Published online September 14, 2017

RESEARCH

Quantitative Trait Locus Mapping of Soybean Maturity Gene *E6*

Xiaoming Li, Chao Fang, Meilan Xu, Fengge Zhang, Sijia Lu, Haiyang Nan, Tong Su, Shichen Li, Xiaohui Zhao, Lingping Kong, Xiaohui Yuan, Baohui Liu,* Jun Abe,* Elroy R. Cober,* and Fanjiang Kong*

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] sensitivity to photoperiod determines adaptation to a specific range of latitudes for soybean cultivars. When temperate-adapted soybean cultivars are grown in low latitude under short day conditions, they flower early, resulting in low grain yield, and consequently limiting their utility in tropical areas. Most cultivars adapted to low-latitude environments have the trait of delayed flowering under short day conditions, and this trait is commonly called long juvenile (LJ). In this study, the *E6* locus, the classical locus conditioning the LJ trait, was molecularly mapped on Gm04 near single-nucleotide polymorphism marker HRM101. Testcross, genetic mapping, and sequencing suggest that the *E6* and *J* loci might be tightly linked. Genetic interaction evaluation between *E6* and *E1* suggests that *E6* has a suppressive effect on *E1* and that the function of *E6* is dependent on *E1*. The tagging markers for *E6* are very useful for molecular breeding for wide adaptation and stable productivity of soybean under low-latitude environments. Molecular identification and functional characterization of the *E6* gene will greatly facilitate the understanding of the genetic and molecular mechanisms underlying the LJ trait.

X. Li, C. Fang, M. Xu, F. Zhang, S. Lu, H. Nan, T. Su, S. Li, X. Zhao, L. Kong, X. Yuan, B. Liu, and F. Kong, The Key Laboratory of Soybean Molecular Design Breeding, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, No. 138 Haping Rd., Nangang District, Harbin 150081, China; S. Lu, H. Nan, X. Zhao, B. Liu, and F. Kong, School of Life Sciences, Guangzhou Univ., Guangzhou 510006, China; X. Li, C. Fang, F. Zhang, T. Su, S. Li, and L. Kong, Univ. of Chinese Academy of Sciences, Beijing 100049, China; X. Li, Heilongjiang Academy of Agricultural Sciences, 368 Xuefu Rd., Harbin 150086, China; J. Abe, Research Faculty of Agriculture, Hokkaido Univ., Sapporo, Hokkaido 060-8589, Japan; E.R. Cober, Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Bldg. #110, Central Experimental Farm, Ottawa, ON K1A 0C6, Canada. X. Li, C. Fang and M. Xu, F. Zhang and S. Lu contributed equally to this work. Received 18 Feb. 2017. Accepted 6 June 2017. *Corresponding author (kongfj@iga.ac.cn; elroy.cober@agr.gc.ca; jabe@res.agr.hokudai.ac.jp; liubh@iga.ac.cn). Assigned to Associate Editor Zenglu Li.

Abbreviations: CJ, conventional juvenile; DAE, days after emergence; HRM, high-resolution melting; Indel, insertion and deletion; LD, long day; LJ, long juvenile; LOD, logarithm of odds; QTL, quantitative trait locus; PCR, polymerase chain reaction; PGH, F_2 population of the cross between Paranagoiana and Harosoy; PGI, F_2 population of the cross between Paranagoiana and PI 159925; PGO, F_2 population of the cross between Paranagoiana and OT94-47; SD, short day; SNP, single-nucleotide polymorphism; SSR, simple sequence repeat.

Soybean [*Glycine max* (L.) Merr.] is an important crop for human consumption, for animal feed, and for use as biodiesel fuel. Time to flowering and maturity significantly affects soybean adaptation and grain yield (Cober and Morrison, 2010). Soybean is a short day (SD) plant, and SD results in early flowering, whereas long day (LD) delays flowering (Watanabe et al., 2012). Soybean, however, can grow over a wide range of latitudes, at

Published in Crop Sci. 57:2547–2554 (2017).
doi: 10.2135/cropsci2017.02.0106

least from 50° N to 35° S (Cao et al., 2016). The wide adaptability of soybean has been created by natural variation in the major genes and quantitative trait loci (QTLs) controlling flowering time (Kong et al., 2014). In soybean, 11 maturity loci (*E1–E10* and *J*) that control flowering time and maturity have been previously identified and characterized at the phenotypic and genetic levels (Bernard, 1971; Buzzell, 1971; Buzzell and Voldeng, 1980; McBlain and Bernard, 1987; Ray et al., 1995; Bonato and Vello, 1999; Cober and Voldeng, 2001; Cober et al., 2010; Kong et al., 2014; Samanfar et al., 2016). Among them, *E1* was cloned by a map-based approach and assumed to be a legume-specific transcription factor that has a putative nuclear localization signal and a B3 distantly related domain (Xia et al., 2012); *E2* was identified as a ortholog of the *Arabidopsis GIGANTEA* gene (Watanabe et al., 2011); and *E3* and *E4* were confirmed as *PHYA* homologs by a map-based cloning (Watanabe et al., 2009) and a candidate gene approach (Liu et al., 2008), respectively. The various allelic combinations at the *E1*, *E3*, and *E4* loci condition soybean flowering time, as well as preflowering and postflowering photoperiod responses, and greatly contribute to the wide adaptability in soybean (Tsubokura et al., 2013; Xu et al., 2013; Jiang et al., 2014). In addition, two *FLOWERING LOCUS T* homologs, *GmFT2a* and *GmFT5a*, are involved in the transition to flowering, and these two *FT* homologs coordinately control flowering in soybean (Kong et al., 2010; Nan et al., 2014). The maturity genes *E1*, *E2*, *E3*, and *E4* downregulate *GmFT2a* (*E9*) and *GmFT5a* expression to delay flowering and maturation under the LD condition, suggesting that *GmFT2a* and *GmFT5a* are the soybean flowering integrators and the major targets in the control of flowering (Kong et al., 2010; Thakare et al., 2011; Watanabe et al., 2011; Xia et al., 2012, Nan et al., 2014).

Soybean sensitivity to photoperiod determines the limits of the sowing period for a specific latitude and prevents adaptation to wider ranges of latitude. When soybean cultivars are grown under SD conditions, cultivars with sensitivity to photoperiod flower early, resulting in low grain yield and consequently limiting their growing area in SD environments (Destro et al., 2001). It is therefore important to understand the genetic control of delayed flowering time under SD environments. This trait was termed the long juvenile (LJ) trait (Parvez and Gardner, 1987; Sinclair and Hinson, 1992; Ray et al., 1995). The LJ trait, which delays flowering under SD conditions, has been identified in tropical soybean cultivars. The introduction of the LJ characteristic in soybean has made its cultivation possible in regions with latitudes lower than 15° S (Destro et al., 2001). The LJ trait plays a pivotal role in extending the range of adaptation of soybean cultivars to lower latitudes and to new management schemes with shifted sowing dates in tropical countries (Destro et al., 2001). It has been reported that the northward expansion of soybean production in South America, where more extensive research has been performed, is dependent on the LJ trait (Spehar, 1995). However, the genetic control mechanism for this trait remains elusive. Two genes, *J* and *E6*, were reported to play an important role in the LJ trait (Ray et al., 1995; Bonato and Vello, 1999). The single locus *J* has been identified in a number of crosses with PI 159925 (Ray et al., 1995). The single locus *E6* was a natural mutation in cultivar Parana and produced the LJ cultivars Paranagoiana and SS-1 (Bonato and Vello, 1999). Recently, an F_2 population resulting from a cross between conventional juvenile (CJ) line OT94-47 and the LJ line Paranagoiana exhibited a 15:1 early/late flowering ratio in 12-h photoperiods. A similar 15:1 ratio was observed in offspring of a cross between CJ line OT94-47 and the LJ line PI 159925 (Cober, 2011). These results suggested that the LJ trait is conditioned by two recessive alleles in PI 159925 and Paranagoiana (Cober, 2011). Other studies of LJ parents also suggested that recessive alleles at two or three loci control the LJ trait (Carpentieri-Pípolo et al., 2000, 2002). The *J* gene has been mapped to the soybean linkage group Gm04 between the simple sequence repeat (SSR) markers Sat_337 and Satt396, where the genetic distance between the *J* allele and the closest marker, Sat_337, is 0.7 cM (Cairo et al., 2002, 2009). Recently, *J* had been molecularly identified as an orthologue of *Arabidopsis EARLY FLOWERING 3* (*ELF3*), and natural variations at *J* locus improved soybean adaptation in low-latitude regions and enhanced soybean yield (Yue et al., 2016; Lu et al., 2017). *J* depends genetically on the legume-specific flowering repressor *E1*, and J protein physically associates with the *E1* promoter to downregulate its transcription, relieving repression of two important *FT* genes and promoting flowering under SD. (Lu et al., 2017). Using a different LJ cultivar from Thailand, a new major QTL, *qFT-J2*, conditioning the LJ trait was identified and mapped on Gm16, where the flowering gene *GmFT2a* is located, but the QTL near the *J* locus was not detected (Lu et al., 2015). This suggested that different genes and QTLs condition the LJ trait in different genetic backgrounds from different geographical regions.

Although variation in the *J* gene clearly plays important role in conferring the LJ trait and has been widely deployed in several major soybean production regions, the existence of many late-flowering lines from low-latitude regions that carry an apparently functional *J* allele suggests that, on a global scale, it is not the only locus responsible for this trait (Lu et al., 2017). The *E6* locus, another classical LJ locus, plays important roles for soybean adaptation and yield improvement in low-latitude regions (Bonato and Vello 1999). In spite of the importance of the *E6* for soybean adaptation and yield productivity in tropical regions, the genetic information regarding *E6* is very

limited. The objectives of this study were (i) to QTL map the *E6* locus, (ii) to study the genetic relationship between the *E6* and *E1* loci, and (iii) to develop molecular markers of the *E6* gene for molecular breeding in tropics.

MATERIALS AND METHODS

Genetic Populations and Growth Conditions

To map the *E6* locus, an F_2 population (hereafter named PGH) of the cross between LJ line Paranagoiana (*e6e6*, PI 628880) and CJ line Harosoy (*E6E6*, PI 548573) was developed. An F_2 population (hereafter named PGO) from a cross between Paranagoiana and OT94-47, developed previously, was also used for *E6* mapping (Cober, 2011). To understand the genetic effect between *J* and *E6* loci, a F_2 population (hereafter named PGI) from the cross between two LJ lines, Paranagoiana (*e6e6*) and PI 159925 (*jj*), was developed.

The F_2 populations PGH and PGI and the parental lines were sown in pots in growth cabinets under SD conditions (12 h light and 12 h dark). Each pot contained four plants for the two F_2 populations. Days to flowering were recorded at the R1 stage (first open flower appeared) for each plant (Fehr et al., 1971). The R1 values reported for the three parents Paranagoiana, PI 159925, and Harosoy are the means from five plants. The R1 data from the cross of PGO were reported previously (Cober, 2011).

Molecular Analysis

DNA was extracted individually from leaves of plants, as described by Kong et al. (2010). Simple sequence repeat markers were selected from those designed and mapped by Cregan et al. (1999). Insertion and deletion (Indel) markers and single-nucleotide polymorphism (SNP) markers (Table 1) were developed in this study according to resequencing data from the parents Harosoy, PI 159925, and Paranagoiana. The SNP markers were developed and detected by the high-resolution melting (HRM) approach (Li et al., 2010). Briefly, HRM SNP marker analysis contained two rounds of polymerase chain reaction (PCR) amplification, the first round of PCR is to amplify the specific fragment containing the SNP site, and the second round of PCR was run in a LightScanner HR196 (Idaho Technology) to detect the SNP signal (Li et al., 2010). The whole-genome resequencing of Harosoy, PI 159925, and Paranagoiana and the Indel analysis using the software of SOAPindel was conducted by BGI-Shenzhen, China, as described previously (Kong et al., 2014). The procedures for PCR and gel electrophoresis followed a standardized procedure, as reported earlier (Kong et al., 2014; Lu et al., 2016). Marker order and distance were determined by Map Manager Program QTXb20 (Lu et al., 2015) using the Kosambi function and a criterion of 0.001 probability (df = 1), and a genetic map was constructed. The multiple QTL model from MapQTL 5.0 (Van Ooijen, 2004), interval mapping from QTL IciMapping (Meng et al., 2015), and composite interval mapping 3.0 from Windows QTL Cartographer 2.5 (Wang et al., 2012) were used for QTL detection. A logarithm of odds (LOD) score of 3.0 was used as a minimum to declare the significance of a QTL in a particular genomic region. One thousand permutations at a 0.05 probability were conducted to identify the genomewide LOD score (Churchill and Doerge, 1994). Sequencing of *E1* and *J* genes was performed according to previous reports (Xia et al., 2012; Lu et al., 2017). Genotyping of $e1^{as}$ and $e1^{fs}$ alleles using their functional markers was conducted as described previously (Xu et al., 2013; Jiang et al., 2014).

RESULTS AND DISCUSSION

To study the *E6* locus, two F_2 populations named PGH and PGO from a cross between a Brazilian LJ cultivar, Paranagoiana (*e6e6*, Table 2), in which the *E6* gene was originally identified (Bonato and Vello, 1999) and two CJ lines, Harosoy (*E6E6*) and OT94–47 (*E6E6*), (Cober, 2011) were used. Under SD conditions (12 h light and 12 h dark), large variations in flowering time (R1 stage) were observed in both F_2 populations (Table 3). The variations of flowering time ranged from 27 to 61 d after emergence (DAE) in PGH and from 25 to 54 DAE in PGO (Table 3). We next constructed linkage maps of both crosses using different molecular markers (Table 1, Supplemental Table 1). In the F_2 population of PGH, 162 polymorphic markers were identified between parents Paranagoiana and Harosoy. Twenty linkage groups were constructed and covered the genetic length of 1681 cM (Table 1). In the F_2 population of PGO, 216 polymorphic markers were used to construct 20 linkage groups, which covered 2364 cM genetic length (Table 1). We used three approaches—multiple QTL mapping from MapQTL 5.0, interval mapping from QTL IciMapping, and composite interval mapping from WinQTL cartographer—to conduct whole-chromosome scans to identify consensus QTLs. In the F_2 population of PGH, a major QTL conditioning flowering time under SD, *qFT-C1*, located on chromosome 4 (Gm04), was consistently identified by all three methods (Fig. 1a, Table 4). The same major QTL of *qFT-C1* was also consistently identified by all three methods in the second F_2 population, PGO (Fig. 1b, Table 4). In addition, a second QTL, *qFT-C2*, coinciding with *E1*, a major soybean maturity locus (Xia et al., 2012), was consistently detected in both crosses PGH and PGO (Fig. 1a

Table 1. Linkage groups obtained from four soybean F_2 populations.

Crosses	Population names	No. of F_2 plants	No. of linked markers†	No. of linkage group	Total map length
					cM
Paranagoiana × Harosoy	PGH	184	162	20	1681
Paranagoiana × OT94-47	PGO	58	216	20	2364
Paranagoiana × PI 159925	PGI	126	163	20	1987

† Markers for linkage map construction are listed in Supplemental Table 1.

Table 2. Soybean materials used in this study.

Line name	Line ID	Genotype	Pedigree	Flower time (12 h light/12 h dark)
				d after emergence
Paranagoiana	PI 628880	*E1*, *E2*, *E3*, *E4*, *e6*	Natural mutation from cultivar Parana	62a†
PI 159925	PI 159925	*E1*, *E2*, *E3*, *E4*, *j*	Landrace collected from Peru	47b
Harosoy	PI 548573	*e1*^{as}, *e2*, *E3*, *E4*, *J*, *E6*	Mandarin (Ottawa) × 2/AK (Harrow)	28c
OT97-47	–	*e1*^{fs}, *e2*, *e3*, *e4*, *J*, *E6*	OT89-5/X2749-K1	27c

† Means within a column followed by the same letter are not significantly different (P = 0.01) according to Tukey's honestly significant difference test.

Table 3. Statistical analysis of the flowering times of F_2 populations in short day environments.

	F_2						
	Flowering time					Parents	
Population†	Min.	Max.	Mean ± SD	Kurtosis‡	Skewness§	Female	Male
	d after emergence						
PGH	27	61	36.6 ± 9.3a¶	−0.16	0.98	62.4 ± 0.5	28.2 ± 0.5
PGO	25	54	30.2 ± 6.9b	4.37	2.28	62.4 ± 0.5	27.6 ± 1.3
PGI	45	71	58.5 ± 6.2c	0.48	0.19	62.4 ± 0.5	47.6 ± 0.8

† PGH, Paranagoiana × Harosoy; PGO, Paranagoiana × OT94-47; PGI, PI 159925 × Paranagoiana.

‡ Kurtosis of the phenotypic trait.

§ Skewness of the phenotypic trait.

¶ Different lowercase letters (a, b, and c) indicate the extremely significant differences in short day environments ($P < 0.01$).

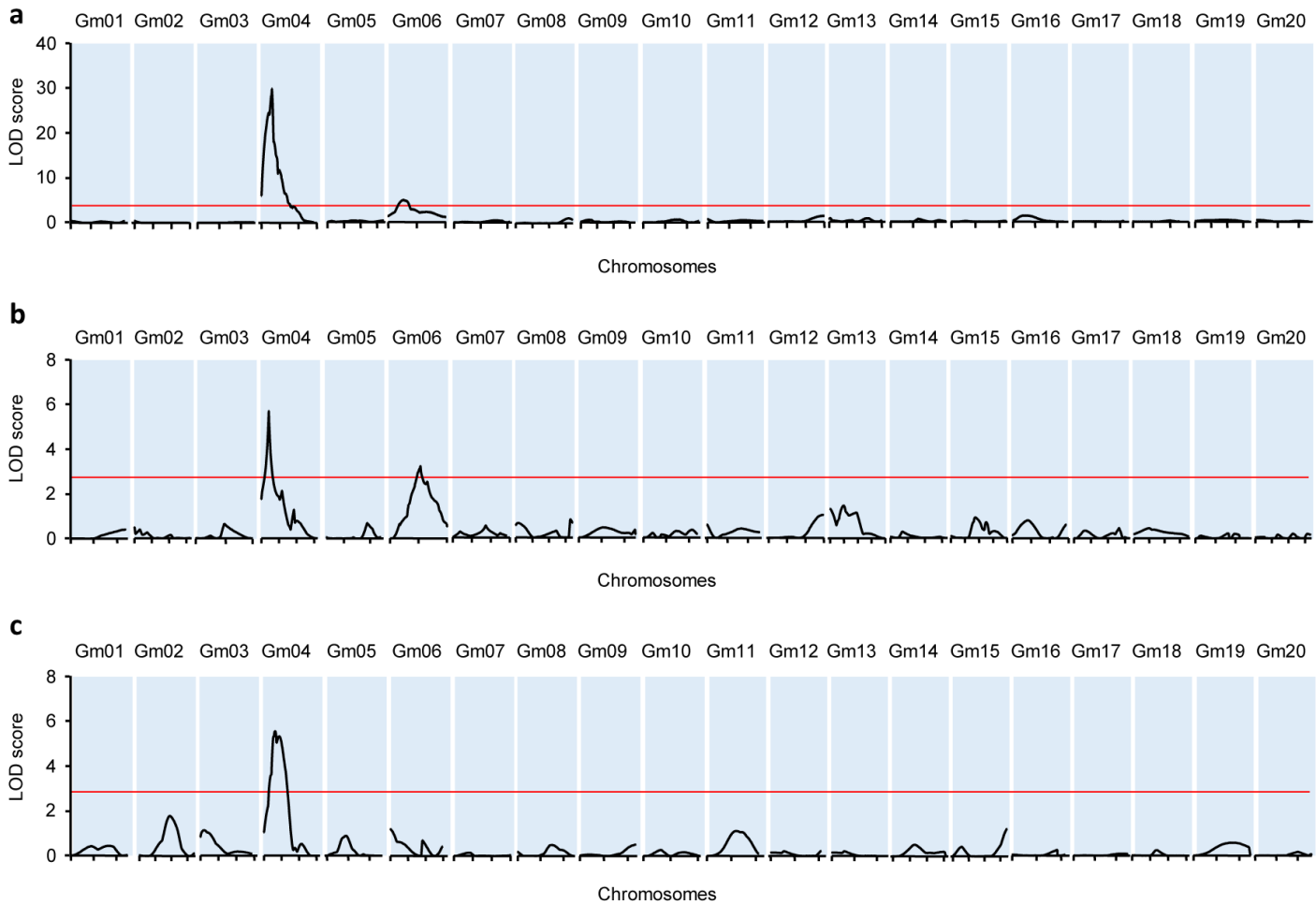

Fig. 1. Quantitative trait locus (QTL) mapping by multiple QTL mapping, implemented by MapQTL 5.0. Whole-chromosome scan of QTLs in three F_2 populations: (a) Paranagoiana × Harosoy, (b) Paranagoiana × OT94-47, and (c) PI 159925 × Paranagoiana. Red lines indicated the threshold of QTL detection. The detailed QTL information is indicated in Table 4. LOD, logarithm of odds.

Table 4. Identification of main-effect quantitative trait loci (QTLs).

Model†	Population‡	QTL name	Chromosome	Marker or interval§	Position¶	LOD#	R^2††	A‡‡	Threshold§§
					cM		%		
MQM	PGH	*qFT-C1*	Gm04	HRM101–ID04101	30.4	30.38	68.80	9.49	3.53
		qFT-C2	Gm06	E1	28.5	3.76	21.25	4.18	
	PGO	*qFT-C1*	Gm04	HRM101–ID04106	26.2	5.73	37.10	5.98	3.02
		qFT-C2	Gm06	E1	56.6	3.27	25.30	4.46	
	PGI	*qFT-C1*	Gm04	ID04090–ID04101	31.5	5.61	45.40	8.86	2.96
IM	PGH	*qFT-C1*	Gm04	HRM101–ID04101	32.0	37.02	51.60	9.67	8.71
		qFT-C2	Gm06	E1-ID06095	29.0	13.22	17.04	3.19	
	PGO	*qFT-C1*	Gm04	BAR04007– HRM101	18.8	16.32	44.67	9.37	4.13
		qFT-C2	Gm06	E1	59.2	5.94	27.14	5.77	
	PGI	*qFT-C1*	Gm04	ID04101–ID04134	40.5	7.50	44.90	4.26	3.74
		qFT-D1b	Gm02	ID02182–ID02230	58.0	4.94	13.47	1.96	
CIM	PGH	*qFT-C1*	Gm04	HRM101–ID04101	27.2	16.30	55.10	8.22	4.62
		qFT-C2	Gm06	E1–ID06095	29.7	5.87	26.45	3.08	
	PGO	*qFT-C1*	Gm04	HRM101–ID04106	29.2	16.29	52.10	8.62	5.80
		qFT-C2	Gm06	E1–ID06106	62.9	7.13	18.68	3.71	
	PGI	*qFT-C1*	Gm04	ID04090–ID04101	32.7	7.77	43.76	6.06	3.75

† Identification of QTLs by multiple QTL mapping (MQM), interval mapping (IM), and composite interval mapping (CIM), implemented by MapQTL 5.0, QTL IciMMapping, and WinQTLcart, respectively.

‡ PGH, Paranagoiana × Harosoy; PGO, Paranagoiana × OT94-47; PGI, PI 159925 × Paranagoiana.

§ Marker or interval: Markers or support intervals on the linkage map in which the LOD is the largest.

¶ Position: The LOD peak for candidate QTL on the genetic linkage map.

LOD, logarithm of odds.

†† Percentage of phenotypic variance explained by the QTL.

‡‡ The additive effects contributed by QTL. A positive value (+) of the additive effect indicates that the allele originated from the female; a negative value (–) of the additive effect indicates that the allele originated from the male.

§§ Significance at 0.05 probability by 1000 permutation tests.

and 1b, Table 4) under SD conditions. Only two major QTLs, *qFT-C1* and *qFT-C2*, were consistently identified from two crosses, PGH and PGO, in which *E6* and *E1* loci were segregating. Sequencing of *E1* genes in Paranagoiana and OT94-47 showed that Paranagoiana possesses the dominant allele of *E1*, whereas OT94-47 possesses the loss of function allele $e1^{fs}$ and Harosoy possesses the recessive weak allele $e1^{as}$ (Xia et al., 2012) (Fig. 2). These results suggested that the QTL *qFT-C2* corresponds to the *E1* locus in both F_2 populations PGH and PGO (Fig. 1a and 1b, Table 4). We therefore consider that the major QTL *qFT-C1* on Gm04 conditioning flowering time under SD corresponds to the *E6* locus.

Previously, the *J* locus had been mapped on Gm04 between SSR markers Sat_337 and Satt396 (Cairo et al., 2009; Yue et al., 2016; Lu et al., 2017), and the *E6* locus was mapped in the same position in two F_2 populations, PGH and PGO, in this study. This raised the question whether *J* and *E6* are the same gene. To investigate this question, we generated a F_2 population (named PGI) from a cross between Paranagoiana (*e6e6*) and PI 159925 (*jj*), in which the *J* allele was original identified (Ray et al., 1995). The flowering time in F_2 population PGI under SD segregated from 45 to 71 DAE (Table 3). We also generated 20 linkage groups by integrating 163 polymorphic markers in the F_2 population PGI (Table 1, Supplemental Table 1). Whole-chromosome QTL scans by the above mentioned three methods identified the same major QTL *qFT-C1* located in Gm04 (Fig. 1c, Table 4). The LOD scores were from 5.61 to 7.77, and the allele from Paranagoiana had an additive effect of 4.26 to 8.86 over that from PI 159925 (Table 4), which suggests that the allele from Paranagoiana delayed flowering under SD condition, in contrast with the allele from PI 159925. Since *E6* and *J* loci were comapped in the same position on Gm04 and *J* was the *Arabidopsis* flowering gene *ELF3* (Yue et al., 2016; Lu et al., 2017), this suggests that *E6* might be the *J* (*ELF3*) gene. We therefore sequenced the *J* gene in Paranagoiana, but there are no sequence polymorphisms between Paranagoiana and Harosoy. These results suggested that *E6* and *J* might be different genes but tightly linked together. In addition, a second QTL *qFT-D1b* on Gm02 conditioning the LJ trait was also identified by QTL IciMapping (Table 4), confirming that the LJ trait is a quantitative trait and conditioned by multiple loci (Carpentieri-Pípolo et al., 2000, 2002). Identification of the responsible genes of *E6* and *qFT-D1b* loci will facilitate the understanding of molecular mechanisms underlying LJ trait.

E1 is the legume-specific transcription factor and is the core soybean flowering suppressor that downregulates two soybean *FLOWERING LOCUS T* genes, *FT2a* (*E9*) and *FT5a* (Xia et al., 2012). Two major QTLs for

```
                    *        20         *        40         *        60         *        80
Paranagoiana : ATGAGCAACCCTTCAGATGAAAGGGAGCAGTGTCAAAAGAAGACGAAATCCACCATATGCGAAGCCTCTAACTTTAGGACATCAAGGA :  88
Harosoy      : ATGAGCAACCCTTCAGATGAAAGGGAGCAGTGTCAAAAGAAGACCAAATCCACCATATGCGAAGCCTCTAACTTTAGGACATCAAGGA :  88
OT94-47      : ATGAGCAACCCTTCAGATGAAAGGGAGCAGTGTCAAAAGAAGAGG-AATCCACCATATGCGAAGCCTCTAACTTTAGGACATCAAGGA :  87

                  *       100         *       120         *       140         *       160         *
Paranagoiana : GAAGATTCTGCAGCAACAACAAAAATGAAGAGGAGATGAACAATAAGGGAGTTTCAACAACACTGAAGCTTTACGATGATCCTTGGAA : 176
Harosoy      : GAAGATTCTGCAGCAACAACAAAAATGAAGAGGAGATGAACAATAAGGGAGTTTCAACAACACTGAAGCTTTACGATGATCCTTGGAA : 176
OT94-47      : GAAGATTCTGCAGCAACAACAAAAATGAAGAGGAGATGAACAATAAGGGAGTTTCAACAACACTGAAGCTTTACGATGATCCTTGGAA : 175

                   180         *       200         *       220         *       240         *       260
Paranagoiana : GATCAAGAAGACGCTAACCGATAGCGATTTGGGAATCCTAAGTAGACTCTTGCTGGCTGCAGATTTGGTGAAGAAACAAATTTTGCCT : 264
Harosoy      : GATCAAGAAGACGCTAACCGATAGCGATTTGGGAATCCTAAGTAGACTCTTGCTGGCTGCAGATTTGGTGAAGAAACAAATTTTGCCT : 264
OT94-47      : GATCAAGAAGACGCTAACCGATAGCGATTTGGGAATCCTAAGTAGACTCTTGCTGGCTGCAGATTTGGTGAAGAAACAAATTTTGCCT : 263

                       *       280         *       300         *       320         *       340         *
Paranagoiana : ATGTTGGGTGCATATCATGCAAGAGCTGCAGAAACTGAAGGGACCCCAGTTAGAGTTTGGGACATGGACACCAAATCCATGCACCAAC : 352
Harosoy      : ATGTTGGGTGCATATCATGCAAGAGCTGCAGAAACTGAAGGGACCCCAGTTAGAGTTTGGGACATGGACACCAAATCCATGCACCAAC : 352
OT94-47      : ATGTTGGGTGCATATCATGCAAGAGCTGCAGAAACTGAAGGGACCCCAGTTAGAGTTTGGGACATGGACACCAAATCCATGCACCAAC : 351

                     360         *       380         *       400         *       420         *       440
Paranagoiana : TCGTTCTAAAGCGATGGTCTTCATCCAAGAGCTATGTTCTTATTGGAAAGTGGAACCAAGATTTCGTCAGAAGAAGAGATCTCAGGAA : 440
Harosoy      : TCGTTCTAAAGCGATGGTCTTCATCCAAGAGCTATGTTCTTATTGGAAAGTGGAACCAAGATTTCGTCAGAAGAAGAGATCTCAGGAA : 440
OT94-47      : TCGTTCTAAAGCGATGGTCTTCATCCAAGAGCTATGTTCTTATTGGAAAGTGGAACCAAGATTTCGTCAGAAGAAGAGATCTCAGGAA : 439

                       *       460         *       480         *       500         *       520
Paranagoiana : AGGTGATGAGATCGGATTTCATTGGGATCCATATAATTGCGTTTTCAATTTCTGTGTCCTTAAACAAGCTATGCCAGAGAATTAA : 525
Harosoy      : AGGTGATGAGATCGGATTTCATTGGGATCCATATAATTGCGTTTTCAATTTCTGTGTCCTTAAACAAGCTATGCCAGAGAATTAA : 525
OT94-47      : AGGTGATGAGATCGGATTTCATTGGGATCCATATAATTGCGTTTTCAATTTCTGTGTCCTTAAACAAGCTATGCCAGAGAATTAA : 524
```

Fig. 2. Sequence alignment of *E1* in Paranagoiana, Harosoy, and OT94-47. Rectangular windows indicate different *E1* alleles, the $e1^{as}$ allele (G44C) and $e1^{fs}$ allele (A46-).

flowering time corresponding to *E6* and *E1* loci under SD were consistently identified from two crosses in PGH and PGO (Table 4). To understand the genetic effects between QTLs of *E6* and *E1*, we classified the allelic combinations using the tagging marker HRM101 of *E6* and *E1* functional markers in the two F_2 populations PGH and PGO (Fig. 3). In the population PGH, in homozygous dominant *E6E6* lines, there were no flowering time differences whether the *E1* allele was dominant or recessive, which suggested that *E6* had a suppressive effect on *E1* (Fig. 3a). In homozygous recessive *e6e6* lines, the partial loss of function allele $e1^{as}$ reduced the suppressive effect of *E6* on *E1*, which suggested that *E6* is dependent on *E1* (Fig. 3a). The suppressive effect of *E6* on *E1* on flowering

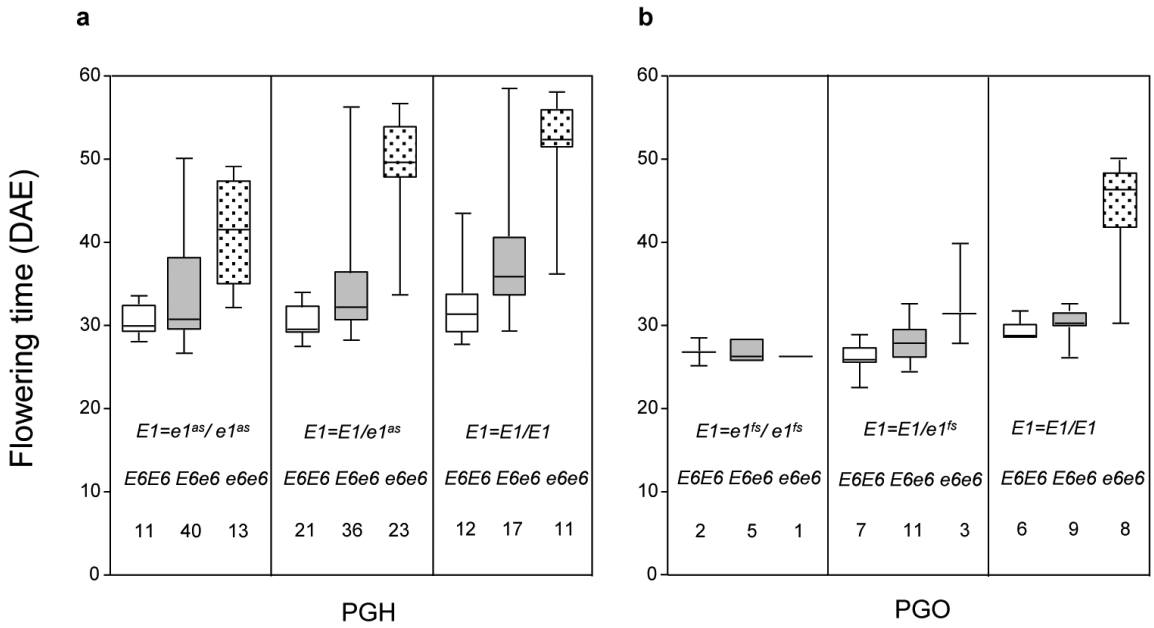

Fig. 3. Allelic effects on flowering time of quantitative trait loci (QTLs) of *E6* and *E1* in two F_2 populations, (a) PGH (Paranagoiana × Harosoy) and (b) PGO (Paranagoiana × OT94-47). Allelic combinations of *E6* and *E1* loci are indicated in each column. The numbers indicate the plants tested for each allelic combination of *E6* and *E1*. The genotypes of the *E6* allele were analyzed by tagging marker HRM101 (Supplemental Table 1). The *E1* allele was genotyped by its functional markers, E1-dCAPs. DAE, days after emergence.

time was further confirmed in the F_2 population PGO (Fig. 3b). In the homozygous loss of function $e1^{fs}e1^{fs}$ lines, *E6* completely lost an effect on flowering time control, suggesting that *E6* is fully dependent on *E1* function (Fig. 3b). Molecular cloning of the *E6* gene will further facilitate understanding the regulatory relationships between *E6* and *E1* and their molecular mechanisms controlling flowering time and the LJ trait.

In summary, the maturity gene *E6* was molecular mapped on Gm04 adjacent to marker HRM101. The mapping results indicate that *E6* and *J* genes might be the two genes that are tightly linked. The *E1* gene has a role in control of flowering time under SD conditions. We found that *E1* has an epistatic effect on *E6* and that *E6* has a suppressive effect on *E1*. The two F_2 populations of PGH and PGI will be selfed to advanced generations, and appropriate populations such as heterogeneous inbred family will be selected from residual heterozygous lines using the closest markers of *E6* for positional cloning of *E6*. Molecular identification and functional characterization of the *E6* gene will greatly facilitate understanding of the genetic and molecular mechanisms underlying the LJ trait. The markers for *E6* are very useful for molecular breeding for wide adaptation and stable productivity of soybean under low-latitude environments.

Conflict of Interest

The authors declare that there is no conflict of interest.

Supplemental Material Available

Supplemental material for this article is available online.

Acknowledgments

We thank Randall L. Nelson (Pathology and Genetics Research Unit and Department of Crop Sciences, Soybean/Maize Germplasm, USDA-ARS, University of Illinois, Urbana, IL) for providing some soybean germplasm. This work was funded by the National Natural Science Foundation of China (31430065, 31071445, 31501330, 31371643, and 31571686); the National Key Research and Development Program (no. 2016YFD0100401); the Open Foundation of the Key Laboratory of Soybean Molecular Design Breeding, Chinese Academy of Sciences; the "Hundred Talents" Program of the Chinese Academy of Sciences; and the Strategic Action Plan for Science and Technology Innovation of the Chinese Academy of Sciences (XDA08030108).

References

Bernard, R.L. 1971. Two major genes for time of flowering and maturity in soybeans. Crop Sci. 11:242–244. doi:10.2135/cropsci1971.0011183X001100020022x

Bonato, E.R., and N.A. Vello. 1999. *E6*, a dominant gene conditioning early flowering and maturity in soybeans. Genet. Mol. Biol. 22:229–232. doi:10.1590/S1415-47571999000200016

Buzzell, R.I. 1971. Inheritance of a soybean flowering response to fluorescent-daylength conditions. Can. J. Genet. Cytol. 13:703–707.

Buzzell, R.I., and H.D. Voldeng. 1980. Inheritance of insensitivity to long daylength. Soybean Genet. Newsl. 7:26–29.

Cairo, C.A., M.V. Cambursano, and E.N. Morand. 2009. Molecular mapping of the juvenile locus in soybean. In: Proceedings of the World Soybean Research Conference VIII, Beijing, China. 10–15 Aug. 2009. Crop Sci. Soc. China, Beijing. p. 141.

Cairo, C.A., J. Stein, L. Delgado, S. Bortolotti, S.A. Guelman, J.P.A. Ortiz, and E.N. Morandi. 2002. Tagging the juvenile locus in soybean [*Glycine max* (L.) Merr.] with molecular markers. Euphytica 124:387–395. doi:10.1023/A:1015767623038

Cao, D., R. Takeshima, C. Zhao, B. Liu, J. Abe, and F. Kong. 2016. Molecular mechanisms of flowering under long days and stem growth habit in soybean. J. Exp. Bot. doi:10.1093/jxb/erw394

Carpentieri-Pípolo, V., L.A.d. Almeida, and R.A.S. Kiihl. 2002. Inheritance of a long juvenile period under short-day conditions in soybean. Genet. Mol. Biol. 25:463–469.

Carpentieri-Pípolo, V., L.A.d. Almeida, R.A.S. Kiihl, and C.A. Rosolem. 2000. Inheritance of long juvenile period under short day conditions for the BR80–6778 soybean (*Glycine max* (L.) Merrill) line. Euphytica 112:203–209. doi:10.1023/A:1003927817278

Churchill, G.A., and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.

Cober, E.R. 2011. Long juvenile soybean flowering responses under very short photoperiods. Crop Sci. 51:140–145. doi:10.2135/cropsci2010.05.0262

Cober, E.R., S.J. Molnar, M. Charette, and H.D. Voldeng. 2010. A new locus for early maturity in soybean. Crop Sci. 50:524–527. doi:10.2135/cropsci2009.04.0174

Cober, E.R., and M.J. Morrison. 2010. Regulation of seed yield and agronomic characters by photoperiod sensitivity and growth habit genes in soybean. Theor. Appl. Genet. 120:1005–1012. doi:10.1007/s00122-009-1228-6

Cober, E.R., and H.D. Voldeng. 2001. A new soybean maturity and photoperiod-sensitivity locus linked to *E1* and *T*. Crop Sci. 41:698–701. doi:10.2135/cropsci2001.413698x

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler et al. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. 39:1464–1490. doi:10.2135/cropsci1999.3951464x

Destro, D., V. Carpentieri-Pípolo, R.A.S. Kiihl, and L.A.d. Almeida. 2001. Photoperiodism and genetic control of the long juvenile period in soybean: A review. Crop Breed. Appl. Biotechnol. 1:72–92. doi:10.13082/1984-7033.v01n01a10

Fehr, W.R., C.E. Caviness, D.T. Burmood, and J.S. Pennington. 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. Crop Sci. 11:929–931. doi:10.2135/cropsci1971.0011183X001100060051x

Jiang, B., H. Nan, Y. Gao, L. Tang, Y. Yue, S. Lu et al. 2014. Allelic combinations of soybean maturity loci *E1*, *E2*, *E3* and *E4* result in diversity of maturity and adaptation to different latitudes. PLoS One 9:e106042. doi:10.1371/journal.pone.0106042

Kong, F., H. Nan, D. Cao, Y. Li, F. Wu, J. Wang et al. 2014. A new dominant gene conditions early flowering and maturity in soybean. Crop Sci. 54:2529–2535. doi:10.2135/cropsci2014.03.0228

Kong, F., B. Liu, Z. Xia, S. Sato, B.M. Kim, S. Watanabe et al. 2010. Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. Plant Physiol. 154:1220–1231. doi:10.1104/pp.110.160796

Li, Y., Z. Chu, X. Liu, H. Jing, Y.G. Liu, and D. Hao. 2010. A cost-effective high-resolution melting approach using the EvaGreen dye for DNA polymorphism detection and genotyping in plants. J. Integr. Plant Biol. 52:1036–1042. doi:10.1111/j.1744-7909.2010.01001.x

Liu, B., A. Kanazawa, H. Matsumura, R. Takahashi, K. Harada, and J. Abe. 2008. Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. Genetics 180:995–1007. doi:10.1534/genetics.108.092742

Lu, S., Y. Li, J. Wang, H. Nan, D. Cao, X. Li et al. 2016. Identification of additional new QTLs for flowering time by removing the major effect of the maturity gene *E1* in soybean. J. Integr. Agric. 151:60345–60357.

Lu, S., Y. Li, J. Wang, P. Srinives, H. Nan, D. Cao et al. 2015. QTL mapping for flowering time in different latitude in soybean. Euphytica. doi:10.1007/s10681-015-1501-5

Lu, S., X. Zhao, Y. Hu, S. Liu, H. Nan, X. Li et al. 2017. Natural variation at the soybean *J* locus improves adaptation to the tropics and enhances yield. Nat. Genet. 49:773–779. doi:10.1038/ng.3819

McBlain, B.A., and R.L. Bernard. 1987. A new gene affecting the time of flowering and maturity in soybeans. J. Hered. 78:160–162. doi:10.1093/oxfordjournals.jhered.a110349

Meng, L., H. Li, L. Zhang, and J. Wang. 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. The Crop J. 3:269–283. doi:10.1016/j.cj.2015.01.001

Nan, H., D. Cao, D. Zhang, Y. Li, S. Lu, L. Tang et al. 2014. GmFT2a and GmFT5a redundantly and differentially regulate flowering through interaction with and upregulation of the bzip transcription factor GmFDL19 in soybean. PLoS One 9:e97669. doi:10.1371/journal.pone.0097669

Parvez, A., and F. Gardner. 1987. Daylength and sowing date responses of soybean lines with "juvenile" trait. Crop Sci. 27:305–310. doi:10.2135/cropsci1987.0011183X002700020037x

Ray, J.D., K. Hinson, J.E.B. Mankono, and M.F. Malo. 1995. Genetic control of a long-juvenile trait in soybean. Crop Sci. 35:1001–1006. doi:10.2135/cropsci1995.0011183X003500040012x

Samanfar, B., S.J. Molnar, M. Charette, A. Schoenrock, F. Dehne, A. Golshani, F. Belzile, and E.R. Cober. 2016. Mapping and identification of a potential candidate gene for a novel maturity locus, *E10*, in soybean. Theo. Appl. Genet. 130:377–390. doi:10.1007/s00122-016-2819-7

Sinclair, T.R., and K. Hinson. 1992. Soybean flowering in response to the long-juvenile trait. Crop Sci. 32:1242–1248. doi:10.2135/cropsci1992.0011183X003200050036x

Spehar, C.R. 1995. Impact of strategic genes in soybean on agricultural development in the Brazilian tropical savannahs. Field Crops Res. 41:141–146. doi:10.1016/0378-4290(95)00007-D

Thakare, D., S. Kumudini, and R.D. Dinkins. 2011. The alleles at the E1 locus impact the expression pattern of two soybean *FT*-like genes shown to induce flowering in Arabidopsis. Planta 234:933–943. doi:10.1007/s00425-011-1450-8

Tsubokura, Y., H. Matsumura, M. Xu, H. Nakashima, B. Liu, T. Anai et al. 2013. Genetic variation in soybean at the maturity locus *E4* is involved in adaptation to long days at high latitudes. Agronomy 3:117–134. doi:10.3390/agronomy3010117

Van Ooijen, J. 2004. MapQTL5. Software for the mapping of quantitative trait loci in experimental populations. Kyazma BV, Wageningen, the Netherlands.

Wang, S., C.J. Basten, and Z.-B. Zeng. 2012. Windows QTL Cartographer 2.5. Dep. of Stat., North Carolina State Univ., Raleigh, NC. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm (accessed 6 Dec. 2016).

Watanabe, S., K. Harada, and J. Abe. 2012. Genetic and molecular basis of photoperiod responses of flowering in soybean. Breed. Sci. 61:531–543. doi:10.1270/jsbbs.61.531

Watanabe, S., R. Hideshima, Z. Xia, Y. Tsubokura, S. Sato, Y. Nakamoto et al. 2009. Map-based cloning of the gene associated with the soybean maturity locus *E3*. Genetics 182:1251–1262. doi:10.1534/genetics.108.098772

Watanabe, S., Z. Xia, R. Hideshima, Y. Tsubokura, S. Sato, N. Yamanaka, R. Takahashi et al. 2011. A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. Genetics 188:395–407. doi:10.1534/genetics.110.125062

Xia, Z., S. Watanabe, T. Yamada, Y. Tsubokura, H. Nakashima, H. Zhai et al. 2012. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. Proc. Natl. Acad. Sci. USA 109:E2155–E2164.

Xu, M., Z. Xu, B. Liu, F. Kong, Y. Tsubokura, S. Watanabe et al. 2013. Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. BMC Plant Biol. 13:91. doi:10.1186/1471-2229-13-91

Yue, Y., N. Liu, B. Jiang, M. Li, H. Wang, Z. Jiang et al. 2016. A single nucleotide deletion in *J* encoding *GmELF3* confers long juvenility and is associated with adaption of tropic soybean. Mol. Plant 10:656–658. doi:10.1016/j.molp.2016.12.004