Dissecting four correlated growth period traits using a genome-wide association study approach in soybean

Yanfei Li · Ying-hui Li · Huilong Hong · Yansong Ma · Yongzhe Gu · Zhangxiong Liu · Scott A. Jackson · Li-juan Qiu

Received: 19 January 2021 / Accepted: 29 October 2021 / Published online: 15 November 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract Soybean [Glycine max (L.) Merr.] is a short-day, photoperiod-sensitive crop. The vegetative growth period (V), reproductive growth period (R), whole growth period (V + R) and the ratio of growth period structure (R/V) are photoperiod traits critical to adaptability and yield in soybean. To thoroughly dissect the genetic basis of these growth period traits and their correlations, a panel of 277 soybean accessions were genotyped with a single nucleotide polymorphisms (SNPs) panel and phenotyped at three field environments over three years. A genome-wide association study found 28 SNPs on fifteen chromosomes with 33 significant marker-trait associations, eleven, eight, four and ten for V, R, V + R and R/V, respectively. Two SNPs, Map-6077 and -6502, showed pleiotropy in three and two traits, respectively. Map-6077, associated with V, R, and V + R, was a synonymous SNP of the GmMFT, a member of PEBP family that has been reported to inhibit seed germination and regulate oil content. These results provide candidate SNP markers and genes for further functional experiments to advance molecular marker-assisted breeding.

Keywords Genome-wide association study · Soybean · SNPs · Growth period traits · Photoperiod

Y. Li · Y. Li (✉) · H. Hong · Y. Ma · Y. Gu · Z. Liu · L. Qiu (✉)
The National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI) / Key Lab of Germplasm Utilization, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China
e-mail: liyinghui@caas.cn

L. Qiu
e-mail: qiulijuan@caas.cn

H. Hong
College of Agriculture, Northeast Agricultural University, Harbin 150090, Heilongjiang, China

Y. Ma
Beidahuang KenFeng Seed Co., Ltd, Heilongjiang 150090, Harbin, China

S. A. Jackson
Center for Applied Genetic Technologies, University of Georgia, Athens, GA 30602, USA
Introduction

As a typical short-day crop, soybean (Glycine max (L.) Merr.) is sensitive to changes in photoperiod, which varies with latitude and planting season (Liu et al. 2017). This sensitivity limits the introduction, cultivation and large-scale planting of elite soybean cultivars. Growth period traits are influenced by the response of the plant to the relative length of day and night. This is important not only in the context of extending the geographical range of cultivation of the crop but also to increase yield.

Although the whole growth period (V + R) includes vegetative growth (V) and reproductive growth (R), they are relatively independent developmental stages (Fehr and Caviness 1977) and are used as evaluation traits (Yang et al. 1994; Zhang et al. 2015). R is related to yield, quality and resistance to stresses from environment, and several genetic loci have been associated with R besides V and V + R (Cheng et al. 2011). Time to flower (R1), maturity (R8), and reproductive period (RP) have also been defined as three separate traits (Kong et al. 2018). Genetic loci differentiating DTF, days to flowering, (V) and DTM, days to maturity, (V + R) as well as DTFM, days from flowering to maturity, (R) have been found indicating that there may be major and minor-effect loci underlying V and V + R (Zhang et al. 2015). The ratio of R and V (R/V) is used as an index of growth period structure and the relationships between two stages (Jian 1990; Wang et al. 2015a). These traits are all photoperiod and important to the adaptability and yield of soybean (Garner and Allard 1927; Han et al. 2006). Pleiotropy is due to a QTP (SNP or other polymorphism affecting a quantitative trait) affecting two or more traits (Chen and Lubberstedt 2010). For example, a few genes have been found to associate with multiple agronomic traits in crops, for example, Ghd7 regulated heading date and yield potential in rice (Xue et al. 2008).

Growth period traits of soybean are complex quantitative traits, controlled by multiple genes, their interactions (Chen and Lubberstedt 2010), and the environment. Two major approaches, linkage and association analysis, have been used to dissect the genetic basis of complex traits (Purcell et al. 2003). With the development of high-throughput sequencing and SNP genotyping techniques, the genome-wide association study (GWAS) has become an efficient method to detect variants underlying complex traits. GWAS was used to identify ZmVPP1 for drought tolerance in maize seedlings (Wang et al. 2016) and OsSPL13 which controls grain size in cultivated rice (Si et al. 2016). Previously, GWAS on growth period traits in soybean was conducted using 113 wild accessions (Hu et al. 2014), 139 short-season accessions (Sonah et al. 2015), 309 early mature accessions (Zhang et al. 2015), 91 cultivars (Mao et al. 2017), and 421 landraces and 388 cultivars (Fang et al. 2017) planted at 32.20 to 45.57°N. However, growth period traits are sensitive to different environments and the genetic characteristics of the target population. This study selected accessions from the mini core collections of the Chinese cultivated soybean, which represent the 94.5% phenotypic diversity and 63.5% genetic diversity (Qiu et al. 2009, 2013), that were planted in three locations with latitudes ranging from 18.25 to 40.22°N.

Previously, 362 bi-parental QTLs (quantitative trait loci) and 242 GWAS QTLs involved in growth period traits were identified by linkage and association analysis (summarized at: http://soybase.org). Several major genes [E1 (Glyma.06g207800) (Bernard 1971; Xia et al. 2012), E2 (Glyma.10g221500, GmGia) (Bernard 1971; Watanabe et al. 2011), E3 (Glyma.19g224200, GmPhyA3) (Buzzell 1971; Watanabe et al. 2009), E4 (Glyma.20g090000, GmPhyA2) (Buzzell and Voldeng 1980; Matsumura et al. 2008; Saindon et al. 1989a, 1989b), E9 (Glyma.16g150700, GmFT2a) (Kong et al. 2014; Zhao et al. 2016), E10 (Glyma.08g363100, GmFT4) (Samanfar et al. 2017), J (Glyma.04g050200, GmELF3) (Lu et al. 2017; Ray et al. 1995; Yue et al. 2017), qDTF-J (Glyma.16g044100, GmFT5a) (Takeshima et al. 2016), and GmPrr3b (Glyma.12g073900) (Li et al. 2020)] have been reported to contribute to time of flowering and/or maturity through linkage analysis. We expect that there are many more loci or genes yet to be discovered.

Here, we conducted a genome-wide study of vegetative growth period (V), reproductive growth
period (R), whole growth period (V + R), and the ratio of growth period structure (R/V) separately using 277 soybean accessions and 1161 SNPs, which found 32 significant marker-trait associations (MTAs). This study provides a better understanding of the underlying genetic basis of the growth period traits and provides functional SNPs as tools for molecular soybean breeding programs aimed at improving the adaptation and yield of soybean to each latitude.

Materials and methods

Plant materials

The germplasm panel was comprised of 277 cultivated accessions with 192 landraces and 85 improved cultivars collected worldwide. Half (135) are members of the Chinese soybean mini-core collections (Qiu et al. 2009, 2013) (Table S1). The remaining 142 originated in China (120), United States (11), Canada (1), Poland (1), North Korea (1), Denmark (1), Germany (1), Japan (1), India (1), Russia (1), France (1), Italy (1) and Romania (1) (Fig. 1; Table S1). All 277 accessions were collected between 18.15°N and 55.45°N (Fig. 1).

Measurement of phenotypes

The 277 accessions were planted in a randomized complete block design with three replications at three field environments (BJSp, Beijing Spring; BJSu, Beijing Summer; SY, Sanya) over two years 2011–2012 (BJSp, BJSu) or three years 2010–12 (SY). For BJSu, BJSp, and SY, the average photoperiod (day length) decreased gradually due to planting in the field without any treatment. The field environments were defined as combinations of each location and planting season. BJSp and BJSu are from the Changping and Shunyi district of Beijing (40.22°N, 116.2°E; 40.13°N, 116.65°E) with two different planting seasons: Spring (planted in mid-May) and Summer (planted in mid-June). SY is from Sanya, Hainan Province (18.25°N, 109.5°E), nearly 22° south of Beijing, planted in early November. At the Beijing site, each accession was planted in a 1.5 m long row with an interplant distance of ten centimeters (cm) and an inter-row distance of 55 cm. At Sanya, the row length was uniformly one meter, the interplant distance was six cm, and the inter-row distance was 45 cm. The following traits were scored: date of emergence of the cotyledons as VE, the date when 50% of the plants within the row first flowered as R1 and the date when 50% of the plants within the row had reached maturity (95% of pods maturity for one plant) as R8 (Fehr et al. 1971; Qiu et al. 2006). Four traits were analyzed: V = R1-VE, R = R8-R1, V + R = R8-VE, R/V = (R8-R1) / (R1-VE). V and R represent the vegetative and reproductive stages as described previously (Fehr et al. 1971). Measurements were averaged over the three replicates, and subsequently over the three years in each environment, using a best linear unbiased prediction (BLUP) (Henderson 1975).

Phenotype data analysis and Heritability estimation

Plants that did not reach maturity were not included and treated as missing data. The field data was used to get V, R, V + R and R/V phenotype data, and extreme values that due to artificial error or plant died from virus or other reasons were filtered. BLUP values for each independent environment used in GWAS, phenotypic description analysis, and trait correlations analysis were estimated through the R environment package lme4 package (Bates et al. 2014) to evaluate the BLUPs with the linear mixed model, lmer(y ~ (1|Accession) + (1|Year) + (1|Accession:Year)). The broad-sense heritability for each trait was evaluated by lme4 package using the formula $H^2 = V_g/(V_g + V_{GL}/L + V_{GY}/Y + V_e/(L*Y))$, where $V_g$, $V_{GL}$, $V_{GY}$, $V_e$ represent variance of the genotype, genotype and location, genotype and year, and residual variance, respectively. Descriptive analysis of the phenotypic data was performed using R software (https://www.r-project.org).

SNP genotyping and filtering

A total of 1236 SNPs were used for genotyping 277 cultivated soybean accession. These SNPs included 1150 SNPs from a reported SNP genotyping array which selected from random distribution, homologs and reported QTL based on 55 whole-genome resequenced genome and a transcript map (Li et al. 2015). An additional panel of 86 SNPs obtained by Sanger sequencing that around known QTL of growth period traits or distributed among chromosomes. The
methods for SNP genotyping of 1150 SNPs and 86 SNPs were based on the Illumina BeadArray platform (Illumina Inc., San Diego, CA, USA) and Sequenom MassARRAY platform (Gabriel et al. 2009). After obtaining genotype data, we filtered 75 SNPs with missing rate (> 10% of samples), or apparent

![Diagram](https://example.com/diagram.png)

**Fig. 1** The geographical distribution and population structure analysis of 277 soybean accessions. NR, North region; CR, Central region; SR, South region. NR, CR, SR, and mixture correspond to the color of red, yellow, blue, and gray. (a) Geographical distribution of 277 soybean accessions. The x-axis represents longitude, and the y-axis represents latitude. One point represents one accession. (b) Genome-wide principal components analysis (PCA) of 277 soybean accessions. Principal Component 1 (PC1) is plotted against 2 (PC2) and 3 (PC2). (c) Evolutionary relationships of 277 soybean accessions. The evolutionary history was inferred using the Neighbor-Joining method. The branch length of the tree indicates the evolutionary distances which computed using the p-distance method. (d) Three subpopulations inferred from structure result. Each color represents one cluster, and the length of the colored segment shows the accession’s estimated proportion of membership in that cluster as calculated by STRUCTURE software. (Color figure online)
heterozygosity (> 20%), or low minor allele frequency (< 5%). The final data set consisted of 1161 SNPs (Table S2).

Population structure analysis

TASSEL 5.0 software was used to calculate the Minor Allele Frequency, Proportion Missing, and Proportion Heterozygous for each SNP locus (Bradbury et al. 2007). A subset of 800 SNPs distributed across twenty chromosomes were selected to determine population structure through the software package STRUCTURE 2.1 (Falush et al. 2007) using Bayesian Markov Chain Monte Carlo approach. The admixture and independent allele frequency model were adopted, with cluster number (K) from one to ten. Twenty runs performed with a 100,000 burn-in length and 100,000 iterations for each value of K. The most likely number of sub-clusters was obtained through the online Structure Harvester (Earl and vonHoldt 2011). The program executed the Evanno method (Evanno et al. 2005) to detect the number of K clusters (with the peak delta K value) that best fit the data.

Genome-wide association study and candidate gene analysis

GWAS was conducted using a developed model selection algorithm, FarmCPU (the fixed and random model circulating probability unification) in R package (Liu et al. 2016). The confounding problem between covariates and test SNP is considered in the algorithm by using both a Fixed Effect Model (FEM) and a Random Effect Model (REM). The first three calculated principal components were used as covariates. A sum of 1,161 SNPs, of which MAF > 5% and missing rate < 0.1, was used to perform GWAS. We used FarmCPU.P_THRESHOLD function to obtain the recommended p-value threshold. The relationship with the genotypes was broken by the phenotypes permutation. The permutation time was 500. A vector of minimum p-value is produced, of which the 95% quantile value is recommended. For the significantly associated MTAs, the allelic effects were determined by representing phenotype data carrying different alleles as box-plots and were confirmed through the Student’s t-test.

All the annotated high-confidence genes located in the 300 kb genomic regions centered on the each significant MTA were selected for candidate gene identification. The gene annotation information and the expression profiles were obtained from SoyBase database (https://www.soybase.org/). These genes, of which annotation information were related to growth period traits and/or expressed in young leave, were identified as candidate genes.

Results

Phenotypic evaluation

A total of 277 soybean accessions from worldwide (Fig. 1a) were phenotyped at three environments, including Sanya (SY), Beijing Spring (BJSp) and Beijing Summer (BJSu). BLUP values among two or three years were used in the following analysis as the variance of year was not significant. The descriptive statistics of V, R, V + R and R/V, revealed wide variation both within and between the three environments, providing a diverse phenotypic dataset with which to dissect the complex genetic architecture of the target traits. Among these three environments, SY had the smallest variation range from 1.29 (V) to 1.57 (R/V) while BJSp had the largest range from 1.99 (R) to 4.65 (V). For V, the mean value in BJSp was 0.25- and 1.08-fold higher than BJSu and SY, and the differences decreased for R and V + R as the fold differences were 0.19 and 0.38, 0.19 and 0.53, respectively. In contrast to R/V, the mean value in BJSp was the same as BJSu but 0.21-fold lower than SY. A different trend was observed for V, R, and V + R where they decreased from BJSp to BJSu, then to SY, while increasing for R/V (Table 1). The ANOVA indicate that the effects of genotype and the genotype and environment interaction were both significant for V, R, V + R, and R/V at 0.001 level.

The correlation coefficients of the four traits were calculated for the three environments (BJSp, BJSu, and SY corresponding to Fig. 2a–c). Similar correlation patterns were observed for BJSp and BJSu, significant positive correlations between R and V + R, and R and R/V, significant negative correlations between V and R, V and R/V, and V + R and R/V, and little between V and V + R. However, SY differed in that V and R had a significant negative correlation, r = −0.47, and a significant positive correlation between R and V + R with r = 0.92,
while other traits had no significant correlations. The high correlation coefficient between R and V + R in Sanya suggested that the length of V + R depended on R. The correlation coefficients between BJSp, BJSu, and SY for four traits were shown in Fig. S1. BJSp and BJSu had high positive correlation coefficients (r = 0.63–0.94) between them for the four traits. But SY showed lower positive coefficients with BJSp and BJSu for V and R (r = 0.38–0.45) and no relationship for V + R and R/V. In all, the Sanya environment exhibited a large difference from BJSp and BJSu while BJSp and BJSu had similar trends.

Table 1  Descriptive analysis of the phenotype data of 277 soybean accessions

| Trait | Location | Range (days) | Variation fold | Mean ± SD (days) | CV(%) | Heritability | ANOVA |
|-------|----------|--------------|----------------|-----------------|-------|--------------|-------|
| V     | BJSp     | 26.2–121.8   | 4.65           | 56.6 ± 22.8     | 40.3  | 0.64         | 127.1*** 0 | 136.2*** |
|       | BJSu     | 22.7–96.5    | 4.26           | 45.3 ± 16.0     | 35.3  |              |        |
|       | SY       | 24.6–31.9    | 1.29           | 27.2 ± 1.4      | 5     |              |        |
| R     | BJSp     | 52.5–155.9   | 2.97           | 76.9 ± 14.4     | 18.7  | 0.55         | 20*** 0.77 | 16.52*** |
|       | BJSu     | 54.1–75.9    | 1.4            | 64.4 ± 4.2      | 6.5   |              |        |
|       | SY       | 45.5–62.7    | 1.38           | 55.6 ± 3.1      | 5.5   |              |        |
| V + R | BJSp     | 80.6–160.7   | 1.99           | 126.3 ± 16.4    | 13    | 0.56         | 56.93*** 0 | 82.1***  |
|       | BJSu     | 79.3–130.4   | 1.64           | 106.5 ± 10.4    | 9.7   |              |        |
|       | SY       | 71.6–88.3    | 1.23           | 82.8 ± 2.6      | 3.2   |              |        |
| R/V   | BJSp     | 0.74 ~ 3     | 4.05           | 1.7 ± 0.5       | 29.7  | 0.64         | 0.1*** 0  | 0.05***  |
|       | BJSu     | 0.9–2.7      | 2.91           | 1.7 ± 0.4       | 24.9  |              |        |
|       | SY       | 1.6–2.5      | 1.57           | 2.1 ± 0.2       | 8.6   |              |        |

Each significance level is associated to a symbol: p-values (0.001, 0.01, 0.05) < = > symbols ("***", "**", "+") SD standard deviation, CV the coefficient of variation, V vegetative growth period, R reproductive growth period, R/V the value of growth period structure, V + R whole growth period, BJSp Beijing Spring, BJSu Beijing Summer, SY Sanya

Fig. 2  The correlations between four growth period traits for three environments. The correlation matrix for the environments of BJSp (a), BJSu (b), and SY (c). The upper, diagonal, and lower parts show the correlation coefficients, frequency distribution histograms plot and density curve, and scatterplots and LOESS smoothed fits of the four growth period traits, respectively. Three environments: BJSp, Beijing Spring; BJSu, Beijing Summer; SY, Sanya. Four growth period traits: V, vegetative growth period; R, reproductive growth period; V + R, whole growth period; R/V, the ratio of growth period structure. The method of Pearson’s was used to calculate the coefficient of correlation for pairwise comparison of four traits. Each significance level is associated to a symbol: p-values (0.001, 0.01, 0.05) < = > symbols ("***", "**", "+")
which resulted from the very different photoperiodic environments (Beijing, 40.22°N; Sanya, 18.25°N).

Population structure

A total of 1161 SNPs remaining after filtering (Table S2) were used in the ensuing analyses. Population structure analysis was conducted with STRUCTURE (Evanno et al. 2005; Kauffner et al. 2007) using 800 evenly distributed SNPs on the twenty chromosomes. The results indicated three clusters, NR (Northern Region), CR (Central Region), and SR (Southern Region), concordant with a shared allele-based neighbor-joining tree and PCA (principal components analysis) (Fig. 1b–d). Subpopulation SR was the largest group with 105 accessions, containing 76 accessions from south China; subpopulation NR consisted of 90 accessions, dominated by the accessions of 63 from north and northeast and 16 from outside China; and subpopulation CR had 34 accessions, 25 of which were from North and Huanghuai (Table S1). The derivative clusters of these soybean accessions referred to Q value from population structure analysis consistent with the geographical origin (Fig. 1; Table S1).

Marker-trait association analysis for V, R, V + R and R/V

We used the fixed and random model through FarmCPU to conduct the GWAS and added the first three PCs in this model to avoid false positives from non-genetic effects associated with population structure (NR, CR, and SR). The GWAS revealed 29 SNP markers representing 33 significant MTAs above the threshold \( p < 6.70\times10^{-05} \) for V (10), R (8), V + R (4) and R/V (10) distributed across fourteen soybean chromosomes except for Gm01, 02, 04, 11, 12, and 19 (Fig. 3, Table 2). Two SNPs were associated with more than one trait and/or identified in two environments, for example Map-6502 associated with V (BJSp, \( p = 3.80\times10^{-07} \); BJSu, \( p = 9.05\times10^{-07} \)) and R/V (BJSp, \( p = 5.46\times10^{-05} \)). Most, 27 of the 28 SNPs were located in coding regions of annotated genes. The genetic effects of 27 SNPs were also determined: non-synonymous polymorphisms (17), synonymous polymorphisms (9) and a stop-gain codon (Table 2). Twelve of the 28 SNPs were close (300 kb, kilobase) or within (Map-6077) of reported flowering-related genes, 12 were in known QTLs related to growth period traits, and six matched both conditions (Table 2).

Vegetative growth period (V)

For V (Fig. 3), 11 significant MTAs were identified on chromosomes Gm07 (2), 10 (2), 14 (2), 16, and 18 (4). Map-6173, was covered by five QTL of growth period traits including V (Suo et al. 2016; Wang et al. 2015b), was located at 172 kb downstream of Gm.07g049400, a homolog of APRR5 (pseudo-response regulator 5) belonging to the APRR1/TOC1 quintet implicated in circadian rhythm, control of flowering time, and early photomorphogenesis in Arabidopsis thaliana (Watanabe et al. 2012; Yamamoto et al. 2003). Map-6384 is a synonymous polymorphism within a B-box (BBX) protein gene GmCOL7a (Glyma.10g274300) (Sedivy et al. 2017). BBX proteins are key factors in regulatory networks controlling growth and developmental processes that included seedling photomorphogenesis, photoperiodic regulation of flowering (Gangappa and Botto 2014). Map-3175 was located at 129 kb downstream of Glyma.16g196300 (GmTFL), a member of PEBP (phosphatidylethanolamine binding protein) family protein, which participates in the photoperiod pathway to regulate flowering time as a floral repressor in soybean (Liu et al. 2010).

Reproductive growth period (R)

For R, we found eight significant MTAs on chromosomes Gm03, 05 (2), 07, 08 (2), 10, and 18. Map-6214 was located at 19 kb downstream of Glyma.07g089000, which was homologous to the A. thaliana VRN5 gene (At3g24440) (Suo et al. 2016) involved in both vernalization (Sung et al. 2006) and photoperiod pathways (Greb et al. 2007), and its expression is down-regulated significantly in the V4 stage in the photoperiod pathway in soybean (Zeng et al. 2018). The peak SNP, Map-6285, was associated with R by the most significant association \( p = 4.76\times10^{-11} \) in BJSu. Glyma.08g141000, located at 69 kb upstream of the peak SNP Map-6285, is a homolog of the A. thaliana gene AT4G14540, the Nuclear Factor Y subunits NF-YB3, a candidate flowering gene in long-day photoperiod under the influence of E3 in soybean (Haider 2014). Map-6383
was located at 268 kb upstream of Glyma.10g221500 (E2/GmGIa), which was validated by functional experiments to regulate soybean maturity and flowering time (Tsubokura et al. 2014; Watanabe et al. 2012; Watanabe et al. 2011) and covered by four QTLs of growth period traits including two of R.

Whole growth period (V + R)

For V + R, we found four significant MTAs on chromosomes Gm08, 15, 18, and 20. Map-6541 was covered by three QTLs related to growth period traits (two of V + R). Glyma.15g162300 (GmZTL3) was located at 262 kb upstream of Map-6541 and has been proposed to be a photoreceptor and play a role in the

---

**Fig. 3** Manhattan and QQ Plot of four growth period traits. a, c, e, g The Manhattan plots GWAS results. The legend on the right shows the density of the SNP markers, the color from light to dark represent the density from low to high. b, d, f, h The QQ plots of GWAS results. The x axis displays the twenty chromosomes in soybean genome, and the y axis displays the logarithm of p value. Different colors represent the 12 GWAS analyses that combine by four traits and three environments. Three environments: BJSp, Beijing Spring; BJSu, Beijing Summer; SY, Sanya. Four growth period traits: V, vegetative growth period; R, reproductive growth period; V + R, whole growth period; R/V, the ratio of growth period structure. BLUP Data over years were used. (Color figure online)
| Marker name | Chr: position (bp) | MTAs | p Value | r2 | Wm82.a2.v1 Location: effect | Reported QTL [a] | Reported genes (gene name; distance(kb); reference) |
|-------------|-------------------|------|---------|----|---------------------------|-----------------|--------------------------------------------------|
| V Map-6173  | 7: 4378553        | V_BJSp | 5.53E−05 | 0.008 | Glyma.07g050700 CDS: A --- S | First flower 6–1; Reproductive stage length 4–3; Pod maturity 10–2; 14–4; 18–4 | Glyma.07g049400 (GmPRR5; 172; Watanabe et al., 2012) |
| Map-1899 | 10: 44738812 | V_BJSp | 5.17E−05 | 0.144 | Glyma.10g215300 CDS: A --- S | Glyma.10g274300 (GmCOL7a; within gene; Sedivy et al., 2017) |
| Map-6502 | 14: 10917924 | V_BJSp; V_BJSu | 3.80E−07 | 0.03; | Glyma.14g105900 CDS: V --- A |
| Map-3175 | 16: 35908252 | V_BJSp | 2.24E−05 | 0.159 | Glyma.16g197900 CDS: W --- X | Photoperiod insensitivity 5–4 | Glyma.16g196300 (GmTFL; 129; Suo et al., 2016) |
| Map-6646 | 18: 86492 | V_BJSu | 2.24E−05 | 0.159 | Glyma.18g000500 CDS: S --- T |
| Map-6708 | 18: 2813504 | V_BJSp; V_BJSu | 1.37E−05 | 0.021; | Glyma.18g036100 CDS: T --- N |
| Map-6743 | 18: 9082263 | V_BJSp | 2.71E−05 | 0.013 | Glyma.18g091100 CDS: R --- S | Pod maturity 16–2; Barc-013665-01226 5: 33789326 R_SY 5.07E-05 0.002 | Glyma.049219800 (130; Jung et al., 2012); Glyma.05g244100 (GmMFT; 129; Suo et al., 2016); Glyma.07g090100 (-110; Jung et al., 2012) |
| Map-6051 | 3: 42449846 | R_BJSu | 5.27E−06 | 0 | Glyma.03g221400 CDS: S --- R |
| BARC-013665-01226 | 5: 33789326 | R_SY | 5.09E−05 | 0.003 | Glyma.05g143900 CDS |
| Map-6077 | 5: 41853190 | R_SY | 5.17E−05 | 0.041 | Glyma.05g244100 CDS |
| Map-6214 | 7: 8323908 | R_BJSu | 6.14E−06 | 0 | Glyma.07g089200 CDS: S --- N | First flower 2–2; SeedFill 2–2; Pod maturity 2–1 | Glyma.07g089000 (GmVRN5; 19; Zeng et al., 2018); Glyma.07g090100 (-110; Jung et al., 2012) |
| Map-6285 | 8: 10308599 | R_SY | 4.76E−11 | 0.019 | Glyma.08g134500 CDS: L --- P |

[a] [a] Reported QTL [a] | Reported genes (gene name; distance(kb); reference)
| Marker name   | Chr: position (bp) | MTAs       | p Value       | r2       | Wm82.a2.v1         | Location: effect                                                                 | Reported QTL [a]                                                                 | Reported genes (gene name; distance(kb); reference) |
|--------------|--------------------|------------|---------------|----------|--------------------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------|-----------------------------------------------------|
| Map-6303     | 8: 11089154        | R_SY       | 2.44E-05      | 0        | Glyma.08g145800    | CDS: H —> Y                                                                      |                                                                                 |                                                                     |
| Map-6383     | 10: 45026735       | R_BJSu     | 1.58E-07      | 0        | Glyma.10g218400    | CDS                                                                              |                                                                                 |                                                                     |
| Map-3569     | 18: 25723766       | R_BJSu     | 8.59E-08      | 0.069    | Intergenic         | First flower 10–2; Reproductive period 1–3; Pod maturity 34–5; Photoperiod insensitivity 2–2 | First flower 24–4; Reproductive stage length 8–2; 9–2; Pod maturity 26–3       | Gm10g2221500 (E2/GnGla; 268; Watanabe et al., 2011) |
| V + R        |                    |            |               |          |                    |                                                                                 |                                                                                 |                                                                     |
| Map-6312     | 8: 11647122        | V + R_SY   | 2.16E-07      | 0.001    | Glyma.08g151600    | CDS: T —> A                                                                      |                                                                                 |                                                                     |
| Map-6541     | 15: 13587270       | V + R_SY   | 1.79E-05      | 0.001    | Glyma.15g160300    | CDS                                                                              |                                                                                 |                                                                     |
| Map-6733     | 18: 4240398        | V + R_SY   | 5.00E-05      | 0        | Glyma.18g049300    | CDS: S —> N                                                                      |                                                                                 |                                                                     |
| Map-6803     | 20: 25300503       | V + R_SY   | 3.29E-06      | 0        | Glyma.20g071000    | CDS                                                                              |                                                                                 |                                                                     |
| R/V          |                    |            |               |          |                    |                                                                                 |                                                                                 |                                                                     |
| Map-6143     | 6: 15873865        | R/V_BJSp   | 2.02E-05      | 0.033    | Glyma.06g183900    | CDS                                                                              |                                                                                 |                                                                     |
| Map-6146     | 6: 22813330        | R/V_BJSp; R/V_BJSp | 2.79E-09; 5.09E-06 | 0.035; 0.026 | Glyma.06g217000 | CDS                                                                              |                                                                                 |                                                                     |
| Map-6267     | 8: 9438024         | R/V_BJSp   | 4.12E-06      | 0.155    | Glyma.08g122400    | CDS                                                                              |                                                                                 |                                                                     |
| Map-6293     | 8: 10706843        | R/V_BJSp   | 6.20E-05      | 0.022    | Glyma.08g139900    | CDS: E —> Q                                                                      |                                                                                 |                                                                     |
| Map-1722     | 9: 47977625        | R/V_BJSp   | 2.49E-07      | 0.001    | Glyma.09G261900    | CDS: D —> G                                                                      |                                                                                 |                                                                     |
| Map-6413     | 13: 34928184       | R/V_BJSp   | 1.08E-05      | 0.032    | Glyma.13g239000    | CDS: S —> R                                                                      |                                                                                 |                                                                     |

123
control of flowering time in soybean (Jung et al. 2012; Zhou et al. 2016).

The ratio of growth period structure (R/V)

For R/V, we found 10 significant MTAs on chromosomes Gm06 (3), 08 (2), 09, 13, 14, 15 and 17. Map-6267 was located at 129 kb downstream of Glyma.08g120600 (GmMADS), one of the MADS-box gene family, which plays an important role in soybean growth and floral development (Shu et al. 2013). Glyma.08g148200 was located at 212 kb upstream of Map-6293 (R/V). It was a homolog to the A. thaliana gene AT1G08970, Nuclear Factor Y subunits NF-YC9, and related to seed development and maturation in soybean.

Glyma.15g176300, a homolog of ELF4-L3 (EARLY FLOWERING 4-like3), functioned as a signaling intermediate in circadian clock function, and photoperiod perception in A. thaliana (Marcolino-Gomes et al. 2017), was located at 31 kb downstream of Map-7187. The second significant SNP Map-6146 was found in BJSu (p = 2.79E-09) and BJSp (p = 5.09E-06). Although it was covered by 19 growth period QTLs including seven of V → R, seven of V, and three of Photoperiod insensitivity (Table 2), no candidate genes with growth period traits were found nearby.

### Pleiotropy on V, R, V + R, and R/V in soybean

Two SNPs (Map-6502 and -6077) showed pleiotropy on two or more traits of V, R, V + R, and R/V. Map-6502 was associated with two traits, V (BJSp, p = 3.80E-07; BJSu, p = 9.05E-07) and R/V (BJSp, p = 5.46E-05). Further statistical analysis exhibited the significant difference between allele A and G in the defined cluster of SR for V and difference in the HR and SR for R/V but failed to calculate significance because the small number of statistics of 7 (allele A in HR) and 6 (allele G in SR) in Beijing Spring and 8 (allele A in HR) and 6 (allele G in SR) in Beijing Summer (Fig. S2). Map-6077 was related with R (SY, p = 2.69E-03) and V + R (SY, p = 9.82E-04) where the p value below the threshold, and further confirmed by significant difference analysis (Fig. S3). The days significantly decreased for V and V + R while increased for R due to the favorable allele A of Map-6077 in BJSp and
BJSu (Fig. S2). Map-6077 was a synonymous polymorphism within *Glyma.05g244100*, named as *GmMFT* in soybean, a member of the PEBP family that proposed to be a suppressor of seed germination (Li et al. 2014), also associated with seed oil content (Li et al. 2018; Zhang et al. 2018).

**Discussion**

Soybean is a short-day crop and the timing of transition from vegetative to reproductive period is regulated by environmental cues. We explored the genetic basis for these developmental phases using an association approach of cultivated germplasm grown at different environments. It had been reported causal genes playing different roles on DPF (before flowering) and DFM (after flowering) in different planting seasons (Yang et al. 1994). Thus, we picked two locations (Beijing and Sanya) with a latitudinal difference of 21.97 degrees and two planting seasons (Spring and Summer) to form three diverse field environments (BJSp, BJSu, and SY). The major soybean loci (*E1*-8, and *J*) were reported to be related to both time of flowering and maturity (Watanabe et al. 2012) and were photoperiod-sensitive loci (Cober and Morrison 2010). Photoperiod effects on genes have been shown to be dependent on growth period stages, affecting *E1* and *E4* during V and R, respectively, and *E3* during V + R (Wang et al. 2008).

The four traits, V, R, V + R, and R/V had relationship with each other. In BJSp and BJSu, a higher correlation was observed for V and V + R (*r* = 0.56, 0.55) as compared to R and V + R (*r* = 0.37, 0.39), indicating that V had greater impact on V + R than R. However, in SY, the correlation coefficient of R and V + R was large, *r* = 0.92, while V and V + R was negligible, *r* = 0.08, indicating that R had greater impact on V + R than V. The negative correlation between V and R (*r* = -0.47 ~ -0.07) indicating that a shorter V might imply a longer R, which may be beneficial to the accumulation of more dry matter in seeds while V, R, and V + R were all positively related with yield (Orf et al. 1999). SNP Map-6077 for V and R showed opposite effects, indicating an underlying genetic explanation for the negative correlation between V and R. We speculated that the distinct photoperiodic environments contributed to the obvious difference for V, R, R/V and V + R in Beijing (BJSp and BJSu, long-day (LD) with average day length of 13.5 h in summer) and Sanya (SY, short-day (SD) with average day length of 11.3 h). Take flowering time (V) for example, soybean is a typical SD plant that flowering early under SD but not LD condition. As a complex quantitative trait, flowering time is regulated by a sophisticated regulatory network in response to the change of photoperiodic environments (Lin et al. 2021). At least 12 flowering time loci (*E1* to *E11* and *J*) involved in this network have been identified with epistasis relationships by genetic analysis (Kim et al. 2020; Watanabe et al. 2012; Yang et al. 2017).

Several GWAS have been conducted to identify causal loci or genes. Mao et al. (2017) using 91 cultivars through four photoperiod treatments (SD, LD, SP, and SU) taken at 39°54′N, 116°46′ (Beijing China) identified six makers associated with flowering time close to (within 23 kb) or overlapping with other GWAS results (Fang et al. 2017; Hu et al. 2014; Mao et al. 2017; Sonah et al. 2015; Zhang et al. 2015). The first SSR marker was Satt564 identified in Mao’ result and Hu’s GWAS, which used 113 wild accessions planted in summer at Nanjing (32°12′N 118°37′E), and Nanyang (38°7′N 110°34′E), China (Hu et al. 2014). And the second SNP marker was ss715619034, 17.50 kb upstream of ss715619036 identified in Zhang’s GWAS, which collected phenotypic data of 309 early mature accessions in Aurora (44°17′N 96°41′W), Brookings (44°22′N, 96°47′W) and Watertown (44°54′N, 97°7′W), United States (Zhang et al. 2015). The other four SNP markers, Chr02_12316110, Chr05_40012587, Chr11_31718686, and Chr15_13133351, were 18.33, 22.11, 18.86, and 16.37 kb far from ss715581063, ss715591933, ss715610231, and ss715620540 that discovered in Fang’s research (Fang et al. 2017; Mao et al. 2017). Fang planted 421 landraces and 388 cultivars in six environments, including Beijing, (40°22′N, 116°23′E), Mudanjiang (44°58′N, 129°60′E), Zhoukou (33°62′N, 114°65′E), China, over three years. In total, there were six markers identified between Mao with Hu (one), with Zhang (one), and with Fang (four). The highly consistency in Mao and Fang’s GWAS could be caused by the similar photoperiodic environments.

To better understand the adaptation of soybean to more extreme latitudinal differences, we used GWAS to validate and find genes or loci during the growth
We selected 277 accessions including 135 from Chinese mini core collections (Qiu et al. 2009, 2013) and planted them in three locations with latitudes from 18.25 to 40.22°N. Three SNPs, Map-6077, Map-1899, and Map-6383, from our study were close to previously reported loci. Map-6077 related with V, R, and V + R in our study was 1.23 kb upstream of Chr05_41854422 associated with V + R (Fang et al. 2017) and 140.77 kb downstream of ss715599242 (Gm05_38636402, Glyma1.1) (Fang et al. 2017; Mao et al. 2017). Map-1899 and Map-6383, associated with V and R individually, were 27.87, and 1.58 kb downstream of Chr10_44710944 and Chr10_45025153 from Fang et al. (2017). The remaining 26 SNP loci were not found in other reported GWAS studies. Furthermore, we found that six of 33 significant MTAs were identified both in two environments, while most of which (27/33) were identified in different photoperiodic environments, implying the genotype by environment effects. These photoperiodic environment related MTAs will play an important role in the promotion of soybean accessions spreading to different latitudes.

The pleiotropy and trait correlations in this study are different growth stages, different from the general pleiotropy, such as the correlation between seed oil content and protein content. The last two traits (V + R and R/V) are derived from the first two and their relationships with other traits are determined by the variance and covariance of V and R and the calculation formula. Deciphering the correlations among V, R, V + R, and R/V could facilitate soybean improvement. A positive correlation could simultaneously increase the trait values, and negative correlation raised the trait value for one while decreasing other traits.

Map-6077 exhibited extreme significant difference between phenotypic variations carrying A and G alleles in all three environments (BJSp, BJSu, and SY) for V, R, and V + R (Fig. S2). Improved cultivars had a higher proportion of favorable allele A which had higher R, shorter V and V + R, as compared to landraces with 0.84 to 0.51, implying that the soybean accessions with early flowering (V) and maturity (R) and a comparatively longer reproductive period (R) were selected during improvement (Fig. S2). Map-6077 was a synonymous SNV in GmMFT, expressed at the stage of seed 10DAF and increased at seed 42DAF stage. It has been reported to be a negative regulator of seed germination (Li et al. 2014) and a potential causal candidate gene controlling oil content (Li et al. 2018) in soybean. There was another SNP in GmMFT associated with R8 full maturity as shown by GWAS (Fang et al. 2017). We proposed that GmMFT is a pleiotropic gene and Map-6077 is a pleiotropic SNP associated with V, R, and V + R. The analysis above implied that Map-6077 and -6502 as the pleiotropic SNP could be used as linkage makers for targeted traits. With the development of molecular marker technology, SNPs could be converted to CAPS/dCAPS or KASP makers easily. For most researchers and breeders, it is feasible to used developed markers to screen large numbers of samples followed the routine procedure. The elite materials carrying with favorable allele will be selected and used for new modern cultivars cultivation.

Conclusions

We identified 28 SNPs related to V, R, V + R or R/V. Fourteen candidate genes were found, including the E2 gene. Although the number of SNPs (1161) used in this study was not large, the low LD decay rate in soybean made it feasible for us to discover trait-associated SNPs through GWAS. Because of the decrease in costs associated with obtaining high-density genome-wide marker sets and improvements of statistical methods for GWAS, future challenges lay in innovation of accurate, low-cost, high-throughput, and reproducible phenotyping, which together may launch a new wave of marker-assisted breeding for crops. It is straightforward for crop breeders to provide the pleiotropic SNP allele with desirable effects on two or more correlated traits. In this study, two SNPs (Map-6077 and -6502) have pleiotropic effects the measured traits, but further functional experiments are needed. It was exciting to find that GmMFT, reported to inhibit the seed germination and related with seed oil content, was associated with V, R, and V + R in this study. Our critical mission in the future is to provide candidate genes and favorable alleles to produce better soybean cultivars more rapidly and efficiently.

Acknowledgements This research was supported by the National Key R & D Program for Crop Breeding (2020YFE002300 and 2016YFD0100201), the Platform of
# References

Bates D, Mächler M, Bolker B, Walker S (2014) Fitting linear mixed-effects models using lme4. arXiv preprint XXXaXXXrXiXXX:v:14065823

Bernard RL (1971) Two major genes for time of flowering and maturity in soybeans. I. Crop Sci 11:242–244. https://doi.org/10.2135/cropsci1971.0011183X001100020022x

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Buzzell RI, Voldeng H (1980) Inheritance of insensitivity to long day length. Soybean Genet News 7:26–29

Buzzell RI (1971) Inheritance of a soybean flowering response to fluorescent-daylight conditions. Can J Genet Cytol 13:703–707. https://doi.org/10.1139/g71-100

Chen Y, Lubberstedt T (2010) Molecular basis of trait correlations. Trends Plant Sci 15:454–461. https://doi.org/10.1016/j.tplants.2010.05.004

Cheng L et al (2011) Genetic analysis and QTL detection of reproductive period and post-flowering photoperiod responses in soybean. Theor Appl Genet 123:421–429. https://doi.org/10.1007/s00122-011-1394-8

Cober ER, Morrison MJ (2010) Regulation of seed yield and agronomic characters by photoperiod sensitivity and growth habit genes in soybean. Theor Appl Genet 120:1005–1012. https://doi.org/10.1007/s00122-009-1228-6

Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574–578. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Fang C et al (2017) Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. Genome Biol 18:161. https://doi.org/10.1186/s13059-017-1289-9

Fehr WR, Caviness CE (1977) Stages of soybean development. Special Report. 87

Fehr WR, Caviness CE, Burmood DT, Pennington JS (1971) Stage of development descriptions for soybeans. Glycine Max (L.) Merrill 1. Crop Sci 11:929–931. https://doi.org/10.2135/cropsci1971.0011183X001100020022x

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. Current Protocols in Human Genetics 60:2.12. 11–12.12

Gangappa SN, Botto JF (2014) The BBX family of plant transcription factors. Trends Plant Sci 19:460–470. https://doi.org/10.1016/j.tplants.2014.01.010

Garner WW, Allard HA (1927) Effect of short alternating periods of light and darkness on plant growth. Science 66:40–42. https://doi.org/10.1126/science.66.1697.40

Greb T, Mylne JS, Crevillen P, Geraldo N, An H, Gendall AR, Dean C (2007) The PHD finger protein VRN5 functions in the epigenetic silencing of Arabidopsis FLC. Curr Biol 17:73–78. https://doi.org/10.1010/j.cub.2006.11.052

Haider W (2014) Exploring flowering gene networks in soybean and arabidopsis through transcriptome analysis. University of Illinois at Urbana-Champaign

Han T, Wu C, Tong Z, Mentreddy RS, Tan K, Gai J (2006) Postflowering photoperiod regulates vegetative growth and reproductive development of soybean. Environ Exp Bot 55:120–129. https://doi.org/10.1016/j.envexpbot.2004.10.006

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423–447

Hu Z, Zhang D, Zhang G, Kan G, Hong D, Yu D (2014) Development and implementation of background linkage disequilibrium. Heredit (edinh) 99:374–380. https://doi.org/10.1038/sj.hdy.6801010

Jian SZTPW (1990) Preliminary study on the growth period structure of soybean varieties in the northeast of China. Soybean Science 3

Jung CH, Wong CE, Singh MB, Bhalla PL (2012) Comparative genomic analysis of soybean flowering genes. PLoS ONE 7:e38250. https://doi.org/10.1371/journal.pone.0038250

Kaeuffer R, Reale D, Coltman DW, Pontier D (2007) Detecting population structure using STRUCTURE software: effect of background linkage disequilibrium. Heredit (edinh) 99:374–380. https://doi.org/10.1038/sj.hdy.6801010

Kim KH et al (2020) Genome-wide association and epistatic interactions of flowering time in soybean cultivar. PloS one 15:e0228114
Kong LJ et al. (2010) The soybean stem growth habit gene Dt1 is an effector involved in regulation of seed germination. PLoS ONE 5:e99642. https://doi.org/10.1371/journal.pone.0099642

Li LY et al. (2018) Genome-wide association mapping of QTL underlying seed oil and protein contents of a diverse panel of soybean accessions. Plant Sci 266:95–101. https://doi.org/10.1016/j.plantsci.2017.04.013

Li YH et al. (2018) Identification of a soybean MOTHER OF FT AND TFL1 homolog involved in regulation of seed germination. PLoS ONE 9:e99642. https://doi.org/10.1371/journal.pone.0099642

Lin X, Liu B, Weller JL, Abe J, Kong F (2015) Targeted association mapping demonstrating the complex molecular genetics of fatty acid formation in soybean. BMC Genomics 16:841. https://doi.org/10.1186/s12864-015-2049-4

Liu B et al. (2010) The soybean stem growth habit gene Dt1 is an ortholog of Arabidopsis TERMINAL FLOWER1. Plant Physiol 153:198–210. https://doi.org/10.1111/j.1365-3109.2009.04067

Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. PLoS Genet 12:e1005767. https://doi.org/10.1371/journal.pgen.1005767

Liu X et al. (2017) Genetic variation of world soybean maturity date and geographic distribution of maturity groups. Breed Sci 67:221–232. https://doi.org/10.1270/jsbbs.16167

Lu S et al. (2017) Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. Nat Genet 49:773–779. https://doi.org/10.1038/ng.3819

Mao T et al. (2017) Association mapping of loci controlling genetic and environmental interaction of soybean flowering time under various photo-thermal conditions. BMC Genomics 18:415. https://doi.org/10.1186/s12864-017-3778-3

Marcolino-Gomes J et al. (2017) Functional characterization of a putative Glycine max ELF4 in transgenic Arabidopsis and its role during flowering control. Front Plant Sci 8:618. https://doi.org/10.3389/fpls.2017.00618

Matsumura H, Liu B, Abe J, Takahashi R (2008) AFLP mapping of soybean maturity gene E4. J Hered 99:193–197. https://doi.org/10.1093/jhered/esn114

Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, Lark KG (1999) Genetics of soybean agronomic traits: I. comparison of three related recombinant inbred populations. Crop Sci 39:1642–1651. https://doi.org/10.2135/cropsci1999.3961642x

Porcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19:149–150. https://doi.org/10.1093/bioinformatics/19.1.149

Qiu L, Chang R, Liu Z, Guan R, Li Y (2006) Descriptors and data standard for soybean (Glycine spp.). China Agriculture Press, Beijing

Qiu L-J, Li Y-H, Guan R-X, Liu Z-X, Wang L-X, Chang R-Z (2009) Establishment, representative testing and research progress of soybean core collection and mini core collection. Acta Agron Sin 35:571–579. https://doi.org/10.3724/SP.J.1006.2009.00571

Qiu LJ, Xing LL, Guo Y, Wang J, Jackson SA, Chang RZ (2013) A platform for soybean molecular breeding: the utilization of core collections for food security. Plant Mol Biol 83:41–50. https://doi.org/10.1007/s11103-013-0076-6

Ray JD, Hinson K, Mankono JEB, Malo MF (1995) Genetic control of a Long-Juvenile trait in soybean. Crop Sci 35:1001–1006. https://doi.org/10.2135/cropsic1995.0011183X002900060021x

Saindogn, Beversdorff WD, Voldeng HD (1989a) Adjustment of the soybean phenology using the E 4 locus. Crop Sci 29:1361–1365. https://doi.org/10.2135/cropsci1989.0011183X002900060021x

Saindogn, Voldeng HD, Beversdorff WD, Buzzell RI (1989b) Genetic control of long daylength response in soybean. Crop Sci 29:1346–1349. https://doi.org/10.2135/cropsci1989.0011183X002900060021x

Samanfar B et al. (2017) Mapping and identification of a potential candidate gene for a novel maturity locus, E10, in soybean. Theor Appl Genet 130:377–390. https://doi.org/10.1007/s00122-016-2819-7

Sedivy EY, Wu F, Hanzawa Y (2017) Soybean domestication: the origin, genetic architecture and molecular bases. New Phytol 214:539–553. https://doi.org/10.1111/nph.14418

Shu Y, Yu D, Wang D, Guo D, Guo C (2013) Genome-wide survey and expression analysis of the MADS-box gene family in soybean. Mol Biol Rep 40:3901–3911. https://doi.org/10.1007/s11033-012-2438-6

Si L et al. (2016) OsSPL13 controls grain size in cultivated rice. Nat Genet 48:447–456. https://doi.org/10.1038/ng.3518

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

Sung S, Schmitz RJ, Amasino RM (2006) A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis. Genes Dev 20:3244–3248. https://doi.org/10.1103/gad.1493306

Si L et al. (2016) OsSPL13 controls grain size in cultivated rice. Nat Genet 48:447–456. https://doi.org/10.1038/ng.3518

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

Sung S, Schmitz RJ, Amasino RM (2006) A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis. Genes Dev 20:3244–3248. https://doi.org/10.1103/gad.1493306

Suo H et al. (2016) The AtDREB1A transcription factor up-regulates expression of a vernalization pathway gene, GmVRN1-like, delaying flowering in soybean. Acta Physiol Planta 38:137

Takeshima R et al. (2016) A soybean quantitative trait locus that promotes flowering under long days is identified as FT3a, a FLOWERING LOCUS T ortholog. J Exp Bot 67:5247–5258. https://doi.org/10.1093/jxb/erw283
Tsubokura Y et al (2014) Natural variation in the genes responsible for maturity loci $E_1$, $E_2$, $E_3$ and $E_4$ in soybean. Ann Bot 113:429–441. https://doi.org/10.1093/aob/mct269

Wang X et al (2016) Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings. Nat Genet 48:1233–1241. https://doi.org/10.1038/ng.3636

Wang Y, Cheng LR, Leng JT, Wu CX, Shao GH, Hou WS, Han TF (2015a) Genetic analysis and quantitative trait locus identification of the reproductive to vegetative growth period ratio in soybean (Glycine max (L.) Merr.). Euphytica 201:275–284. https://doi.org/10.1007/s10681-014-1209-y

Wang Y, Wu C-X, Zhang X-M, Wang Y-P, Han T-F (2008) Effects of soybean major maturity genes under different photoperiods. Acta Agron Sin 34:1160–1168

Wang Z et al (2015b) Functional evolution of phosphatidylethanolamine binding proteins in soybean and Arabidopsis. Plant Cell 27:323–336. https://doi.org/10.1105/tpc.114.135103

Watanabe S, Harada K, Abe J (2012) Genetic and molecular bases of photoperiod responses of flowering in soybean. Breed Sci 61:531–543. https://doi.org/10.1270/jsbsbs.61.531

Watanabe S et al (2009) Map-based cloning of the gene associated with the soybean maturity locus $E_3$. Genetics 182:1251–1262. https://doi.org/10.1534/genetics.108.098772

Watanabe S et al (2011) A map-based cloning strategy employing a residual heterozygous line reveals that the GIGANTEA gene is involved in soybean maturity and flowering. Genetics 188:395–407. https://doi.org/10.1534/genetics.110.125062

Xia Z et al (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus $E_1$ that regulates photoperiodic flowering. Proc Natl Acad Sci U S A 109:E2155–2164. https://doi.org/10.1073/pnas.1117982109

Xue WY et al (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. Nat Genet 40:761–767. https://doi.org/10.1038/ng.143

Yamamoto Y et al (2003) Comparative genetic studies on the APRR5 and APRR7 genes belonging to the APRR1/TOC1 quintet implicated in circadian rhythm, control of flowering time, and early photomorphogenesis. Plant Cell Physiol 44:1119–1130. https://doi.org/10.1093/pcp/pcg148

Yang G et al (2017) QTL effects and epistatic interaction for flowering time and branch number in a soybean mapping population of Japanese×Chinese cultivars. J Integr Agric 16:1900–1912

Yang Y, Gai J, Ma Y (1994) Differential inheritance of crop period traits under different planting seasons in soybean. Zhongguo Nongye Kexue 27:1–6

Yue Y et al (2017) A Single nucleotide deletion in J encoding GmELF3 confers long juvenility and is associated with adaption of tropic soybean. Mol Plant 10:656–658. https://doi.org/10.1016/j.molp.2016.12.004

Zeng X et al (2018) Soybean MADS-box gene GmAGL1 promotes flowering via the photoperiod pathway. BMC Genomics 19:51. https://doi.org/10.1186/s12864-017-4402-2

Zhang J, Song Q, Cregan PB, Nelson RL, Wang X, Wu J, Jiang GL (2015) Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (Glycine max) germplasm. BMC Genomics 16:217. https://doi.org/10.1186/s12864-015-1441-4

Zhang X et al (2018) Genome-wide association studies of soybean seed hardness in the Chinese Mini Core Collection. Plant Mol Biol Report 36:605–617. https://doi.org/10.1007/s11105-018-1102-2

Zhao C et al (2016) A recessive allele for delayed flowering at the soybean maturity locus $E_9$ is a leaky allele of FT2a, a FLOWERING LOCUS T ortholog. BMC Plant Biol 16:20. https://doi.org/10.1186/s12870-016-0704-9

Zhou L et al (2016) Identification and validation of candidate genes associated with domesticated and improved traits in soybean. Plant Genome. https://doi.org/10.3835/plantgenome2015.09.0090

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.