

Integrating GWAS, linkage mapping and gene expression analyses reveals the genetic control of growth period traits in rapeseed (*Brassica napus* L.)

tengyue wang

Southwest University

Lijuan Wei

Southwest University

Jia wang

Southwest University

Ling Xie

Southwest University

Yangyang Li

Southwest University

Shuyao Ran

Southwest University

Lanyang Ren

Southwest University

Kun Lu

Southwest University

Jiana Li

Southwest University

Michael P. Timko

University of Virginia

Liezhao Liu (✉ liezhao@swu.edu.cn)

Research

Keywords: Brassica napus, growth period traits, GWAS, linkage mapping, RNA sequencing

Posted Date: June 4th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-32662/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on August 3rd, 2020. See the published version at <https://doi.org/10.1186/s13068-020-01774-0>.

Abstract

Background: *Brassica napus* is one of the most important oilseed crops, and also an important biofuel plant due to its low air pollution and renewability. Growth period are important traits that affect yield and are crucial for its adaptation to different environments in *B. napus*.

Results: To elucidate the genetic basis of growth period traits, genome-wide association analysis (GWAS) and linkage mapping were employed to detect the quantitative trait loci (QTL) for days to initial flowering (DIF), days to final flowering (DFF), days from initial flowering to final flowering (FP), days from final flowering to maturity (MT), and whole growth period (GP). A total of 146 SNPs were identified by association mapping, and 83 QTLs were identified by linkage mapping using the RIL population. Among these QTLs, 19 were pleiotropic SNPs related to multiple traits, and six (q18DFF.A03-2, q18MT.A03-2, q17DFF.A05-1, q18FP.C04, q17DIF.C05 and q17GP.C09) were consistently detected using both mapping methods. Additionally, we performed RNA sequencing to analyze the differential expression of gene (DEG) transcripts between early- and late-flowering lines selected from the RIL population, and the DEGs were integrated with association mapping and linkage analysis to confirm their roles in the growth period. Consequently, twelve candidate genes associated with growth period traits were identified in *B. napus*. Among these genes, seven have polymorphic sites in the coding sequence and the upstream 2-kb sequence based on the resequencing data. The haplotype BnaSOC1.A05-HapIb and BnaLNK2.C06-HapIa showed more favorable phenotypic traits.

Conclusions: The candidate genes identified in this study will contribute to our genetic understanding of growth period traits and can be used as targets for target mutations or marker assisted breeding for rapeseed adapted to different environments.

Introduction

Brassica napus is one of the most important oilseed crops and an important source of protein-rich livestock feed in the world. At the same time, due to the energy crisis, low air pollution and renewability of rapeseed oil, an increasing number of people regard rapeseed oil as an ideal green energy source. This species is an allopolyploid (AACC) that evolved from an interspecies cross between *Brassica rapa* (genome AA, $n = 10$) and *Brassica oleracea* (genome CC, $n = 9$) approximately 7500 years ago [1]. The species originated in Europe and spread worldwide (to East Asia, Australia, and North America). The growth period of rapeseed can be divided into vegetative growth period (the duration from seedling to flowering), and reproductive growth period (the duration from flowering to maturity). Flowering is a major trait in the plant's growth period, as it represents the transition from the vegetative stage to the reproductive stage. Flowering time plays an indispensable role in adaptation to specific environments. It has been reported that the growth stages show high correlations with plant height, pod number and grain yield [2]. Therefore, understanding the genetic network underlying growth period traits provides a theoretical and practical basis for developing new cultivars adapted to different geographical environments.

Linkage mapping based on bi-parental populations has been widely applied to detect QTLs for flowering time traits in *B. napus* [3–9]. Long et al. (2007) identified 36 significant level (SL-QTL) and 6 micro real QTL (MR-QTL) in the doubled haploid (DH) population (Tapidor Ningyou7) and its derived F2 population. Raman et al. (2013) identified at least 20 loci localized on ten different chromosomes in a DH population (Skipton Ag-Spectrum). Raman et al. (2014) constructed an integrated genetic linkage map of the SASDH population and up to 40.2% of phenotypic variation for flowering time was identified. Liu et al. (2016) found 22 QTLs for days to flowering (DIF) in the spring-type DH population in multiple environments.

Genome-wide association studies (GWAS) using natural populations provide higher mapping resolution than linkage mapping based on bi-parental segregating populations and have greater cost-effectiveness. Based on the GWAS website (<https://bigd.big.ac.cn/gwas/>), approximately 8373 associations related to nineteen growth period traits have been identified across six crops (cotton, maize, rapeseed, rice, sorghum, and soybean) using GWAS technology [10]. Among these associations, 198 associations for seven growth period traits were identified in rapeseed. Xu et al. (2015) found 41 SNPs correlated with flowering time with a diversity panel comprising 523 *B. napus* cultivars, and 12 SNPs were consistent with previously identified QTLs [11]. Niklas et al. (2016) found six QTLs for beginning of flowering (BOF) and one QTL for end of flowering (EOF) in 405 *B. napus* inbred lines [12]. Zhou et al. (2017) identified 131 SNPs strongly linked to four earliness traits (initial flowering day, maturity time, final flowering day and flowering period), 40 of which fell into or were physically close to published flowering time SNPs [13]. Wei et al. (2017) detected 12 SNPs significantly associated with flowering time using 3,318,6 high-quality SNPs from 327 accessions [14]. From the above results, it can be seen that these studies focused mainly on the initial flowering time. Other growth period traits, such as final flowering stage (DFF), flowering period (FP) and maturity time (MT), have rarely been reported by mapping methods.

RNA sequencing has been widely applied to examine differences in global gene expression with the benefit of high sensitivity and is more cost-effective than microarray analysis. However, transcriptome analysis often obtained a large number of differentially expressed genes (DEGs) between samples. Therefore, the integration of DEGs with GWAS or linkage mapping has been considered to be an effective way to identify candidate genes related to complex traits [15–18].

In this study, using a panel of 588 accessions, we carried out a GWAS for five growth period traits, the number of days to initial flowering (DIF), the number of days to final flowering (DFF), the flowering period (FP), the number of days to maturity (DM), and the growth period (GP). Furthermore, the SNPs associated with DIF, DFF, FP, MT and GP were detected in the recombinant inbred line (RIL) population. In addition, we performed transcriptomic analysis of the leaves of early-flowering and late-flowering accessions at different developmental stages (vegetative and reproductive). The main objectives of our study are: (1) to dissect the genetic mechanism of growth period traits by GWAS using a natural population and QTL mapping using an RIL population derived from GH06 and P174; (2) to dissect the molecular mechanism of floral development by characterizing the DEGs of leaves for early-flowering and late-flowering lines using RNA sequencing technology; and (3) integrating the QTLs and DEGs to identify candidate genes that control growth period traits for further verification.

Results

Phenotypic variations, correlations, and ANOVA for growth period traits

Five growth period traits (DIF, DFF, FP, MT and GP) of 588 rapeseed lines were investigated in three years. Significant variation among the genotypes for five growth period traits was observed: for instance, DIF ranged from 138 d to 207 d with an average of 157.98 d in 2017, from 149 d to 174 d with an average of 157.34 d in 2018, and from 137 d to 189 d with an average of 151.87 d in 2019 (Table 1). The coefficient variation (CV) of FP and MT was the highest (over 10%) followed by DIF and DFF, and GP was the lowest. In addition, the phenotypic frequency distributions of five traits in the three years all showed approximately continuous and normal distributions, which showed that the group was suitable for association analysis (Fig. 1). To test the effects of genotype (G), environment (E) and their interactions (GE), we conducted analysis of variance (ANOVA) for each trait (Table S1a). Significant variations were observed among environments and genotypes for all five traits. The DFF has the highest broad-sense heritability (h^2) of 92.66%, the FP has relatively low heritability (74.11%). Overall, these findings indicated that all five traits were stably inherited.

To determine whether there are any relationships between five growth period traits in *B. napus*, the average value of three years for each trait was used for correlation analysis (Table 2). DIF showed a significant positive correlation with DFF and GP but was negatively correlated with FP and MT and DFF. DFF was significantly correlated with FP, MT, and GP with correlation coefficients of 0.148, -0.664 and 0.760, respectively. These results are consistent with a previous study reported by Zhou et al. [13]. It can be concluded that a shorter flowering time corresponds to a shorter growth period and a longer maturity time.

Genome-wide association analysis for growth period traits in three years

Through a set of processes of library construction, paired-end sequencing and SNP calling, a series of 3,856,91 highly consistent and locus-specific SNPs (minor allele frequency > 0.05 and call frequencies > 0.9) were retained for the following analysis. Fig. S1 shows the density distribution of SNP markers on different chromosomes. The 3,856,91 SNPs were covered and unevenly distributed on all 19 *B. napus* chromosomes.

To avoid false negative associations, three general linear models (GLM), naïve, PCA and Q models, and three mixed linear models (MLM), K, Q + K and PCA+ K models were chosen to evaluate the effects of population structure (Q, PC) and relative kinship (K). According to the Q-Q plots of the six models, the MLM model can control false positives well for each trait in three years (Fig. S2). To minimize the effect of environmental variation, best linear unbiased predictor (BLUP) value for each line were also calculated

for each trait using the R package lme4 [19]. Therefore, the MLM models were conducted for GWAS of five growth period traits with three years' data and BLUP values. We mainly focused on the QTLs detected in at least two environments. In total, 146 SNP loci significantly associated ($-\log_{10}(P) \geq 5.58$) with five growth-related traits were identified, including 13 related to DIF, 23 to DFF, 6 to FP, 9 to MT and 85 to GP for the whole panel of accessions (Table S2 and Fig. 2).

Identification of candidate genes for growth period traits in *B. napus*

Among the 146 SNP loci significantly associated with growth period traits, 60 SNPs were divided into 19 genomic regions using a haplotype block estimation, with sizes ranging from 7 bp to 270.31 kb, while the remaining 86 SNP loci were not present in the LD blocks (Table S3 and Fig. S3). Then, we obtained the genes within the same LD block or within 300 kb to either side of the significant SNPs using *B. napus* 'Darmor v4.1' as the reference genome. Finally, a total of 101 candidate genes were identified as orthologous to Arabidopsis flowering genes reported in the Flowering Interactive Database (FLOR-ID), most of which were involved in six flowering pathways of aging, autonomous pathway, vernalization, photoperiod, GA, and circadian clock; other genes functioned as flower development and meristem identity and flowering time integrator (Table S4). The flowering time integrators *FLOWERING LOCUS C* (*FLC*) control the transition from vegetative to reproductive meristem by integrating the signals from six pathways and then precisely regulating the expression of specific flower meristem identity genes *APETALA2* (*AP2*) and *FRUITFULL* (*FUL*) [20].

By comparing the SNP regions, 28 SNP loci (within 300 kb) associated with growth period traits, were located in or near the QTL regions identified in previous studies: 23 were reported by Raman et al., 7 were reported by Wang et al., and 4 were reported by Zhou et al. [3, 13, 21, 22] (Table S3 and Fig. 3). Among these loci, S1_5075025, S2_5719334 and S15_5922896 were reported in 2 studies, and S3_13708544 and S7_10803897 were detected simultaneously in three studies. At the same time, the candidate genes we identified in this study were also reported in previous studies [13, 21–27]. The flowering time integrator (*BnaFLC.C02*, *BnaFUL.A03* and *BnaSVP.A09*) were reported in three or more previous studies (Table S3). Overall, the comparison of QTLs and candidate genes further validates the reliability of the GWAS mapping and help us mine the novel QTL.

QTL linkage mapping of growth period traits in two years

The RIL population of GH06 P174 grown in 2017 and 2018 was evaluated for DIF, DFF, FP, MT and GP phenotypes and then used for QTL analysis. The flowering time of GH06 was later than that of P174 in the two investigated environments. A wide range of variation (Table 3) and the normal phenotypic distribution for the growth period traits (Fig. S4) were observed in the RIL population, indicating a

quantitative inheritance suitable for QTL mapping. ANOVA was performed on the phenotypic data from two years, and h^2 ranged from 49% (DFF) to 53% (GP) (Table S1b).

A total of 17, 25, 7, 21 and 13 QTLs for DIF, DFF, FP, MT and GP were detected in two years and were located on all *B. napus* chromosomes except C02 (Table S5 and Fig. S5). These QTLs for DIF, DFF, FP, MT and GP explained ~17.48%, ~15.92%, ~7.95%, ~18.30% and ~9.43% of the phenotypic variance, and the additive effect varied from -0.43 to 2.30, -1.12 to 1.79, -0.83 to 0.71, -2.09 to 1.13, and from -0.70 to 1.10, respectively. The QTLs have the overlapped confidence intervals with the same direction of additive effect are considered to be the same QTL. Among the 83 QTLs, 14 QTLs were associated with at least two traits. The QTL q18DIF.A01-2, q18DFF.A01 and q18MT.A01-1 have the same QTL region and explained the highest total phenotypic variation of each trait. However, the additive values of DIF and DFF were positive, and the additive value of MT was negative, which indicates that the QTL from GH06 increased DIF and DFF, while the QTL from P174 increased MT. The QTL q17DIF.A06-1, q17DFF.A06-2, q17MT.A06-3 and the QTL q17DIF.A07-2, q17DFF.A07-2, q17FP.A07, q17MT.A07-3 also has the same region and opposite additive effect. These results indicate that the same QTL may have opposite effects on the different traits. Based on the rapeseed genome annotation and the physical locations of these QTL regions, a total of 75 Arabidopsis flowering homolog genes were identified (Table S6), and their location relative to the adjacent SNP markers is shown in Fig. 4.

Identification of DEGs for floral transition and flower development by RNA-sequencing

To identify candidate DEG transcripts controlling flowering time, we sequenced four RNA samples from leaves of early-flowering cultivar 18Z134 and late-flowering cultivar 18Z88 sampled at vegetative and reproductive development stages (EV, ER, LV and LR). Every sample had two replicates. 18Z134 and 18Z88 were selected from the above RIL population based on six years of field data. Their flowering time differences are shown in Table S7. The total reads, mapped and unique mapped reads to the reference *B. napus* genome are shown in Table S8. After removing low-quality sequences, 34,862,287 (86.32%) ~ 37,786,256 (90.79%) clean reads were successfully mapped to the genome using TopHat. Of these clean reads, 33,089,439 (81.93%) ~ 38,684,877 (85.93%) were uniquely mapped.

Two types of comparisons were performed: (1) to identify DEGs between the two extreme lines at each development stage: LR vs. ER and LV vs. EV; (2) to identify expression changes between different stages in each line: EV vs. ER and LV vs. LR. The false discovery rate (FDR) ≤ 0.05 and absolute value of $|\log_2(\text{fold change})| \geq 1$ were used as thresholds to judge the DEGs between the two groups. In total, 3727 DEGs were identified in LR vs. ER (2421 upregulated, 1306 downregulated), 2327 DEGs were identified in LV vs. EV (1557 upregulated, 770 downregulated), 3750 DEGs were identified in EV vs. ER (2135 upregulated, 1615 downregulated), and 2038 DEGs were identified in LV vs. LR (1281 upregulated, 757 downregulated). In addition, 1662 and 683 DEGs were common to LR vs. ER/LV vs. EV and EV vs. ER/LV vs. LR, respectively (Fig. S6 and Table S9).

GO and KEGG analysis of differentially expressed genes

As the first criterion, we analyzed DEGs between vegetative and reproductive stages of development (EV vs. ER and LV vs. LR) to identify the key phase transition-associated genes. The 683 common DEGs between EV vs. ER and LV vs. LR were subjected to an enrichment analysis for GO annotation terms. The top 20 significantly enriched GO terms are shown in Fig. S7a and Table S10a. Among these terms, the salicylic acid biosynthetic process (GO:0009697), response to cold (GO:0009409), negative regulation of floral organ abscission (GO:0060862), intracellular auxin transport (GO:0080162) and long-day photoperiodism, flowering (GO:0048574) were involved in the phase transition process.

To further determine the metabolic pathways in the phase transition process, we performed KEGG enrichment analysis (Fig. S7b and Table S10b). Circadian rhythm–plant (ko04712) and plant hormone signal transduction (ko04075) were significantly enriched and participate in plant development and flowering processes. As the second criterion, we looked for DEGs between early-flowering and late-flowering lines that result in flowering time variation. The 2327 DEGs between LV vs. EV and 3727 DEGs between LR vs. ER were subjected to GO and KEGG enrichment analysis (Fig. S8). DEGs in the circadian rhythm were enriched in both GO and KEGG analysis between LR vs. ER, suggesting that the different expression of genes involved in circadian rhythm could result in flowering differences.

Identification of TFs involved in floral transition and flower development

Transcription factors (SBP, AP2, MADS) play a central role in flowering development [20]. The MADS-box transcription factor is a major group of regulators controlling floral transition, floral organ specification and floral development in flowering plants [28]. In total, 50 TF-encoding DEGs, including the basic helix-loop-helix (bHLH, five members), ERF (eight members), MIKC_MADS (eight members), NAC (three members) and WRKY (four members), were identified in the early and late-flowering lines at two developmental stages. Fig. S9 shows the overall expression trend in the four samples, and most TF families were significantly upregulated in the reproductive stages.

Identification of DEGs involved in flowering-related hormone pathways

Plant hormones, such as auxin, cytokinin (CK), abscisic acid (ABA), gibberellic acid (GA), salicylic acid (SA), and jasmonic acid (JA), are involved in the floral transition process and plant flowering [29–31]. The effects of GA on flowering have been extensively studied in Arabidopsis [32, 33]. In total, 20 genes in the ABA signaling pathway, 23 genes in the auxin signaling pathway, five genes in the GA signaling pathway, five genes in the cytokinin signaling pathway, four genes in the JA signaling pathway and five genes in the SA signaling pathway were identified in our transcriptome data (Fig. S9b).

Identification of floral transition- and flowerdevelopment-related genes

According to the annotation of unigenes in Arabidopsis, a total of 125 DEGs related to flowering time were identified in the comparisons LR vs. ER, LV vs. EV, EV vs. ER and LV vs. LR (Table S11). These genes mainly included the photoperiod, circadian rhythms, vernalization, GA signaling, aging, flowering time integrator and flower meristem identity genes, and the expression value of these genes in the four samples is shown in Fig. 5.

In the photoperiod pathway, 36 unigenes homologous to *EARLY FLOWERING 3 (ELF3)*, *EARLY FLOWERING 4 (ELF4)*, *CONSTANS-like 9 (COL9)*, *cycling DOF factor 1 (CDF1)* and *LATE ELONGATED HYPOCOTYL (LHY)* were identified; the circadian rhythm, as an internal timekeeper, controls daily and seasonal changes, is an important part of the photoperiod pathway and plays a key role in controlling plant flowering [34]. *Circadian clock associated 1 (CCA1)* is a key component of the Arabidopsis circadian oscillator, and it interacts with *LATE ELONGATED HYPOCOTYL (LHY)* and *TIMING OF CAB EXPRESSION 1 (TOC1)* to inhibit transcription of the Evening Complex (EC) proteins ELF4 and ELF3. In the circadian clock pathway, two *CCA1* genes, three *TOC1* genes, two *pseudo response regulator 7 (PRR7)* genes and three *pseudo response regulator 9 (PRR9)* genes were included; For the aging pathway, eight unigenes, including *SQUAMOSA PROMOTER-BINDING-LIKE PROTEIN 4 (SPL4)*, *SQUAMOSA PROMOTER-BINDING-LIKE PROTEIN 5 (SPL5)*, *SQUAMOSA PROMOTER-BINDING-LIKE PROTEIN 15 (SPL15)* and *TARGET OF EARLY ACTIVATION TAGGED 1 (TOE1)*, were found. Seven homologous genes of the GA signaling pathway were also identified, including *GA2 oxidase (GA2ox, three unigenes)*, *GA2 oxidase 1 (GA2ox1)*, *gibberellin 20-oxidase 3 (GA20OX3)*, *DELLA protein RGA-like 2 (RGL2)*, and the GA receptor *GA INSENSITIVE DWARF1A (GID1A)*; Additionally, five unigenes were annotated in the vernalization pathway, which included the *AGAMOUS-like 19 (AGL19)*, *INDUCER OF CBF EXPRESSION 1 (ICE1)*, *vernalization5/VIN3-like (VEL1)* and *REDUCED VERNALIZATION RESPONSE 1 (VRN1)*. Furthermore, ten floral pathway integrator genes related to *FLC*, *FLOWERING LOCUS T (FT)*, *AGAMOUS-like 20 (AGL20)*, *SHORT VEGETATIVE PHASE (SVP)* and *TWIN SISTER OF FT (TSF)* and ten flowering meristem-identifying genes, such as *AGAMOUS-like 8 (AGL8)*, *AGAMOUS-like 14 (AGL14)*, *AGAMOUS-like 24 (AGL24)* and *APETALA 2 (AP2)*, were all identified in our transcriptome database. All these DEGs are important resources for the further study of floral transition and floral development in *B. napus*.

Identification of candidates for growth period traits by integrating QTLs with DEGs

To further understand the roles of these DEGs in regulating floral transition and flower development, the DEGs were integrated with the significant QTLs identified in either association analysis or linkage mapping. Therefore, the DEGs were considered candidate genes if they were located within the confidence interval (CI) of the QTLs identified by GWAS, linkage mapping, or both. According to the above

criteria, a total of twelve DEGs located in the CI of sixteen significant loci were identified as candidate genes of growth period traits. The loci were *BnaC04g15640D*, *BnaA03g40160D*, *BnaC06g15270D*, *BnaC09g05250D*, *BnaA04g22640D*, *BnaA05g05560D*, *BnaA05g05000D*, *BnaC09g48370D*, *BnaC02g36310D*, *BnaC09g23670D*, *BnaA03g39820D* and *BnaA05g05010D*. The twelve DEGs were known to regulate floral development by affecting aging, photoperiod/circadian clock, GA, vernalization, flower development and meristem identity and flowering time integrator (Table 4 and Fig. 6). *BnaAGL6.A05*, which encodes a MADS-box transcription factor, negatively regulates the *FLC/MAF* clade genes and positively regulates *FT* in Arabidopsis [35]. Among them, *BnaC06g15270D* (*BnaLNK2.C06*), *BnaC09g48370D* (*BnaGA200X3.C09*) and *BnaA03g39820D* (*BnaFUL.A03*) were also identified as EDGs between BBCH20 and BBCH50 (vernalized and nonvernalized) associated with flowering time and yield QTLs [26]. In addition, an ortholog of the circadian clock pathway, *BnaTOC1.A03* and floral meristem identity gene *BnaFUL.A03* were identified in both GWAS and linkage analysis[36]. Thus, we considered these five genes to be our most promising candidate genes for future prospects.

To analyze the polymorphism and explore their relationship with five growth traits, the sequence variation was analyzed in the genomic sequence including the complete coding sequence and 2000 bp upstream of the ATG translational start codon in the twelve candidate genes in the 558 lines based on our resequencing data. Among twelve genes, seven genes had polymorphic sites: one, fifteen, four, five, one, three and ten SNPs were identified in the *BnaFUL.A03*, *BnaELF4.A04*, *BnaSOC1.A05*, *BnaELF4.A05*, *BnaLNK2.C06*, *BnaTOC1.C09* and *BnaGA200X3.C09*, respectively, and formed different types of haplotypes. Detailed haplotype information is listed in Table S12.

The association of each haplotype with five growth periods was then analyzed in the GWAS population. We compared the phenotypic variations of different haplotypes for the above agronomic traits (Fig. 7). In general, for *BnaELF.A04*, accessions with *BnaELF.A04*-Haplc accounted for 94.2% and showed a significantly shorter MT period in 2018 and 2019 compared to *BnaELF.A04*-Hapla. For *BnaELF.A05*, varieties with *BnaELF.A05*-Hapla accounted for 90.8% and showed the shortest FP compared to the other two haplotypes in 2017 and 2019. For *BnaSOC1.A05*, most accessions have *BnaSOC1.A05*-Haplb (97.5%) and exhibited shorter DFF and longer MT over three years compared to the Hapla. For *BnaKNK2.C06*, most varieties have *BnaKNK2.C06*-Hapla (93.7%) and showed lower DFF and GP than Haplb. For *BnaTOC1.C09*, accessions with *BnaTOC1.C09*-Haplc accounted for 94.1% and had lower DFF, lower FP and longer MT over three years. For *BnaGA200X3.C09*, the two haplotypes did not exhibit significant differences between the five growth period traits. Zhou et al. (2017) reported that SNP alleles with positive effects on decreases in the values of DIF and DFF or on increases in FP were defined as 'favorable alleles' for earliness. Thus, the haplotype *BnaSOC1.A05*-Haplb and *BnaLNK2.C06*-Hapla showed more favorable phenotypic traits and may be used for further earliness molecular breeding.

Confirmation of candidate gene expression using qRT-PCR

To verify the accuracy and reproducibility of the transcriptome analysis, nine candidate genes listed in Table 6 were randomly selected for qRT-PCR analysis. As shown in Fig. 8, the expression of nine genes were consistent with the RNA-Seq results between four samples. These results demonstrated the reliability of the RNA-sequencing results.

Discussion

Growth period are important traits that can influence crop yield and quality. Identifying the genetic loci for growth period traits may help to elucidate the genetic basis underlying the growth period and be valuable for developing cultivars adapted to different geographical regions. Linkage analysis and association mapping are the two most common methods used for mapping complex traits. Linkage mapping can exclude false positives of associated loci caused by high linkage disequilibrium, and association mapping can be used to narrow the confidence interval of linkage analysis when its QTL regions are large. The combination of two mapping methods further improves mapping efficiency and accuracy [37–39]. In the current study, we detected the SNP loci associated with five growth period traits (DIF, DFF, FP, MT, GP) via GWAS in the natural population in three-year environments and detected QTLs that influence growth in a RIL population over two years in *B. napus*.

A total of 146 SNP loci were found to be associated with growth period traits (13 related to DIF, 23 to DFF, 6 to FP, 9 to MT and 85 to GP), which were detected in at least two environments (including BLUP). Among these SNP loci, 28 were consistent with at least one QTL identified in one or more previous studies. In addition, 101 candidate genes were identified as flowering orthologous genes based on the LD block or 300 kb to either side of the significant SNPs, most of which were also reported in previous studies. In particular, the flowering time integrator and floral meristem identify genes (*BnaFLC.C2*, *BnaFUL.A3*, and *BnaSVPA9*) were reported in at least three studies [13, 21–27]. These results indicated that the results in the present study are notably reliable. A total of 83 QTLs were observed relating to DIF, DFF, FP, MT and GP by linkage mapping in two years. These QTLs could explain over 10% of the phenotypic variance for DIF, DFF, FP, MT and GP. However, most QTLs were only detected in a single-year environment, which indicates that the rapeseed growth period was largely affected by environmental variation. Finally, six QTLs were associated with growth period traits using both association and linkage mapping (Table S13). S19_17595451 and q17GP.C09 were associated with GP. However, the other five QTLs with overlapping confidence intervals were related to different traits in linkage analysis and association analysis. The results were also reported in Zuo et al [40].

SNP markers with pleiotropic effects

Some GWAS and linkage mapping studies have indicated that one locus can control multiple highly correlated traits [38, 41]. We detected significant correlations between the five growth period traits (Table 2), therefore, it is possible that some SNPs can affect these traits simultaneously. Five QTLs found in the GWAS and fourteen QTLs identified in the linkage mapping were co-associated with at least two growth

period traits, which coincided with significant phenotypic correlations among the traits (Table S14a, b). It has been reported that co-mapping of QTLs for correlated traits may result from either tight linkage of multiple genes [42] or single gene pleiotropy [43]. In our research, we found that a single gene can control multiple traits (Fig. S10). This finding further confirmed the above explanation. The co-mapping of QTLs can help breeders identifying favorable alleles for multiple traits simultaneously in marker-assisted breeding[44].

Identification of flowering DEGs

A KEGG pathway enrichment analysis indicated that the common DEGs between EV vs. ER and LV vs. LR were enriched for circadian rhythm—plant and plant hormone signal transduction (Fig. S7), indicating that plant hormones play an important role in the *B. napus* floral transition, which was also reported in the sugar apple and *Rosa chinensis* [45, 46]. We also observed significant enrichment for genes in the circadian rhythm between LR vs. ER, suggesting that the difference in flowering time between early- and late-flowering lines was primarily due to the expression differences of circadian rhythm genes (Fig. S8).

IFs have been reported to play crucial roles in the reproductive development of flowering plants [47]. In the present study, bHLH, ERF, MIKC_MADS, NAC and WRKY IFS were identified. Most IF genes were upregulated during the floral transition of *B. napus* (Fig. S9), suggesting their pivotal roles in the induction of floral transition. Some TFs are also involved in reproductive development in other species [48–50].

In this study, we also detected 125 flowering and flower development-related homologous genes based on BLAST analysis using sequences of 306 known flowering genes in Arabidopsis. Some of these genes are related to five major flowering pathways, including circadian rhythms/photoperiod, vernalization, GA signaling and aging, while others encode regulators functioning as flowering integrators or floral meristem identification (Fig. 5 and Table S11) [51].

Candidate genes for growth period traits

The integration of QTL mapping and transcriptome sequencing is a highly useful and popular strategy for discovering candidate genes of complex traits [17]. By integrating DEGs from RNA-seq with GWAS and linkage mapping, twelve genes mainly related to circadian clock, hormone, flowering time integrator and flower meristem identity were identified as candidate genes regulating growth period traits in *B. napus*.

Of the twelve candidate genes detected in this study, five genes are involved in the circadian clock/photoperiod pathways: *BnaA03g40160D* and *BnaC09g05250D* are homologous to *TOC1* and interact with *LHY*, and *CCA1* forms the negative feedback loop of the circadian clock [34]; *BnaA04g22640D* and *BnaA05g05560D* are homologous to *EARLY FLOWERING 4 (ELF4)*, which is a part of a corepressor complex consisting of *ELF4*, *ELF3*, and *LUX* involved in the transcriptional regulation of

APRR9 and *elf4* mutations result in early flowering in non-inductive photoperiods[52]; *BnaC06g15270D* is homologous to *NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED 2 (LNK2)*, *LNK1* and *LNK2* integrate early light signals with core oscillator components (*TOC1*, *PRR*) to keep track of seasonal changes in day length[53]. The up-regulation of *TOC1* and *ELF4* in the late-flowering lines (LR) was in accordance with the mutant phenotype in Arabidopsis; *BnaC02g36310D* homologous to *ICE1* is involved in vernalization pathways. Cold-activated *ICE1* directly induces the expression of *FLC*, which represses *SOC1* expression, resulting in delayed flowering [54]; *BnaC09g48370D* is homologous to *GA20OX3*, which encodes a gibberellin 20-oxidase involved in the GA (gibberellin) pathway. Both *GA20ox* and *GA 2-oxidase* genes are regulated by several transcription factors, including *SHORT VEGETATIVE PHASE (SVP)* and *SOC1*[55–57]; *BnaC04g15640D* encodes a gene homologous to *TOE1*, which is a type of *AP2* transcription factor and the target gene of miR172. *TOE1* and its microRNA regulator miR172 play an important role in plant growth and phase change in different crops [58–60]; *BnaA05g05000D* is homologous to *AGL6* (encoding a MADS-box transcription factor). *AGL6* regulates flowering time through control the transcription of two key regulators of flowering time: *FLOWERING LOCUS C (FLC)* and *FT*. The *agl6-1D* mutant, in which *AGL6* was activated by the 35S enhancer, showed early flowering under both LD and SD conditions [61]; The floral integrator genes play important roles in activating floral meristem formation genes to induce flowering. *BnaA05g05010D* was identified as a *SOC1* paralog, which is a floral activator and integrates signals from the photoperiod, vernalization, and autonomous pathways in Arabidopsis [62, 63]. Overexpression of *SOC1* could suppress the late flowering and delayed phase transitions during the vegetative stages [64]; *BnaA03g39820D* was homologous to *FRUITFULL (FUL)* of Arabidopsis, a MADS-box transcription factor negatively regulated by *APETALA1(AP1)*, which mediates the vegetative and meristem identity transitions by forming *FUL-SVP* and *FUL-SOC1* heterodimers. *FUL* has been reported to connect several flowering pathways as a downstream target gene of age, photoperiod, and ambient temperature pathways [36]; *BnaC09g23670D* is homologous to *AGL14*, and *AGL14* promotes flowering transition and participates in flower meristem maintenance and determinacy by positively regulating *TERMINAL FLOWER 1 (TFL1)* expression [65, 66].

In conclusion, we demonstrated the power of transcriptome analysis in combination with GWAS and linkage mapping to identify candidate genes regulating growth period traits in *B. napus*.

Conclusions

In this context, we conducted systematic research (including GWAS, linkage mapping and gene expression analysis) to demonstrate the natural variation of growth period traits in *B. napus*. In summary, 146 SNPs and 83 QTLs associated with the five growth period traits (DIF, DFF, FP, MT and GP) were identified by association and linkage mapping techniques. Six QTLs were associated with growth period traits using both mapping methods. Five QTLs found in the GWAS, and fourteen QTLs identified in the linkage mapping were identified as pleiotropic SNPs. RNA-Seq analysis demonstrated that the genes mediated by photoperiod/circadian rhythms, vernalization, GA signaling, aging, flowering time integrator and flower meristem identity may regulate floral transition and flower development in *B. napus*. An integrated analysis of QTLs and DEGs identified twelve candidate genes for growth period traits.

BnaTOC1.A03 and *BnaFUL.A03* were identified in both GWAS and linkage analysis; *BnaLNK2.C06*, *BnaGA200X3.C09* and *BnaFUL.A03* have been previously reported. Among twelve genes, seven genes have polymorphic sites in the coding sequence and 2000 bp upstream of the TSS. According to phenotype analysis, the haplotype BnaSOC1.A05-HapIb and BnaLNK2.C06-HapIa showed more favorable phenotypic traits. Therefore, our data suggest that these genes may play important roles as regulators of flowering time and growth period in rapeseed.

Materials And Methods

Plant material and phenotypic evaluation

A diversity panel consisting of 588 rapeseed inbred lines (74 winter types, 428 semi winter types, and 86 spring types) was used for the GWAS analysis in the present study [23]. These lines were grown at the Southwest University of Beibei, Chongqing, China, in 2017, 2018 and 2019. Each variety was planted in a plot with two rows (30-cm line width and 20-cm plant distance), and each row had 10 plants. A randomized complete block design with two replications was employed. The RIL populations of 172 RILs derived from GH06 P174 were previously described in Liu et al, 2013[67]. The genetic linkage map contains 9164 SNP markers covering 1832.9 cM, and 2795 SNPs were applied for QTL mapping. The population was planted in the experimental field of Southwest University, Chongqing, China, in 2017 and 2018. The planting method is consistent with the GWAS population.

Five growth period traits, DIF (the number of days from sowing to the date when 25% plants opened first flower), DFF (the number of days from sowing to the date when 75% plants had stopped blooming), FP (the difference between DFF and DIF), MT (the number of days from DFF to the date when 75% pods were yellow) and the GP (the number of days from sowing to the pod ripen) were recorded in the field trials, respectively [13]. Analysis of variance (ANOVA) was performed using the GLM procedure of SAS.

Genome-wide association analysis

GWAS for the growth period traits were performed using Tassel 5.0 software by general linear models (GLM) and mixed linear models (MLM) methods. Population structure, relative kinship, and LD analysis were completed in previous studies [23]. A negative log (1/n) was used as a threshold for significant association SNPs with traits, where n represents the SNP number used in the GWAS. Significant markers in the same LD block were viewed as one QTL region. The QQ plot and Manhattan plot were displayed using qqman [68] and CMplot software.

QTL analysis

Windows QTL Cartographer Version 2.5 (WinQTLcart2.5) software was used in QTL mapping via the composite interval mapping method [69]. The logarithm of the odds (LOD) threshold for a significant QTL

was calculated with 1000 permutations at a significance level of $p = 0.05$. A confidence interval for each QTL was defined by LOD change from the peak position. The contribution rate (R^2) and additive effect of a putative QTL were also calculated from WinQTLcart2.5. If the phenotypic variance explained was larger than 10%, we classified the QTL as a major QTL. QTLs were assigned names by adding the prefix of trait abbreviation, and their locations, including the chromosome number, were determined according to McCouch et al. [70]. If more than one QTL was detected on the same chromosome for a trait, QTLs were serially numbered. MapChart software was used to draw the QTL position [71].

Gene expression analysis

Early-flowering material (18Z134) and late-flowering material (18Z88) were selected based on six years' field data. The juvenile leaves at the vegetative and reproductive stages were collected respectively and immediately placed in liquid nitrogen and stored at $-80\text{ }^{\circ}\text{C}$. Total RNA was extracted using RNAprep Pure Tissue Kit (Tiangen, China) according to the manufacturer's instructions. RNA was pooled to yield four RNA samples, EV (leaves of early-flowering material at vegetative stage), LV (leaves of late-flowering material at vegetative stage), ER (leaves of early-flowering material at reproductive stage) and LR (leaves of late-flowering material at reproductive stage). Each sample has two biological replications. The sequencing data have been deposited in NCBI Sequence Read Archive (SRA, <https://www.ncbi.nlm.nih.gov/sra>) with the accession number PRJNA540020.

DEGs between the two samples were identified based on the criteria false discovery rate (FDR) < 0.05 and $|\log_2(\text{Fold change})| > 1$. Gene ontology (GO) enrichment analysis of the DEGs was performed using Blast2GO, and the significantly enriched GO terms ($P < 0.05$) were displayed using the online tool WEGO (<http://wego.genomics.org.cn>). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the DEGs was performed using the KOBAS2.0 website (<http://kobas.cbi.pku.edu.cn/home.do>). The transcript abundance was calculated based on the fragments per kilobase of exon model per million mapped reads (FPKM). The \log_2 -transformed (FPKM+1) values were used to study gene expression.

Candidate gene predication

According to the decay of LD, genes within 300 kb up and downstream of the significantly associated SNP loci were mined [11, 72]. The QTL intervals were aligned to the *B. napus* reference genomes (<http://www.genoscope.cns.fr/blat-server/cgi-bin/colza/webBlat>). Based on the physical positions of the flanking marker, genes located in the QTL confidence interval were extracted. The prediction of candidate genes referred to the two conditions: genes of known function in *B. napus* or genes with function-known orthologs in Arabidopsis. Genes within the confidence interval and meeting one of the above conditions were used for gene expression analysis.

Validation of candidate genes by qRT-PCR

For qRT-PCR, the samples used for RNA-seq were used to synthesize cDNA using the QuantiTect® Reverse Transcription Kit (Qiagen), and qRT-PCR experiments were performed with SYBR qPCR Mix (Bio-Rad) according to the manufacturer's specification. At least three biological replicates along with two technical replicates were analyzed for expression levels. BnActin7 (5'-GGAGCTGAGAGATTCGGTTG-3' and 5'-GAACCA CCACTGAGGACGAT-3') was used as an internal control. The $2^{-\Delta\Delta C_t}$ method was used to calculate the normalized expression of target genes [73]. The candidate gene-specific primers are shown in Table S15.

Abbreviations

DIF: Days to initial flowering; DFF: Days to final flowering; FP: days from initial flowering to final flowering; MT: Days from final flowering to maturity; GP: whole growth period; GWAS: Genome-wide association study; GLM: General linear models; MLM: Mixed linear models; SNP: Single nucleotide polymorphism; MAF: Minor allele frequency; QTL: Quantitative trait locus; BLUP: Best linear unbiased prediction; EV: leaves of early-flowering material at vegetative stage; LV: leaves of late-flowering material at vegetative stage; ER: leaves of early-flowering material at reproductive stage; LR: leaves of late-flowering material at reproductive stage; RNA-seq: RNA sequencing; DEGs: Differentially expressed genes; FPKM: Fragments per kilobase of exon per million reads mapped; FDR: False discovery rate; qRT-PCR: Quantitative real-time PCR; TOE1: Target of early activation tagged 1; TOC1: Timing of cab expression 1; LNK2: Night light-inducible and clock regulated 2; ELF4: Early flowering 4; AGL6: AGAMOUS-like 6; GA20OX3: Gibberellin 20-oxidase 3; ICE1: Inducer of CBF expression 1; AGL14: AGAMOUS-Like14; FUL: FRUITFULL; SOC1: Suppressor of overexpression of CO 1

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All experimental materials are available on request. All data generated during this study are included in this article and its additional files.

Competing interests

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

Funding

This research was supported by the National Natural Science Foundation of China (31971902, 31771830, 31701335), National Key Research and Development Program of China (2016YFD0100202), the Chongqing Science and Technology Commission (cstc2016shmszx80083) and the “111” Project (B12006). Thanks to the China Scholarship Council (CSC) for its support.

Authors' contributions

TYW and LZL designed the experiments. TYW and LX collected the phenotypic data. TYW conducted the data analysis and wrote the manuscript. LJW, JW and YYL helped analyze the data. SYR and LYR performed qRT-PCR. JNL, KL and Timko participated in drafting partial manuscript and revised the paper. All authors read and approved the final manuscript.

Acknowledgements

We are grateful to the reviewers and editors for their constructive review and suggestions for this paper.

References

1. Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *science*. 2014; 345(6199):950-953.
2. Chen X, Xie H, Li X, Yang J, Han T. Studies on correlation of development stages and agronomic traits of summer sowing soybean. *Fen zi zhi wu yu zhong*. 2004; 2(2):247-252.
3. Udall JA, Quijada PA, Lambert B, Osborn TC. Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 2. Identification of alleles from unadapted germplasm. *Theor Appl Genet*. 2006; 113(4):597-609.
4. Long Y, Shi J, Qiu D, Li R, Zhang C, Wang J, Hou J, Zhao J, Shi L, Park B-S. Flowering time quantitative trait loci analysis of oilseed Brassica in multiple environments and genomewide alignment with Arabidopsis. *Genetics*. 2007; 177(4):2433-2444.
5. Mei D, Wang H, Hu Q, Li Y, Xu Y, Li Y. QTL analysis on plant height and flowering time in *Brassica napus*. *Plant Breed*. 2009; 128(5):458-465.
6. Raman H, Raman R, Eckermann P, Coombes N, Manoli S, Zou X, Edwards D, Meng J, Prangnell R, Stiller J. Genetic and physical mapping of flowering time loci in canola (*Brassica napus* L.). *Theor*

- Appl Genet. 2013; 126(1):119-132.
7. Raman H, Dalton-Morgan J, Diffey S, Raman R, Alamery S, Edwards D, Batley J. SNP markers-based map construction and genome-wide linkage analysis in *Brassica napus*. Plant Biotechnol J. 2014; 12(7):851-860.
 8. Liu H, Du D, Guo S, Xiao L, Zhao Z, Zhao Z, Xing X, Tang G, Xu L, Fu Z. QTL analysis and the development of closely linked markers for days to flowering in spring oilseed rape (*Brassica napus* L.). Mol Breed. 2016; 36(5):52.
 9. Luo Z, Wang M, Long Y, Huang Y, Shi L, Zhang C, Liu X, Fitt BD, Xiang J, Mason AS. Incorporating pleiotropic quantitative trait loci in dissection of complex traits: seed yield in rapeseed as an example. Theor Appl Genet. 2017; 130(8):1569-1585.
 10. Tian D, Wang P, Tang B, Teng X, Li C, Liu X, Zou D, Song S, Zhang Z. GWAS Atlas: a curated resource of genome-wide variant-trait associations in plants and animals. Nucleic Acids Res. 2019; 48(D1):D927-D932.
 11. Xu L, Hu K, Zhang Z, Guan C, Chen S, Hua W, Li J, Wen J, Yi B, Shen J. Genome-wide association study reveals the genetic architecture of flowering time in rapeseed (*Brassica napus* L.). DNA Res. 2015; 23(1):43-52.
 12. Körber N, Bus A, Li J, Parkin IA, Wittkop B, Snowdon RJ, Stich B. Agronomic and seed quality traits dissected by genome-wide association mapping in *Brassica napus*. Front Plant Sci. 2016; 7:386.
 13. Zhou Q, Han D, Mason AS, Zhou C, Zheng W, Li Y, Wu C, Fu D, Huang Y. Earliness traits in rapeseed (*Brassica napus*): SNP loci and candidate genes identified by genome-wide association analysis. DNA Res. 2017; 25(3):229-244.
 14. Wei D, Cui Y, He Y, Xiong Q, Qian L, Tong C, Lu G, Ding Y, Li J, Jung C. A genome-wide survey with different rapeseed ecotypes uncovers footprints of domestication and breeding. J Exp Bot. 2017; 68(17):4791-4801.
 15. Lu K, Peng L, Zhang C, Lu J, Yang B, Xiao Z, Liang Y, Xu X, Qu C, Zhang K. Genome-wide association and transcriptome analyses reveal candidate genes underlying yield-determining traits in *Brassica napus*. Front Plant Sci. 2017; 8:206.
 16. Ye J, Yang Y, Chen B, Shi J, Luo M, Zhan J, Wang X, Liu G, Wang H. An integrated analysis of QTL mapping and RNA sequencing provides further insights and promising candidates for pod number variation in rapeseed (*Brassica napus* L.). BMC Genomics. 2017; 18(1):71.
 17. Li S, Zhu Y, Varshney RK, Zhan J, Zheng X, Shi J, Wang X, Liu G, Wang H. A systematic dissection of the mechanisms underlying the natural variation of silique number in rapeseed (*Brassica napus* L.) germplasm. Plant Biotechnol J. 2019.
 18. Wei L, Jian H, Lu K, Filardo F, Yin N, Liu L, Qu C, Li W, Du H, Li J. Genome-wide association analysis and differential expression analysis of resistance to Sclerotinia stem rot in *Brassica napus*. Plant Biotechnol J. 2016; 14(6):1368-1380.
 19. Bates D, Sarkar D, Bates MD, Matrix L. The lme4 package. R package version. 2007; 2(1):74.

20. Bluemel M, Dally N, Jung C. Flowering time regulation in crops—what did we learn from *Arabidopsis*? *Curr Opin Biotechnol*. 2015; 32:121-129.
21. Wang N, Chen B, Xu K, Gao G, Li F, Qiao J, Yan G, Li J, Li H, Wu X. Association mapping of flowering time QTLs and insight into their contributions to rapeseed growth habits. *Front Plant Sci*. 2016; 7:338.
22. Raman H, Raman R, Qiu Y, Yadav AS, Sureshkumar S, Borg L, Rohan M, Wheeler D, Owen O, Menz I. GWAS hints at pleiotropic roles for FLOWERING LOCUS T in flowering time and yield-related traits in canola. *BMC Genomics*. 2019; 20(1):636.
23. Lu K, Wei L, Li X, Wang Y, Wu J, Liu M, Zhang C, Chen Z, Xiao Z, Jian H. Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. *Nature communications*. 2019; 10(1):1154.
24. Raman H, Raman R, Coombes N, Song J, Prangnell R, Bandaranayake C, Tahira R, Sundaramoorthi V, Killian A, Meng J. Genome-wide association analyses reveal complex genetic architecture underlying natural variation for flowering time in canola. *Plant, Cell Environ*. 2016; 39(6):1228-1239.
25. Schiessl S, Iniguez-Luy F, Qian W, Snowdon RJ. Diverse regulatory factors associate with flowering time and yield responses in winter-type *Brassica napus*. *BMC Genomics*. 2015; 16(1):737.
26. Shah S, Weinholdt C, Jedrusik N, Molina C, Zou J, Große I, Schiessl S, Jung C, Emrani N. Whole-transcriptome analysis reveals genetic factors underlying flowering time regulation in rapeseed (*Brassica napus* L.). *Plant, Cell Environ*. 2018; 41(8):1935-1947.
27. Wu D, Liang Z, Yan T, Xu Y, Xuan L, Tang J, Zhou G, Lohwasser U, Hua S, Wang H. Whole-genome resequencing of a worldwide collection of rapeseed accessions reveals the genetic basis of ecotype divergence. *Molecular plant*. 2019; 12(1):30-43.
28. Becker A, Theißen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol*. 2003; 29(3):464-489.
29. Bartrina I, Otto E, Strnad M, Werner T, Schmülling T. Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*. *Plant Cell*. 2011; 23(1):69-80.
30. Davis SJ. Integrating hormones into the floral-transition pathway of *Arabidopsis thaliana*. *Plant, Cell Environ*. 2009; 32(9):1201-1210.
31. Martínez C, Pons E, Prats G, León J. Salicylic acid regulates flowering time and links defence responses and reproductive development. *The Plant Journal*. 2004; 37(2):209-217.
32. Yu S, Galvão VC, Zhang Y-C, Horrer D, Zhang T-Q, Hao Y-H, Feng Y-Q, Wang S, Schmid M, Wang J-W. Gibberellin regulates the *Arabidopsis* floral transition through miR156-targeted SQUAMOSA PROMOTER BINDING-LIKE transcription factors. *Plant Cell*. 2012; 24(8):3320-3332.
33. Mutasa-Göttgens E, Hedden P. Gibberellin as a factor in floral regulatory networks. *J Exp Bot*. 2009; 60(7):1979-1989.
34. McClung CR. Plant circadian rhythms. *Plant Cell*. 2006; 18(4):792-803.

35. Dreni L, Zhang D. Flower development: the evolutionary history and functions of the AGL6 subfamily MADS-box genes. *J Exp Bot.* 2016; 67(6):1625-1638.
36. Balanzà V, Martínez-Fernández I, Ferrándiz C. Sequential action of FRUITFULL as a modulator of the activity of the floral regulators SVP and SOC1. *J Exp Bot.* 2014; 65(4):1193-1203.
37. He Y, Wu D, Wei D, Fu Y, Cui Y, Dong H, Tan C, Qian W. GWAS, QTL mapping and gene expression analyses in *Brassica napus* reveal genetic control of branching morphogenesis. *Scientific reports.* 2017; 7(1):15971.
38. Han K, Lee HY, Ro NY, Hur OS, Lee JH, Kwon JK, Kang BC. QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in Capsicum. *Plant Biotechnol J.* 2018; 16(9):1546-1558.
39. Sun F, Liu J, Hua W, Sun X, Wang X, Wang H. Identification of stable QTLs for seed oil content by combined linkage and association mapping in *Brassica napus*. *Plant Sci.* 2016; 252:388-399.
40. Zuo Q, Hou J, Zhou B, Wen Z, Zhang S, Gai J, Xing H. Identification of QTL s for growth period traits in soybean using association analysis and linkage mapping. *Plant Breed.* 2013; 132(3):317-323.
41. Crowell S, Korniliev P, Falcao A, Ismail A, Gregorio G, Mezey J, McCouch S. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nature communications.* 2016; 7:10527.
42. Sandhu D, Champoux JA, Bondareva SN, Gill KS. Identification and physical localization of useful genes and markers to a major gene-rich region on wheat group 1S chromosomes. *Genetics.* 2001; 157(4):1735-1747.
43. Xiao J, Li J, Yuan L, Tanksley S. Identification of QTLs affecting traits of agronomic importance in a recombinant inbred population derived from a subspecific rice cross. *Theor Appl Genet.* 1996; 92(2):230-244.
44. Liu Z, Li H, Fan X, Huang W, Yang J, Li C, Wen Z, Li Y, Guan R, Guo Y. Phenotypic characterization and genetic dissection of growth period traits in soybean (*Glycine max*) using association mapping. *PLoS one.* 2016; 11(7):e0158602.
45. Liu K, Feng S, Pan Y, Zhong J, Chen Y, Yuan C, Li H. Transcriptome analysis and identification of genes associated with floral transition and flower development in sugar apple (*Annona squamosa* L.). *Front Plant Sci.* 2016; 7:1695.
46. Guo X, Yu C, Luo L, Wan H, Zhen N, Xu T, Tan J, Pan H, Zhang Q. Transcriptome of the floral transition in *Rosa chinensis* 'Old Blush'. *BMC Genomics.* 2017; 18(1):199.
47. Smaczniak C, Immink RG, Muiño JM, Blanvillain R, Busscher M, Busscher-Lange J, Dinh QP, Liu S, Westphal AH, Boeren S. Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. *Proc Natl Acad Sci.* 2012; 109(5):1560-1565.
48. Singh VK, Garg R, Jain M. A global view of transcriptome dynamics during flower development in chickpea by deep sequencing. *Plant Biotechnol J.* 2013; 11(6):691-701.
49. Vimolmangkang S, Han Y, Wei G, Korban SS. An apple MYB transcription factor, MdMYB3, is involved in regulation of anthocyanin biosynthesis and flower development. *BMC Plant Biol.* 2013; 13(1):176.

50. Wellmer F, Riechmann JL, Alves-Ferreira M, Meyerowitz EM. Genome-wide analysis of spatial gene expression in Arabidopsis flowers. *Plant Cell*. 2004; 16(5):1314-1326.
51. Srikanth A, Schmid M. Regulation of flowering time: all roads lead to Rome. *Cell Mol Life Sci*. 2011; 68(12):2013-2037.
52. Doyle MR, Davis SJ, Bastow RM, McWatters HG, Kozma-Bognár L, Nagy F, Millar AJ, Amasino RM. The ELF4 gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. *Nature*. 2002; 419(6902):74.
53. Rugnone ML, Soverna AF, Sanchez SE, Schlaen RG, Hernando CE, Seymour DK, Mancini E, Chernomoretz A, Weigel D, Más P. LNK genes integrate light and clock signaling networks at the core of the Arabidopsis oscillator. *Proc Natl Acad Sci*. 2013; 110(29):12120-12125.
54. Lee JH, Jung JH, Park CM. INDUCER OF CBF EXPRESSION 1 integrates cold signals into FLOWERING LOCUS C-mediated flowering pathways in Arabidopsis. *The Plant Journal*. 2015; 84(1):29-40.
55. Tenreira T, Lange MJP, Lange T, Bres C, Labadie M, Monfort A, Hernould M, Rothan C, Denoyes B. A specific gibberellin 20-oxidase dictates the flowering-running decision in diploid strawberry. *Plant Cell*. 2017; 29(9):2168-2182.
56. Tao Z, Shen L, Liu C, Liu L, Yan Y, Yu H. Genome-wide identification of SOC1 and SVP targets during the floral transition in Arabidopsis. *The Plant Journal*. 2012; 70(4):549-561.
57. Andrés F, Porri A, Torti S, Mateos J, Romera-Branchat M, García-Martínez JL, Fornara F, Gregis V, Kater MM, Coupland G. SHORT VEGETATIVE PHASE reduces gibberellin biosynthesis at the Arabidopsis shoot apex to regulate the floral transition. *Proc Natl Acad Sci*. 2014; 111(26):E2760-E2769.
58. Poethig RS. Small RNAs and developmental timing in plants. *Curr Opin Genet Dev*. 2009; 19(4):374-378.
59. Wu G, Park MY, Conway SR, Wang J-W, Weigel D, Poethig RS. The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell*. 2009; 138(4):750-759.
60. Schwab R: The roles of miR156 and miR172 in phase change regulation. In: *MicroRNAs in plant development and stress responses*. Springer; 2012: 49-68.
61. Yoo SK, Wu X, Lee JS, Ahn JH. AGAMOUS-LIKE 6 is a floral promoter that negatively regulates the FLC/MAF clade genes and positively regulates FT in Arabidopsis. *The Plant Journal*. 2011; 65(1):62-76.
62. Lee J-H, Park C-M. Integration of photoperiod and cold temperature signals into flowering genetic pathways in Arabidopsis. *Plant signaling & behavior*. 2015; 10(11):e1089373.
63. Jung JH, Ju Y, Seo PJ, Lee JH, Park CM. The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in Arabidopsis. *The Plant Journal*. 2012; 69(4):577-588.
64. Borner R, Kampmann G, Chandler J, Gleißner R, Wisman E, Apel K, Melzer S. A MADS domain gene involved in the transition to flowering in Arabidopsis. *The Plant Journal*. 2000; 24(5):591-599.

65. Pérez-Ruiz RV, García-Ponce B, Marsch-Martínez N, Ugartechea-Chirino Y, Villajuana-Bonequi M, De Folter S, Azpeitia E, Dávila-Velderrain J, Cruz-Sánchez D, Garay-Arroyo A. XAANTAL2 (AGL14) is an important component of the complex gene regulatory network that underlies Arabidopsis shoot apical meristem transitions. *Molecular plant*. 2015; 8(5):796-813.
66. Guo Y, Hans H, Christian J, Molina C. Mutations in single FT-and TFL1-paralogs of rapeseed (*Brassica napus* L.) and their impact on flowering time and yield components. *Front Plant Sci*. 2014; 5:282.
67. Liu L, Qu C, Wittkop B, Yi B, Xiao Y, He Y, Snowdon RJ, Li J. A high-density SNP map for accurate mapping of seed fibre QTL in *Brassica napus* L. *PloS one*. 2013; 8(12):e83052.
68. Turner SD. qqman: an R package for visualizing GWAS results using QQ and manhattan plots. *Biorxiv*. 2014:005165.
69. Luciano Da Costa ES, Wang S, Zeng Z-B: Composite interval mapping and multiple interval mapping: procedures and guidelines for using Windows QTL Cartographer. In: *Quantitative trait loci (QTL)*. Springer; 2012: 75-119.
70. McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M: Microsatellite marker development, mapping and applications in rice genetics and breeding. In: *Oryza: From Molecule to Plant*. Springer; 1997: 89-99.
71. Voorrips R. MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered*. 2002; 93(1):77-78.
72. Qian L, Qian W, Snowdon RJ. Sub-genomic selection patterns as a signature of breeding in the allopolyploid *Brassica napus* genome. *BMC Genomics*. 2014; 15(1):1170.
73. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *methods*. 2001; 25(4):402-408.

Figures

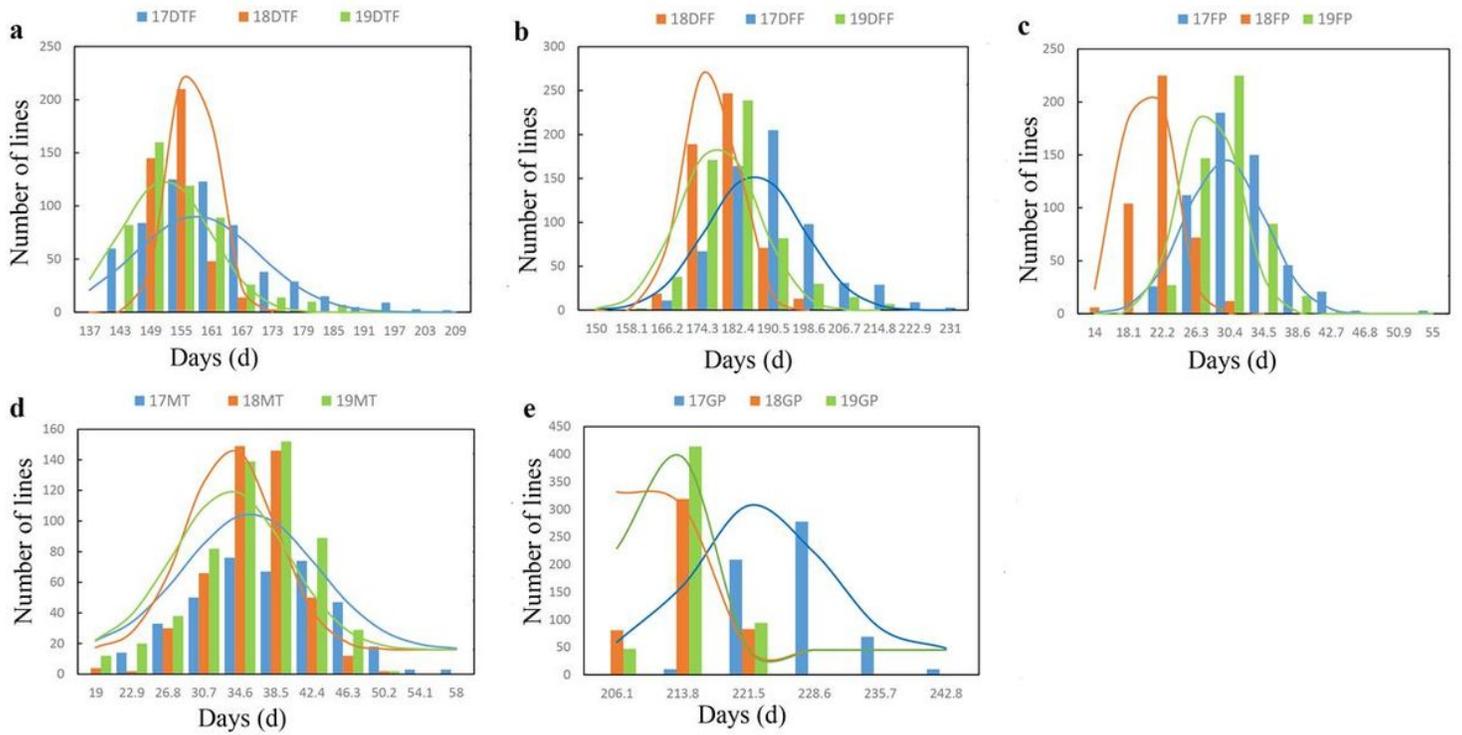


Figure 1

Frequency distribution of five growth period traits in GWAS population in three years (2017, 2018 and 2019).

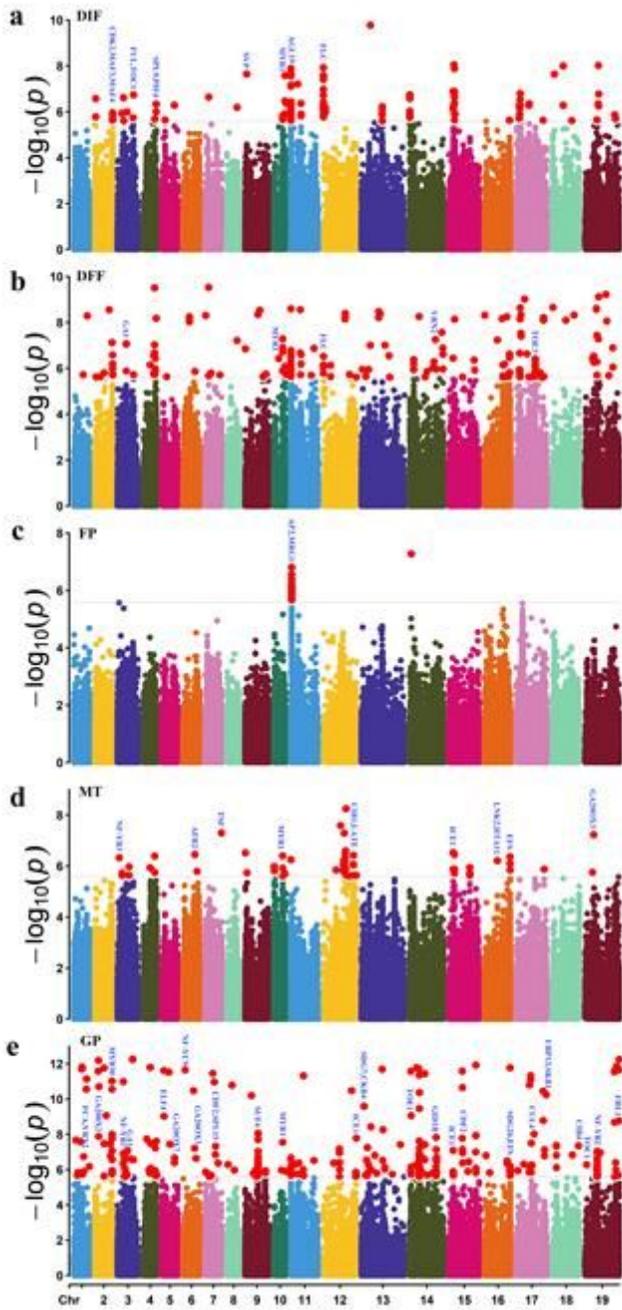


Figure 2

The Manhattan plots for five growth period traits using the BLUP value. Significant associations - $\log_{10}(P)$ values of ≤ 5.58 are shown with a horizontal line. The position of candidate flowering genes that located in the vicinity of the significant SNP is shown.

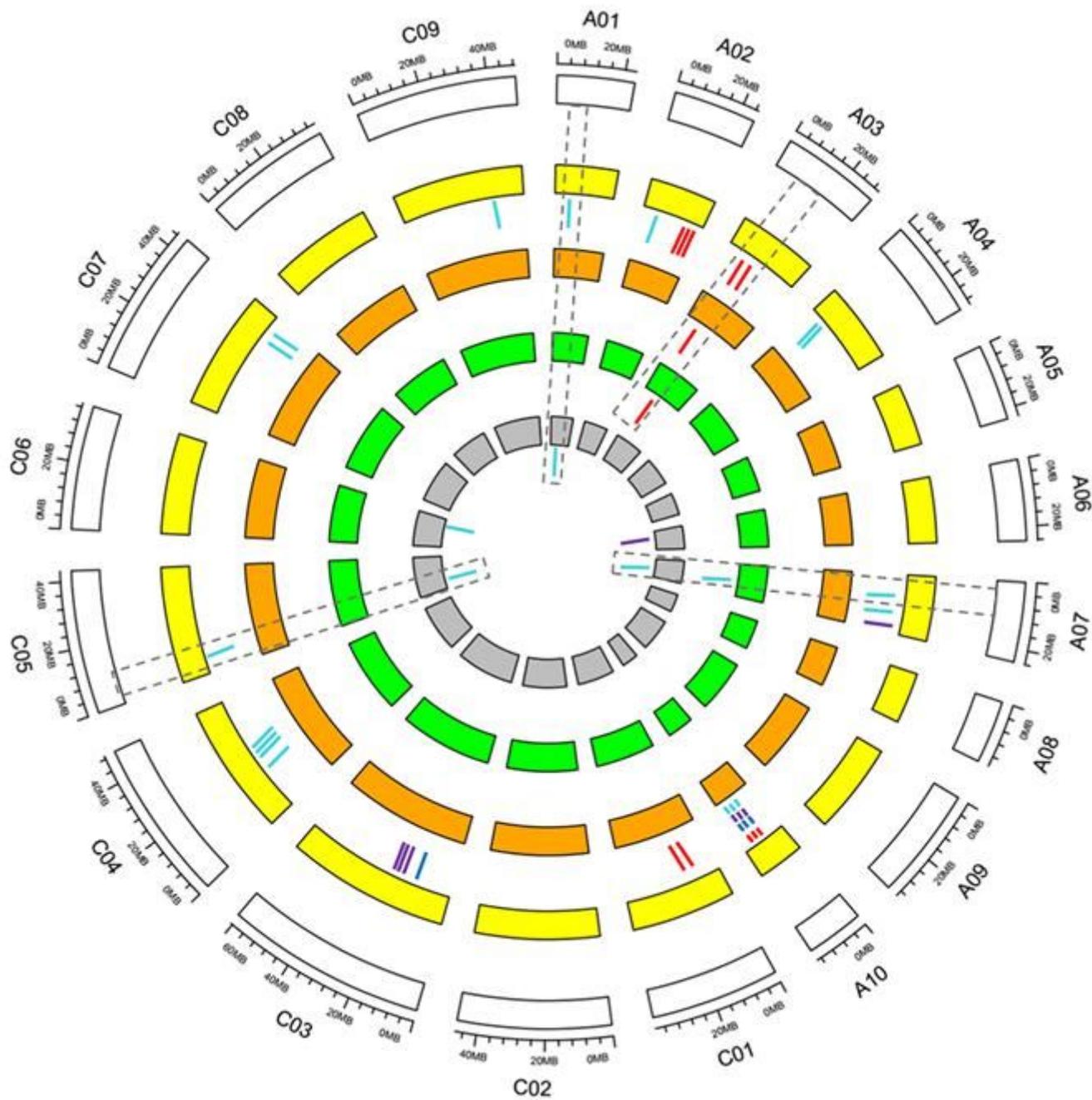


Figure 3

Distribution of consensus QTLs for five traits among different populations. QTLs obtained in GWAS population are shown on the outside circle. The inside circles represent SNPs projected from Raman, Udall, Zhou and Wang, respectively. Different traits are indicated by bars with various backgrounds (Red bars, QTLs for DIF; blue bars, DFF; purple bars, MT; bright green bars, GP). The dotted lines indicated that the SNP loci detected in different populations were co-localized.

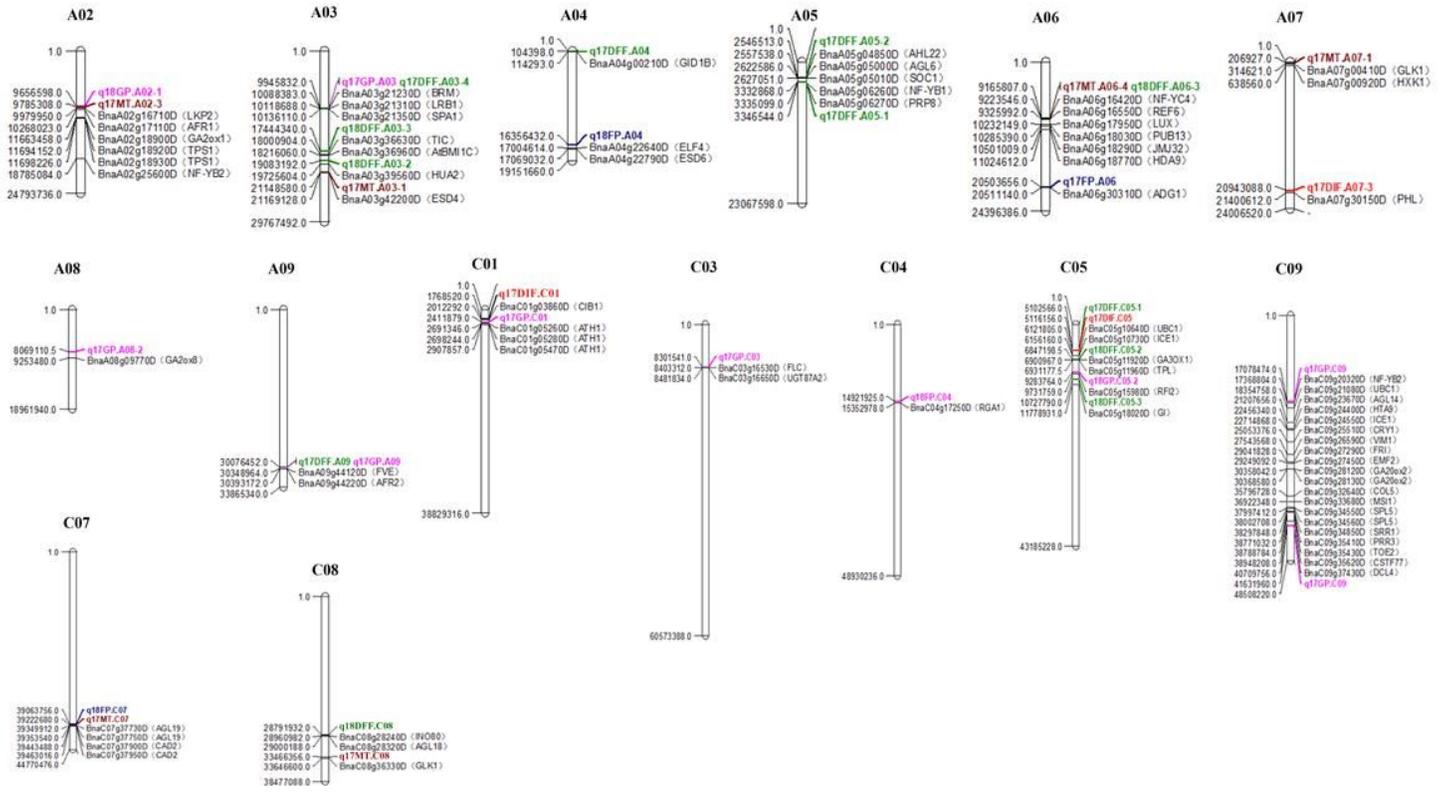


Figure 4

The distribution of candidate flowering genes and their corresponding SNPs detected in the RIL population.

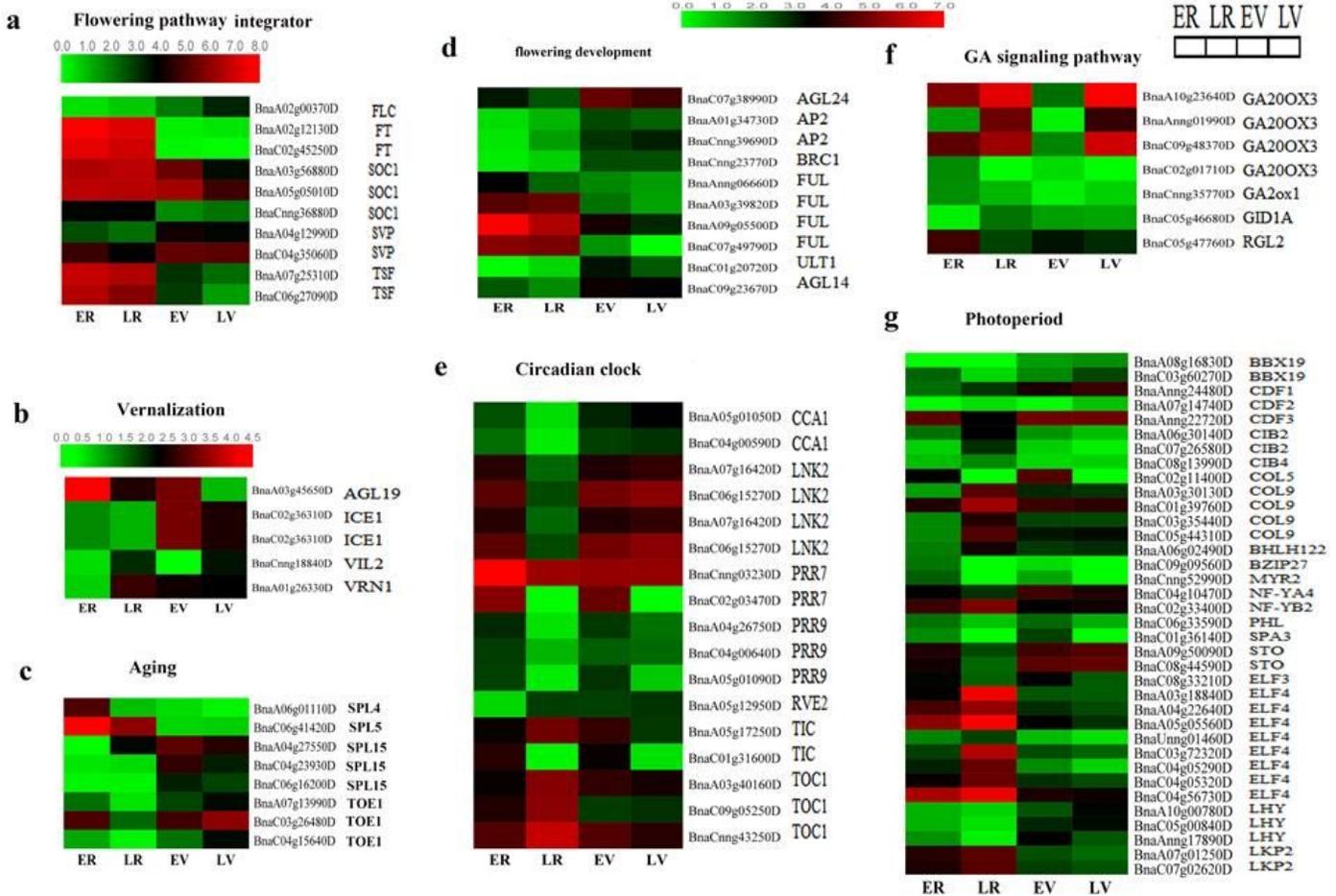


Figure 5

Heatmap showing expression patterns of DEGs involved in the flowering development including flowering pathway integrator (a), vernalization pathway(b), aging pathway(c), non-classified flowering regulators(d), circadian clock (e), GA signaling pathway(f) and photoperiod pathway(g). Gene expression levels were transformed with $\log_2(\text{FPKM} + 1)$.

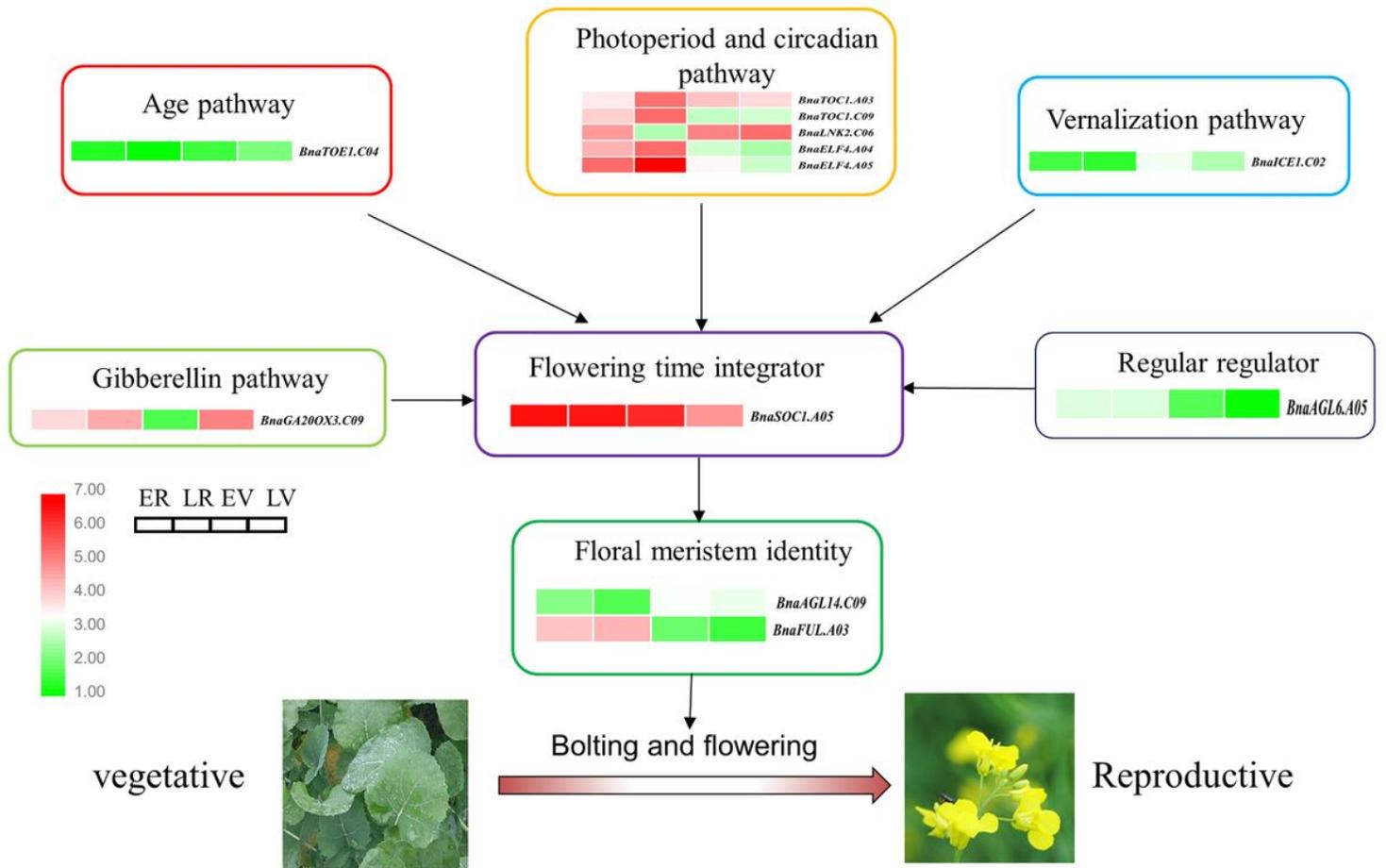


Figure 6

The putative model of flowering regulatory network of twelve candidate genes associated with flowering time in *B.napus*. The heatmap represents the expression levels in four samples.

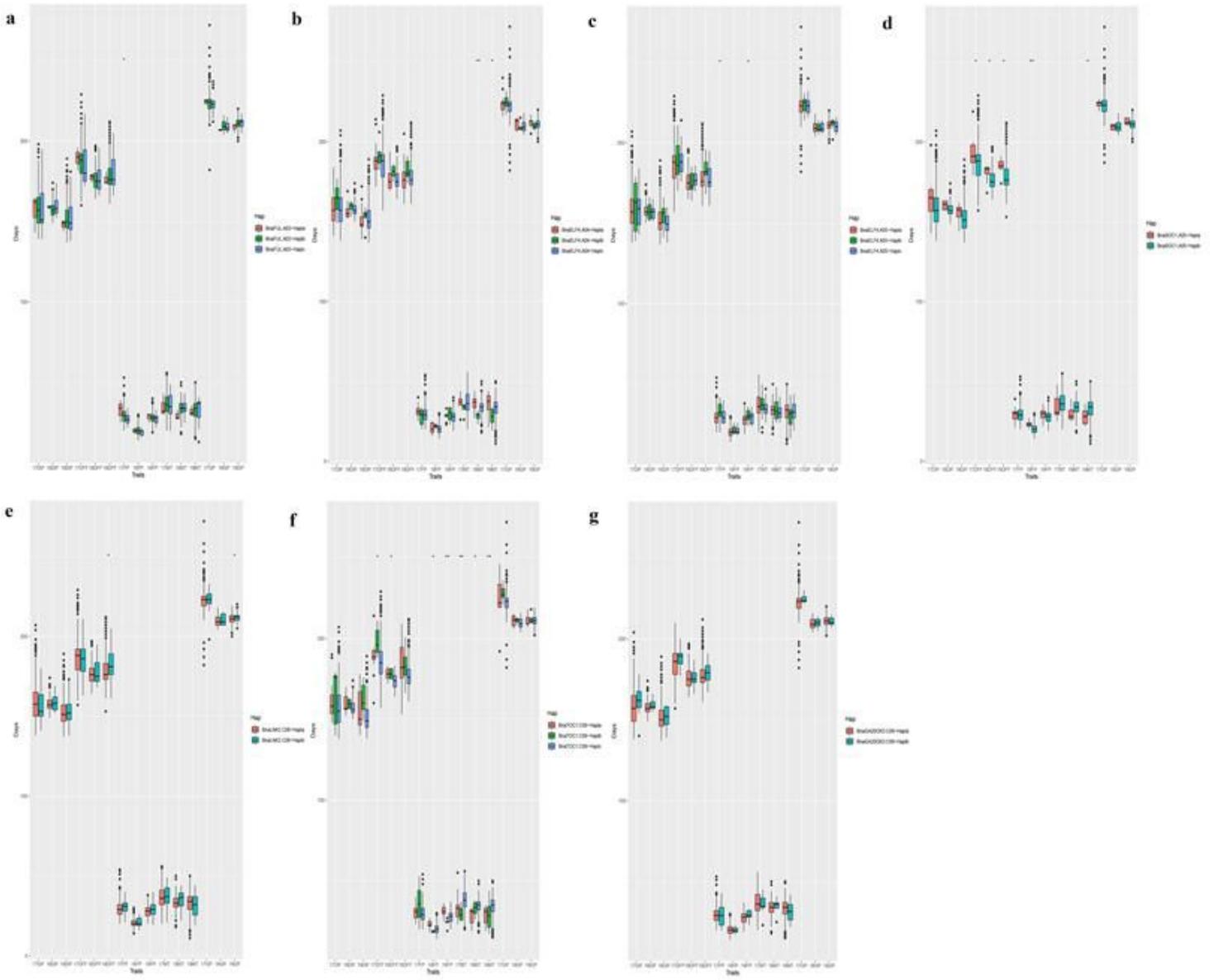


Figure 7

Comparison among different haplotypes of seven candidate genes on the basis of five growth periods in three years.

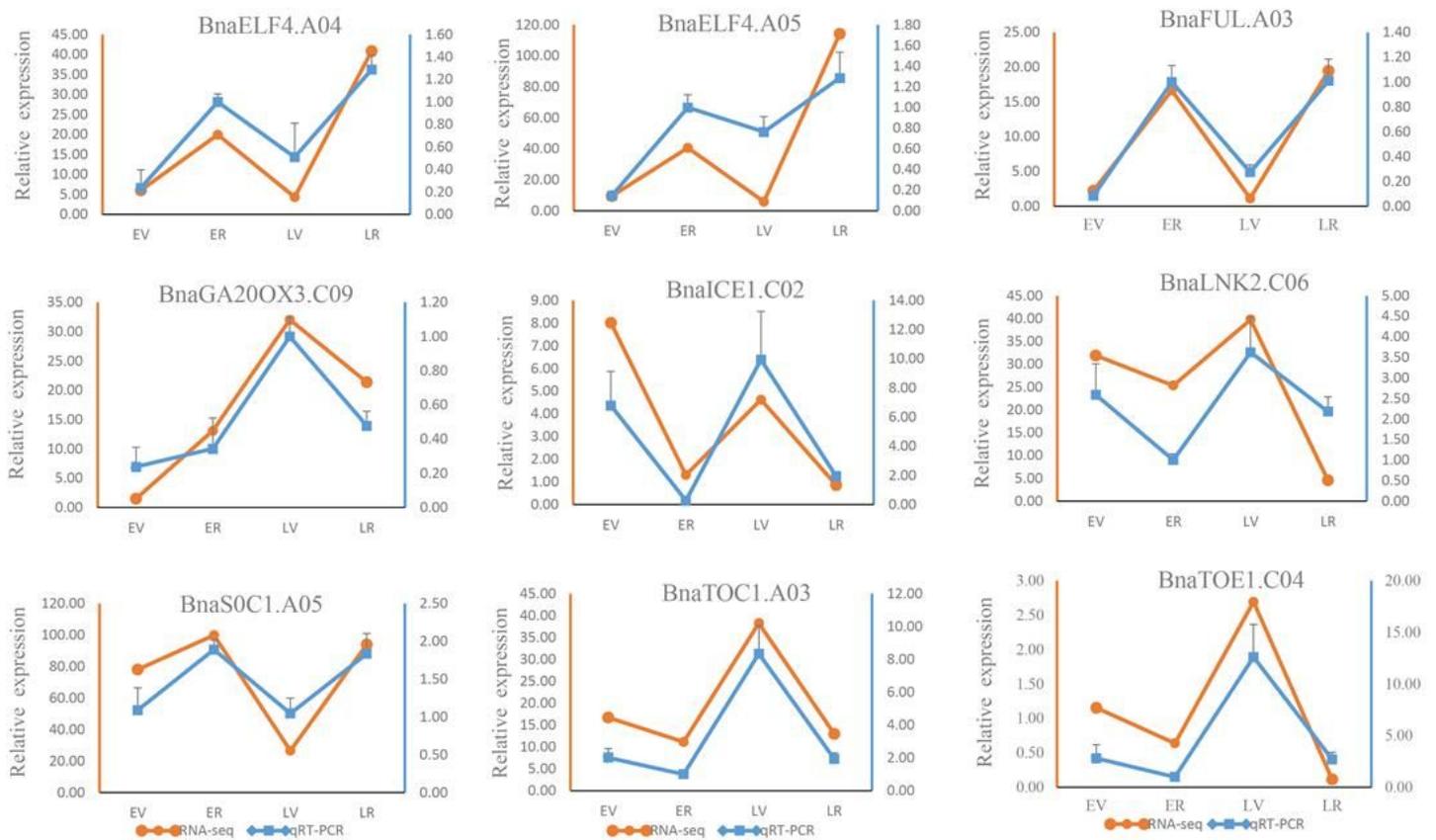


Figure 8

qRT-PCR validation of the expression patterns of nine candidate genes in four samples. The orange line represents the RNA-Seq results, the blue line represents the qRT-PCR results and data are shown as mean \pm s.e.m.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [FigureS4.jpg](#)
- [TableS4.xlsx](#)
- [TableS3.xlsx](#)
- [FigureS5.jpg](#)
- [FigureS6.jpg](#)
- [TableS6.xlsx](#)
- [TableS5.xlsx](#)

- [TableS7.xlsx](#)
- [TableS8.xlsx](#)
- [FigureS8.jpg](#)
- [FigureS7.jpg](#)
- [FigureS9.jpg](#)
- [TableS9.xlsx](#)
- [FigureS10.jpg](#)
- [TableS10.xlsx](#)
- [TableS11.xlsx](#)
- [TableS12.xlsx](#)
- [TableS13.xlsx](#)
- [TableS15.xlsx](#)
- [TableS14.xlsx](#)