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,†† Toyoaki Anai,‡ Satoshi Tabata,§ Keisuke Kitamura** and Kyuya Harada*‡†

*National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan, †Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150040, People’s Republic of China, ‡Faculty of Agriculture, Saga University, Saga 840-8502, Japan, §Kazusa DNA Research Institute, Kisarazu, Chiba 292-0812, Japan, ‡‡Japan International Research Center for Agricultural Sciences, Tsukuba, Ibaraki 305-8686, Japan, ††National Institute of Crop Science, Tsukuba, Ibaraki 305-8518, Japan and ‡‡‡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 “Misuzu-daizu” (j2/j2, 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 Misuzu-daizu 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.

FLOWERING 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...
protein degradation increases under LD conditions and induces the expression of FT in leaves (Yano and Kay 2002; Valverde et al. 2004; Liu et al. 2008). The *GI*GANTEA* (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, Hdh1, 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*, Hdh1 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 Hdh1 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*; Lischitz et al. 2006), onion (*Allium cepa*; Taylor et al. 2010), cucurbitas (*Cucurbita*), 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.phytozone.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 *GmFT1*, *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. Tunestra 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 (Njitti et al. 1998; Meksim 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<sub>8:10</sub>) 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 f1, fβ, and f3, 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 was 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 Chubu University, Matsudo (35°27′ N, 139°10′ E). One pair of NILs, with contrasting alleles for FT2, 6-8/FT2<sup>+</sup> and 6-8/FT2<sup>−</sup>, 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<sup>+</sup> and 168 plants for 6-8/FT2<sup>−</sup>), 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 GmGI (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; PI548533), Clark (E2/E2; PI548533), and the isolate 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<sub>E2</sub>Mn, with the letters E and M denoting the EcoRI and MseI primers, respectively, and the subsequent n and s 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 Hiyashi et al. (2001). Bulked segregant analysis (BSA) was used to identify the locus affecting segregration 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 advant Genetic Analyzer using a BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Japan, Tokyo) according to the manufacturer’s instructions. Primer3 (Rozen and Skallecky 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 (WbB353C13, MibB300H101, WbB225514, MibB319A04, and MibB093C03) 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° C for 30 sec, 58° C for 30 sec, and 72° C 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
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 | CACCTGGTTAGTTGTCTCCAAGCC |
| Marker 2<sup>a</sup> | BAC end | GMWBb35C13M4 | Reverse | CGAAGATCCACAGGCCCTAAG |
| Marker 3 | BAC end | GMJMiB300H01RV | Forward | GAACTCGGACCTTCTCAGAA |
| Marker 4<sup>a</sup> | AFLP (SCAR) | E60M58 | Reverse | AGCCCCATATGGCAGCCATAC |
| Marker 5 | BAC end | GMJMiB300H01FW | Forward | GAGAGGAGGTATTAGGATGA |
| Marker 6 | BAC end | GMJMiB319A04RV | Reverse | GAGACGGCCATTGATGTA |
| Marker 7 | BAC end | GMJMiB114F08RV | Forward | CCTTTGTTGTCTCTACGCA |

<sup>a</sup>Digestion with the restriction enzymes MseI, MspI, or EcoRI 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.

Expression analysis of soybean florigen gene: Two combinations of NILs, NILs 6–8, Bay and Bay-27 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 trifoliate leaf was collected, with three replications, at 7:00 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 29 cycles at 96°C for 30 sec, 58°C for 30 sec, and 72°C for 3 min. PCR products were isolated and cloned into the pGEM T-easy vector system and sequenced. To evaluate the null allele of GmFla, specific primers for the detection of the premature stop codon were 5′-AAGCCCTATGCCAAGTATGTATT-3′ and 5′-GAAGCCCTATGCCAAG-3′ for GmFlb were P1b (5′-CATCTGATCCCATGTCGA-3′) and P2. The PCR conditions using Ex-Taq consisted of 30 cycles at 96°C for 30 sec, 58°C for 30 sec, and 72°C for 3 min. PCR products were digested with Dral (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% EMS solution for 12 hr and then rinsed in tap water for 3 Gy/min. 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% EMS solution for 12 hr and then rinsed in tap water for 3 Gy/min. 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% EMS solution for 12 hr and then rinsed in tap water for 3 Gy/min. 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% EMS solution for 12 hr and then rinsed in tap water for 3 Gy/min.

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-ex10F1 (ATCGGGCCACGTTGCTCAGTCA) and GmGla-ex10R1 (GTGACAAAGATGGCCCATGCACT). The PCR reactions, using Pfu DNA polymerase and 0.5% DMSO, consisted of 40 cycles at 95°C for 30 sec, 70°C for 30 sec, and 72°C for 2 min. CEL I was purified from celery stalks as described previously (YANG et al. 2000).

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-ex10F1 (ATCGGGCCACGTTGCTCAGTCA) and GmGla-ex10R1 (GTGACAAAGATGGCCCATGCACT). The PCR reactions, using Pfu DNA polymerase and 0.5% DMSO, consisted of 40 cycles at 95°C for 30 sec, 70°C for 30 sec, and 72°C for 2 min. CEL I was purified from celery stalks as described previously (YANG et al. 2000).

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 (F2) and showed phenotypic segregation in the subsequent generation (F3). A region covering ~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 ± 1.72 days (n = 168) for 6-8-ft2 and 67.6 ± 1.56 (n = 210) days for 6-8-FT2. These results indicated that the progeny derived from RHLs 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 RHL 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 (Wb8382C13, MiB300H01, Wb2225N14, MiB319A04, and MiB3099C03) 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 ~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 RHL 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 ± 1.08), heterozygous (n = 420, 63.9 ± 1.95), and ft2 homozygous early flowering genotype (n = 210, 58.1 ± 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 -51.17 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 44,845 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
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. McClain 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 ± 1.16; n = 8) than the wild type (47.6 ± 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,
rice, and pea (Hayama et al. 2003; Mizoguchi et al. 2005; Hecht et al. 2007). We examined the relationship between the GmGIa (E2) gene and the soybean flowering 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 GmGIa 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 allele 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 (GxE; 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 GxE 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...
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 \( C(n, k) \cdot p^k \cdot (1 - p)^{n-k} \) on the basis of a binomial distribution. In the case of an F7 generation of RILs, the ratio of heterozygosity \( (p) \) is 1.56%, and with a population size of 200 \( (n) \), the probability 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 F6–F8 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 Lemma gibba (Dunford

---

**TABLE 2**

Results of progeny test

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

---

**Figure 3.**—Phylogenic analysis of GIGANTEA proteins. A phylogenic 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* (ABF38988), and *Lemma paucicostata* (BAD97864) were used.
In these studies, the expression of Gm 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 GmGANTEA 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 GmGla 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 GmGla, 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 GmGla 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 GmGls 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; Ober 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 GmGls. Premature stop codon mutations in the GI genes, induced by various mutants, were found to have severe effects on flowering as well as pleiotropic effects on other traits in Arabidopsis (Araki and Komeida 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
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 (Emert 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    | DF  | MS   | 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) | 95.9   | 1081 | 46.9  | 18.3    | ***          |
| Error           | 2,772.9 | 1086 | 2.6   |         |              |
| Total           | 94,857.3 | 1086 | 87.3  |         |              |

*aSum of squares.
*bDegree of freedom.
*cMean square.
***p < 0.001.
of the soybean GI 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, 1995 Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. 8: 135–141.

Akai, T., and Y. Komiya, 1993 Analysis of the role of the late-flowering locus, GI, 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 conferring 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 cold stress response and in Arabidopsis. Plant Growth Regul. 48: 261–270.

CarmONA, M. J., M. CalonJé, and J. M. Martinez-Zapater, 2007 The FT/FTL1 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 $E2$ and $T$. Crop Sci. 41: 608–701.

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.

Doki, K., T. Izawa, T. Fuse, U. Yamanouchi, T. Kurbo et al., 2004 Eth1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of $Hd1$. Genes Dev. 18: 926–936.

Duclos, J. 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. LEE 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. Ferrandez, 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 Schoor, L. C. Liew, S. E. Jones et al., 2007 pea late bloomer is a GIGANTEA orthologue with roles in photoperiodic flowering, dehiscence, 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.

Hug, E. M., J. M. Tepperman and P. H. Quail, 2000 GIGANTEA is a nuclear protein involved in photchrome singnaling 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 photoperiod-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: 295–297.

Jailon, O., J. M. Aury, B. Noel, A. Policriti, C. Clefet et al., 2007 The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449: 465–467.

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 Ty3/copia-like retrotransposon in soybean. J. Mol. Evol. 69: 164–175.

Kojima, S., Y. Takahashi, Y. Kobayashi, L. Monna, T. Sasaki et al., 2002 Hdl3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hdl under short-day conditions. Plant Cell Physiol. 43: 1106–1105.

Kong, F., B. Liu, Z. Xia, S. Sato, B. M. Kim et al., 2010 Two coordinate 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. Blankenship-de Vries, C. J. Hanhart and A. J. Peeters, 1998 Genetic interactions among late-flowering mutants of Arabidopsis. Genetics 148: 885–889.

Kover, P. X., W. Valdar, J. Trakalo, N. Scarcelli, I. M. Ehrenreich et al., 2000 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 MEGA4: 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.

Lipschitz, E., T. Evtatar, A. Rozman, A. Shaltit, A. Goldsmith 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 florigen signal in the cucurbits. Plant Cell 19: 1488–1506.

Liu, B., A. Kanazawa, H. Matsunuma, T. Takahashi, K. Harada et al., 2008 Genetic redundancy in soybean responses 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. Creemans and J. F. Korczak, 1986 A rapid procedure to identify genes affecting maturity using soybean isolate 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. D. Doubler, K. Chancharenchais, N. Njiti S. J. Chang et al., 1999 Clusterering 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. Fujisawa, 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. Charrette 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.

Nishimura, Y., M. Ueda, F. Fuse, M. Yano and Y. Ohasha, 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. Doubrle, 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. Saka, Y. Shigemoto, K. Yamauchi et al., 2006 The Rice Annotation Project Database (RAP-DB): hub for Oryza sativa ssp. japonica genome information. Nucleic Acids Res. 34: D741–D744.

Olivero, 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. F., 1962 Supervital mutants of Arabidopsis. Genetics 47: 443–469.

Rozen, S., and H. J. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers, pp. 356–386 in Bioinformatics Methods and Protocols: Methods in Molecular Biology, edited by S. Krawetz and S. Misener. Humana Press, Totowa, NJ.

Saidon, G., W. D. Beversdorf and H. D. Voldeng, 1989 Adjusting of the soybean phenology using the E4 loci. Crop Sci. 29: 1361–1365.
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. BLECH, 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. KODALL, 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. KATAYOSI, 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.

YANOFSKY, M. J., and S. A. KAY, 2002 Molecular basis of seasonal time measurement in Arabidopsis. Nature 419: 308–312.

ZHANG, X. Y., M. S. LIU, J. R. LI, C. M. GUAN and X. S. ZHANG, 2005 The wheat TaGI1, involved in photoperiodic flowering, encodes an Arabidopsis GI ortholog. Plant Mol. Biol. 58: 53–64.

Communicating editor: A. H. Paterson
A Map-Based Cloning Strategy Employing a Residual Heterozygous Line Reveals that the \textit{GIGANTEA} Gene Is Involved in Soybean Maturity and Flowering

Satoshi Watanabe, Zhengjun Xia, Rumiko Hideshima, Yasutaka Tsubokura, Shusei Sato, Naoki Yamanaka, Ryoji Takahashi, Toyoaki 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 Misuzudaiuz, the Moshidou Gong 503 and heterozygous alleles, respectively. The first screening of AFLP markers was performed using NIL6-8 (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.
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 GmGIa. (A) The position of the stop codon mutation, the primer sequences and the specific DraI 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 DraI restriction enzyme digestion, are compared. Arrows indicate the specific band originating from e2 allele (GmGIa).
**Figure S4.**—Screening and characterization of a GmGIA 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