010; 26: 3133–3134. doi: [10.1093/bioinformatics/btq580](https://doi.org/10.1093/bioinformatics/btq580) PMID: [20956241](https://pubmed.ncbi.nlm.nih.gov/20956241/)
44. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* 2007; 23: 2633–2635. PMID: [17586829](https://pubmed.ncbi.nlm.nih.gov/17586829/)
45. Yu J, Holland JB, McMullen MD, Buckler ES. Genetic design and statistical power of nested association mapping in maize. *Genetics.* 2008; 138: 539–551.
46. Roberts EH, Qi A, Ellis RH, Summerfield RJ, Lawn RJ, Shanmugasundaram S. Use of field observations to characterize genotypic flowering responses to photoperiod and temperature: a soybean exemplar. *Theor Appl Genet.* 1996; 93: 519–533. doi: [10.1007/BF00417943](https://doi.org/10.1007/BF00417943) PMID: [24162343](https://pubmed.ncbi.nlm.nih.gov/24162343/)
47. Cober ER, Molnar SJ, Charette M, Voldeng HD. A new locus for early maturity in soybean. *Crop Sci.* 2010; 50: 524–527.

48. Han Y, Li D, Zhu D, Li H, Li X, Teng W, et al. QTL analysis of soybean seed weight across multi-genetic backgrounds and environments. *Theor Appl Genet.* 2012; 125: 671–683. doi: [10.1007/s00122-012-1859-x](https://doi.org/10.1007/s00122-012-1859-x) PMID: [22481120](https://pubmed.ncbi.nlm.nih.gov/22481120/)
49. Palomeque L, Liu LJ, Li WB, Hedges B, Cober ER, Rajcan I. QTL in mega-environments: I. Universal and specific 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: 417–427. doi: [10.1007/s00122-009-1049-7](https://doi.org/10.1007/s00122-009-1049-7) PMID: [19462148](https://pubmed.ncbi.nlm.nih.gov/19462148/)
50. Vieira A, de Oliveira D, Soares T, Schuster I, Piovesan N, Martinez C, et al. Use of the QTL approach to the study of soybean trait relationships in two populations of recombinant inbred lines at the F7 and F8 generations. *Braz J PI Phys.* 2006; 18: 281–290.
51. Chen QS, Zhang ZC, Liu CY, Xin DW, Qiu HM, Shan DP, et al. QTL analysis of major agronomic traits in soybean. *Sci Agric Sin.* 2007; 6: 399–405.
52. Specht JE, Chase K, Macrander M, Graef GL, Chungd J, Markwella JP, et al. Soybean response to water. *Crop Sci.* 2001; 41: 493–509.
53. Sun Y, Pan J, Shi X, Du X, Wu Q, Qi Z, et al. Multi-environment mapping and meta-analysis of 100-seed weight in soybean. *Mol Biol Rep.* 2012; 39: 9435–9443. doi: [10.1007/s11033-012-1808-4](https://doi.org/10.1007/s11033-012-1808-4) PMID: [22740134](https://pubmed.ncbi.nlm.nih.gov/22740134/)
54. Li D, Pfeiffer T, Cornelius P. Soybean QTL for yield and yield components associated with *Glycine soja* alleles. *Crop Sci.* 2008; 48: 571–581.
55. Kabelka E, Diers B, Fehr W, LeRoy A, Baianu I, You T, et al. Putative alleles for increased yield from soybean plant introductions. *Crop Sci.* 2004; 44: 784–791.
56. Sun D, Li W, Zhang Z, Chen Q, Ning H, Qiu L, et al. Quantitative trait loci analysis for the developmental behavior of soybean (*Glycine max* L. Merr.). *Theor Appl Genet.* 2006; 112: 665–673. PMID: [16365761](https://pubmed.ncbi.nlm.nih.gov/16365761/)
57. Guzman P, Diers B, Neece D, St. Martin S, LeRoy A, Grau C, et al. QTL associated with yield in three backcross-derived populations of soybean. *Crop Sci.* 2007; 47: 111–122.
58. Liu B, Abe J. QTL mapping for photoperiod insensitivity of a Japanese soybean landrace Sakamoto-wase. *J Hered.* 2010; 101: 251–256. doi: [10.1093/jhered/esp113](https://doi.org/10.1093/jhered/esp113) PMID: [19959597](https://pubmed.ncbi.nlm.nih.gov/19959597/)
59. Githiri S, Yang D, Khan N, Xu D, Komatsuda T, Takahashi R. QTL analysis of low temperature-induced browning in soybean seed coats. *J Hered.* 2007; 98: 360–366. PMID: [17621588](https://pubmed.ncbi.nlm.nih.gov/17621588/)
60. Zhang W, Wang Y, Luo G, Zhang J, He C, Wu X, et al. QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr) genetic map and their association with EST markers. *Theor Appl Genet.* 2004; 108: 1131–1139. PMID: [15067400](https://pubmed.ncbi.nlm.nih.gov/15067400/)
61. Du W, Wang M, Fu S, Yu D. Mapping WTLs for seed yield and drought susceptibility index in soybean (*Glycine max* L.) across different environments. *J Genet Genom.* 2009; 36: 721–731.
62. Reinprecht Y, Poysa VW, Yu KF, 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: 1510–1527. PMID: [17426766](https://pubmed.ncbi.nlm.nih.gov/17426766/)
63. Panthee D, Pantalone V, Saxton A, West D, Sams C. Quantitative trait loci for agronomic traits in soybean. *Plant Breed.* 2007; 126: 51–57.
64. Zuo QM, Hou JF, Zhou B, Wen ZX, Zhang SY, Gai JY, et al. Identification of QTLs for growth period traits in soybean using association analysis and linkage mapping. *Plant Breed.* 2013; 1–7.
65. Eskandari M, Cober ER, Rajcan I. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased seed yield. *Theor Appl Genet.* 2013; 126: 1677–1687. doi: [10.1007/s00122-013-2083-z](https://doi.org/10.1007/s00122-013-2083-z) PMID: [23536049](https://pubmed.ncbi.nlm.nih.gov/23536049/)
66. Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR, et al. Seed quality QTL in a prominent soybean population. *Theor Appl Genet.* 2004; 109: 552–561. PMID: [15221142](https://pubmed.ncbi.nlm.nih.gov/15221142/)
67. Rossi M, Orf J, Liu L, Dong Z, 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. doi: [10.1007/s00122-013-2094-9](https://doi.org/10.1007/s00122-013-2094-9) PMID: [23595202](https://pubmed.ncbi.nlm.nih.gov/23595202/)
68. Gai J, Wang Y, Wu X, Chen S. A comparative study on segregation analysis and QTL mapping of quantitative traits in plants-with a case in soybean. *Front of Ag in China.* 2007; 1: 1–7.
69. Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, et al. Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. *Crop Sci.* 1999; 39: 1642–1651.
70. Reyna N, Sneller CH. Evaluation of marker-assisted introgression of yield QTL alleles into adapted soybean. *Crop Sci.* 2001; 41: 1317–1321.
71. Fan C, Hu R, Zhang X, Wang X, Zhang W, Zhang Q, et al. Conserved CO-FT regulons contribute to the photoperiod flowering control in soybean. *BMC Plant Biol.* 2014; 14: 9. doi: [10.1186/1471-2229-14-9](https://doi.org/10.1186/1471-2229-14-9) PMID: [24397545](https://pubmed.ncbi.nlm.nih.gov/24397545/)

72. Komatsu K, Hwang T, Takahashi M, Sayama T, Funatsuki H, Oki N, et al. Identification of QTL controlling post-flowering period in soybean. *Breed Sci.* 2012; 61: 646–652. doi: [10.1270/jsbbs.61.646](https://doi.org/10.1270/jsbbs.61.646) PMID: [23136504](https://pubmed.ncbi.nlm.nih.gov/23136504/)
73. Molnar SJ, Rai S, Charette M, Cober ER. Simple sequence repeat (SSR) markers linked to *E1*, *E3*, *E4*, and *E7* maturity genes in soybean. *Genome.* 2003; 46: 1024–1036. PMID: [14663521](https://pubmed.ncbi.nlm.nih.gov/14663521/)
74. Funatsuki H, Kawaguchi K, Matsuba S, Sato Y, Ishimoto M. Mapping of QTL associated with chilling tolerance during reproductive growth in soybean. *Theor Appl Genet.* 2005; 111: 851–861. PMID: [16059730](https://pubmed.ncbi.nlm.nih.gov/16059730/)
75. Khan N, Githiri S, Benitez E, Abe J, Kawasaki S, Hayashi T, et al. QTL analysis of cleistogamy in soybean. *Theor Appl Genet.* 2008; 117: 479–487. doi: [10.1007/s00122-008-0792-5](https://doi.org/10.1007/s00122-008-0792-5) PMID: [18506418](https://pubmed.ncbi.nlm.nih.gov/18506418/)
76. Poomprapan P, Wasee S, Toojinda T, Abe J, Chanprame S, Srinives P. Molecular marker analysis of days to flowering in vegetable soybean (*Glycine max* (L.) Merrill). *Kase J.* 2006; 40: 573–581.
77. Csanadi G, Vollmann J, Stift G, Lelley T. Seed quality QTLs identified in a molecular map of early maturing soybean. *Theor Appl Genet.* 2001; 103: 912–919.
78. Kantolic AG, Slafer GA. Reproductive development and yield components in indeterminate soybean as affected by post-flowering photoperiod. *Field Crop Res.* 2005; 93: 212–222.
79. Sayama T, Hwang TY, Yamazaki H, Yamaguchi N, Komatsu K, Takahashi M, et al. Mapping and comparison of quantitative trait loci for soybean branching phenotype in two locations. *Breed Sci.* 2010; 60: 380–389.
80. Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* 1996; 36: 1327–1336.
81. Zhai H, Lü S, Wang Y, Chen X, Ren H, Yang J, et al. Genes and transcriptional abundance of the *E1* gene are associated with flowering time and maturity of soybean cultivars. *PLoS One.* 2014; 9: 1–16.
82. Raychaudhuri S. Mapping rare and common causal alleles for complex human diseases. *Cell.* 2011; 147: 57–69. doi: [10.1016/j.cell.2011.09.011](https://doi.org/10.1016/j.cell.2011.09.011) PMID: [21962507](https://pubmed.ncbi.nlm.nih.gov/21962507/)
83. Ladouceur M, Dastani Z, Aulchenko YS, Greenwood CM, Richards JB. The empirical power of rare variant association methods: results from sanger sequencing in 1,998 individuals. *PLoS Genet.* 2012; 8: e1002496. doi: [10.1371/journal.pgen.1002496](https://doi.org/10.1371/journal.pgen.1002496) PMID: [22319458](https://pubmed.ncbi.nlm.nih.gov/22319458/)