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 $E_6$ 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 $E_6$ and $J$ loci might be tightly linked. Genetic interaction evaluation between $E_6$ and $E_1$ suggests that $E_6$ has a suppressive effect on $E_1$ and that the function of $E_6$ is dependent on $E_1$. The tagging markers for $E_6$ 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 $E_6$ gene will greatly facilitate the understanding of the genetic and molecular mechanisms underlying the LJ trait.
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 Gardener, 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 (Spear, 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 F2 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-Pipolo 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 an 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

www.crops.org crop science, vol. 57, september–october 2017
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 *F₂* population (hereafter named PGH) of the cross between LJ line Paranagoiana (*e6e6, PI 628880*) and CJ line Harosoy (*E6E6, PI 548573*) was developed. An *F₂* 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 *F₂* population (hereafter named PGI) from the cross between two LJ lines, Paranagoiana (*e6e6*) and PI 159925 (*jj*), was developed.

The *F₂* 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 *F₂* 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 *e1e1* and *e1e0* 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 *F₂* 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 *F₂* 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 *F₂* 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 *F₂* 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 *F₂* 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 *F₂* 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 *F₂* populations.

<table>
<thead>
<tr>
<th>Crosses</th>
<th>Population names</th>
<th>No. of <em>F₂</em> plants</th>
<th>No. of linked markers†</th>
<th>No. of linkage group</th>
<th>Total map length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paranagoiana × Harosoy</td>
<td>PGH</td>
<td>184</td>
<td>162</td>
<td>20</td>
<td>1681</td>
</tr>
<tr>
<td>Paranagoiana × OT94-47</td>
<td>PGO</td>
<td>58</td>
<td>216</td>
<td>20</td>
<td>2364</td>
</tr>
<tr>
<td>Paranagoiana × PI 159925</td>
<td>PGI</td>
<td>126</td>
<td>163</td>
<td>20</td>
<td>1987</td>
</tr>
</tbody>
</table>

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

CROP SCIENCE, VOL. 57, SEPTEMBER–OCTOBER 2017  WWW.CROPS.ORG  2549
Table 2. Soybean materials used in this study.

<table>
<thead>
<tr>
<th>Line name</th>
<th>Line ID</th>
<th>Genotype</th>
<th>Pedigree</th>
<th>Flower time (12 h light/12 h dark) d after emergence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paranagoiana</td>
<td>PI 628880</td>
<td>$E_1, E_2, E_3, E_4, e_6$</td>
<td>Natural mutation from cultivar Parana</td>
<td>62a†</td>
</tr>
<tr>
<td>PI 159925</td>
<td>PI 159925</td>
<td>$E_1, E_2, E_3, E_4, j$</td>
<td>Landrace collected from Peru</td>
<td>47b</td>
</tr>
<tr>
<td>Harosoy</td>
<td>PI 548573</td>
<td>$e^{1h}, e_2, E_3, E_4, J, E_6$</td>
<td>Mandarin (Ottawa) × 2/AK (Harrow)</td>
<td>28c</td>
</tr>
<tr>
<td>OT97-47</td>
<td>–</td>
<td>$e^{1h}, e_2, e_3, e_4, J, E_6$</td>
<td>OT89-5/X2749-K1</td>
<td>27c</td>
</tr>
</tbody>
</table>

† 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.

<table>
<thead>
<tr>
<th>Population†</th>
<th>$F_2$ Flowering time</th>
<th>Parents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min.</td>
<td>Max.</td>
</tr>
<tr>
<td>PGH</td>
<td>27</td>
<td>61</td>
</tr>
<tr>
<td>PGO</td>
<td>25</td>
<td>54</td>
</tr>
<tr>
<td>PGI</td>
<td>45</td>
<td>71</td>
</tr>
</tbody>
</table>

† 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.
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 \( e6 \) and Harosoy possesses the recessive weak allele \( e6w \) (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 \textit{Arabidopsis} flowering gene \textit{ELF3} (Yue et al., 2016; Lu et al., 2017), this suggests that \( E6 \) might be the \( J \) (\textit{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 \textit{FLOWERING LOCUS T} genes, \textit{FT2a} (\( E9 \)) and \textit{FT5a} (Xia et al., 2012). Two major QTLs for
flowering time corresponding to $E_6$ and $E_1$ loci under SD were consistently identified from two crosses in PGH and PGO (Table 4). To understand the genetic effects between QTLs of $E_6$ and $E_1$, we classified the allelic combinations using the tagging marker HRM101 of $E_6$ and $E_1$ functional markers in the two F$_2$ populations PGH and PGO (Fig. 3). In the population PGH, in homozygous dominant $E_6E_6$ lines, there were no flowering time differences whether the $E_1$ allele was dominant or recessive, which suggested that $E_6$ had a suppressive effect on $E_1$ (Fig. 3a). In homozygous recessive $e_6e_6$ lines, the partial loss of function allele $e_1$ as reduced the suppressive effect of $E_6$ on $E_1$, which suggested that $E_6$ is dependent on $E_1$ (Fig. 3a). The suppressive effect of $E_6$ on $E_1$ on flowering time to $E_6$ and $E_1$ loci under SD were consistently identified from two crosses in PGH and PGO.

Fig. 3. Allelic effects on flowering time of quantitative trait loci (QTLs) of $E_6$ and $E_1$ in two F$_2$ populations, (a) PGH (Paranagoiana × Harosoy) and (b) PGO (Paranagoiana × OT94-47). Allelic combinations of $E_6$ and $E_1$ loci are indicated in each column. The numbers indicate the plants tested for each allelic combination of $E_6$ and $E_1$. The genotypes of the $E_6$ allele were analyzed by tagging marker HRM101 (Supplemental Table 1). The $E_1$ allele was genotyped by its functional markers, E1-dCAPs. DAE, days after emergence.
time was further confirmed in the F\textsubscript{2} population PGO (Fig. 3b). In the homozygous loss of function e\textsuperscript{P}Fe\textsuperscript{P} lines, E\textsubscript{6} completely lost an effect on flowering time control, suggesting that E\textsubscript{6} is fully dependent on E\textsubscript{1} function (Fig. 3b). Molecular cloning of the E\textsubscript{6} gene will further facilitate understanding the regulatory relationships between E\textsubscript{6} and E\textsubscript{1} and their molecular mechanisms controlling flowering time and the LJ trait.

In summary, the maturity gene E\textsubscript{6} was molecular mapped on Gm04 adjacent to marker HRM101. The mapping results indicate that E\textsubscript{6} and J genes might be the two genes that are tightly linked. The E\textsubscript{1} gene has a role in control of flowering time under SD conditions. We found that E\textsubscript{1} has an epistatic effect on E\textsubscript{6} and that E\textsubscript{6} has a suppressive effect on E\textsubscript{1}. The two F\textsubscript{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 E\textsubscript{6} for positional cloning of E\textsubscript{6}. Molecular identification and functional characterization of the E\textsubscript{6} gene will greatly facilitate understanding of the genetic and molecular mechanisms underlying the LJ trait. The markers for E\textsubscript{6} 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**


Jiang, B., H. Nan, Y. Gao, L. Tang, Y. Yue, S. Lu et al. 2014. Allelic combinations of soybean maturity loci E\textsubscript{1}, E\textsubscript{2}, E\textsubscript{3} and E\textsubscript{4} result in diversity of maturity and adaptation to different latitudes. PLoS One 9:e106042. doi:10.1371/journal.pone.0106042


