

# A Map-Based Cloning Strategy Employing a Residual Heterozygous Line Reveals that the *GIGANTEA* Gene Is Involved in Soybean Maturity and Flowering

Satoshi Watanabe,\* Zhengjun Xia,\*<sup>†</sup> Rumiko Hideshima,<sup>‡</sup> Yasutaka Tsubokura,\*  
Shusei Sato,<sup>§</sup> Naoki Yamanaka,\*<sup>\*</sup> Ryoji Takahashi,<sup>††</sup> Toyoaki Anai,<sup>‡</sup>  
Satoshi Tabata,<sup>§</sup> Keisuke Kitamura<sup>††</sup> and Kyuya Harada\*<sup>1</sup>

\*National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan, <sup>†</sup>Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150040, People's Republic of China, <sup>‡</sup>Faculty of Agriculture, Saga University, Saga 840-8502, Japan, <sup>§</sup>Kazusa DNA Research Institute, Kisarazu, Chiba 292-0812, Japan, <sup>\*\*</sup>Japan International Research Center for Agricultural Sciences, Tsukuba, Ibaraki 305-8686, Japan, <sup>††</sup>National Institute of Crop Science, Tsukuba, Ibaraki 305-8518, Japan and <sup>†††</sup>Graduate School of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan

Manuscript received November 10, 2010

Accepted for publication March 3, 2011

## ABSTRACT

Flowering is indicative of the transition from vegetative to reproductive phase, a critical event in the life cycle of plants. In soybean (*Glycine max*), a flowering quantitative trait locus, *FT2*, corresponding to the maturity locus *E2*, was detected in recombinant inbred lines (RILs) derived from the varieties “Misuzudaizu” (*ft2/ft2*; JP28856) and “Moshidou Gong 503” (*FT2/FT2*; JP27603). A map-based cloning strategy using the progeny of a residual heterozygous line (RHL) from the RIL was employed to isolate the gene responsible for this quantitative trait locus. A *GIGANTEA* ortholog, *GmGla* (Glyma10g36600), was identified as a candidate gene. A common premature stop codon at the 10th exon was present in the Misuzudaizu allele and in other near isogenic lines (NILs) originating from Harosoy (*e2/e2*; PI548573). Furthermore, a mutant line harboring another premature stop codon showed an earlier flowering phenotype than the original variety, Bay (*E2/E2*; PI553043). The *e2/e2* genotype exhibited elevated expression of *GmFT2a*, one of the florigen genes that leads to early flowering. The effects of the *E2* allele on flowering time were similar among NILs and constant under high (43°N) and middle (36°N) latitudinal regions in Japan. These results indicate that *GmGla* is the gene responsible for the *E2* locus and that a null mutation in *GmGla* may contribute to the geographic adaptation of soybean.

**F**LOWERING represents the transition from the vegetative to reproductive phase in plants. Various external cues, for example, photoperiod and temperature, are known to initiate plant flowering under the appropriate seasonal conditions. Of these cues, photoperiod

sensitivity is one of the important keys that enables crops to adapt to a wide range of latitudes (JUNG and MULLER 2009). To understand the molecular mechanism of flowering, extensive studies have been performed using model plant species, such as *Arabidopsis thaliana* and rice (*Oryza sativa*), and these have revealed the numerous regulatory network components associated with flowering (JUNG and MULLER 2009). Several steps are involved in perceiving the external light signals and integrating these into floral stimuli in the plant leaf organ, and many of these components, including photoreceptors such as phytochromes, as well as clock genes, clock-associated proteins, and transcriptional factors, have been well characterized.

In *A. thaliana*, flowering is promoted under long-day (LD) conditions through the *CONSTANS* (*CO*)-mediated photoperiodic flowering pathway that regulates the expression of the *FLOWERING LOCUS T* (*FT*) encoding florigen (SAMACH *et al.* 2000; SUAREZ-LOPEZ *et al.* 2001; ABE *et al.* 2005). The stability of *CO* mediated by phytochrome A and cryptochrome avoiding proteasome

Supporting information is available online at <http://www.genetics.org/cgi/content/full/genetics.110.125062/DC1>.

Available freely online through the author-supported open access option.

Sequence data described in this article have been deposited at the DNA Data Bank in Japan (DDBJ) Data Libraries under accession nos. AB554196–AB554222. The first 5 sequences and subsequent 20 sequences correspond to AFLP fragments and 20 BAC end sequences, respectively. Putative coding sequences of *GmGla*, originating from Moshidou Gong 503, and *GmGlb*, originating from the Misuzudaizu allele, have been deposited under accession nos. AB554221 and AB554222, respectively. Sequences of the five BAC clones have been deposited under accession nos. AP011822 (WBb35C13), AP011811 (MiB300H01), AP011821 (WBb225N14), AP011813 (MiB319A04), and AP011810 (MiB039C03), respectively.

<sup>1</sup>Corresponding author: Soybean Genome Research Team, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan. E-mail: haradaq@nias.affrc.go.jp

protein degradation increases under LD conditions and induces the expression of *FT* in leaves (YANOVSKY and KAY 2002; VALVERDE *et al.* 2004; LIU *et al.* 2008). The *GIGANTEA* (*GI*) mutant, discovered half a century ago, shows the late-flowering phenotype in LD conditions (REDEI 1962). *GI* encodes a nuclear-localized membrane protein that functions upstream of *CO* and *FT* (KOORNNEEF *et al.* 1998; FOWLER *et al.* 1999; HUQ *et al.* 2000; MIZOGUCHI *et al.* 2005). *GI* coupled with the blue light receptor protein, FLAVIN binding KELCH REPEAT F-BOX with PAS/LOV domains (FKF), forms a blue-light-dependent complex that degrades the CYCLING DOF FACTOR 1 (CDF) protein that binds the promoter region of *CO* and thereby induces *CO* expression (NELSON *et al.* 2000; IMAIZUMI *et al.* 2003, 2005; SAWA *et al.* 2007). Moreover, *GI* regulates another pathway that controls *FT* expression via microRNAs that cooperate with another transcriptional factor (JUNG *et al.* 2007).

The rice *CO* ortholog, *Hd1*, was discovered as a quantitative trait locus (QTL) affecting the heading date in the cross between *japonica* and *indica* rice cultivars (YANO *et al.* 2000; KOJIMA *et al.* 2002). In contrast to *CO* in *A. thaliana*, *Hd1* is considered to promote the expression of *FT* orthologs under inductive short-day (SD) conditions and to repress it under LD conditions (YANO *et al.* 2000). Overexpression of *OsGI*, an ortholog of the Arabidopsis *GI* gene, in transgenic rice increased *Hd1* expression, whereas expression of the *FT* ortholog was suppressed and resulted in a late-flowering phenotype (HAYAMA *et al.* 2003). Many other QTL for heading date in rice have been identified and some of the genes were found to encode novel transcriptional factors with no orthologs in the *A. thaliana* genome (DOI *et al.* 2004; XUE *et al.* 2008). Although some components, for example, *CO*, *GI*, and *FT*, for the flowering response are conserved in many crops, such as wheat (*Triticum aestivum*; NEMOTO *et al.* 2003; ZHAO *et al.* 2005; YAN *et al.* 2006), barley (*Hordeum vulgare*; DUNFORD *et al.* 2005; YAN *et al.* 2006), pea (*Pisum sativum*; HECHT *et al.* 2007), tomato (*Solanum lycopersicum*; LIFSCHITZ *et al.* 2006), onion (*Allium cepa*; TAYLOR *et al.* 2010), cucurbitas (LIN *et al.* 2007), and grapevine (*Vitis vinifera*; CARMONA *et al.* 2007), the mechanisms controlling the flowering pathway may differ between species, and these therefore must be elucidated at the species level.

Several maturity loci, designated as *E* loci, have been characterized by classical methods in soybean [*Glycine max* (L.) Merrill], which is classified as an SD plant. These loci are *E1* and *E2* (BERNARD 1971), *E3* (BUZZELL 1971), *E4* (BUZZELL and VOLDENG 1980), *E5* (MCBLAIN and BERNARD 1987), *E6* (BONATO and VELLO 1999), *E7* (COBER and VOLDENG 2001), and *E8* (COBER *et al.* 2010). Using molecular markers, the *E1*, *E2*, *E3*, *E4*, *E7*, and *E8* loci were mapped to soybean chromosomes Gm06, Gm10, Gm19, Gm20, Gm06, and Gm03, respectively (CREGAN *et al.* 1999; COBER and VOLDENG 2001; ABE *et al.* 2003; MOLNAR *et al.* 2003; COBER *et al.* 2010).

Of these, the *E1*, *E3*, and *E4* loci are considered as photoperiod sensitivity loci under various light conditions (SAIDON *et al.* 1989; COBER *et al.* 1996; ABE *et al.* 2003). The *E3* (Glyma19g41210) and *E4* (Glyma20g22160) genes encode phytochrome A (*PhyA*) homologs (LIU *et al.* 2008; WATANABE *et al.* 2009; <http://www.phytozome.net/>), and the recessive alleles of these loci confer an early flowering phenotype under LD conditions extended by incandescent (*e4*) or fluorescent lighting (*e3*), respectively (ABE *et al.* 2003; LIU *et al.* 2008; WATANABE *et al.* 2009). The expression of *GmFTs*, *FT* orthologs of soybean, that promote flowering in transgenic Arabidopsis plants (KONG *et al.* 2010) cannot be suppressed in the recessive alleles of both *PhyA* genes under LD conditions (KONG *et al.* 2010). Although other flowering gene orthologs have been analyzed using information from EST libraries and by expression analyses (TASMA and SHOEMAKER 2003; HECHT *et al.* 2005), the association between other *E* alleles, including *E2*, and flowering genes is still limited.

The soybean genome sequence based on the American cultivar "Williams82" has been available since January 2008 and was recently published (SCHMUTZ *et al.* 2010). The availability of such information in public databases can greatly accelerate the identification of genes controlling agronomically important traits. However, reliable and efficient mapping methods are a prerequisite. Although comparison of near isogenic lines (NILs) is the most effective way of evaluating a locus that affects a phenotype, developing NILs by backcrossing is a time-consuming and laborious process. Alternatively, using a mapping population derived from a residual heterozygous line (RHL) has been proposed (YAMANAKA *et al.* 2004). An RHL selected from a recombinant inbred line (RIL) population harbors a heterozygous region where the target QTL is located, but contains a homozygous background for most other regions of the genome. The progeny of the RHL is expected to show a simple phenotypic segregation based on the effects of the target QTL at the heterozygous region. TUINSTRA *et al.* (1997) used a similar term, heterogeneous inbred family, for progenies of RHLs to identify the QTL associated with seed weight in sorghum. The RHL strategy has already been used to identify loci underlying pathogen resistance in soybean (NJITI *et al.* 1998; MEKSEM *et al.* 1999; TRIWITAYAKORN *et al.* 2005).

In previous studies, using the same population as in this study, three flowering-time QTL, *FT1*, *FT2*, and *FT3* loci, corresponding to maturity loci *E1*, *E2*, and *E3*, respectively, were identified (YAMANAKA *et al.* 2001; WATANABE *et al.* 2004). Although *FT2* has a larger effect on flowering time than *FT3* (*E3*), the responsible gene for this QTL is still unknown. The responsible gene for the *E3* locus was isolated using an RHL-derived map-based cloning strategy (WATANABE *et al.* 2009). So far, however, there are no other successful examples of the isolation of responsible genes for QTL using RHLs in soybean. The aims of this study were

therefore to confirm the utility of the RHL strategy for map-based cloning of a gene for a QTL; to identify the gene responsible for the *FT2* (*E2*) locus; to characterize the *E2* gene, including the relationship with soybean florigen genes; and to determine the genetic stability of the *E2* allele under differential environmental conditions.

## MATERIALS AND METHODS

**Plant materials and flowering-time investigation:** A population of 156 RILs ( $F_{8,10}$ ) that was derived from a cross between the two varieties, Misuzudaizu (accession no. JP28856 in the National Institute of Agrobiological Sciences Genebank) and Moshidou Gong 503 (JP27603), was used. This population had been used previously for linkage map construction and QTL analysis of agronomic traits (WATANABE *et al.* 2004). Three QTL for flowering time, *FT1*, *FT2*, and *FT3*, were identified at LG C2 (chromosome 6), LG O (chromosome 10) and LG L (chromosome 19), respectively. The late-flowering alleles *FT1*, *FT2*, and *FT3* are partially dominant over the early-flowering alleles *ft1*, *ft2*, and *ft3*, respectively. Misuzudaizu harbored the late-flowering allele of the *FT1* and *FT3* loci, whereas Moshidou Gong 503 carried the late-flowering alleles of the *FT2* locus.

Screening methods for identifying candidate lines for RHLs with large phenotypic variance for flowering time were performed as previously described (WATANABE *et al.* 2009). RIL 6-8 were found to be heterozygous for the *FT2* locus. Seeds of RHL 6-8 were sown on May 21, 2003, and seedlings were grown under natural conditions at Chiba University, Matsudo (35°78' N, 139°90' E). One pair of NILs, with contrasting alleles for *FT2*, 6-8-*FT2*, and 6-8-*ft2*, was selected from the progenies of this RHL. Additionally, several plants heterozygous for the *FT2* locus were screened from the progeny of RHL 6-8 using molecular markers linked to this QTL. All seeds obtained from heterozygous plants were bulked to develop a large segregating population for fine mapping. This population consisted of 888 plants, and together with NILs 6-8 (211 plants for 6-8-*FT2* and 168 plants for 6-8-*ft2*), were sown on May 30, 2006, at the Japan International Research Center for Agricultural Sciences, Hachimandai, Tsukuba (36°03' N, 140°04' E), and the seedlings were transplanted to the field on June 12 and grown under natural conditions. A total of 21 recombinants were screened using new DNA markers tightly linked to the *FT2* locus from this population. The progenies, consisting of 24–74 plants derived from these recombinants, and the NILs 6-8 were sown on May 24, 2007, at Hokkaido University, Sapporo (43°07' N, 141°39' E), and seedlings were transplanted on June 7 to the field and grown under natural conditions.

A mutant line, carrying a premature stop codon mutation in the *GmGla* (*E2*) gene, and the original variety [Bay; PI553043, plant introduction (PI) number deposited in the National Plant Germplasm System in the United States] were sown on June 23, 2008, at the National Institute of Agrobiological Sciences, Tsukuba (36°02' N, 140°11' E) and grown under natural day-length conditions.

For phenotypic investigations, the dates from sowing to the first flowering, corresponding to the R1 developmental stage, were scored (FEHR *et al.* 1971). The known genotypes of Harosoy (*e2/e2*; PI548573), Clark (*E2/E2*; PI548533), and the isoline Harosoy-*E2* (*E2/E2*; PI 547768) were used to validate the identity of the *FT2* and *E2* genes.

**DNA isolation:** Genomic DNA was extracted from fresh trifoliate leaves of 2-week-old seedlings using the standard cetyltrimethyl ammonium bromide (CTAB) method (MURRAY and THOMPSON 1980) and used for amplified fragment length

polymorphism (AFLP) analysis, fine mapping, progeny tests, and confirmation of the *E2* allele.

**AFLP analysis:** Templates for the AFLP reaction were prepared based on the method of Vos *et al.* (1995) using 150 ng DNA for restriction enzyme digestion with *EcoRI* and *MseI*. Selective amplification was performed using combinations of *EcoRI* (E) primers and *MseI* (M) primers, each with three selective nucleotides. Nomenclature for the AFLP markers were expressed as  $En_1Mn_2$ , with the letters E and M denoting the *EcoRI* and *MseI* primers, respectively, and the subsequent n1 and n2 codes representing the three selective nucleotides for each primer. The amplification procedures, conditions for polyacrylamide gel electrophoresis, and methods for detection and scoring of the polymorphic bands followed those described by HAYASHI *et al.* (2001). Bulked segregant analysis (BSA) was used to identify the locus affecting segregation of flowering time in a candidate of RHL. Two DNA pools, an early flowering bulk and a late-flowering bulk from the RIL 6-8 subpopulation, were used as DNA templates. Additional polymorphic AFLP markers were identified using NILs 6-8 from the rest of the primers of all possible 4096 AFLP primer combinations.

**Development of sequence characterized amplified region markers for fine mapping of the *FT2* locus:** The detected polymorphic bands were isolated using the pGEM T-easy vector system (Promega KK, Tokyo) and sequenced with an ABI PRISM 3100 avant Genetic Analyzer using a BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems Japan, Tokyo) according to the manufacturer's instructions. Primer3 (ROZEN and SKALETSKY 2000) was used to design primers for new sequence characterized amplified region (SCAR) markers originating from AFLPs.

**Construction of a physical contig of the *FT2* region:** Two libraries of BAC clones constructed from the genomic DNA of Misuzudaizu and William 82 were used. Library screening and end-sequencing were performed as described previously (XIA *et al.* 2005). Several BAC clones were identified using the SCAR markers. No positive clone was obtained for the SCAR marker originating from AFLP E37M27/47. The flanking region of this AFLP fragment was amplified using the adapter ligation method described previously (TSUCHIYA *et al.* 2009).

The nucleotide sequences of five BAC clones (WBb35C13, MiB300H01, WBb225N14, MiB319A04, and MiB039C03) were determined according to the bridging shotgun method described previously (SATO *et al.* 2001). The generated sequences were assembled using Phred-Phrap programs (Philip Green, University of Washington, Seattle). A lower threshold of acceptability for the generation of consensus sequences was set at a Phred score of 20 for each base. These BAC sequences were used to develop new DNA markers and to predict the candidate gene for the *FT2* locus.

**DNA marker analysis:** For fine mapping, seven DNA markers developed from the AFLP and BAC sequences listed in Table 1 were used to determine the genotypes of the progenies of the RHL. Genomic DNA (20–30 ng) was used as template, and PCR reactions were performed using Ex-Taq (Takara Bio, Shiga, Japan) with 30 cycles at 96° for 30 sec, 58° for 30 sec, and 72° for 1 min. Some markers required restriction enzyme digestion to detect a polymorphism (Table 1). PCR products were separated by 10% (w/v) polyacrylamide gel electrophoresis and visualized with ethidium bromide (EtBr).

**Analysis of *GmGI* transcripts and diagnostic marker for the *e2* allele:** Total RNA from leaves of NILs 6–8 was extracted by the Trizol method (Invitrogen Japan K.K., Tokyo). Primers designed from the predicted start codon to stop codon region, based on the BAC sequence, were used to obtain the *GmGI* transcript. For the RT-PCR reaction, 1 µg of total RNA was used for first-strand cDNA synthesis with ReverTra Ace

**TABLE 1**  
**DNA markers used in this study**

Marker name	Type of marker	Clone name	Direction	Sequence (5'–3')
Marker 1 <sup>a</sup>	AFLP (SCAR)	E37M27/47 <sup>b</sup>	Forward Reverse	CACTGTGGTAGTGTCTCCAAGC CGAAGATTCCACAGCCCTAA
Marker 2 <sup>a</sup>	BAC end	GMWb35C13M4	Forward Reverse	GGAATTATCATGGTCCCTGAATC AATACAACCTCACATCTCAAGCCTCT
Marker 3	BAC end	GMJMiB300H01RV	Forward Reverse	CATAGCCGACCTTCTCCAAA AGCCCAATATGGCAGCATACT
Marker 4 <sup>a</sup>	AFLP (SCAR)	E60M38	Forward Reverse	CAGTGTTCGCCAGGCTTAGT GCTTGGGTAAACATCCCAAAA
Marker 5	BAC end	GMJMiB300H01FW	Forward Reverse	GAGAGCAGGGTTATTGGATGA GCCACTGTGCCACATTACAC
Marker 6	BAC end	GMJMiB319A04RV	Forward Reverse	AGAAGGTGGCCGGAGATTAT CACCTTTCATCCTACACTTGTTT
Marker 7	BAC end	GMJMiB114F08RV	Forward Reverse	TTGTGGTGGCATTGTGATTT CCTCTTGTCTTCACCAA

<sup>a</sup>Digestion with the restriction enzymes *Mse*I, *Msp*I, or *Eco*RI is needed to detect polymorphism with markers 1, 2, and 4, respectively.

<sup>b</sup>Two AFLP bands, E37M27 originating from the Misuzudaizu allele and E37M47 from the Moshidou Gong 503 allele, share the same sequence and are converted to a single DNA marker.

(TOYOBO, Osaka, Japan) and a standard oligo(dT<sub>20</sub>) primer according to the manufacturer's instructions. The cDNAs were then diluted twofold with PCR-grade water, and 2- $\mu$ l aliquots were used as RT-PCR templates. Specific primers for RT-PCR of *GmGla* were P1a (5'-TGTCGTCATCTTCGTCTTCG-3') and P2 (5'-CCAGAGCAGAGTCACAAGCA-3') and for *GmGlb* were P1b (5'-CATCGTTTTCAACCACTGAGA-3') and P2. The PCR conditions using Ex-Taq consisted of 30 cycles at 96° for 30 sec, 58° for 30 sec, and 72° for 3 min. PCR products were isolated and cloned into the pGEM T-easy vector system and sequenced. To evaluate the null allele of *GmGla*, specific primers for the detection of the premature stop codon were 5'-AAGCCTATGCCAGCTAGGTATTT-3' and 5'-GAAGCCCA TCAGAGGCATGTCTTATT-3'. The PCR conditions using Ex-Taq consisted of 30 cycles at 96° for 30 sec, 58° for 30 sec, and 72° for 1 min. PCR products were digested with *Dra*I (Takara Bio) and then separated by 10% (w/v) polyacrylamide gel electrophoresis and visualized by EtBr staining.

**Construction of soybean mutant libraries:** Seeds of the soybean cultivar, Bay, were independently treated with two different mutagens (X ray or EMS). For X-ray treatment, dry seeds were irradiated with 200 Gy X rays at an exposure rate of 3 Gy/min. For EMS treatment, seeds were soaked in a 0.35% (w/v) EMS solution for 12 hr and then rinsed in tap water for 8 hr. M2 seeds were obtained from self-fertilized M1 plants. Green leaves were harvested from individual M2 plants for DNA preparation. Genomic DNAs were purified using diatomaceous earth columns, followed by CTAB extraction. Pooled DNAs from eight individuals were used for mutant screening.

**Mutant screening:** A screening for targeting-induced local lesions in genomes (TILLING) was essentially carried out as previously described (McCALLUM *et al.* 2000), except that 2% (w/v) agarose gel electrophoresis was used to separate the CEL I-digested DNA fragments with GelRed stain (Biotium, Hayward, CA). Primers for TILLING screening were GmGla-ex10-F1 (ATCGGGCTGAACCTGTCCATGCTCA) and GmGla-ex10-R1 (GTACAACAGATGTCCGCCAGCAT). The PCR reactions, using Pfu DNA polymerase and 0.5% DMSO, consisted of 40 cycles at 95° for 30 sec, 70° for 30 sec, and 72° for 2 min. CEL I was purified from celery stalks as described previously (YANG *et al.* 2000).

**Expression analysis of soybean florigen gene:** Two combinations of NILs, NILs 6–8, Bay and Bay-*e2* mutant were used, with eight plants for each line. Seeds were sown on June 13, 2008, in a greenhouse under natural day-length conditions. A piece of a fully expanded trifoliolate leaf was collected, with three replicates, at 9:00 AM 4 weeks after sowing. The methods for phenotypic investigation, RNA extraction, and cDNA synthesis were essentially as described above. The RT-PCR conditions using Ex-Taq consisted of 25 cycles at 96° for 30 sec, 58° for 30 sec, and 72° for 30 sec with *GmFT2a*-specific primers 5'-ATCCCGATGCACC TAGCCCA-3' and 5'-ACACCAAACGATGAATCCCA-3'. The *GmTubulin* gene was used as an internal control and amplified for 28 cycles at 96° for 30 sec, 58° for 30 sec, and 72° for 30 sec with primers 5'-TCTTGGACAACGAAGCCATCT-3' and 5'-AAG CCTATGCCAGCTAGGTATTT-3'. PCR products were separated by 13% (w/v) polyacrylamide gel electrophoresis and visualized by SYBR Green (Invitrogen) staining. The levels of *GmTubulin* expression were used to calculate the relative expression levels of genes using Image J software (<http://rsb.info.nih.gov/ij/>). Three independent experiments were performed.

**Data analysis:** Data were analyzed using R software (<http://cran.r-project.org/>) for one-way or two-way classification analysis of variance (ANOVA) without the assumption of equal variance. Multiple regression analysis was applied to estimate the additive and dominance effects of the *FT2* locus, details of which were described previously (WATANABE *et al.* 2004). Phylogenetic analysis of the GIGANTEA protein was performed using the neighbor-joining (NJ) method with the program MEGA 4.0 (KUMAR *et al.* 2008).

## RESULTS

**Development of an RHL for the *FT2* locus:** Most RILs showed a small variance in their flowering phenotype segregation in lines with an average standard deviation (S.D.) of 1.33 (supporting information, Table S1 and Table S2) similarly to the parents [estimated S.D. values for Misuzudaizu (Mi) and Moshidou Gong 503 (Mo) were 1.20 and 1.59, respectively (Table S3)], whereas RIL 6-8 exhibited a larger S.D. value of 3.52. We

expected that RIL 6-8 would harbor some heterozygous region, including the QTL for flowering time in its genome. Using BSA analysis, a polymorphic AFLP marker, E7M19, was detected between the early-flowering bulk and the late-flowering bulk derived from the progeny of RILs 6–8. Mapping results and QTL analysis for flowering time using the RILs suggested that this marker was located close to the LOD peak position of the QTL assigned *FT2* (Figure 1). This indicated that RIL 6-8 harbored a heterozygous region, including the *FT2* locus, in the previous generation ( $F_7$ ) and showed phenotypic segregation in the subsequent generation ( $F_8$ ). A region covering  $\sim 10$  cM, including the *FT2* locus, was found through DNA marker analysis to have segregated in the progeny of RIL 6-8 (Figure 1). The plants heterozygous for this region, designated as RIL 6-8, generated NILs 6-8-*FT2* and -*ft2* among its progeny. The difference between NILs for flowering time was highly significant ( $P < 0.001$ ) with  $57.5 \pm 1.72$  days ( $n = 168$ ) for 6-8-*ft2* and  $67.6 \pm 1.56$  ( $n = 210$ ) days for 6-8-*FT2*. These results indicated that the progeny derived from RILs 6–8 would be suitable for fine mapping of the *FT2* locus. We developed additional DNA markers tightly linked to this QTL using NILs 6-8.

**Screening of AFLP markers and construction of a physical contig around the *FT2* locus:** From the progeny of RIL 6-8, we selected several NILs harboring the narrower heterozygous region, including the *FT2* locus for the development of new AFLP markers tightly linked to this locus. Figure S1 shows the scheme used to screen for the AFLP marker. Among the products amplified from all possible 4096 primer pair combinations, only five polymorphic bands showed constant polymorphism between the contrasting genotypes of *FT2/FT2* and *ft2/ft2* in NILs 6-8. We confirmed the genetic positions of these markers using the RILs (Figure S1). These polymorphic bands were excised from the gel, sequenced, and converted to SCAR markers. Three SCAR markers, originating from these AFLP bands, were developed. Using these SCAR markers, 10 BAC clones were identified after screening of two independent genomic DNA libraries. A physical contig covering this region was constructed on the basis of the results of PCR analysis using the BAC end sequences (Figure S2). Five of the 10 BAC clones (Wb35C13, MiB300H01, Wbb225N14, MiB319A04, and MiB039C03) were then subjected to shotgun sequence analysis. Each BAC clone was separately analyzed and assembled, and the sequence information was then combined using overlapping sequences. The total length covered by the five clones was  $\sim 430$  kbp. A total of seven DNA markers, including two AFLP-derived markers (markers 1 and 4) and five PCR-based markers developed from these BAC sequences (markers 2, 3, 5, 6, and 7), were used in the fine-mapping experiments to precisely restrict the *FT2* locus (Table 1).

**Fine mapping of the *FT2* locus:** A population consisting of 888 plants, derived from several RIL 6-8

plants, was used for fine mapping of the *FT2* locus. Markers 1, 4, 6, and 7 were used for preliminary screening (Table S4 and Table S5). Recombination between these markers was not observed in 822 plants but was found in 21 other plants. The remaining 45 individuals were omitted from the analysis because of missing data for phenotypes or genotypes (Table S5). The numbers of *FT2* homozygous late-flowering genotype ( $n = 213$ ,  $68.5 \pm 1.08$ ), heterozygous ( $n = 420$ ,  $63.9 \pm 1.95$ ), and *ft2* homozygous early flowering genotype ( $n = 210$ ,  $58.1 \pm 2.50$ ) fitted well with a 1:2:1 segregation ratio (Table S6), and differences in flowering time among genotypes based on marker 4 were significant ( $P < 0.001$ ). The additive effect and the dominance effect of this QTL were estimated to be  $-5.17$  days and  $0.57$  days, respectively. Furthermore, the ratio of genetic variance explained by the *FT2* locus accounted for 87.9% of the total variance, indicating that the variation observed in this population was largely controlled by the single-QTL effect. On the other hand, the 21 plants harboring recombination between these markers were analyzed in more detail using additional DNA markers (Figure 2A). One-way ANOVA of the phenotypic data of each recombinant classified by genotypes of DNA markers enabled us to estimate the location of the QTL more precisely (Figure 2B). The highest statistic was obtained at the locus of marker 4 ( $F = 91.25$ ,  $p < 0.001$ ), indicating that the QTL was close to this marker. Of these recombinants, line #060501 had a recombination point between marker 3 and marker 4. Another line, #060120, generated a recombination between marker 4 and marker 5 (Figure 2). Markers 3 and 5 originated from the end sequences of MiB300H01 (Table 1). The phenotypic segregation patterns in progenies of lines #060120 and #060501 were evaluated to restrict the precise position of the *FT2* locus since both lines had a crossover within a single BAC clone, MiB300H01. The NILs (6–8-*FT2* and 6–8-*ft2*) and two homozygous lines (#060452 and #060528) showed small levels of phenotypic segregation compared with lines #060120 and #060501 (Table 2 and Table S7). A clear phenotypic segregation classified by the genotype of marker 4 was observed in progenies of lines #060120 and #060501 (Table 2). Considering the recombination points in each line and cosegregation of phenotypes and genotypes of marker 4 in the lines, this result indicated that the *FT2* locus was restricted to a single BAC clone, MiB300H01. To identify the gene responsible for this QTL, the nucleotide sequence of this BAC clone was investigated.

**Candidate gene for the *FT2* locus:** The complete sequence of the physical contig, assembled from the five BAC clones, corresponded to a 427-kbp sequence from 44,415 to 44,843 kbp in Gm10 of the soybean genome sequence available in the public database. The 94-kbp sequence of MiB300H01 coincided with the sequence from 44,693 to 44,787 kbp in Gm10. The average

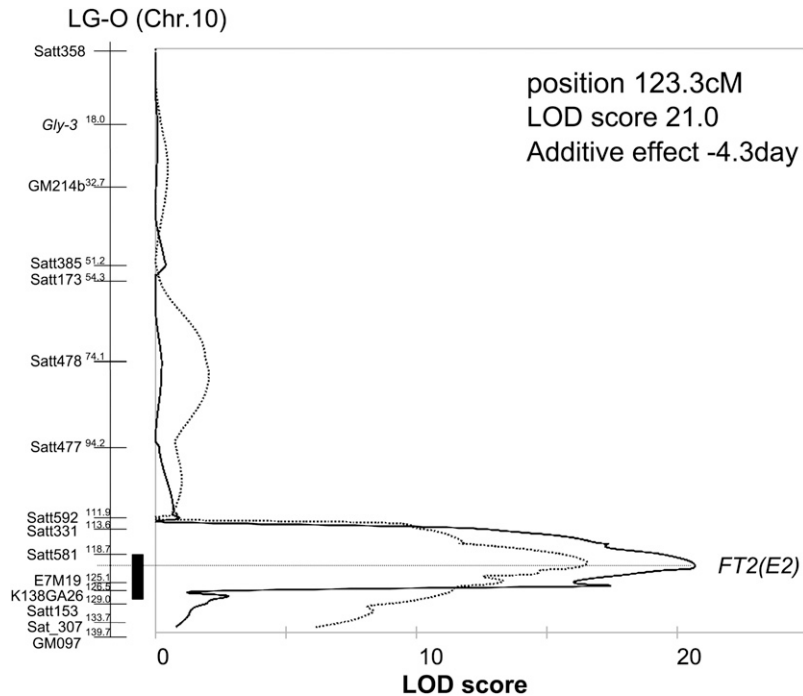


FIGURE 1.—Detection of the QTL designated *FT2* using RILs and the identification of the heterozygous region of RHL 6-8 determined by DNA markers mapped on LG-O [chromosome (Chr.) 10]. Genetic linkage map, marker names, and genetic distances (cM) are displayed on the left side of the graph. Solid (2001) and dotted (1999) lines indicate the LOD scores for QTL analysis of flowering time, using 2 years of data on RILs and calculated by composite interval mapping (WATANABE *et al.* 2004). Representative information relating to the LOD peak position corresponding to the *FT2* locus, the additive effect of the Misuzudaizu allele, and the phenotypic variance explained by the QTL (PVE) are summarized in the upper-right section of the graph. The heterozygous region of RHL 6-8 is indicated by the solid bar.

similarity between the sequences of Misuzudaizu and Williams 82 in this region was 99.6%. Nine annotated genes (Glyma10g36580–36670) were predicted in this region. One of these genes, Glyma10g36600, with a high level of similarity with the *GI* gene, was considered a strong candidate for the *FT2* locus, since the loss of function of *GI* is known to cause drastic changes in the flowering phenotype of other plant species (FOWLER *et al.* 1999; HECHT *et al.* 2007). On the basis of this assumption, we isolated the complete predicted coding region using an RNA sample extracted from leaves of NILs 6–8-*FT2*. We hereafter refer to this gene as *GmGla*, since another *GI* gene, *GmGlb*, was also obtained from the same RNA sample. *GmGI* genes isolated from the Mo late-flowering allele (*GmGla-Mo* and *GmGlb-Mo*) were found to encode proteins consisting of 1170 and 1168 amino acids, respectively. Amino acid identity between the two *GmGI* proteins was close to 97.3%. The predicted amino acid sequences of *GmGla* and *GmGlb* were completely identical to Glyma10g36600 and Glyma20g30980, respectively. Phylogenetic analysis showed that these soybean *GI* proteins displayed a high level of similarity (71–91%) to *GI* proteins from dicots and monocots (Figure 3). The presence of paralog genes with high levels of similarity attracted our attention. However, as there were no experimental data for *GmGlb*, we focused on the *GmGla* gene in the present study.

The coding sequence of *GmGla* was extended to a 20-kbp genomic region and contained 14 exons (Figure 4A). Marker 4, which cosegregated with the QTL genotypes and originated from the AFLP marker, E60M38, was located in the 5th intron (Figure 4A). Compared to *GmGla-Mo*, the Misuzudaizu early-flowering allele,

*GmGla-Mi*, showed four single nucleotide polymorphisms (SNPs) in its coding sequence (Figure 4B). One of these SNPs, detected in 10th exon, introduced a premature stop codon mutation that led to a truncated 521-aa *GI* protein in the *GmGla-Mi* allele. This stop codon mutation was considered a candidate for a functional nucleotide polymorphism in *GmGla*. A derived cleaved amplified polymorphic sequence (dCAPs) marker was developed to examine the identity of this stop codon mutation in other NILs for the *E2* locus. McBLAIN *et al.* (1987) performed testcross experiments using various combinations of NILs originating from Harosoy (*e2/e2*) and Clark (*E2/E2*). The genotypes of all isolines tested coincided well with the genotype of this diagnostic dCAPs marker (Figure S3). This result indicated that the gene responsible for the QTL of the *FT2* was probably identical to the *E2* gene and that a conserved mutation might have caused the early flowering phenotype in their recessive allele. To validate the significance of the mutation in the *GmGla* gene, we screened a mutant line from the mutagen-treated libraries. The sequence of *GmGla* in the wild-type Bay cultivar was completely identical to the *E2* allele. One mutant line harboring a deletion in the 10th exon that caused a truncated protein (735 aa; Figure 4B and Figure S4) showed a significantly earlier flowering phenotype ( $39.8 \pm 1.16$ ;  $n = 8$ ) than the wild type ( $47.6 \pm 1.06$ ;  $n = 29$ ) under natural day-length conditions. These results indicated that the mutation in *GmGla* causes the early flowering phenotype in soybean.

*GI* has the conserved function of controlling the expression of the *FT* (florigen) gene in Arabidopsis,

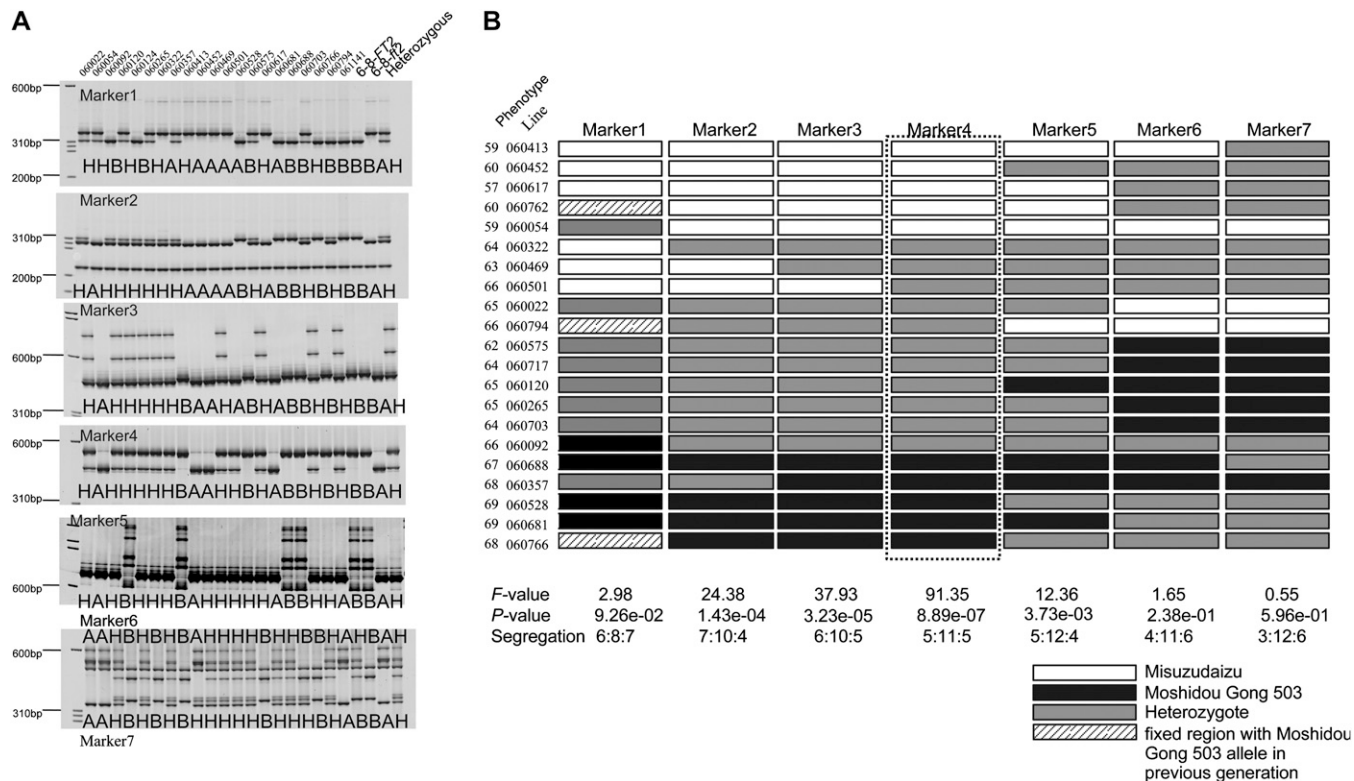


FIGURE 2.—The genotype of each recombinant plant and estimation of QTL among the marker loci. (A) The genotypes of each plant were analyzed using the marker located on this physical region. The genotypes of A, B, and H indicate the Misuzudaizu homozygous and the Moshidou Gong 503 homozygous and heterozygous genotypes, respectively. (B) One-way classification ANOVA was performed to estimate the position of the QTL. Phenotypic values of each plant investigated in the experimental field in 2006 are shown at the left side of the line name. The representation of each marker genotype is summarized at the bottom. The statistical values and probability calculated using ANOVA are shown below the graphical genotypes of each marker. The highest statistical value is obtained for marker 4 surrounded by the dotted line.

rice, and pea (HAYAMA *et al.* 2003; MIZOGUCHI *et al.* 2005; HECHT *et al.* 2007). We examined the relationship between the *GmGla* (*E2*) gene and the soybean florigen homologous genes *GmFT2a* (Glyma16g26660) and *GmFT5a* (Glyma16g04830). The diurnal expression profile of *GmFT2a/5a* showed that the highest level of expression was at 4 hr after dawn under SD conditions (KONG *et al.* 2010). We analyzed the expression of *GmFT2a* at 9:00 AM 4 weeks after sowing using *E2* (*FT2*) NILs grown under natural day-length conditions whose photoperiod changed from LD to SD for soybean because of the sowing in midsummer. A clear association between the *GmFT2a* expression level and the early flowering phenotype was observed in both NILs (Figure 5). However, there was no significant difference in the *GmFT5a* expression levels between these NILs (data not shown). These results suggest that *GmGla* probably controlled flowering time through the regulation of *GmFT2a*. The recessive alleles, whether generated naturally or artificially, were perhaps unable to suppress *GmFT2a* expression and resulted in the early flowering phenotype. In addition to studying the molecular mechanisms of *E2* gene functions, we considered it important to evaluate the genetic stability of the *E2* allele and its interaction with

the environment, and we therefore compared the genetic effects of the *E2* gene under different latitudinal conditions.

**Environmental stability of the genetic effects of the *E2* allele:** We performed fine-mapping experiments and progeny tests at two different locations—Tsukuba (36° 03' N, 140°04' E) and Sapporo (43°07' N, 141°39' E). Differences between these experimental locations provided us with information about the environmental stability of the *E2* allele. Two-way ANOVA was used to estimate the variance components associated with genetic and environmental effects and interactions between the genetic and environmental conditions (G×E; Table 3). All factors were statistically significant, but most of the variance components were occupied with the environmental (77.5%) and genetic (15.2%) variances. This indicated that differences in photoperiod, caused by the differences in latitude, probably had significant effects on suppressing the initiation of the flowering process. On the other hand, a small G×E interaction (0.1%) indicated that the genetic effect of the *E2* allele was invariable between the different environments. Comparisons of the genetic effect of the *E2* allele in three different experiments supports the stability of the *E2* allele because the additive effects observed

**TABLE 2**  
**Results of progeny test**

Line	Segregation pattern in each family				Phenotype classified with the genotype of marker 4 in the heterozygous family			
	Genotype of <i>FT2</i>	<i>n</i>	Average	SD	Genotype of marker 4	<i>n</i>	Average	SD
6-8- <i>FT2</i>	<i>FT2/FT2</i>	27	88.8	1.18				
6-8- <i>ft2</i>	<i>ft2/ft2</i>	22	79.2	1.94				
#060120	<i>FT2/ft2</i>	69	83.5	3.21	Mi homozygous ( <i>ft2/ft2</i> )	16	79.3	1.96
					Heterozygous ( <i>FT2/ft2</i> )	37	84.1	2.16
					Mo homozygous ( <i>FT2/FT2</i> )	16	86.5	1.59
#060452	<i>ft2/ft2</i>	19	79.4	2.11				
#060501	<i>FT2/ft2</i>	74	82.9	3.84	Mi homozygous ( <i>ft2/ft2</i> )	16	78.8	2.56
					Heterozygous ( <i>FT2/ft2</i> )	41	82.9	2.89
					Mo homozygous ( <i>FT2/FT2</i> )	17	86.9	2.49
#060528	<i>FT2/FT2</i>	33	88.2	1.85				

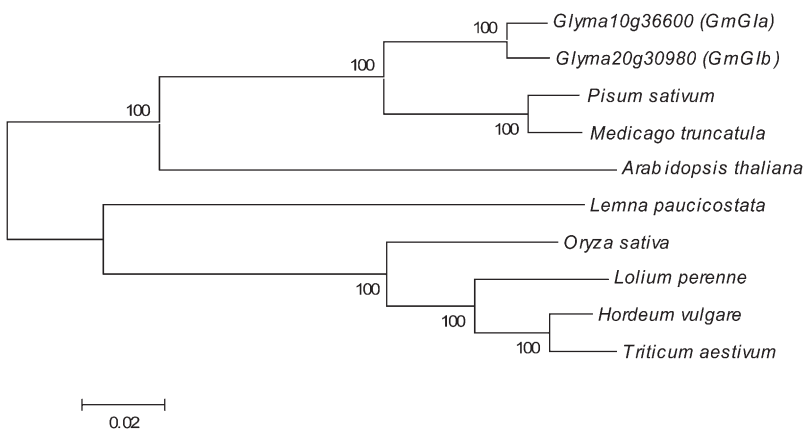
in Tsukuba, in Sapporo, and in the mutant analysis were 5.2, 4.3, and 3.9 days, respectively.

#### DISCUSSION

In this study, we have isolated the gene responsible for the soybean maturity locus *E2* using a RHL-derived map-based cloning strategy. We used the same strategy to identify the gene responsible for the target QTL of another maturity locus, *E3*, in a previous study (WATANABE *et al.* 2009). These results indicate that the RHL strategy is an extremely useful method for determining the position of a QTL and for identifying the responsible gene. The probability of discovering RHLs for a target QTL depends on the ratio of heterozygosity in a population and the size of the population. If  $p$  is the probability of heterozygosity of any population with size  $n$ , then the probability of detecting  $k$  individuals with a heterozygous genotype is  $C(n, k) p^k (1 - p)^{n-k}$  on the basis of a binomial distribution. In the case of an  $F_7$  generation of RILs, the ratio of heterozygosity ( $p$ ) is 1.56%, and with a population size of 200 ( $n$ ), the prob-

ability of detecting at least one RHL is  $>0.95$ . In our case, we discovered one and two RHLs from RILs, consisting of 156 lines, for fine mapping of the genes responsible for *E2* and *E3* (WATANABE *et al.* 2009), respectively. We therefore propose that a combination of QTL analysis, using the  $F_6$ - $F_8$  RIL population, and the RHL strategy is useful for dissecting QTL for agronomic traits in crops where backcrossing is both time-consuming and laborious. Indeed, the homozygous ratio is sufficiently high to evaluate traits with replication, and the heterozygosity ratio is not so low and will allow the identification of a sufficient number of RHLs.

Our results demonstrate that the gene responsible for the soybean maturity locus *E2* is a *GI* ortholog. *GI* plays an important role in flowering through the control of *CO* and *FT* mRNA expression levels under inductive conditions in a wide range of plant species, including monocots and dicots such as rice and Arabidopsis (KOORNNEEF *et al.* 1998; FOWLER *et al.* 1999; HAYAMA *et al.* 2003; MIZOGUCHI *et al.* 2005). Moreover, many studies have characterized the expression profile of *GI* and its associations with that of circadian clock genes in barley, *Brachypodium distachyon*, and *Lemna gibba* (DUNFORD



**FIGURE 3.**—Phylogenetic analysis of GIGANTEA proteins. A phylogenetic tree of GIGANTEA proteins, constructed by the NJ method with a  $p$ -distance and a 1000-replication bootstrap confidence value, is shown. Amino acid sequences of *A. thaliana* (AAT80910; FOWLER *et al.* 1999), *P. sativum* (ABP81863; HECHT *et al.* 2007), *Medicago truncatula* (ABE81212), *O. sativa* (NP\_001042220; OHYANAGI *et al.* 2006), *H. vulgare* (AAW66946; DUNFORD *et al.* 2005), *T. aestivum* (AAQ11738; ZHAO *et al.* 2005), *Lolium perenne* (ABF83898), and *Lemna paucicostata* (BAD97864) were used.



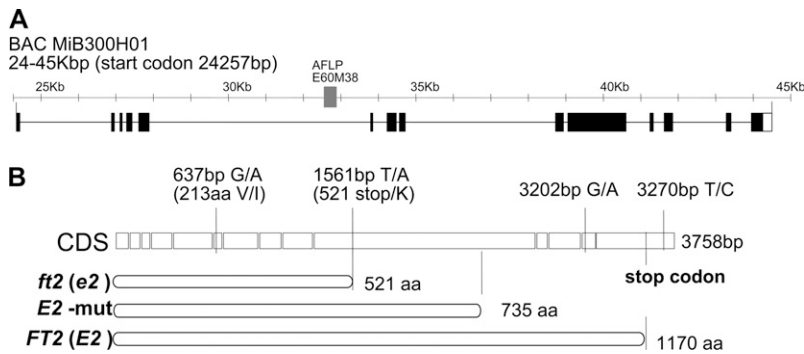


FIGURE 4.—Gene structure of *GmGIa* in NIL 6-8-*ft2* (*ft2/ft2*), NIL 6-8-*FT2* (*FT2/FT2*), and the *GmGIa* mutant line (*E2*-mut). (A) Exons, a part of the 3' UTR, and introns of the *GmGIa* gene in the 24 to 45-kbp region of MiB300H01 are indicated by solid boxes, an open box and lines, respectively. Units on the nucleotide sequence scale shown by the vertical lines represent 1 kbp. The location in the 5th intron of marker 4, originating from AFLP E60M38, is represented by the shaded bar. (B) The substitution of nucleotides and amino acid residues (in parentheses) in the coding sequence are displayed above the boxes that represent each exon. Nucleotide and amino

acid changes in the case of nonsynonymous and nonsense mutations from the *ft2* to the *FT2* allele, on the basis of the start codon position, are shown. In the early flowering *ft2* allele, a single nucleotide substitution causing a premature stop codon was discovered in the 10th exon. The mutant line, harboring a nucleotide deletion in the same exon as the *ft2* allele, had a truncated GI protein that similarly caused the early flowering phenotype under natural day-length conditions. The truncated protein lengths of the three alleles are compared at the bottom.

*et al.* 2005; SERIKAWA *et al.* 2008; HONG *et al.* 2010). In these studies, the expression of *GI* orthologs showed circadian rhythms. THAKARE *et al.* (2010) compared the expression of flowering-time gene orthologs in soybean with *E1* NILs that have the genetic background of Harosoy. Although the *E1* gene is a major maturity gene, the responsible gene has not yet been isolated (BERNALD 1971; YAMANAKA *et al.* 2004). The authors used Glyma10g36600, the gene responsible for *E2*, as a *GIGANTEA* ortholog in their expression experiments. Glyma10g36600 showed a circadian rhythm in both LD and SD conditions regardless of the *e2/e2* genotype of Harosoy, and there were no significant differences in the expression patterns between the *E1* NILs (THAKARE *et al.* 2010). This indicates that the circadian rhythm of *GmGIa* was not affected by the premature stop codon in the *e2* allele. KONG *et al.* (2010) identified functional soybean *FT* homologs included in the PhyA-mediated photoperiodic response under SD conditions. In our experiment, the early flowering phenotype, caused by the loss of function of *GmGIa*, was probably related to the expression level of *GmFT2a* (Figure 5). On the other hand, the PhyA-mediated pathway showed larger differences in the *GmFT* expression levels in comparison of NILs for the *PhyA* genes (*e3* or *e3e4*; KONG *et al.* 2010). This was probably due to differences in the pathways controlled by *GmPhyA* and *GmGI* that promote *GmFT* expression, to differences in the experimental conditions, or to differences in the genetic background used as experimental materials. Considering the similarities in amino acid sequence and functionality of *GmGIa* and *GmFTs* with other *GIs* and *FTs*, respectively, the mechanisms controlling the photoperiodic pathways are most probably conserved in soybean and other plant species. In contrast, some other studies have reported that the genetic interactions between the *E1* and *E2* alleles and the recessive *e1* allele suppress or weaken the effects of the *E2* allele (BERNALD 1971; YAMANAKA *et al.* 2000; WATANABE *et al.* 2004). Further investigations are therefore needed to clearly elucidate

the relationships between functional *GmGIs* and other genetic components, including the *E1* gene.

A redundancy in highly similar genes, which originate from genome duplication, can increase the chance of evolving subfunctionalized genes (ADAMS and WENDEL 2005). Indeed, hexaploid wheat has a greater potential for successful adaptation to a wide range of environmental conditions than tetraploid wheat (DUBCOVSKY and DVORAK 2007). The soybean genome is a typical paleopolyploid species, and the duplicated regions extend over the whole genome (SHOEMAKER *et al.* 1996; TSUBOKURA *et al.* 2008; SCHMUTZ *et al.* 2010). The soybean genome is considered to originate from two duplication events: the soybean lineage-specific paleotetraploidization [13–15 million years ago (Mya)] and the early duplication in legumes that occurred near the origin of the papilionoid lineage (44–59 Mya; SCHLUETER *et al.* 2004; SCHMUTZ *et al.* 2010). Representative examples of such gene redundancy that affect the same agronomic trait are the two photoreceptor genes, *GmPhyA2* and -3, which affect the photoperiod-sensitive maturity loci, *E4* and *E3*, respectively (LIU *et al.* 2008; WATANABE *et al.* 2009). In some cases, the *E3* and *E4* alleles are found to have overlapping functions in flowering responses depending on photo quality and length (COBER *et al.* 1996; ABE *et al.* 2003), and the geographical distribution of the recessive *e4* allele is found to be restricted to the high-latitude regions of Japan (KANAZAWA *et al.* 2009). These facts indicate that duplication and subfunctionalization of *PhyA* genes contribute to the divergence of flowering time and maturity in soybean. Such duplicated gene phenomena may well explain the situation of the *GmGIs*. Premature stop codon mutations in the *GI* genes, induced by various mutagens, were found to have severe effects on flowering as well as pleiotropic effects on other traits in *Arabidopsis* (ARAKI and KOMEDA 1993; FOWLER *et al.* 1999) and pea (HECHT *et al.* 2007). *GI* is associated not only with circadian rhythms (MIZOGUCHI *et al.* 2005; MARTIN-TRYON *et al.* 2007) but also with several phenotypes and

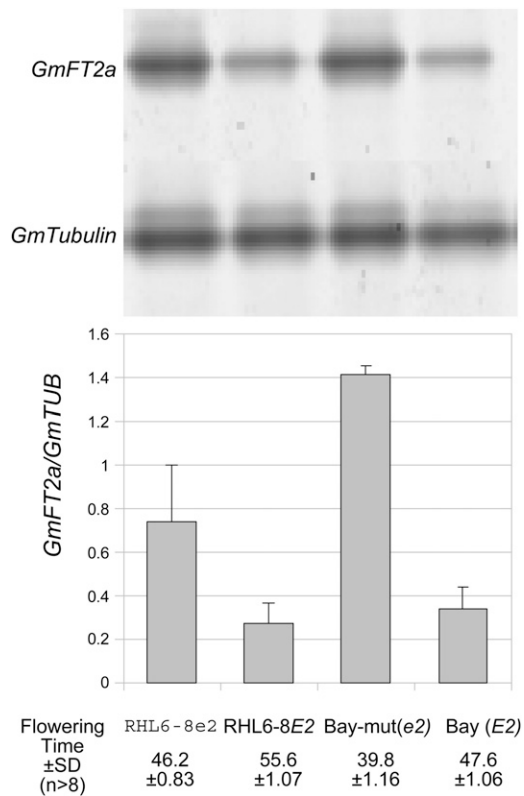


FIGURE 5.—Expression analysis of *GmFT2a* between NILs of *E2* under natural day-length conditions. Two pairs of NILs were used. Quantitative RT-PCR experiments for *GmFT2a* were performed using RNA isolated from trifoliolate leaves sampled at 9:00 AM 4 weeks after sowing. A representative polyacrylamide gel demonstrating the expression levels is shown. Each bar indicates the average expression level of *GmFT2a* compared with that of the *GmTubulin* gene with three independent replications and time to flowering; standard deviations of each line are shown below the bar graph.

physiological processes, including photomorphogenesis (PALTIEL *et al.* 2006; OLIVERIO *et al.* 2007), cold stress response (CAO *et al.* 2005), oxidative stress tolerance (KUREPA *et al.* 1998; CAO *et al.* 2006), and starch accumulation (EIMERT *et al.* 1995). Association analysis using the flowering gene in *Arabidopsis* has shown that the *FRIGIDA* and *FLC* genes, which are related to the vernalization pathway, contribute to geographic adaptation.

However, SNPs in the *GI* gene, not related to the null allele, had only small effects on flowering time under LD conditions in a natural accession (EHRENREICH *et al.* 2009) or in inbred lines derived from intercrosses with multiple parental lines (KOVER *et al.* 2009). Although *GI* overexpressing transgenic rice plants and specific RNA interference lines have been evaluated, a natural loss-of-function *GI* rice mutant has not yet been discovered. In barley, the correlation between the *GI* locus and some QTL related to flowering traits has also not yet been reported (DUNFORD *et al.* 2005). Compared with species that harbor a single *GI* gene in their genome, the loss of function of *E2* in soybean has little influence on growth except for phenotypes directly associated with flowering. The high level of sequence similarity (97%) between *GmGla* (Glyma10g36600) and *GmGlb* (Glyma20g30980) indicates that these genes originated from soybean lineage-specific duplication. Although we have no experimental data regarding the functions of *GmGlb*, it is likely that these duplicated soybean *GI* genes cooperate or functionally compensate for each other and that the loss of function of *GmGla* in soybean varieties contributes to their geographic adaptation, as with the *GmPhyAs*. Studies of such duplicated genes, represented by *GmGls* and *GmPhyAs*, especially in comparison with *Vitis vinifera* that contains no recent genome duplication (JAILLON *et al.* 2007) or with plants that have complex genomes, such as *Zea mays*, *T. aestivum*, and *Populus trichocarpa*, will provide further novel insights into the functions of these genes.

In this study, we have demonstrated, through a map-based cloning strategy using the progeny of an RHL, that the gene responsible for the soybean maturity locus *E2* is an ortholog of *GI*, *GmGla*. The *e2/e2* genotype caused early flowering by inducing the expression of the soybean florigen gene homolog, *GmFT2a*, whereas the effect of the *E2* allele on flowering under different environments was stable. These results indicate that the functions of *GI* in flowering are conserved in soybean and that null mutations in the *GI* gene may be useful resources for adapting plants with complex genomes, such as soybean, to a wide range of geographic regions. More information regarding the biochemical functions

TABLE 3

Analysis of the effect and interaction between gene and experimental location with two-way classification ANOVA

Factor	SS <sup>a</sup>	DF <sup>b</sup>	MS <sup>c</sup>	F value	Significance
Genotype of QTL	14,382.7	2	7191.4	2,803.5	***
Location	73,465.3	1	73465.3	28,640.4	***
Interaction (G×E)	93.9	2	46.9	18.3	***
Error	2,772.9	1081	2.6		
Total	94,857.3	1086	87.3		

<sup>a</sup>Sum of squares.

<sup>b</sup>Degree of freedom.

<sup>c</sup>Mean square.

\*\*\**p* < 0.001.

of the soybean *G1* genes and their interactions with other maturity genes are required, and these will provide new and intriguing insights into the flowering network in plants.

We thank R. L. Nelson at the U.S. Department of Agriculture/Agricultural Research Service, University of Illinois, for supplying seeds of the Harosoy and Harosoy-*E2* isolines. This study was partly supported by the Program for the Promotion of Basic Research Activities for Innovative Biosciences, Research Fellowships from the Japan Society for the Promotion of Science for Young Scientists (17-2554), a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Genomics for Agricultural Innovation, DD-2040), and a Grant-in-Aid for Scientific Research A (18208001).

#### LITERATURE CITED

- ABE, J., D. XU, A. MIYANO, K. KOMATSU, A. KANAZAWA *et al.*, 2003 Photoperiod-insensitive Japanese soybean landraces differ at two maturity loci. *Crop Sci.* **43**: 1300–1304.
- ABE, M., Y. KOBAYASHI, S. YAMAMOTO, Y. DAIMON, A. YAMAGUCHI *et al.*, 2005 FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* **309**: 1052–1056.
- ADAMS, K. L., and J. F. WENDEL, 2005 Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* **8**: 135–141.
- ARAKI, T., and Y. KOMEDA, 1993 Analysis of the role of the late-flowering locus, *G1*, in the flowering of *Arabidopsis thaliana*. *Plant J.* **3**: 231–239.
- BERNARD, R. L., 1971 Two major genes for time of flowering and maturity in soybeans. *Crop Sci.* **11**: 242–244.
- 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.
- 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 day length. *Soybean Genet. Newsl.* **7**: 26–29.
- CAO, S., M. YE and S. JIANG, 2005 Involvement of *GIGANTEA* gene in the regulation of the cold stress response in *Arabidopsis*. *Plant Cell Rep.* **24**: 683–690.
- CAO, S., S. JIANG and R. ZHANG, 2006 The role of *GIGANTEA* gene in mediating the oxidative stress response and in *Arabidopsis*. *Plant Growth Regul.* **48**: 261–270.
- CARMONA, M. J., M. CALONJE and J. M. MARTINEZ-ZAPATER, 2007 The *FT/TFL1* gene family in grapevine. *Plant Mol. Biol.* **63**: 637–650.
- 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.
- COBER, E. R., J. W. TANNER and H. D. VOLDENG, 1996 Soybean photoperiod-sensitivity loci respond differentially to light quality. *Crop Sci.* **36**: 606–610.
- 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.
- CREGAN, P. B., T. JARVIK, A. L. BUSH, R. C. SHOEMAKER, K. G. LARK *et al.*, 1999 An integrated genetic linkage map of the soybean genome. *Crop Sci.* **39**: 1464–1490.
- DOI, K., T. IZAWA, T. FUSE, U. YAMANOUCHI, T. KUBO *et al.*, 2004 *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT*-like gene expression independently of *Hdl*. *Genes Dev.* **18**: 926–936.
- DUBCOVSKY, J., and J. DVORAK, 2007 Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **316**: 1862–1866.
- DUNFORD, R. P., S. GRIFFITHS, V. CHRISTODOULOU and D. A. LAURIE, 2005 Characterisation of a barley (*Hordeum vulgare* L.) homologue of the *Arabidopsis* flowering time regulator *GIGANTEA*. *Theor. Appl. Genet.* **110**: 925–931.
- EHRENREICH, I. M., Y. HANZAWA, L. CHOU, J. L. ROE, P. X. KOVER *et al.*, 2009 Candidate gene association mapping of *Arabidopsis* flowering time. *Genetics* **183**: 325–335.
- EIMERT, K., S.-M. WANG, W. LUE and J. CHEN, 1995 Monogenic recessive mutation causing both late floral initiation and excess starch accumulation in *Arabidopsis*. *Plant Cell* **7**: 1703–1712.
- 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.
- FOWLER, S., K. LEE, H. ONOUCHI, A. SAMACH, K. RICHARDSON *et al.*, 1999 *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. *EMBO J.* **18**: 4679–4688.
- HAYAMA, R., S. YOKOI, S. TAMAKI, M. YANO and K. SHIMAMOTO, 2003 Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* **422**: 719–722.
- HAYASHI, M., A. MIYAHARA, S. SATO, T. KATO, M. YOSHIKAWA *et al.*, 2001 Construction of a genetic linkage map of the model legume *Lotus japonicus* using an intraspecific F2 population. *DNA Res.* **8**: 301–310.
- HECHT, V., F. FOUCHER, C. FERRANDIZ, R. MACKNIGHT, C. NAVARRO *et al.*, 2005 Conservation of *Arabidopsis* flowering genes in model legumes. *Plant Physiol.* **137**: 1420–1434.
- HECHT, V., C. L. KNOWLES, J. K. VANDER SCHEER, L. C. LIEW, S. E. JONES *et al.*, 2007 Pea *LATE BLOOMER1* is a *GIGANTEA* ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. *Plant Physiol.* **144**: 648–661.
- HONG, S. Y., S. LEE, P. J. SEO, M. S. YANG and C. M. PARK, 2010 Identification and molecular characterization of a *Brachypodium distachyon* *GIGANTEA* gene: functional conservation in monocot and dicot plants. *Plant Mol. Biol.* **72**: 485–497.
- HUQ, E., J. M. TEPPERMAN and P. H. QUAIL, 2000 *GIGANTEA* is a nuclear protein involved in phytochrome signaling in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **97**: 9789–9794.
- IMAIZUMI, T., H. G. TRAN, T. E. SWARTZ, W. R. BRIGGS and S. A. KAY, 2003 FKF1 is essential for photoperiodic-specific light signalling in *Arabidopsis*. *Nature* **426**: 302–306.
- IMAIZUMI, T., T. F. SCHULTZ, F. G. HARMON, L. A. HO and S. A. KAY, 2005 FKF1 F-box protein mediates cyclic degradation of a repressor of *CONSTANS* in *Arabidopsis*. *Science* **309**: 293–297.
- JAILLON, O., J. M. AURY, B. NOEL, A. POLICRITI, C. CLEPET *et al.*, 2007 The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**: 463–467.
- JUNG, C., and A. E. MULLER, 2009 Flowering time control and applications in plant breeding. *Trends Plant Sci.* **14**: 563–573.
- JUNG, J. H., Y. H. SEO, P. J. SEO, J. L. REYES, J. YUN *et al.*, 2007 The *GIGANTEA*-regulated microRNA172 mediates photoperiodic flowering independent of *CONSTANS* in *Arabidopsis*. *Plant Cell* **19**: 2736–2748.
- KANAZAWA, A., B. LIU, F. KONG, S. ARASE and J. ABE, 2009 Adaptive evolution involving gene duplication and insertion of a novel *Ty1/copia*-like retrotransposon in soybean. *J. Mol. Evol.* **69**: 164–175.
- KOJIMA, S., Y. TAKAHASHI, Y. KOBAYASHI, L. MONNA, T. SASAKI *et al.*, 2002 *Hd3a*, a rice ortholog of the *Arabidopsis* *FT* gene, promotes transition to flowering downstream of Hd1 under short-day conditions. *Plant Cell Physiol.* **43**: 1096–1105.
- KONG, F., B. LIU, Z. XIA, S. SATO, B. M. KIM *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.
- KOORNNEEF, M., C. ALONSO-BLANCO, H. BLANKESTIJN-DE VRIES, C. J. HANHART and A. J. PEETERS, 1998 Genetic interactions among late-flowering mutants of *Arabidopsis*. *Genetics* **148**: 885–892.
- KOVER, P. X., W. VALDAR, J. TRAKALO, N. SCARCELLI, I. M. EHRENREICH *et al.*, 2009 A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* **5**: e1000551.
- KUMAR, S., M. NEI, J. DUDLEY and K. TAMURA, 2008 MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* **9**: 299–306.

- KUREPA, J., J. SMALLE, M. V. MONTAGU and D. INZE, 1998 Oxidative stress tolerance and longevity in *Arabidopsis*: the late-flowering gigantea is tolerant to paraquat. *Plant J.* **14**: 759–764.
- LIFSCHITZ, E., T. EVIATAR, A. ROZMAN, A. SHALIT, A. GOLDSCHMIDT *et al.*, 2006 The tomato *FT* ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proc. Natl. Acad. Sci. USA* **103**: 6398–6403.
- LIN, M. K., H. BELANGER, Y. J. LEE, E. VARKONYI-GASIC, K. TAOKA *et al.*, 2007 *FLOWERING LOCUS T* protein may act as the long-distance florigenic signal in the cucurbits. *Plant Cell* **19**: 1488–1506.
- LIU, B., A. KANAZAWA, H. MATSUMURA, R. TAKAHASHI, K. HARADA *et al.*, 2008 Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. *Genetics* **180**: 995–1007.
- LIU, L. J., Y. C. ZHANG, Q. H. LI, Y. SANG, J. MAO *et al.*, 2008 COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in *Arabidopsis*. *Plant Cell* **20**: 292–306.
- MARTIN-TRYON, E. L., J. A. KREPS and S. L. HARMER, 2007 GIGANTEA acts in blue light signaling and has biochemically separable roles in circadian clock and flowering time regulation. *Plant Physiol.* **143**: 473–486.
- McBLAIN, B. A., and R. L. BERNARD, 1987 A new gene affecting the time of flowering and maturity in soybean. *J. Hered.* **78**: 160–162.
- McBLAIN, B. A., R. L. BERNARD, C. R. CREMEENS and J. F. KORCZAK, 1987 A procedure to identify genes affecting maturity using soybean isohline testers. *Crop Sci.* **27**: 1127–1132.
- McCALLUM, C. M., L. COMAI, E. A. GREENE and S. HEINKOFF, 2000 Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiol.* **123**: 436–442.
- MEKSEM, K., T. W. DOUBLER, K. CHANCHAROENCHAI, N. NJITI S. J. CHANG *et al.* 1999 Clustering among loci underlying soybean resistance to *Fusarium solani*, SDS and SCN in near-isogenic lines. *Theor. Appl. Genet.* **99**: 1131–1142.
- MIZOGUCHI, T., L. WRIGHT, S. FUJIWARA, F. CREMER, K. LEE *et al.*, 2005 Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in *Arabidopsis*. *Plant Cell* **17**: 2255–2270.
- MOLNAR, S. J., S. RAI, M. CHARETTE and E. R. COBER, 2003 Simple sequence repeat (SSR) markers linked to *E1*, *E3*, *E4*, and *E7* maturity genes in soybean. *Genome* **46**: 1024–1036.
- MURRAY, M. G., and W. F. THOMPSON, 1980 Rapid isolation of high-molecular-weight plant DNA. *Nucleic Acids Res.* **8**: 4321–4325.
- NELSON, D. C., J. LASSWELL, L. E. ROGG, M. A. COHEN and B. BARTEL, 2000 *FKF1*, a clock-controlled gene that regulates the transition to flowering in *Arabidopsis*. *Cell* **101**: 331–340.
- NEMOTO, Y., M. KISAKA, T. FUSE, M. YANO and Y. OGIHARA, 2003 Characterization and functional analysis of three wheat genes with homology to the *CONSTANS* flowering time gene in transgenic rice. *Plant J.* **36**: 82–93.
- NJITI, V. N., T. W. DOUBLER, R. J. SUTTNER, L. E. GRAY, P. T. GIBSON *et al.*, 1998 Resistance to soybean sudden death syndrome and root colonization by *Fusarium solani* f. sp. *glycine* in near-isogenic lines. *Crop Sci.* **38**: 472–477.
- OHYANAGI, H., T. TANAKA, H. SAKAI, Y. SHIGEMOTO, K. YAMAGUCHI *et al.*, 2006 The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic Acids Res.* **34**: D741–D744.
- OLIVERIO, K. A., M. CREPY, E. L. MARTIN-TRYON, R. MILICH, S. L. HARMER *et al.*, 2007 GIGANTEA regulates phytochrome A-mediated photomorphogenesis independently of its role in the circadian clock. *Plant Physiol.* **144**: 495–502.
- PALTIEL, P., R. AMIN, A. GOVER, N. ORI and A. SAMACH, 2006 Novel roles for GIGANTEA revealed under environmental conditions that modify its expression in *Arabidopsis* and *Medicago truncatula*. *Planta* **224**: 1255–1268.
- REDEL, G. P., 1962 Supervital mutants of *Arabidopsis*. *Genetics* **47**: 443–460.
- ROZEN, S., and H. J. SKALETSKY, 2000 Primer3 on the WWW for general users and for biologist programmers, pp. 365–386 in *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, edited by S. Krawetz, and S. Misener. Humana Press, Totowa, NJ.
- SAIDON, G., W. D. BEVERS DORF and H. D. VOLDENG, 1989 Adjusting of the soybean phenology using the *E4* loci. *Crop Sci.* **29**: 1361–1365.
- SAMACH, A., H. ONOUCHI, S. E. GOLD, G. S. DITTA, Z. SCHWARZ-SOMMER *et al.*, 2000 Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. *Science* **288**: 1613–1616.
- SATO, S., T. KANEKO, Y. NAKAMURA, E. ASAMIZU, T. KATO *et al.*, 2001 Structural analysis of a *Lotus japonicus* genome. I. Sequence features and mapping of fifty-six TAC clones which cover the 5.4 mb regions of the genome. *DNA Res.* **8**: 311–318.
- SAWA, M., D. A. NUSINOW, S. A. KAY and T. IMAIZUMI, 2007 FKF1 and GIGANTEA complex formation is required for day-length measurement in *Arabidopsis*. *Science* **318**: 261–265.
- SCHLUETER, J. A., P. DIXON, C. GRANGER, D. GRANT, L. CLARK *et al.*, 2004 Mining EST databases to resolve evolutionary events in major crop species. *Genome* **47**: 868–876.
- SCHMUTZ, J., S. B. CANNON, J. SCHLUETER, J. MA, T. MITROS *et al.*, 2010 Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178–183.
- SERIKAWA, M., K. MIWA, T. KONDO and T. OYAMA, 2008 Functional conservation of clock-related genes in flowering plants: overexpression and RNA interference analyses of the circadian rhythm in the monocotyledon *Lemna gibba*. *Plant Physiol.* **146**: 1952–1963.
- SHOEMAKER, R. C., K. POLZIN, J. LABATE, J. SPECHT, E. C. BRUMMER *et al.*, 1996 Genome duplication in soybean (*Glycine* subgenus *soja*). *Genetics* **144**: 329–338.
- SUAREZ-LOPEZ, P., K. WHEATLEY, F. ROBSON, H. ONOUCHI, F. VALVERDE *et al.*, 2001 *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature* **410**: 1116–1120.
- TASMA, I. M., and R. C. SHOEMAKER, 2003 Mapping flowering time gene homologs in soybean and their association with maturity (*E*) loci. *Crop Sci.* **43**: 319–328.
- TAYLOR, A., A. J. MASSIAH and B. THOMAS, 2010 Conservation of *Arabidopsis thaliana* photoperiodic flowering time genes in onion (*Allium cepa* L.). *Plant Cell Physiol.* **51**: 1638–1647.
- THAKARE, D., S. KUMUDINI and R. D. DINKINS, 2010 Expression of flowering-time genes in soybean *E1* near-isogenic lines under short and long day conditions. *Planta* **231**: 951–963.
- TRIWITAYAKORN, K., V. N. NJITI, M. J. IQBAL, S. YAEGASHI, C. TOWN *et al.*, 2005 Genomic analysis of a region encompassing *QRfs1* and *QRfs2*: genes that underlie soybean resistance to sudden death syndrome. *Genome* **48**: 125–138.
- TSUBOKURA, Y., R. ONDA, S. SATO, Z. XIA, M. HAYASHI *et al.*, 2008 Characterization of soybean genome based on synteny analysis with *Lotus japonicus*. *Breed. Sci.* **58**: 157–167.
- TSUCHIYA, T., N. KAMEYA and I. NAKAMURA, 2009 Straight walk: a modified method of ligation-mediated genome walking for plant species with large genomes. *Anal. Biochem.* **388**: 158–160.
- TUNSTRA, M. R., G. EJETA and P. B. GOLDSBROUGH, 1997 Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative traits loci. *Theor. Appl. Genet.* **95**: 1005–1011.
- VALVERDE, F., A. MOURADOV, W. SOPPE, D. RAVENSCROFT, A. SAMACH *et al.*, 2004 Photoreceptor regulation of *CONSTANS* protein in photoperiodic flowering. *Science* **303**: 1003–1006.
- VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE *et al.*, 1995 AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**: 4407–4414.
- WATANABE, S., T. TADJUDDIN, N. YAMANAKA, M. HAYASHI and K. HARADA, 2004 Analysis of QTLs for reproductive development and seed quality traits in soybean using recombinant inbred lines. *Breed. Sci.* **54**: 399–407.
- WATANABE, S., R. HIDESHIMA, Z. XIA, Y. TSUBOKURA, S. SATO *et al.*, 2009 Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics* **182**: 1251–1262.
- XIA, Z., H. SATO, S. WATANABE, S. KAWASAKI and K. HARADA, 2005 Construction and characterization of a BAC library of soybean. *Euphytica* **141**: 129–137.
- XUE, W., Y. XING, X. WENG, Y. ZHAO, W. TANG *et al.*, 2008 Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* **40**: 761–767.
- YAMANAKA, N., Y. NAGAMURA, Y. TSUBOKURA, K. YAMAMOTO, R. TAKAHASHI *et al.*, 2000 Quantitative trait locus analysis of flowering time in soybean using a RFLP linkage map. *Breed. Sci.* **50**: 109–115.
- YAMANAKA, N., S. NINOMIYA, M. HOSHI, Y. TSUBOKURA, M. YANO *et al.*, 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., S. WATANABE, K. TODA, M. HAYASHI, H. FUCHIGAMI *et al.*, 2004 Fine mapping of *FT1* locus for soybean flowering time, using a residual heterozygous line derived from a recombinant inbred line. *Theor. Appl. Genet.* **110**: 634–639.
- YAN, L., D. FU, C. LI, A. BLECHL, G. TRANQUILLI *et al.*, 2006 The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc. Natl. Acad. Sci. USA* **103**: 19581–19586.
- YANG, B., X. WEN, N. S. KODALI, C. A. OLEYKOWSKI, C. G. MILLER *et al.*, 2000 Purification, cloning, and characterization of the CEL I nuclease. *Biochemistry* **39**: 3533–3541.
- YANO, M., Y. KATAYOSE, M. ASHIKARI, U. YAMANOUCHI, L. MONNA *et al.*, 2000 *Hdl*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. *Plant Cell* **12**: 2473–2484.
- YANOVSKY, M. J., and S. A. KAY, 2002 Molecular basis of seasonal time measurement in *Arabidopsis*. *Nature* **419**: 308–312.
- ZHAO, X. Y., M. S. LIU, J. R. LI, C. M. GUAN and X. S. ZHANG, 2005 The wheat *TaGIL*, involved in photoperiodic flowering, encodes an Arabidopsis *GI* ortholog. *Plant Mol. Biol.* **58**: 53–64.

Communicating editor: A. H. PATERSON

# GENETICS

## Supporting Information

<http://www.genetics.org/cgi/content/full/genetics.110.125062/DC1>

### **A Map-Based Cloning Strategy Employing a Residual Heterozygous Line Reveals that the *GIGANTEA* Gene Is Involved in Soybean Maturity and Flowering**

**Satoshi Watanabe, Zhengjun Xia, Rumiko Hideshima, Yasutaka Tsubokura, Shusei Sato, Naoki Yamanaka, Ryoji Takahashi, Toyooki Anai, Satoshi Tabata, Keisuke Kitamura and Kyuya Harada**

Copyright © 2011 by the Genetics Society of America  
DOI: 10.1534/genetics.110.125062

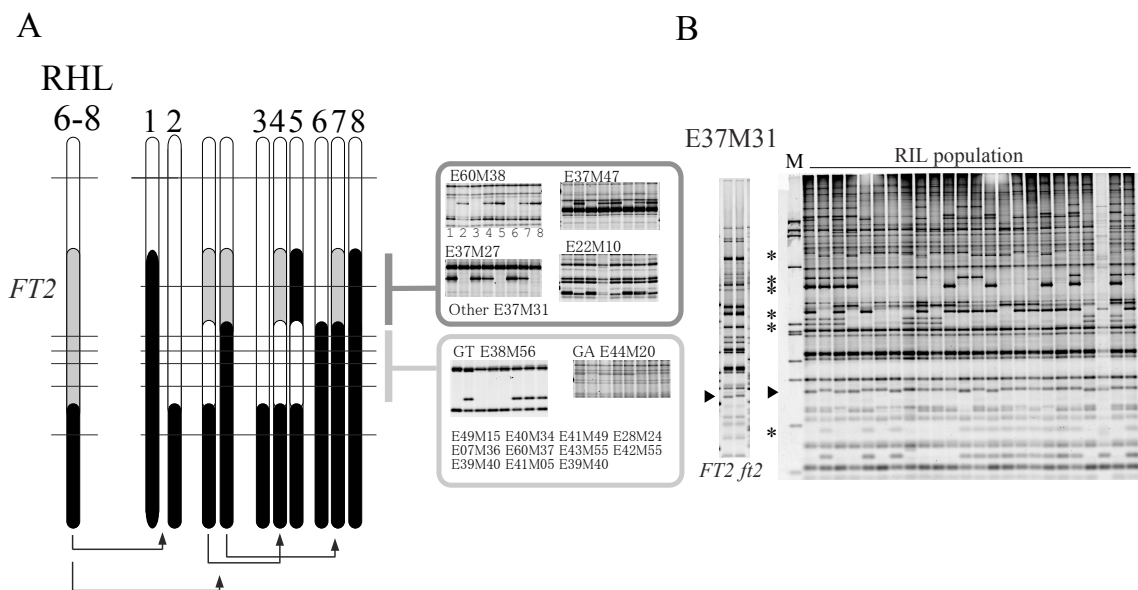


FIGURE S1.—Screening method of an AFLP marker tightly linked to the *FT2* locus using RHL derived from RIL6-8. (A) Graphical genotype of RHL6-8 and NILs are shown on the left side of the figure. The white, black and gray chromosomal segments represent the Misuzudaizu, the Moshidou Gong 503 and heterozygous alleles, respectively. The first screening of AFLP markers was performed using NILs6-8 (No 1 and 2). Polymorphic markers detected between NILs were reanalyzed using sub-families selected from the progenies of RHL6-8 (No 3-8). Five AFLP fragments tightly linked to the *FT2* locus were highly reproducible and were used in subsequent fine mapping experiments. Representative segregation patterns of each fragment among the NILs and RIL population using one AFLP marker, E37M31, are shown in the box and in B, respectively. Arrowhead indicates polymorphic band detected between NILs and asterisks indicate the polymorphic markers detected between Misuzudaizu and Moshidou Gong 503 originating from other genetic loci.

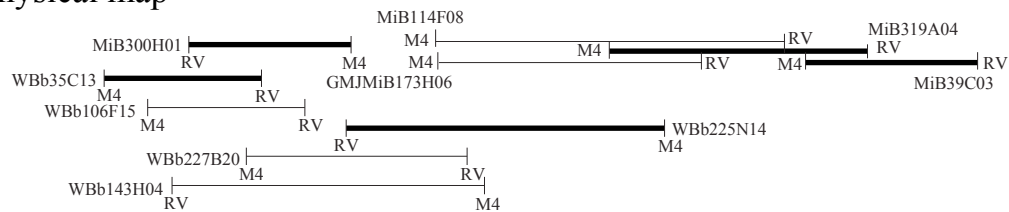
**A****Genetic map****B****Physical map**

FIGURE S2.—The physical contig covering the FT2 locus. SCAR markers originating from AFLP fragments identified 10 BAC clones in this study. Left and right sequences of each BAC clone were deposited in DDBJ under accession numbers AB554201-AB554220. Positional relationships among the DNA markers using fine mapping experiments (A) and the physical contig (B) are displayed. Five BAC clones, indicated by bold lines, were subject to shotgun sequence analysis and their nucleotide sequences deposited in DDBJ under accession numbers AP011822 (WBb35C13), AP011811 (MiB300H01), AP011821 (WBb225N14), AP011813 (MiB319A04) and AP011810 (MiB039C03), respectively.



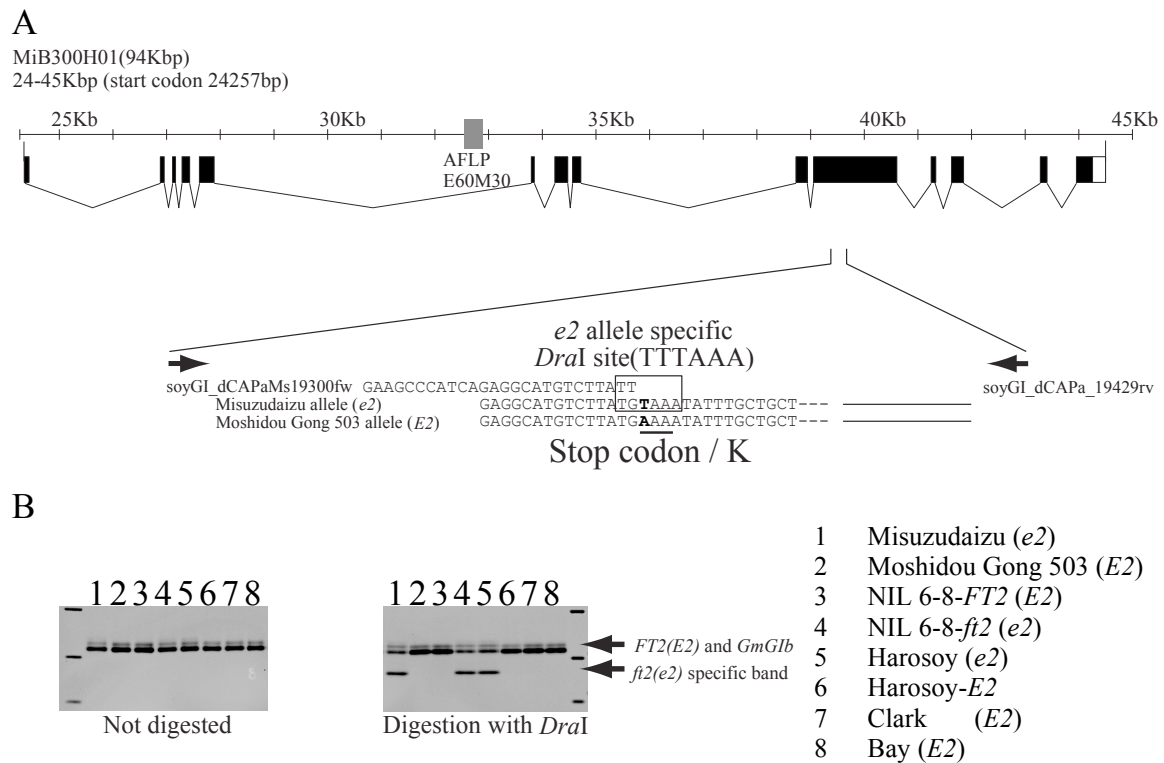


FIGURE S3.—Diagnostic derived cleaved amplified polymorphic sequences (dCAPs) marker detection of a premature stop codon in *GmG1a*. (A) The position of the stop codon mutation, the primer sequences and the specific *Dra*I site amplified with the dCAPs primers are displayed. Detailed primer sequences and experimental conditions of this dCAPS marker are shown in the Materials and methods. (B) Representative NILs mentioned in this study were analyzed (No. 1-8). PCR products, without or following *Dra*I restriction enzyme digestion, are compared. Arrows indicate the specific band originating from *e2* allele (*GmG1a*).

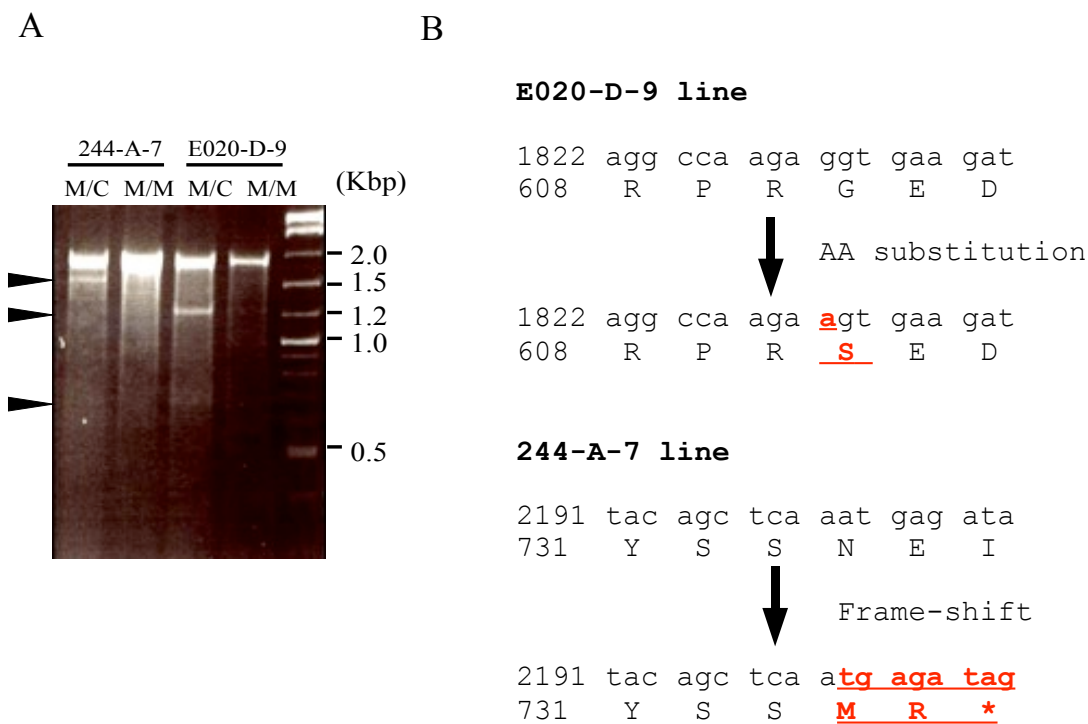


FIGURE S4.—Screening and characterization of a GmG1a mutant line. (A) A mutant library comprising of approximately thirty one thousand of M2 plants were examined. Formation of a heteroduplex and following CEL I digestion of the PCR fragment, amplified from a mixed template of the mutant and wild type DNA (M/C), exhibits a specific truncated DNA band because of mismatch sequences as determined by agarose gel electrophoresis. On the other hand, a non-mixed template (M/M) does not show the specific band. Mutation sites detected in each mutant was validated by direct sequencing of the PCR fragment. (B) One mutant line, 244-A-7, harboring a single nucleotide deletion in the 10th exon, and causing a premature stop codon, was used in the phenotypic evaluation (see text).

**TABLES S1-S7**

Tables S1-S7 are available for download as an Excel file at <http://www.genetics.org/cgi/content/full/genetics.110.125062/DC1>.

Table S1: Data for standard deviation in RILs

Table S2: Raw data for flowering time in RILs

Table S3: Phenotypic variation in parental lines and RHL6-8

Table S4: Raw data for fine mapping experiments

Table S5: Frequency of flowering time classified with the line and QTL genotype

Table S6: Average values for fine mapping and NIL evaluation

Table S7: Raw data for progeny test