Euphytica (2013) 191:23–33
DOI 10.1007/s10681-012-0840-8

Fine mapping of a major flowering time QTL on soybean chromosome 6 combining linkage and association analysis

Dan Zhang · Hao Cheng · Zhenbin Hu · Hui Wang · Guizheng Kan · Chunying Liu · Deyue Yu

Received: 21 April 2012 / Accepted: 18 November 2012 / Published online: 27 November 2012

Abstract Flowering time is a key trait in the plant life cycle and an important selection criterion for soybean. Here, we combine the advantages of genome-wide association and linkage mapping to identify and fine map quantitative trait loci (QTLs) associated with flowering time. Linkage mapping was performed using 152 recombinant inbred lines and a major QTL, *qFT6*, affecting flowering time was found on chromosome 6. To refine the *qFT6*, the 192 natural accessions were genotyped using eight new simple sequence repeats and 10 single nucleotide polymorphisms markers covering the *qFT6* region Haplotype analysis showed that the haplotype between markers BARC-014947-01929 and Satt365 could explain more phenotypic variation (26.5 %) than any other combination of markers. These results suggested that the target flowering time gene was located in ~300 kb between BARC-014947-01929 and Satt365, including three predicted genes. High-resolution map in *qFT6* region will be useful not only for marker-assisted selection of flowering time but also for further positional cloning of the target gene. These results indicate that combining association and linkage mapping provides an efficient approach for fine mapping of soybean genes.

Keywords Fine mapping · Flowering time · Association mapping · Linkage mapping · Soybean

Electronic supplementary material The online version of this article (doi:10.1007/s10681-012-0840-8) contains supplementary material, which is available to authorized users.

D. Zhang · H. Cheng · Z. Hu · H. Wang · G. Kan · C. Liu · D. Yu (✉)
National Key Laboratory of Crop Genetics and Germplasm Enhancement, National Center for Soybean Improvement, Nanjing Agricultural University, Nanjing 210095, China
e-mail: dyyu@njau.edu.cn

D. Zhang
Department of Agronomy, Henan Agricultural University, Zhengzhou 450002, China

Introduction

Flowering time is an important agronomic trait in plants, as it refers to the resources accumulated in vegetative storage tissues that will be reallocated to the seeds. Optimizing reproduction requires that the flowering date matches environmental conditions so that seeds can mature and disperse when appropriate conditions appear (Brachi et al. 2010). In soybean, flowering time is closely related to yield, quality and tolerances to environmental stresses (Cheng et al. 2011). Genetic modulation of the flowering time has been suggested as an effective means for improving yield, since it has a positive correlation with many yield-determining factors (Hanson 1985; Curtis et al. 2000; Kantolic and Slafer 2007). Natural variation in flowering time is considerable, but is a complex trait involving multiple genes. Recent developments in

molecular biology and statistical methodologies for quantitative trait loci (QTLs) mapping have made it possible to identify genetic factors affecting flowering time. Such developments have the potential to significantly increase the rate of genetic improvement of plant species, through marker-assisted selection (MAS) of specific loci, genome-wide selection, gene introgression and positional cloning (Andersson 2001). In the past decades, most studies aiming to unravel the genetics of flowering time variation were performed by conventional linkage mapping in soybean. Since the first QTL experiment on growth stage traits was reported two decades ago (Keim et al. 1990), more than forty loci for flowering time and more loci for maturity have been detected on soybean A2, B1, C1, C2, F1, F2, E, G, H, I, J, K, L, M and N2 linkage groups (LG) in different populations, explaining 3.80–69.7 % of the phenotypic variation (Mansur et al. 1993, 1996; Orf et al. 1999; Wang 2001; Tasma et al. 2001; Zhang et al. 2004; Yamanaka et al. 2005). These reports, however, usually using intercross-derived populations, such as F_2 and recombinant inbred lines (RILs), have some limitations in genome resolution because relatively few new recombination events are generated in the single generation separating parents. Therefore, additional strategies are required to locate QTL more precisely. The use of near isogenic lines (NILs) that differ at a single QTL is an effective approach and some of the flowering-related genes were cloned (Molnar et al. 2003; Liu et al. 2008; Watanabe et al. 2009). However, the development of NILs through repeated backcrossing is time-consuming and laborious (Tuinstra et al. 1997).

At present, fortunately, these limitations can be overcome by association mapping, which enables increased mapping resolution from a QTL interval to the candidate gene level. Recent years, many QTL studies were reported using association mapping in plants (Stracke et al. 2007; Jun et al. 2008; Hu et al. 2011; Li et al. 2011). Use of linkage-disequilibrium information increases the precision of QTL mapping because it exploits the entire number of recombinations accumulated since the original mutation generating the new QTL allele occurred (Meuwissen et al. 2002). Combined association and linkage mapping make it possible to exploit recombinations occurring both within and outside the pedigree and genotyped population. It also gives a clearer signal for QTL positions compared with association and linkage mapping alone. Additionally, the approach (combined association and linkage mapping) reduces the risk of false-positive QTL identification caused by accidental marker-phenotype associations when association and linkage mapping are used separately, and also increases the power and resolution of QTL mapping by combining all available information (Goddard 2005). This joint strategy for dissecting complex quantitative traits has been successfully employed, using linkage analysis to narrow suspected genes to centimorgan-scale regions, followed by an association analysis to fine map the genetic variation in regions showing linkage (Nemri et al. 2010; Lu et al. 2010).

Although several studies have reported the presence of one or more QTL affecting flowering time on chromosome 6, the results differ among studies with respect to the mapping procedure; the number of QTL detected, their positions, and the extent are affected by these QTL. The present study aimed at confirming and refining the previously identified flowering time-related QTLs, and to define the candidate region and to identify the polymorphisms underlying natural variation of flowering time. To achieve these, we first identified QTLs controlling flowering time using a segregating population (RIL) derived from a cross between 'Bogao' and 'Nannong 94-156'. To further narrow down the QTL region, natural accessions were phenotyped under field conditions. Fine mapping is achieved by combining genome wide association statistical methods that control for population structure with primary mapping using RILs. Together, these methods enabled us to fine mapping the QTL and detecting the candidate genes affecting flowering time in soybean.

Materials and methods

Plant materials

To map QTLs for flowering time, a soybean segregating population, consisting of 152 $F_{8:10}$ RILs was used. These RILs were derived from the cross between 'Bogao' and 'Nannong 94-156'. For association mapping, the 192 soybean accessions, including landraces, cultivars, and breeding lines collected in China from latitude 53 to 24°N and longitude 134 to 97°E was used. The population of accessions was selected not only to represent all six ecological

regions of soybean in China, but also soybeans with diverse reproductive related traits. Seeds of all accessions were obtained from the Germplasm Storage facility in the National Center for Soybean Improvement (Nanjing, China).

Field experiments and measure of flowering time

For linkage mapping, two-year trials using 152 RILs to map QTLs for flowering time were carried out at Jiangpu Station (32.0°N, 118.8°E), Nanjing, China in 2006 (sowing on 25 June) and 2007 (sowing on 20 June). There were two replicates with six plants in each replicate. Genotypes were sown in a 10 L pot with 4 kg dry soil, with two plants per pot. For association mapping, three field experiments using 192 accessions were carried out at Jiangpu Station, Nanjing (sowing on 15 June), and Maozhuang Station (34.4°N, 113.4°E), Zhengzhou (sowing on 25 June), Henan, China in 2008, and a repeat experiment was carried out in Zhengzhou, Henan in 2010 (sowing on 17 June). These experiments were in a completely randomized design with three replications, using 2 m in row length, 0.4 m in row space and 0.1 m in plants spacing. The flowering time was calculated as the number of days from the germination to the beginning bloom (R1, 50 % of the plants in a plot had an open flower at one of the top nodes with a fully expanded leaf) (Fehr et al. 1971).

Population genotypic data analysis

The 192 soybean accessions were genotyped with 80 unlinked SSRs, providing an even coverage of the soybean genome. The employed SSR markers are available at http://bldg6.arsusda.gov/pooley/soy/cregan/soymap1.html. The number of alleles and the polymorphism information content (PIC) per locus were calculated using the POWERMAKER 3.25 software (Liu and Muse 2005). The population structure was inferred from the SSR data using the STRUCTURE software version 2.2 (Falush et al. 2007). Six independent runs were carried out using with the following parameters: number of populations (K) from 1 to 10, burn-in time and Markov-chain Monte Carlo replication number both set to 500,000, model of admixture and correlated allele frequencies (See the Structure 2.2 documentation at http://pritch.bsd.uchicago.edu/software). The likely number of sub-populations present was estimated following Evanno et al. (2005), in which the number of sub-groups (Δk) is maximized. STRUCTURE produces a Q matrix that lists the estimated membership coefficients for each individual in each cluster and this was used in the subsequent association analysis.

QTL mapping

For linkage mapping, the composite interval mapping (CIM) program of WinQTLCart version 2.5 (Wang et al. 2005) was used to detect QTLs for flowering time, using the 152 soybean RILs. Empirical thresholds were computed using the permutation test (1,000 permutations, overall error level 5 %) for CIM (Churchill and Doerge 1994). Confidence intervals were set as the map interval corresponding to a 1-LOD decline on either side of the LOD peak. A genetic map, which has been constructed in our laboratory, comprising 306 markers was used for linkage mapping (Zhang et al. 2009). For association mapping, the 192 natural accessions, eight newly developed SSR markers and 10 SNP markers based on the physical map (http://soybase.org/gbrowse/cgi-bin/gbrowse/gmax1.01/#search) were used to further narrow the major QTL region (*qFT6*). LD between pairs of markers in *qFT6* was estimated by the TASSEL software version 2.1 (Bradbury et al. 2007). Squared allele frequency correlations (D' and r^2) were chosen for LD calculations. The significance of LD between markers was determined by a Fisher's exact test. The general linear model (GLM, not considering population structure) analysis and GLM + Q (considering population structure) in TASSEL was employed to identify associations. Those polymorphisms with $P < 0.05$ were considered significantly associated to the trait.

Gene prediction in the target genomic region was conducted based on the soybean genome information and bioinformatics. The flowering related candidate genes were identified after submitting the predicted genes to a BLASTP query of the UniRef database (http://www.ebi.ac.uk/uniref/) and synteny comparisons between soybean and other dicotyledonous plants.

Results

Phenotypic variation and population structure

The means, standard deviation, skew, broad sense heritability and percentage of phenotypic variation

explained by population structure of flowering time are shown in Table 1. The phenotypic values of the parents of showed flowering time ranged from 41 to 46 days and from 37 to 50 days in RILs over two years, respectively. There was obvious transgressive segregation for flowering time and the broad heritability in parents and RIL families averaging 91.8 and 92.2 % across two years (Table 1, Fig. 1). On the other hand, highly significant variation was observed among the 192 natural accessions, with flowering time per accession averaging 55–56 days across different years/sites and ranged from 27 to 77 days (Table 1, Fig. 1). The distribution of flowering time for these natural accessions is roughly normal, and its range exceeds that observed among the RILs at both sides of the distribution (Fig. 1). The ANOVA showed that the genotype effect and the genotype × environmental interaction effects were significant ($P < 0.01$) for flowering time (Table 1). In addition, The 80 SSR markers examined produced a total of 1,422 alleles among the 192 accessions assayed (Table S1). The average number of alleles per locus was 17.78 (ranging from 5 to 37). The mean PIC value was 0.81. Overall, it was clear that soybean accessions exhibited considerable natural variation in flowering time and showed very high genetic diversity.

The evaluation of the population structure of the 192 soybean accessions indicated that there was an increase in LnP(D) value with increasing k value. The ad hoc quantity (Δk) showed a much higher likelihood at $k = 2$ than at $k = 3$–10 (Fig. 2a), suggesting that the population could be clustered into two major subpopulations (Fig. 2b). This result was consistent with previous work, which population structure was evaluated using 1536 SNPs in the same population (Hao et al. 2012). Statistical description for flowering time in the two subpopulations was summarized in Table S2. Nearly threefold difference (27–77 days) for flowering time was found between accessions. While between two subpopulations, a difference of 5.4 days for flowering time was found because of most accessions is early flowering varieties in subpopulation two (S2). The results of one-way analysis of variance (ANOVA) between subpopulations showed that the variation of flowering time was significantly different between subpopulations (Table S2). In addition, the effects of population structure on soybean flowering time were various in the 192 natural accessions, with R^2 average of 3.2–5.3 % (Table 1).

Primary QTL mapping for flowering time by linkage analysis

To identify QTLs for flowering time in soybean, a genetic map comprising 306 markers has been constructed (Zhang et al. 2009) and 152 RILs were used to map QTLs for flowering time in 2006 and 2007, respectively. As a result, five QTLs were detected for flowering time and mapped to five chromosomes based on the pot experiments in two

Table 1 Descriptive statistics, ANOVA, broad sense heritability and percentage of phenotypic variation explained by population structure for flowering time

Plant material	Year/Site[a]	Size	Mean ± SD	(Min–Max)	Skew	G[b]	G × E[c]	H^{2d} (%)	R^{2e} (%)
RIL parents	2006NJ	2	43 ± 9	41–45	0.02	**	**	91.2	/
	2007NJ	2	44 ± 10	42–46	0.04	**	**	92.5	/
RIL families	2006NJ	152	41 ± 10	37–50	0.41	**	**	91.4	/
	2007NJ	144	42 ± 11	38–47	0.45	**	**	93.1	/
Natural accessions	2008NJ	176	55 ± 21	27–77	0.09	**	**	83.7	3.2
	2008HN	167	55 ± 23	31–72	0.02	**	**	84.5	5.3
	2010HN	164	56 ± 24	31–77	0.08	**	**	83.2	3.5

** Significant at $P < 0.01$

[a] NJ and HN denote the field experiments at Nanjin/Henan, China

[b] Genotype across different environments

[c] Genotype × environment

[d] Broad sense heritability

[e] Percentage of phenotypic variation explained by population structure

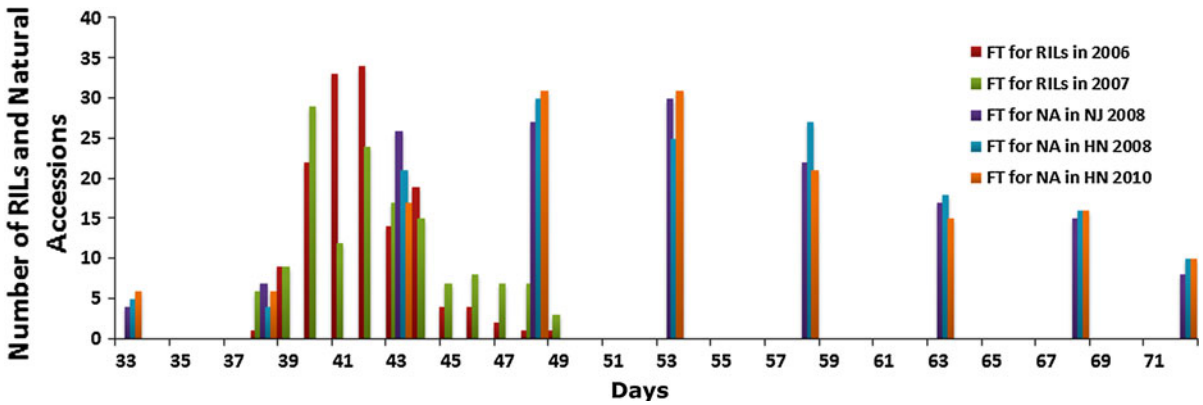

Fig. 1 Distribution of natural variation for flowering time (FT). The *x*-axis gives the days from the germination date to the mean flowering date. The distribution of flowering time scored for the 152 RILs in 2006 (*red*) and 2007 (*green*) and the 192 natural accessions (NA) in 2008 (Nanjing, *purple*; Henan, *blue*) and 2010 (Henan, *orange*)

years (Table S3). Among them, *qFT6*, *qFT8*, and *qFT20* were mapped on chromosome 6, 8 and 20 in 2006, while *qFT6*, *qFT11*, *qFT18*, and *qFT20* were mapped on chromosome 6, 11, 18 and 20 in 2007. The alleles from Bogao at *qFT6*, *qFT18* and *qFT20* loci flowering time shifted earlier by 0.68-1.15 days, whereas those at *qFT8* and *qFT11* from Nannong 94-156 prolonged flowering time by 0.52–0.93 days. From Two main QTLs were co-located flanked by Sat_402-Satt489 on chromosome 6 and Sat_367-Sat_337 on chromosome 20 across two years, indicating a stable genetic effect for flowering time in different environments (Table S3). The QTL, *qFT6*, which explained 23.1 and 27.0 % of phenotypic variation in different years, could be a target region for identifying genes associated with flowering time (Table S3, Fig. 3a). In addition, this locus had been previously identified by Yamanaka et al. (2001) as a QTL controlling soybean flowering.

Fine-mapping for flowering time by the partial scan of association analysis

To narrow down the major QTL region *qFT6*, which stably affected soybean flowering time across years, the 192 soybean accessions, eight previously developed SSR markers and 10 SNP markers inside this region were used to fine map *qFT6* (Fig. 4b). Firstly, we assessed the LD structure across the *qFT6* locus using eight SSR and 10 SNPs markers in the 192 soybean accessions. Two major LD blocks were found and one is including five loci within 2.4 Mb (positions 17.2-19.6 Mb, Table 2, Fig. S1). The pairwise r^2 of these markers in the LD blocks were higher and showed highly significant LD ($P < 0.001$) among the 192 soybean accessions. In addition, the scatter plots of the r^2 against distance also showed that there was a high LD in the *qFT6* locus (Fig. S2). The pairwise r^2 varied from 0.000 to 0.988 with an average of 0.106. LD decay curve showed that LD extends 0.35 Mb in chromosome 6 for the 192 soybean accessions. The short extent of LD across the *qFT6* locus thus provided sufficient genetic resolution for the following association mapping.

Then association mapping was performed by general linear model (GLM) analysis (not considering population structure) and GLM + *Q* (considering population structure) using TASSEL software (Fig. 3b). To reduce both false positive and false negative risks caused by population structure, only the markers that were detected by both analysis methods were taken into account in this study. The results from the partial scan of association analysis showed that 11 markers (Satt286, Satt277, BARC-014947-01929, Satt365, Satt557, BARC-040 213-07685, Satt658, Satt489, BARC-031099-06997, Satt289 and Satt100) were significantly associated with the flowering time in the *qFT6* genomic region (Table 2, Fig. 4b). From Table 2, BARC-014947-01929 and Satt365, which with the highest significant level were closely associated with flowering time ($P = 4.5 \times 10^{-5}$ and $P = 1.2 \times 10^{-7}$). Then the haplotypes analysis between every two markers and regression analysis of the haplotypes to phenotypic data were performed to determine the joint effect of those loci pairs associated with flowering time. Interestingly, the results showed that the haplotype between BARC-014947-01929 and

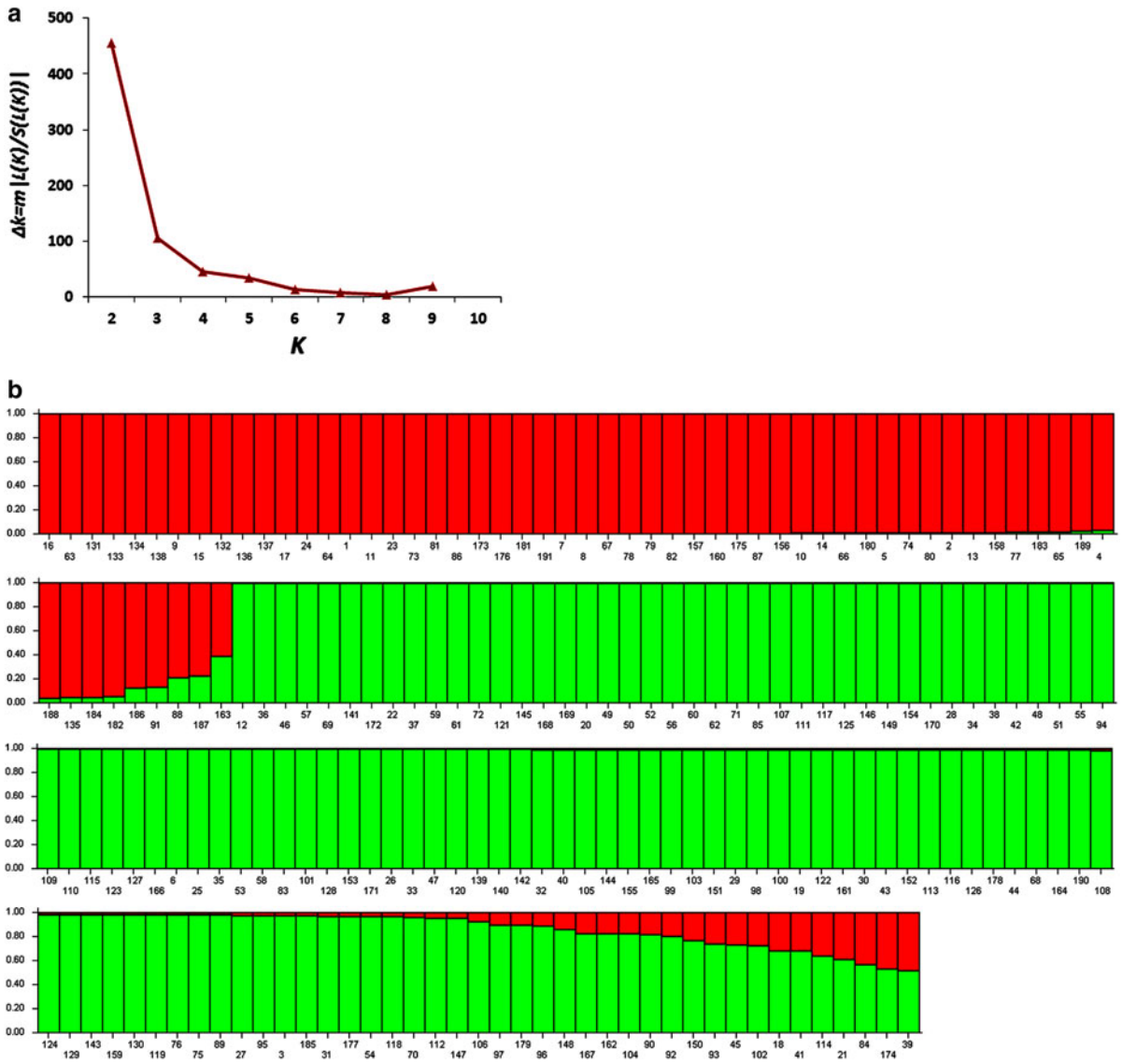

Fig. 2 Population structure of 192 soybean accessions. **a** *Δk* value was over six iterations of running with putative *k* ranging from 1 to 10. **b** Two major subpopulations for 192 soybean accessions, the numbers on the *x*-axis show the accession number and numbers on the *y*-axis show the membership probabilities into the subpopulations. The colors of the *bar* indicate the two subpopulations identified through the STRUCTURE program

Satt365 could explain more phenotypic variation (26.5 %) than any other combination of markers. These results suggested that the target gene was located in ~300 kb between BARC-014947-01929 and Satt365 (Fig. 4c).

Candidate genes predicted for flowering time

Comprehensive analysis (combined with soybean genome information, bioinformatics and genome comparison) of this region predicted 26 annotated genes (Table S4). Among them, three were considered as candidate genes (*Glyma06g22450, Glyma06g22650* and *Glyma06g22680*) related to flowering time or flower development after BLASTP queries of the UniRef database (http://www.ebi.ac.uk/uniref/) and synteny analyses between soybean and other dicotyledonous plants. These included the candidate gene *Glyma06g22450*, encoding a SBP domain, whose homologs play important roles in plant growth and

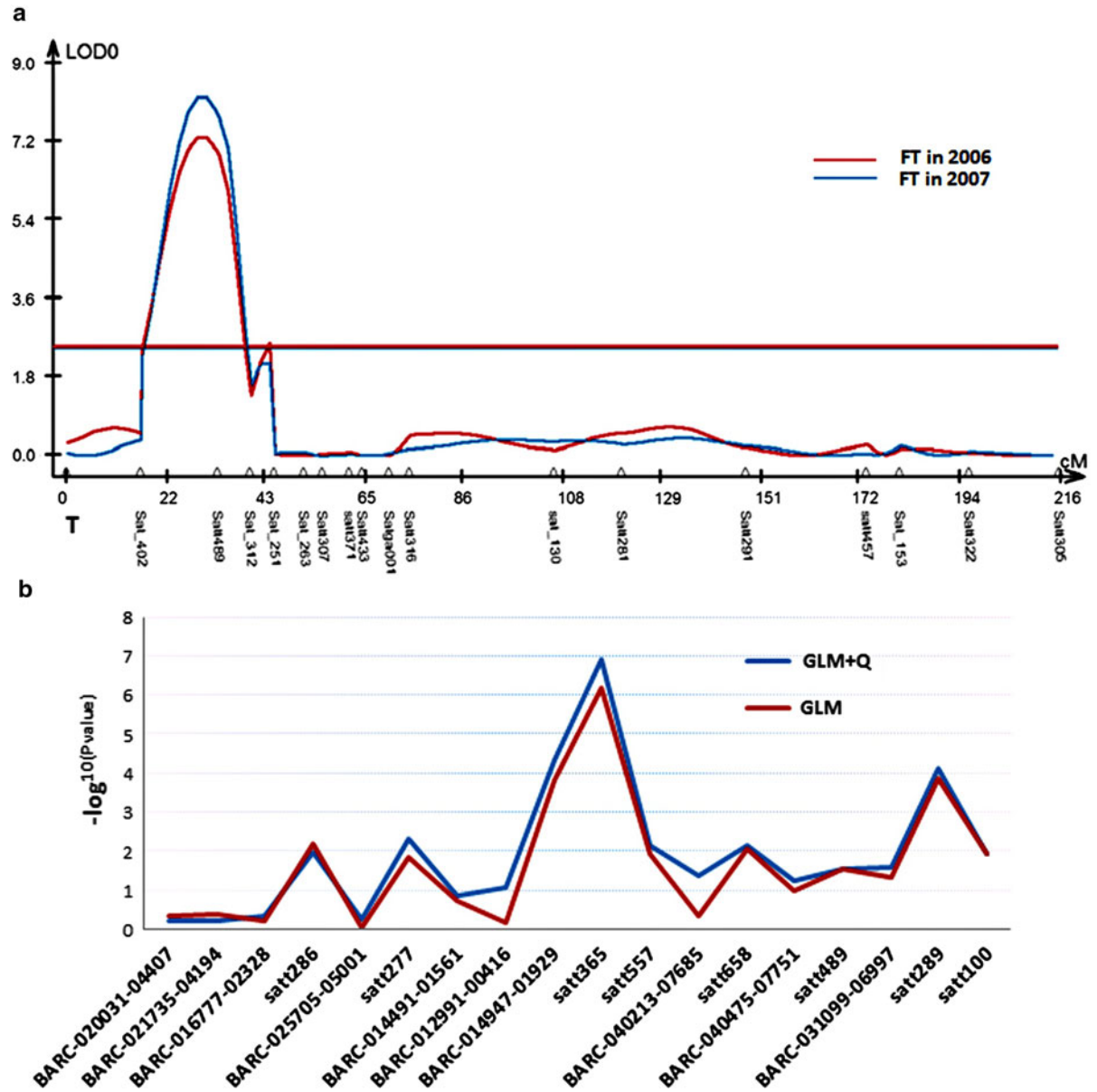

Fig. 3 QTL mapping for flowering time on chromosome 6 using 152 RILs and 192 natural accessions at two-environments. **a** The QTL for flowering time mapped on chromosome 6 using 152 RILs by linkage mapping in 2006 and 2007. **b** Fine mapping centered on Sat_402 and Satt489 in the *qFT6* region on Chromosome 6 by association mapping. GLM denote associations was identified by the general linear model analysis in TASSEL, not considering population structure; GLM + *Q* denote the estimated *Q* matrices were used in the association analysis, considering population structure

flower development in *Arabidopsis*, rice and *Antirrhinum majus* (Klein et al. 1996; Yang et al. 2008). Other studies showed that SBP box gene regulates early flower development because the SBP-box can bind to the promoter of floral meristem identity gene (Yamaguchi et al. 2009). The candidate gene *Glyma 06g22650*, encoding a MADS-Box protein (SRF-type transcription factor), whose homologs play fundamental roles during flower development in plants (Shore and Sharrocks 1995; Pelucchi et al. 2002). The other candidate gene *Glyma06g22680*, encoding a MYB related gene whose homologs play an essential

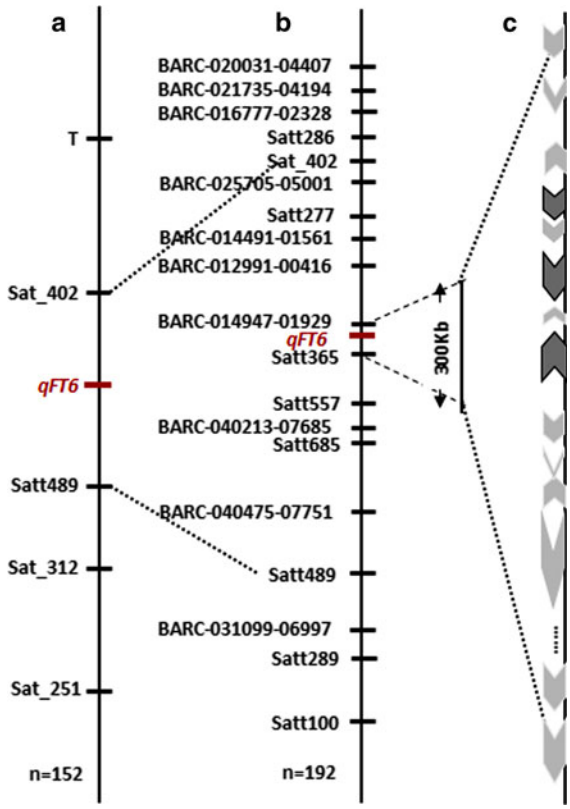

Fig. 4 Fine mapping and candidate genes predict in the *qFT6* region. **a** A flowering time QTL *qFT6* was mapped to the interval between markers Sat_402 and Satt489 on soybean chromosome 6. **b** This QTL *qFT6* was further delimited to ~300 kb region on chromosome 6 using the 192 natural accessions. **c** The *black arrow* indicates the site of the predicted gene between the markers BARC-014947-01929 and Satt365

role in flower development in plants (Preston et al. 2004; Higginson et al. 2003; Li et al. 2006). The candidate genes information identified in this study will be an important resource in marker-assisted selection for flowering time and for cloning the gene.

Discussion

Combined linkage and association mapping is powerful for fine mapping flowering time QTL in soybean

To date, the combined method (combined linkage and association mapping) for fine mapping have been used widely in humans and animals. These studies showed that fine mapping of an identified chromosomal region

Table 2 The markers associated with flowering time (mean value) and their phenotypic variation

Markers	Position (Mb)	P^a	$R^{2\ b}$
BARC-020031-04407	16.0	0.559	0.005
BARC-021735-04194	16.1	0.586	0.011
BARC-016777-02328	16.3	0.417	0.002
Satt286	16.3	0.011	0.110
BARC-025705-05001	16.7	0.507	0.002
Satt277	17.2	0.005	0.106
BARC-014491-01561	17.4	0.134	0.024
BARC-012991-00416	17.6	0.085	0.006
BARC-014947-01929	19.3	0.000	0.091
Satt365	19.6	0.000	0.221
Satt557	20.2	0.007	0.090
BARC-040213-07685	20.6	0.040	0.011
Satt658	20.6	0.007	0.086
BARC-040475-07751	21.7	0.056	0.031
satt489	23.81	0.027	0.067
BARC-031099-06997	27.1	0.025	0.043
Satt289	27.6	0.000	0.102
Satt100	30.5	0.012	0.147

[a] P value from ANOVA analysis of the mean value of flowering time in different years and sites

[b] R^2 showing percentage phenotypic variation explained from analysis of variance (ANOVA)

using the combined method could greatly reduce the QTL interval and was an important step toward identification of the gene and its causative allele (Olsen et al. 2005; Cohen-Zinder et al. 2005). However, the application of this effective method is very limited in plants. As we know, though it is not difficult to map a QTL to a 5–20 cM interval using primary segregating populations (such as F_2, $F_{2:3}$, RILs, etc.), cloning the underlying gene from a large interval is still a big challenge. In the present study, therefore, we combined linkage and association mapping that takes advantage of current and historical recombination events for QTL fine mapping in soybean based on the publicly available soybean genome information (Schmutz et al. 2010). The results showed that the combined method clearly outperforms each method used in isolation. For example, association mapping improves the positioning accuracy and linkage mapping increases our ability to distinguish true from false associations finely mapped by association mapping. In this study, population structure was

not the dominant factor in variation of flowering time, but its effect with a 4 % average effect can be seen in this trait. Therefore, controlling for population structure is necessary for reducing the false associations. This empirical result supports the notion that linkage and association mapping are complementary methods (Liu et al. 2007; Manenti et al. 2009). This may be a solution that is especially powerful for quantitative traits such as flowering time, which variation overlaps with population structure.

Although the significant QTL affecting reproductive traits including flowering and maturity in soybean mapping to chromosome 6 was reported in several previous studies (Yamanaka et al. 2005; Su et al. 2010), in this study, the potential benefit of dual linkage and association mapping would be narrowing QTL intervals in RIL families and greatly increase our power to finely map genomic regions associated with phenotypic variation. At the first stage, primary linkage analysis using a 152 RILs confirmed the existence of the QTL affecting flowering time at a very high significance level ($P < 0.001$) but suggested the resolution was very low (8.6 cM). To refine the position of the QTL, association analysis of a partial scan of the major QTL region (*qFT6*) was then performed. A highly associated marker interval (BARC-014947-01929-Satt365), which can explain more phenotypic variation than any other haplotype combination of markers, was selected. These results indicate that combining linkage and association mapping provides an efficient and precise approach to further elucidate the genetic basis of phenotypes and fine mapping of soybean genes.

Favorable alleles for flowering time in research and soybean breeding

Marker assisted selection (MAS) for natural variation has been limited by resolution and germplasm diversity. The association mapping approach used here allows for rapid generation of selectable markers based on the performance of diverse germplasms. In addition, this provides more relevant markers in a broad genetic background, enabling breeders to search for favorable alleles in locally adapted germplasms. In this study, a causal base transition (A/T) of the SNP marker BARC-014947-01929 was highly associated with soybean flowering time, which leads to an 11.1 days variation of flowering time in different alleles. This polymorphism was corresponding to the trait variance and thus can be considered as the candidate sites for functional molecular markers. In addition, using the method of analyzing a 'null allele' (Breseghello and Sorrells 2006), we found that 20 accessions with the allele A_{283} at the Satt365 produce significantly prolonged flowering time (average 58.8 days) than the lines carrying other alleles. There are eight accessions with the allele A_{358} at the Satt365 shifted earlier for flowering time (average 37.9 days) than the lines carrying other alleles. Although these markers would be useful in MAS, further confirmation is necessary because the marker alleles are correlated with, not entirely predictive of, the gene alleles.

In summary, we identified five QTLs for flowering time; the QTL (*qFT6*) on soybean chromosome 6 was narrowed down to an interval spanning $\sim$300 kb. The markers tightly linked to the QTL in *qFT6* region would help breeders for MAS of appropriate accessions to flowering time. Additionally, the results of this study provide important information for map-based cloning of the causal genes for flowering time.

Acknowledgments We thank Mr. Wang Wen Liang at Henan Agricultural University for assistance in investigating flowering time and performing the field experiments. Dr Zhiwu Zhang from Cornell University is thanked for his critical reading of this manuscript. This work was supported by National Basic Research Program of China (973 Program) (2010CB125906, 2009CB118400); National Natural Science Foundation of China (31000718, 31171573, 31201230), and Jiangsu Provincial Programs (BE2012328, BK2012768).

References

Andersson L (2001) Genetic dissection of phenotypic diversity in farm animals. Nat Rev Genet 2:130–137

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genet 6:e1000940

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

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1172

Cheng L, Wang Y, Zhang C, Wu C, Xu J, Zhu H, Leng J, Bai Y, Guan R, Hou W, Zhang L, Han T (2011) Genetic analysis and QTL detection of reproductive period and post-flowering photoperiod responses in soybean. Theor Appl Genet 123: 421–429

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Cohen-Zinder M, Seroussi E, Larkin DM, Loor JJ, Everts-van der Wind A, Lee JH, Drackley JK, Band MR, Hernandez A, Shani M (2005) Identification of a missense mutation in the bovine *ABCG2* gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res 15:936–944

Curtis D, Luzzi J, BM Hume D (2000) Agronomic and phenological differences of soybean isolines differing in maturity and growth habit. Crop Sci 40:1624–1629

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

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

Fehr W, Caviness C, Burmood D, Pennington J (1971) Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. Crop Sci 11:929–931

Goddard M (2005) The use of association to map quantitative trait loci. Aust J Exp Agr 45:837–845

Hanson W (1985) Association of seed yield with partitioned lengths of the reproductive period in soybean genotypes. Crop Sci 25:525–529

Hao D, Cheng H, Yin Z, Cui S, Zhang D, Wang H, Yu D (2012) Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. Theor Appl Genet 124:447–458

Higginson T, Li SF, Parish RW (2003) *AtMYB103* regulates tapetum and trichome development in Arabidopsis thaliana. Plant J 35:177–192

Hu GL, Zhang DL, Pan HQ, Li B, Wu JT, Zhou XY, Zhang QY, Zhou L, Yao GX, Li JZ (2011) Fine mapping of the awn gene on chromosome 4 in rice by association and linkage analyses. Chinese Sci Bull 56:835–839

Jun TH, Van K, Kim MY, Lee SH, Walker DR (2008) Association analysis using SSR markers to find QTL for seed protein content in soybean. Euphytica 162:179–191

Kantolic AG, Slafer GA (2007) Development and seed number in indeterminate soybean as affected by timing and duration of exposure to long photoperiods after flowering. Ann Bot 99:925–933

Keim P, Diers B, Olson T, Shoemaker R (1990) RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics 126:735–742

Klein J, Saedler H, Huijser P (1996) A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. Mol Gen Genet 250:7–16

Li J, Yang X, Wang Y, Li X, Gao Z, Pei M, Chen Z, Qu LJ, Gu H (2006) Two groups of MYB transcription factors share a motif which enhances trans-activation activity. Biochem Bioph Res Co 341:1155–1163

Li L, Li H, Li Q, Yang X, Zheng D, Warburton M, Chai Y, Zhang P, Guo Y, Yan J (2011) An 11-bp Insertion in Zea mays fatb reduces the palmitic acid content of fatty acids in maize grain. PLoS ONE 6:e24699

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Liu P, Vikis H, Lu Y, Wang D, You M (2007) Large-scale in silico mapping of complex quantitative traits in inbred mice. PLoS ONE 2:e651

Liu B, Kanazawa A, Matsumura H, Takahashi R, Harada K, Abe J (2008) Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. Genetics 180:995–1007

Lu Y, Zhang S, Shah T, Xie C, Hao Z, Li X, Farkhari M, Ribaut JM, Cao M, Rong T (2010) Joint linkage-association mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. P Natl Acad Sci 107:19585–19590

Manenti G, Galvan A, Pettinicchio A, Trincucci G, Spada E, Zolin A, Milani S, Gonzalez-Neira A, Dragani TA (2009) Mouse genome-wide association mapping needs linkage analysis to avoid false-positive loci. PLoS Genet 5: e1000331

Mansur L, Lark K, Kross H, Oliveira A (1993) Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean [*Glycine max* (L.) Merr.]. Theor Appl Genet 86:907–913

Mansur L, Orf J, Chase K, Jarvik T, Cregan P, Lark K (1996) Genetic mapping of agronomic traits using recombinant inbred lines of soybean. Crop Sci 36:1327–1336

Meuwissen THE, Karlsen A, Lien S, Olsaker I, Goddard ME (2002) Fine mapping of a quantitative trait locus for twinning rate using combined linkage and association mapping. Genetics 161:373–379

Molnar SJ, Rai S, Charette M, Cober ER (2003) Simple sequence repeat (SSR) markers linked to E1, E3, E4, and E7 maturity genes in soybean. Genome 46:1024–1036

Nemri A, Atwell S, Tarone AM, Huang YS, Zhao K, Studholme DJ, Nordborg M, Jones JDG (2010) Genome-wide survey of *Arabidopsis* natural variation in downy mildew resistance using combined association and linkage mapping. Proc Natl Acad Sci 107:10302–10307

Olsen HG, Lien S, Gautier M, Nilsen H, Roseth A, Berg PR, Sundsaasen KK, Svendsen M, Meuwissen THE (2005) Mapping of a milk production quantitative trait locus to a 420-kb region on bovine chromosome 6. Genetics 169:275–283

Orf J, Chase K, Jarvik T, Mansur L, Cregan P, Adler F, Lark K (1999) Genetics of soybean agronomic traits: comparison of three related recombinant inbred populations. Crop Sci 39:1642–1651

Pelucchi N, Fornara F, Favalli C, Masiero S, Lago C, Pe EM, Colombo L, Kater MM (2002) Comparative analysis of rice MADS-box genes expressed during flower development. Sex Plant Reprod 15:113–122

Preston J, Wheeler J, Heazlewood J, Li SF, Parish RW (2004) *AtMYB32* is required for normal pollen development in Arabidopsis thaliana. Plant J 40:979–995

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183

Shore P, Sharrocks AD (1995) The MADS-box family of transcription factors. Eur J Biochem 229:1–13

Stracke S, Presterl T, Stein N, Perovic D, Ordon F, Graner A (2007) Effects of introgression and recombination on haplotype structure and linkage disequilibrium surrounding a locus encoding *Bymovirus* resistance in barley. Genetics 175:805–817

Su CF, Lu WG, Zhao TJ, Gai JY (2010) Verification and fine-mapping of QTLs conferring days to flowering in soybean using residual heterozygous lines. Chinese Sci Bull 55:499–508

Tasma I, Lorenzen L, Green D, Shoemaker R (2001) Mapping genetic loci for flowering time, maturity, and photoperiod insensitivity in soybean. Mol Breeding 8:25–35

Tuinstra M, Ejeta G, Goldsbrough P (1997) Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. Theor Appl Genet 95:1005–1011

Wang Y (2001) Establishment and adjustment of RIL population and its application to map construction, mapping genes resistant to SMV and QTL analysis of agronomic and quality traits in soybeans. Nanjing Agricultural University, Doctor Dissertation

Wang S, Basten C, Zeng Z (2005) Windows QTL cartographer version 2.5. Statistical genetics. North Carolina State University, Raleigh

Watanabe S, Hideshima R, Xia Z, Tsubokura Y, Sato S, Nakamoto Y, Yamanaka N, Takahashi R, Ishimoto M, Anai T (2009) Map-based cloning of the gene associated with the soybean maturity locus E3. Genetics 182:1251–1262

Yamaguchi A, Wu MF, Yang L, Wu G, Poethig RS, Wagner D (2009) The microRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of *LEAFY*, *FRUITFULL*, and *APETALA1*. Dev Cell 17:268–278

Yamanaka N, Ninomiya S, Hoshi M, Tsubokura Y, Yano M, Nagamura Y, Sasaki T, Harada K (2001) An informative linkage map of soybean reveals QTLs for flowering time, leaflet morphology and regions of segregation distortion. DNA Res 8:61–72

Yamanaka N, Watanabe S, Toda K, Hayashi M, Fuchigami H, Takahashi R, Harada K (2005) Fine mapping of the *FT1* locus for soybean flowering time using a residual heterozygous line derived from a recombinant inbred line. Theor Appl Genet 110:634–639

Yang Z, Wang X, Gu S, Hu Z, Xu H, Xu C (2008) Comparative study of SBP-box gene family in Arabidopsis and rice. Gene 407:1–11

Yu J, Pressoir G, Briggs W, Bi I, Yamasaki M, Doebley J, McMullen M, Gaut B, Nielsen D, Holland J (2005) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Zhang WK, Wang YJ, Luo GZ, Zhang JS, He CY, Wu XL, Gai JY, Chen SY (2004) QTL mapping of ten agronomic traits on the soybean [*Glycine max* (L.) Merr.] genetic map and their association with EST markers. Theor Appl Genet 108:1131–1139

Zhang D, Cheng H, Geng L, Kan G, Cui S, Meng Q, Gai J, Yu D (2009) Detection of quantitative trait loci for phosphorus deficiency tolerance at soybean seedling stage. Euphytica 167:313–322