



Title	Molecular Genetic Study on a Nobel Early Flowering Mutant That Mitigates Floral Repression by Cool Temperature in Soybean
Author(s)	張, 静語
Degree Grantor	北海道大学
Degree Name	博士(農学)
Dissertation Number	甲第14207号
Issue Date	2020-09-25
DOI	https://doi.org/10.14943/doctoral.k14207
Doc URL	https://hdl.handle.net/2115/79562
Type	doctoral thesis
File Information	Zhang_Jingyu.pdf



**Molecular Genetic Study on a Novel Early Flowering Mutant That
Mitigates Floral Repression by Cool Temperature in Soybean**

(低温による開花抑制を緩和するダイズの新規早生変異体に関する
分子遺伝学的研究)

**Hokkaido University Graduate School of Agriculture
Division of Agrobiological and Bioresources Doctoral Course**

Jingyu Zhang

Contents

Chapter 1	Introduction	4
Chapter 2	Identification of molecular basis for an early-flowering mutant.....	14
2.1	Background and Purposes	14
2.2	Materials and Methods	15
2.2.1	Flowering time segregation analysis.....	15
2.2.2	DNA extraction and SSR marker analysis	16
2.2.3	Construction of genomic sequences harboring the mutation based on whole-genome resequencing data.....	16
2.2.4	Sequencing of genomic region flanking a deletion.....	17
2.3	Results and Discussion	18
2.3.1	Segregation analysis and mapping of the early flowering mutant	18
2.3.2	Fine-mapping of the mutant gene <i>efl</i>	28
2.3.3	Identification of the molecular basis for <i>efl</i> based on resequencing data	30
Chapter 3	Characterization of the early-flowering habit in the mutant	35
3.1	Background and Purposes	35
3.2	Materials and Methods	35
3.2.1	Evaluation of photo-thermal responses.....	35
3.2.2	RNA extraction and expression analyses	36
3.3	Results and Discussion	37

3.3.1	Effects of <i>efl</i> on flowering	37
3.3.2	Expression profiles of flowering genes under different thermal conditions	41
Chapter 4	Molecular diversity of <i>COL</i> genes	47
4.1	Background and Purposes	47
4.2	Materials and Methods	48
4.2.1	Resequencing data and SNP calling	48
4.3	Results and Discussion	48
Chapter 5	General Discussion	57
Summary		68
References		71
Appendix		94
Acknowledgements		129

Chapter 1 Introduction

Crops have been developed to produce greater yields in growing environments by adjusting the reproductive phenology in response to various external signals (Andrés and Coupland, 2012; King, 2015). Photoperiods and ambient temperatures are two major environmental factors, particularly in summer crops, such as rice (*Oryza sativa*) and soybean (*Glycine max*) (Song et al., 2015; Zhang et al., 2019). The molecular mechanisms of the photo-thermal responses of flowering have been extensively studied in model plant species, *Arabidopsis thaliana* and rice (Kaufmann et al., 2009; Shrestha et al., 2014; Andrés and Coupland, 2012; Zhang et al., 2019). Some of the genes identified in these species have conserved functions across various plant species, while others do species or lineage-specific functions (Kaufmann et al., 2009; Tsuji et al., 2011; Wu and Hanzawa, 2014; Brambilla et al., 2017). Crops have developed their own well-designed networks of flowering-related genes.

Arabidopsis is a long-day (LD) plant, in which flowering is induced in longer daylengths. Under optimum growth temperature (22–23°C) conditions, the transcript abundance of *FLOWERING LOCUS T (FT)*, which functions as a florigen (Corbesier et al., 2007; Jaeger and Wigge, 2007; Wigge, 2011), is low in SD (non-inductive photoperiods) conditions, but highly accumulated in LD (inductive photoperiods) conditions, accelerating the timing of flowering. This photoperiodic flowering mainly relies on the function of the CONSTANS (CO) protein, which acts as an upstream transcriptional activator of *FT* (Samach et al., 2000). Moreover, *GIGANTEA (GI)*, a circadian clock-controlled gene, acts upstream of *CO* (Huq et al., 2000; Mizoguchi et al., 2005). GI forms a complex with FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1) in the late afternoon under LD and triggers the degradation of CYCLING DOF FACTOR 1 and its homologs that repress the *CO* expression, resulting in an increase in the level of *CO* transcripts at the end of the day (Imaizumi et al., 2005; Sawa et al., 2007). The photoreceptors PHYTOCHROME A (PHYA), CRYPTOCHROME 1 (CRY1), and CRYPTOCHROME 2 (CRY2), as well as FKF1, stabilize the CO protein toward the end of the day, leading to the activation of *FT*, which in turn induces floral meristem identity

genes, *APETALA1* (*API*) and *LEAFY* (*LFY*), thereby initiating flowering (Valverde et al., 2004; Song et al., 2012a). The activation of *FT* transcription by the GI calibrated CO protein enables the transition from vegetative to reproductive growths under inductive environmental conditions.

Rice is a typical SD plant, in which the flowering is induced when the daylength shortens below the critical daylength. As in *Arabidopsis*, the *CO* ortholog *Heading date1* (*Hd1*), which is activated by the rice GI ortholog OsGI, triggers the expression of the *FT* homologs *Hd3a* (*Heading date 3a*) and *RFT1* (*Rice FT-like 1*) under flowering-inductive SD conditions (Yano et al., 2000; Hayama et al., 2003; Komiya et al., 2008). However, the rice Hd1 acts as a floral repressor under non-inductive LD conditions; a near-isogenic line (NIL) of cv. Nipponbare for *Hd1* exhibited later heading than Nipponbare (*hd1*) under LD conditions but earlier heading under SD conditions (Lin et al., 2000). The dual function of Hd1 in SD and LD conditions is not observed in *Arabidopsis*, and is controlled by an interaction between Hd1 and the monocot-specific CCT domain protein, Grain number, plant height and heading date 7 (*Ghd7*) (Nemoto et al., 2016; Zhang et al., 2017). *Ghd7* acts as a suppressor for *Early heading date 1* (*Ehd1*), an activator of the *Hd3a* and *RFT1* expressions, by associating with a cis-regulatory region in *Ehd1* (Doi et al., 2004; Xue et al., 2008; Shrestha et al., 2014; Weng et al., 2014; Nemoto et al., 2016). Hd1 essentially acts as a transcriptional activator to upregulate *Ehd1*, *Hd3a* and *RFT1* in both LD and SD conditions (Zhang et al., 2017). However, the function is dependent on the genotype at *Ghd7*; in the presence of functional *Ghd7*, it delays heading under LD conditions (Zhang et al., 2017). Hd1 forms a protein complex with *Ghd7*, the CCT domain of which binds to the transcription-activating domain of Hd1 (Zhang et al., 2017). This interaction blocks or weakens the transcriptional activation activity of Hd1 and releases the transcriptional repression activity of *Ghd7* to repress the expression of *Ehd1* (Nemoto et al., 2016; Zhang et al., 2017). Under SD conditions, the *Ghd7* expression is downregulated (Xue et al., 2008) and thereby without forming the protein complex with *Ghd7*. *Hd1* promotes the *Hd3a* expression either directly or via the upregulation of *Ehd1*, resulting in early heading (Zhang et al., 2017). Accordingly, Hd1 acts as a promoter of heading date, but the protein interaction between *Ghd7* and Hd1 determines photoperiod sensitivity in rice (Nemoto et al., 2016; Zhang et al., 2017).

Soybean, another typical SD plant, possesses two distinct modules, PHYA-E1 and GI-CO modules, in the control of flowering in LD conditions (Cao et al., 2017). Soybean is a major legume crop cultivated worldwide for human food, animal feed, and oil production for edible and industrial uses. Its wide adaptability has been supported by diverse cultivars with different photoperiod sensitivities, despite that the cultivation areas of individual cultivars are restricted to relatively narrow ranges of latitudes. Variation in flowering time among soybean cultivars has been mainly attributed to the genetic variation in photoperiod responses.

Until now, 12 major loci regulating flowering time and maturity have been recognized in soybean: *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), *E8* (Cober et al., 2010), *E9* (Kong et al., 2014; Zhao et al., 2016), *E10* (Samanfar et al., 2017), *E11* (Wang et al., 2019), and *J* (Ray et al., 1995).

E1 encodes a legume-specific putative transcription factor (Glyma.06g207800) that contains a bipartite nuclear localization signal and a region distantly related to the B3 domain; the overexpression strongly represses the expression of two soybean orthologs of *Arabidopsis FT*, *FT2a* (Glyma.16g150700) and *FT5a* (Glyma.16g044100), and inhibits floral initiation (Xia et al., 2012). The *E1* locus consists of multiple alleles, functional *E1*, hypomorph allele *e1-as*, loss-of-function alleles *e1-nl* and *e1-fs* (Xia et al., 2012; Xu et al., 2013; Tsubokura et al., 2014). The expression of *E1* and *E1-like* genes, *E1La* (Glyma.04g156400) and *E1Lb* (Glyma.04g143300), are controlled by the maturity genes *E3* and *E4*; the expression of *E1* and *E1L* genes is induced under LD, but is repressed under SD and in the *e3/e4* double recessive homozygote (Xia et al., 2012; Xu et al., 2015).

E2 was identified as an ortholog of *GI*, *Gla* (Glyma.10g221500) (Watanabe et al. 2011). *E2* has two alleles: a functional dominant allele (*E2*) and a recessive loss-of-function allele (*e2*). The *E2* allele encodes a protein of 1170 amino acids, while the *e2* allele is a nonsense allele encoding a truncated protein of 521 amino acids. *E2* inhibits

flowering under LD by mainly suppressing the *FT2a* expression: the *e2* allele promotes early flowering by upregulating the *FT2a* expression but has no detectable effect on the expression of *FT5a* (Watanabe et al., 2011). The *e2* allele is predominant in early-maturing cultivars (Wang et al., 2016a) and photoperiod insensitive ones (Xu et al., 2013).

E3 and *E4* encode PHYA3 (Glyma.19g224200) and PHYA2 (Glyma.20g090000), respectively (Liu et al. 2008; Watanabe et al. 2009). *E3* and *E4* direct different flowering responses to LD conditions with different red-to-far-red (R:FR) quantum ratios (Cober et al. 1996). *E3* responds to light with a high R:FR ratio, while *E4* does to light with a low R:FR ratio under LD conditions (Buzzell, 1971; Buzzell and Voldeng, 1980; Saindon et al., 1989; Cober et al., 1996; Xu et al., 2015). The double recessive homozygote for loss-of-function alleles *e3* and *e4* conditions the insensitivity to FR-enriched incandescent-lamp induced LD (Buzzell and Voldeng, 1980; Saindon et al., 1989; Cober et al., 1996). In the *e3/e4* genotype, the *E1* expression is not induced even under LD conditions (Xia et al., 2012). However, in the LD conditions with extremely low R:FR ratios, *E1* still retains the inhibitory effect on flowering, possibly due to the homoeolog of *E4*, *PHYA1* (Glyma.10g141400)-mediated response (Cober et al., 1996; Watanabe et al., 2012).

Dissanayaka et al. (2016) carried out the test crosses to identify the molecular basis of *E5*. In the Harosoy-*E5* (*e1/e2/E3/E4/E5*) × Clark-*e2* (*e1/e2/E3/E4/e5*) population, QTLs were found near the *E2* gene, suggesting that *E5* located in the close vicinity of *E2* or *E5* was an allele at the *E2* locus. However, no unique and consistent QTL corresponding to *E5* was detected in the progeny of test crosses between Harosoy-*E5* and Harosoy-*E2* (L74-21 or L74-27) (McBlain and Bernard, 1987) or in the progeny of a cross of Harosoy (*e1/e2/E3/E4/e5*) with PI80837 (*E5*), a donor of *E5*. Amplicon length polymorphism and dCAPS analyses further revealed that Harosoy and PI80837 both had the *e2-ns* allele, whereas Harosoy-*E5* and Harosoy-*E2* had the *E2-dl* and *E2-in* alleles, respectively. The *E2-dl* and the *E2-in* alleles possess the identical coding sequence but have many nucleotide polymorphisms in the 5' upstream region, exons, introns and 3' downstream region (Tsubokura et al., 2014). The *E5* allele in Harosoy-*E5* may therefore

have been generated by unexpected outcrossing with pollen having the *E2-dl* allele. Taken together, they concluded which that a unique *E5* gene may not exist.

E6 and *J* are involved in the long-juvenile trait; dominant alleles at both loci condition earlier flowering in SD conditions (Ray et al., 1995; Bonato and Vello, 1999). The *J* gene encodes an ortholog of *Arabidopsis* EARLY FLOWERING 3 (Glyma.04g050200) (Yue et al., 2017; Lu et al., 2017). Under SD conditions, the J protein directly binds to the promoter of *E1* to inhibit the expression; in the plants homozygous for dysfunctional *j* alleles, the *E1* expression is induced even under SD conditions to strictly inhibit the flowering (Lu et al., 2017). It is reported that *E6* and *J* are tightly linked or allelic to each other (Lu et al., 2017).

E7 and *E8* are identified in the progeny of crosses between early-maturing cultivars with the maturity genotype of *e1-as/e2/e3/e4* (Cober and Voldeng, 2001; Cober et al., 2010). *E7* is linked to *E1* in LG C2 (Chr. 6) (Cober and Voldeng, 2001; Molnar et al., 2003), which is shown to have the weakest effect on delaying flowering among five *E* locis, *E1*, *E2*, *E3*, *E4* and *E7* (Rosenzweig et al., 2008). *E8* is located in LG C1 (Chr. 4), where *E1La* and *E1Lb* are located (Cober et al., 2010).

E9 is identified in segregating populations of a cross between wild and cultivated soybeans (Liu et al., 2007; Kong et al., 2014). By fine-mapping and sequencing analysis, Zhao et al. (2016) revealed that *E9* encodes FT2a; the recessive *e9* allele is a hypomorph allele, the transcript abundance of which is lowered by the allele-specific transcriptional repression due to an insertion of a Ty1/copia-like retrotransposon, *SORE-1*, in the first intron. The *e9* allele has an effect to maintain the vegetative growth, in combination with early-flowering alleles at *E1* locus, such as loss-of-function *e1-nl* or hypomorph *e1-as* alleles (Zhao et al., 2016).

E10 is located on chromosome 8 and its recessive allele *e10* confers the earlier flowering (Samanfar et al., 2017). The mapping study suggests that the candidate gene is *FT4* (Glyma.08g363100) (Samanfar et al., 2017). *E11* is located on chromosome 7 and its dominant allele *E11* promotes flowering under LD conditions (Wang et al., 2019).

In addition to these major genes, many QTLs controlling flowering time have been detected in soybean; a total of 104 QTLs on 17 chromosomes (except Chr. 03, 05 and 17), which were identified until 2014, are recorded on SoyBase (<http://soybase.org>). Some of the QTLs most likely correspond to the known maturity genes, *E1* to *E4* (Watanabe et al., 2004; Funatsuki et al., 2005; Yamanaka et al., 2005; Githiri et al., 2007; Khan et al., 2008; Liu and Abe, 2010; Cheng et al., 2011; Lu et al., 2015; Lu et al., 2017; Yang et al., 2017). The molecular dissection of QTLs is important in better understanding of the molecular mechanisms underlying natural variations of flowering time. Recent studies have revealed the causal genes and DNA polymorphisms responsible for some of the QTLs. Takeshima et al. (2016) identified the insertion/deletions (indels) in the promoter and 3' untranslated region (UTR) of *FT5a* was likely involved in the QTLs for flowering in linkage group J (Chromosome16), designated as *qDTF-J1* (Yamaguchi et al., 2014), which was detected in the crosses between the early-flowering cultivars and breeding lines. Zhu et al. (2019) determined that a loss-of-function allele due to a single-base deletion in the coding region of *EILb* conferred the photoperiod-insensitive early-flowering habit in Far-Eastern Russian soybean cultivars. Sun et al. (2019) demonstrated that a non-synonymous substitution in the exon 4 of *FT2a* is responsible for the delayed flowering allele at the QTL, *qDTF-J2*, from a Thai late-flowering cultivar K3 (Pooprompan et al., 2006; Lu et al., 2015). The sequence analyses further indicated that the nucleotide substitution at *FT2a* was strongly associated with the late-flowering in diverse populations (Jiang et al., 2019; Ogiso-Tanaka et al., 2019). By using allelism and association analyses, Li et al. (2019a) indicated that an ortholog of *PSUDO RESPONSE REGULATOR 7* (*PRR7*; Glyma.12g073900) is the most probable candidate for *qFT12-1*. By using the overexpression of *PRR37* (also called as *PRR7*) and CRISPR/Cas9-induced *prp37* mutants, Wang et al. (2020) concluded that *PRR37* (*PRR7*) is responsible for the major QTL *qFT12-2*, which is overlapped with *qFT12-1*. Through whole-genome resequencing, genome-wide association studies and positional cloning of QTL, Lu et al. (2020) identified two QTLs located in the homoeologous genome regions, *Tof11* and *Tof12*, of which the latter of which corresponds to *qFT12-1/2*, encodes homoeologous PRR proteins (*PRR3A*; Glyma.11g155800 and *PRR3B*; Glyma.12g073900), which may have contributed to the changes in flowering and maturity times during the early stage of

soybean domestication. The effects of *Tof11* and *Tof12* are dependent on the *E1* genotype; in the *E1* background, they suppress the expressions of four *LATE ELONGATED HYPOCOTYL (LHY)* orthologs, *LHY1a* (Glyma.16g015500), *LHY1b* (Glyma.07g046900), *LHY2a* (Glyma.19g245000) and *LHY2b* (Glyma.03g240700), by physically binding to their promoters (Lu et al., 2020). The LHY proteins further associates with the *E1* promoter to suppress the expression, promoting early flowering and maturity (Lu et al., 2020).

In addition to the genetic study with ordinary biparental crossings, the genome-wide association study (GWAS), based on linkage disequilibrium, has emerged as a powerful tool to detect the DNA polymorphisms underlying natural variations of traits noted; it provides higher QTL mapping resolution to overcome the limitations of biparental populations (Malosetti et al., 2011; Weng et al., 2011; Zhao et al., 2011; Verbyla et al., 2014). The GWAS by Zhang et al. (2015) revealed a number of SNPs that are significantly associated with flowering time; some of the SNPs implied an involvement of orthologs to *Arabidopsis* flowering genes, such as *EARLY FLOWERING 8* (Glyma.09g07980) and *SUPPRESSOR OF OVEREXPRESSION OF CO 1* (Glyma.07g08831) or *AGAMOUS LIKE 6* (Glyma.07g08890), in the control of flowering time in soybean. Mao et al. (2017) identified 87 markers (11 SSRs and 76 SNPs) associated with flowering time in 91 soybean cultivars of maturity groups (MG000-VIII), of which three loci, Gm04_4497001 (near *COL3a*; Glyma.04g06240, Wu et al., 2014), Gm16_30766209 (near *FT2a* and *FT2b*; Glyma.16g151000) and Gm19_47514601 (near *E3*), were proximal to key genes in the control of flowering time. Among these, the SNP Gm04_4497001 was likely a core locus with epistatic interaction with other loci for controlling flowering time (Mao et al., 2017). Zatybekov et al. (2017) identified two markers associated with flowering time, Gm06_20370075 (near *PHYA1*) and Gm19_49964637 (near *E3*), in the collection of 113 accessions planted in northern, eastern, and south-eastern regions of Kazakhstan. Li et al. (2019b) identified 37 significant SNPs associated with flowering time and other growth period related traits; five probable candidate flowering genes existed near each of four significant SNPs, Gm5_27111367 (*FRIGIDA*; Glyma.05g101800), Gm11_10629613 and Gm11_10950924 (plant *U-box 13*; Glyma.11g140100 and *MYB59*; Glyma.11g142900),

Gm19_34768458 (*CONSTANS*; Glyma.19g099700, *FUS3*; Glyma.19g100900). Further studies with genetic and/or genome editing approaches are needed to confirm whether the candidate genes indicated by GWAS are responsible for the natural variation of flowering time.

Ambient growth temperature also has a large impact on reproductive development in soybean, like in the other plant species (Hardey et al., 1984; Upadhyay et al., 1994; Cober et al., 2001; Thomas et al., 2010; Kumagai and Sameshima, 2014; Mao et al., 2017; Sun et al., 2019). The thermal response of flowering in soybean varies with photoperiods and maturity genotypes. In photoperiods below the critical daylength, increasing temperatures accelerate flowering up to an optimum temperature (~30°C), but further increase adversely delays the flowering (Hardey et al., 1984; Upadhyay et al., 1994; Cober et al., 2001; Thomas et al., 2010). Cober et al. (2001) found different responses of flowering to high temperature between photoperiod-sensitive and insensitive genotypes; under the 20-h photoperiod, flowering of photoperiod-sensitive genotypes was noticeably delayed in 28°C compared with 18°C, whereas such delay was not observed in insensitive genotypes. Sun et al. (2019) found a delay of flowering under 32°C in a photoperiod-insensitive soybean line with a late-flowering trait, which was introduced from a Thai cultivar, a unique characteristic which has not been reported in ordinary photoperiod-insensitive lines.

PHYTOCHROME INTERACTING FACTOR 4 (PIF4) is as an important factor modulating temperature response in many plants; it binds to the promoter of *FT* to activate the expression in *Arabidopsis* (Wigge, 2013). *CO* is also involved in the thermal regulation of flowering, particularly in response to high temperatures, in *Arabidopsis*. Increasing the ambient temperature from the optimum to 27°C accelerates flowering even under non-inductive SD conditions in *Arabidopsis* (Balasubramanian et al., 2006; Sanchez-Bermejo et al., 2015). This acceleration of flowering by high temperature is mediated by *CO* and PIF4; *CO* forms a protein complex with PIF4 to activate the *FT* expression under high-temperature SD conditions (Fernández et al., 2016).

The transcription of a soybean ortholog of *PIF4* (Glyma.14g032200) is enhanced at 35°C compared to 25°C under SD conditions, suggesting that *PIF4* may have unique temperature adaptation in soybean as well (Arya et al., 2018). Besides, there are several genes associated with flowering under different temperature conditions. A soybean ortholog of GAMYB-binding protein (Glyma.01g008600) promotes flowering and enhances heat-tolerance when overexpressed in tobacco (*Nicotiana tabacum*) and *Arabidopsis* plants (Zhang et al., 2013; Zhao et al., 2013). A soybean ortholog of *VERNALIZATION 1* (Glyma.11g13220, Glyma.11g132200 in *Glycine max* Wm82.a2.v1), a gene of vernalization pathway, can respond to low temperature and may function as a floral repressor during low temperature, which may play an important role in ensuring reproductive success against cold stress (Lü et al., 2015). A soybean ortholog (Glyma.05g148700) of *Arabidopsis* *FLOWERING LOCUS C* is involved in the response to long-term low temperature and as a floral repressor in delaying flowering time, providing protective mechanisms against sporadic and extremely low temperatures (Lyu et al., 2020). However, the genetic variation and its underlying molecular mechanism on thermal regulation of flowering, in particular in response to cool temperatures, remain barely explored in soybean.

As mentioned above, CO is a key player which functions in the photoperiodic and thermal pathways. The soybean genome has 26 *CO-like* (*COL*) genes, of which four *CO* homologs, *COL1a* (Glyma.18g278100), *COL1b* (Glyma.08g255200), *COL2a* (Glyma.13g050300) and *COL2b* (Glyma.19g039000), possess high amino acid similarities to *CO* (Wu et al., 2014). The expression of *COL1a* and *COL1b* is induced by darkness, peaks at dawn, and is suppressed by light, whereas the expression of *COL2a* and *COL2b* is at much lower levels and peak both at dawn and in the late afternoon or at dusk (Fan et al., 2014; Cao et al., 2015). These four CO homologs belong to a single clade connected to *Arabidopsis* CO, COL1 and COL2; they are more similar to COL2 out of the three *Arabidopsis* CO/COL proteins, whereas CO is closer to COL2a/COL2b than to COL1a/COL1b (Wu et al., 2014). When ectopically expressed under the control of the cauliflower mosaic virus 35S promoter, each of the four homologs fully complements the late flowering of the *co-1* mutant in *Arabidopsis*, suggesting that they retain the function of *CO* (Wu et al., 2014). Analyses of overexpression and artificially induced missense

mutants revealed that both *COL1a* and *COL1b* inhibit flowering under LD conditions through the suppression of *FT2a* and *FT5a* and the activation of *FT4* under LD (Cao et al., 2015). However, under SD conditions, the overexpression of *COL1a* did not promote flowering, although it up-regulated the *FT2a* and *FT5a* expressions (Cao et al., 2015). This is in contrast to the dual function of rice Hd1 in SD and LD conditions (Lin et al. 2000). By using RNA interference (RNAi) analyses, Wu et al. (2019) indicated that both *COL1* and *COL2* homologs are involved in the control of flowering through the regulation of *E1* expression; in *COL1* and *COL2* RNAi-plants grown under LD conditions, the peak expression of *E1* was repressed and the expression of *FT2a* and *FT5a* was upregulated. Therefore, the *COL* genes in soybean possess the inhibitory effects on flowering in LD, as in rice, although they retain the same function as CO in the *Arabidopsis* background.

In this study, I focused on an early-flowering mutant generated by the ion-beam mutagenesis and conducted a molecular genetic study to determine the causal DNA polymorphism and characterize the function in the control of flowering in soybean. In Chapter 2, I found that the early-flowering mutation is attributed to the deletion of the genome region of 214 kb in chromosome 19 by mapping and re-sequencing analysis, and that the deletion contained a *COL* gene, *COL2b*. In Chapter 3, I characterized the function of the mutation in the thermal regulation of flowering to chilling temperatures. In Chapter 4, I evaluated the molecular diversity for *COL2b* and other *COL* homologs as possible candidates for the natural variation of flowering time in soybean. Finally, I discussed on a possible involvement of *COL2b* in the thermal regulation of flowering, in particular in response to cool weather environments in soybean.

Chapter 2 Identification of molecular basis for an early-flowering mutant

2.1 Background and Purposes

Mutation breeding using X-ray and gamma-ray irradiations or chemical mutagens has developed soybean cultivars with desirable traits of yield, plant height, flowering/maturity and seed size (Rawlings et al., 1958); seed components (Williams and Hanway, 1961; Takagi et al., 1990; Fehr et al., 1991; Hajika et al., 1991; Kitamura, 1991; Takahashi et al., 1994; Hayashi et al., 1998), nodule formation (Carroll et al., 1985), or herbicide tolerance (Sebastian et al., 1989). Mutation rates after low linear energy transfer (LET) radiation, such as X-ray and gamma-ray irradiations, have been extensively studied in plants; they ranged from 3% to 32% (Tanaka et al., 2010).

Ion beams have been utilized as mutagens since late 1960's (Fujii et al., 1966; Fujii et al., 1967; Mei et al., 1994), because the high LTR created by ion beams causes large structural alterations in the genome such as deletion up to 1.2 Mbp in length, inversion or translocation of a large DNA segment (Shikazono et al., 2005; Tanaka et al., 2010; Hirano et al., 2012). In soybean, Arase et al. (2011) and Mikuriya et al. (2017) obtained various kinds of mutants for flowering time, seed isoflavone contents and leaf chlorophyll contents in a mutant population of a Japanese cultivar 'Nourin No.2' (N2) that was generated by ion-beam irradiation on seeds. Mikuriya et al. (2017) found a number of mutants with altered flowering times; interestingly, most of the early-flowering mutants had yellowish leaves with reduced chlorophyll contents and decreases in isoflavones in seeds. The associated mutations may be caused by the large deletions involving a loss of multiple linked genes (Mikuriya et al., 2017).

In this chapter, I performed the genetic analysis for an early-flowering mutant N2-ef1, which was selected from the same mutant population as Arase et al. (2011) and Mikuriya et al. (2017) surveyed. Here, I demonstrated the results obtained from fine-mapping and re-sequencing analyses for the mutation and discussed the probable causal factor.

2.2 Materials and Methods

2.2.1 Flowering time segregation analysis

An early-flowering mutant (N2-ef1) selected from a mutant population of Nourin No. 2 (N2) generated by ion-beam irradiation (Arase et al., 2011) was used in this study. It was crossed with an early-flowering cultivar ‘Tokachi Nagaha’ (TN) in place of parent, N2, to facilitate mapping of the mutant with simple sequence repeat (SSR) markers. The F₂ population was cultivated in the Institute of Crop Science, Tsukuba, Japan (36.03°N, 140.10°E). The genetic analyses were performed at Hokkaido University, Sapporo, Japan (43.07°N, 141.34°E). Single F₃ seeds from each of 119 F₂ plants, together with seeds of the parents (N2-ef1 and TN), were sown in paper pots (Paperpots No. 2, Nippon Beet Sugar Manufacturing Co., Tokyo, Japan) on 14 June 2013, and were then transplanted 10 days later at the experimental farm of Hokkaido University. A piece of undeveloped young trifoliate leaves was sampled from each F₃ plants, and the individual flowering times (number of days after sowing; DAS) were recorded. The segregation of flowering time was examined in 13 F₃ plants that were selected on the basis of the genotype at a tagging marker (Sat_405, detected from the association test described later in the Methods) for the mutant gene in N2-ef1 in 2014 (12 June sowing date) and in 25 F₄ plants from three F₃ families in 2015 (5 June sowing date). The numbers of plants tested in the progeny were 15 to 56 plants in the F₃ plants, and 10 to 186 plants in F₄ plants. Fine-mapping of the mutant gene was carried out in the progeny of a heterozygous F₄ plant (#46-16) from 2016 to 2019 (28 May to 6 June sowing dates); 15 plants were tested in each progeny.

2.2.2 DNA extraction and SSR marker analysis

Total DNA was extracted individually from leaves sampled from the parents and segregants using the modified CTAB method (Doyle and Doyle, 1990) or from seeds of heterozygous plants using the proteinase K DNA extraction method (Kamiya and Kiguchi, 2003). Genotypes at 160 SSRs of N2-ef1 and TN were examined, and 111 polymorphic

markers (Supplemental Table S1) were genotyped for 119 F₃ plants by Drs. T. Sayama and M. Ishimoto of Institute of Crop Science. One-way analysis of variance was used to detect significant associations between flowering time and marker genotype. In the subsequent generations, the tagging markers and their flanking BARCSOY SSR markers (Song et al., 2010) were used to develop the near-isogenic lines (NILs) and narrow the genomic region that harbored the mutation. For fine mapping, a total of 500 seeds from two heterozygous plants were genotyped using two markers (BARCSOY 19_279 and 19_330). The recombinant seeds in the targeted region were germinated and the resultant plants were grown in the greenhouse during the winter of 2018. The flowering genotypes were estimated from the segregation in the progeny in the summer of 2019. The genomic position of the mutation was determined by comparison of the estimated genotypes with the graphical genotypes constructed using the BARCSOY SSR markers. SSR marker analysis was performed as described previously (Zhao et al., 2016). The primers used in fine mapping are listed in Supplemental Table S2.

2.2.4 Construction of genomic sequences harboring the mutation based on whole-genome resequencing data

DNA sequences of N2 and N2-efl were obtained by next-generation sequencing on a HiSeq XTen sequencer (Illumina, San Diego, CA, USA). Raw reads were aligned to the soybean reference genome Williams 82 (Phytozome v.12.1/*Glycine max* Wm82.a2.v1; <https://phytozome.jgi.doe.gov/pz/portal.html>; Schmutz et al., 2010) in Bowtie 2-2.2.9 software (<https://bioweb.pasteur.fr/packages/pack@bowtie2@2.2.9>; Langmead et al., 2009). The resulting alignment was further processed to remove duplicate reads and to correct mate and its pair information in the Picard tools (<http://broadinstitute.github.io/picard>). Small indels were realigned in GATK v. 3.8 software (<https://gatkforums.broadinstitute.org/gatk/>; McKenna et al., 2010), in which the Unified Genotyper function, which filtered out reads with a mapped base-quality Phred score of <20, called variants (SNPs and indels). Using the reference genome and a SNP dataset for each cultivar, the sequences of the targeted genomic regions were reconstructed using the FastaAlternateReferenceMaker function of GATK and compared

between N2-efl and N2 using the Integrative Genomics Viewer (<https://software.broadinstitute.org/software/igv/>; Robinson et al., 2011). The resequencing data used in this study were submitted under BioProject accession number PRJNA600284 in NCBI.

2.2.5 Sequencing of genomic region flanking of a deletion

The deletion in the target region for the mutation in N2-efl was confirmed by means of PCR amplification and sequencing. Two sets of primers were designed for the boundaries of the deletion, based on the re-sequenced data obtained from the deep-sequencing analyses for N2 and N2-efl. The boundary regions were amplified by PCR with the primers F1 (5'-GAGTGGAAGATGACTAATGCAAGGT-3') and R2 (5'-AGATGGTTTCCGGATGAAATGATTTGGG-3') in N2-efl, and with the primers F1 and R1 (5'-GACATTTTGGGTATGTTTTCTTAG-3') and F2 (5'-GTTAAGCTCATTTAAAGGATCCAAGT-3') and R2 in N2 and TN. The amplified fragments from N2-efl were cloned into a pGEM-T Easy vector (Promega, Durham, NC, USA) and sequenced with a BigDye Terminator v. 3.1 Cycle Sequencing kit and an ABI PRISM 3100 Avant Genetic Analyzer (both from Applied Biosystems, Tokyo, Japan) according to the manufacturer's instructions. I used PCR with three of the primers (F1, R1, and R2) as a deletion marker (DEL214) to determine the presence or absence of the deletion in the segregating population. The PCR products were separated by electrophoresis in 2% agarose gel, stained with ethidium bromide, and visualized under UV light.

2.3 Results and Discussion

2.3.1 Segregation analysis and mapping of the early-flowering mutant

The early-flowering mutant N2-efl flowered ~20 days earlier than the wild-type N2 after sowing at standard times in Sapporo (i.e., late May). It showed no other phenotypic aberration, in contrast to the early-flowering mutants with yellowish leaves with low chlorophyll contents selected by Mikuriya et al. (2017) from the same population.

TN, which was used as a parent of the crossing, had the same maturity genotype as N2 at four major loci (*E1/e2/e3/E4*) but, as in N2-ef1, flowered ~20 days earlier than N2 in Sapporo. Flowering time in the F₃ population varied continuously across the flowering times of the parents (Figure 1). The parents did not differ greatly in their average flowering times: 54.4 DAS in N2-ef1 and 56.0 DAS in TN. In contrast, the F₃ population showed a wide variation, ranging from 48 to 68 DAS, which suggests that in addition to the mutant gene, a number of unknown flowering genes that affect the difference in flowering times between N2 and TN segregated in the population.

One-way analysis of variance detected significant associations between the SSR marker genotype and flowering time for nine of the 111 polymorphic markers tested (Table 1). Five markers on chromosomes (chr) 2, 9, and 11 showed that the TN allele promoted flowering; plants homozygous for the TN allele flowered, on average, 2.6 to 3.7 days earlier than those for the N2-ef1 allele. In contrast, four markers, one each on chr4 and chr15 and two on chr19, had the opposite effect: plants homozygous for the N2-ef1 allele flowered, on average, 2.5 to 7.9 days later than those for the TN allele. Because the marker Sat_405 on chr19 had the largest effect on flowering time and the N2-ef1 allele conditioned the early flowering, I focused on the genetic factor that co-segregated with Sat_405 as a candidate for the early-flowering mutation in the subsequent genetic analysis.

Thirteen-three F₃ plants were selected for the progeny test: three plants homozygous for each of the TN and N2-ef1 alleles and 7 heterozygous plants. The segregations of flowering time in their F₄ progeny were mostly consistent with the expectation from the segregations associated with the Sat_405 genotype in the F₃ population; plants homozygous for the N2-ef1 allele (#36, 71, and 88) produced mainly early-flowering progeny, whereas plants homozygous for the TN allele (#64, 78, and 98) produced mainly late-flowering progeny, although the segregation patterns varied among the F₃ plants in each homozygote. Four plants (#11, 25, 30, and 46) of the seven heterozygous plants produced progeny that segregated widely, from early-flowering plants homozygous for the N2-ef1 allele to late-flowering plants homozygous for the TN allele (Figure 2). The

parent–offspring correlation coefficient for the 13 families ($r = 0.911$) was strong and significant ($P < 0.001$).

Among the segregating families, three (#11, 25, and 46) were selected for further analyses. All three families exhibited close associations between the Sat_405 genotype of the F₄ plants and the segregation pattern for flowering time in the progeny (Figure 3 for #46, Figure 4 for #11 and Figure 5 for #25), as was observed in the progeny of the F₃ plants (Figure 2). The F₄ plants homozygous for the N2-efl allele produced only early-flowering progeny in all families, whereas those for the TN allele produced mainly late-flowering progeny, but their segregation patterns were variable, in particular in families #11 and 25 (Figures 4 and 5), suggesting that several genes in addition to the mutant gene were still segregating in these families. In contrast, in progeny of family #46, flowering time segregated mostly within a limited range of 42 to 58 DAS; three plants homozygous for the N2-efl allele (#46-16, 46-28, and 46-30) and one for the TN allele (#46-04) produced only early-flowering and late-flowering progeny, respectively, whereas heterozygous plants (#46-2, 46-14, 46-17, and 46-20) produced progeny that segregated mostly in the range between the progeny of the homozygous plants for the N2-efl and that for the TN alleles (Figure 3).

Next, I randomly selected 93 plants from the 164 F₅ plants tested in the progeny of F₄ plant #46-14 and tested the segregation of flowering time in the progeny. On the basis of the segregation pattern, the 93 plants were classified into three classes: class 1, plants that produced only early-flowering progeny; class 2, plants that showed segregation of flowering time; and class 3, plants that produced only late-flowering progeny (Figure 6). Twenty-three plants were classified into class 1, 45 into class 2, and 25 into class 3. The segregation ratio of 23:45:25 was in good accordance with an expected 1:2:1 ratio from monogenic Mendelian inheritance ($\chi^2 = 0.18$, $P = 0.91$). Average flowering times in the progeny ranged from 48 to 52 DAS (with an overall average of 51.0) in the class 1 F₅ plants and from 58 to 61 DAS (with an overall average of 60.4) in the class 3 F₅ plants; plants in class 1 thus flowered, on average, approximately 9 days earlier than those in class 3 (Figure 6A). Average flowering times in the progeny of the class 2 F₅ plants ranged from 52 to 59 DAS (with an overall average of 56.2), which slightly overlapped with the

average flowering times of plants in classes 1 and 3. Furthermore, flowering times (46 to 64 DAS) of the F₆ progeny derived from the heterozygous F₅ plants (i.e., heterozygous progeny) overlapped the times (46 to 55 DAS) for the plants that were homozygous for early-flowering (class 1) and the times (57 to 63 DAS) for the plants that were homozygous for late-flowering (class 3); the late-flowering plants, which flowered similarly to those in class 3, segregated with higher frequency than the early-flowering plants, which flowered similarly to those in class 1 (Figure 6B). These results suggest that the segregation of flowering time in the family 46-14 was controlled mainly by a single gene, in which the early flowering from N2-efl was conditioned by a recessive allele (hereafter, tentatively designated *efl*).

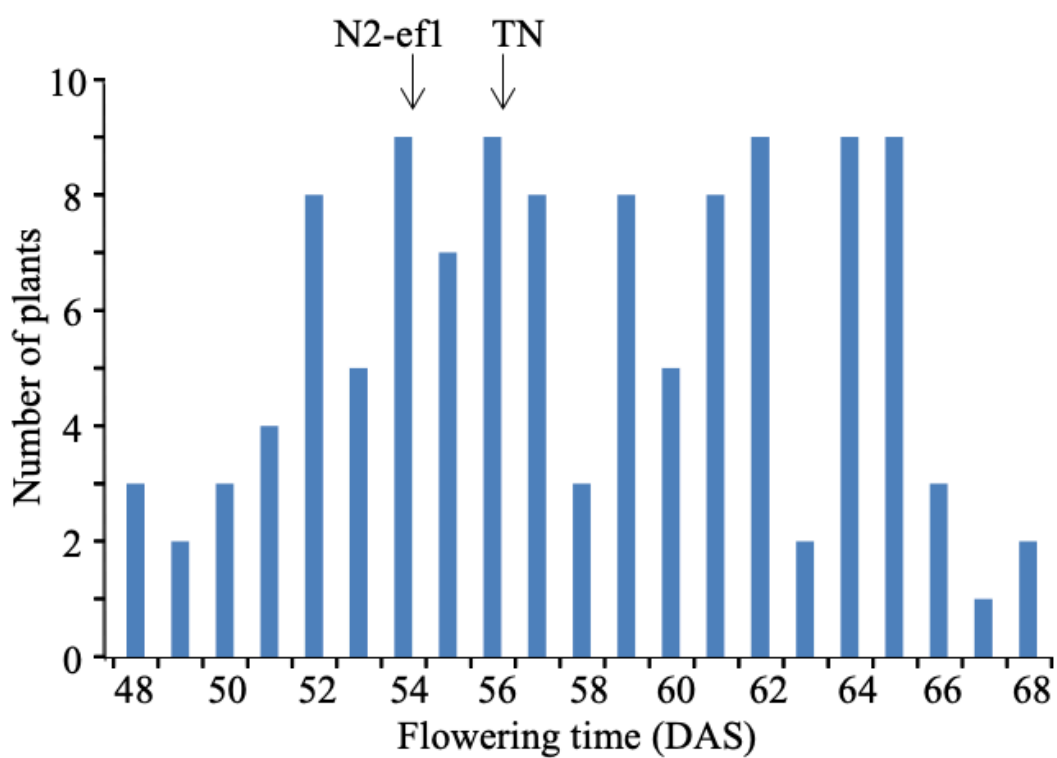


FIGURE 1. Segregation of flowering time in the F₃ population of the cross between ‘Nourin No. 2’ early-flowering mutant 1 (N2-ef1) and ‘Tokachi Nagaha’ (TN) in Sapporo. DAS, number of days after sowing. Arrows represent the mean flowering times of the parent accessions.

Table 1. Association tests between SSR marker genotypes and flowering time in the F₃ progeny of the cross between an early-flowering mutant 1 of Nourin No. 2 (N2-efl) and Tokachi Nagaha (TN).

Chr	LG	SSR loci	Marker genotypes				<i>F</i>	<i>p</i>	Donrer of early- flowering allele
			at SSR						
			AA		BB				
			<i>n</i>	Mean	<i>n</i>	Mean			
2	D1b	Satt459	35	37.5	54	34.9	5.74	0.0187	TN
4	C1	Satt565	41	34.5	41	37.4	6.54	0.0124	N2-efl
9	K	Satt417	42	36.7	51	34.5	4.35	0.0398	TN
9	K	Satt559	57	36.6	43	34.5	4.17	0.0438	TN
9	K	Sat_352	41	37.4	46	34.3	7.17	0.0089	TN
11	B1	Satt197	39	37.7	45	34.1	9.82	0.0024	TN
15	E	Sat_124	36	34.9	51	37.4	4.76	0.0319	N2-efl
19	L	Sat_301	37	32.4	37	38.5	23.74	6.3E-06	N2-efl
19	L	Sat_405	43	31.4	40	39.3	44.79	2.6E-09	N2-efl

One-way analysis of variance was carried out for the mean values of flowering time (days after sowing) between two homozygotes.

AA and BB; homozygotes for the allele from N2-efl (A) and TN (B), respectively.

n; number of plants tested. Chr; chromosome, LG; linkage group

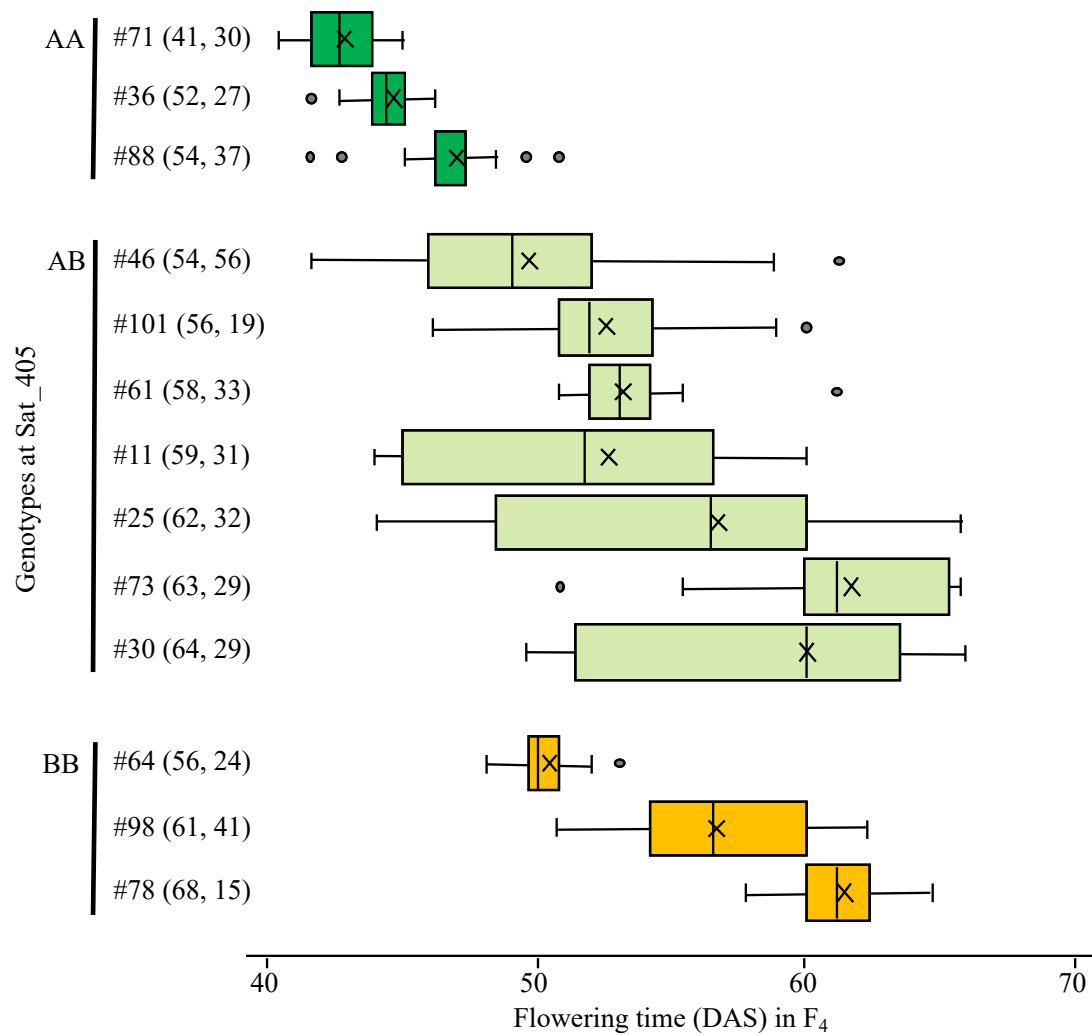


FIGURE 2. Box-plots of segregation of flowering time in the progeny of 13 F₃ plants selected on the basis of the marker genotypes at Sat_405 in a cross between ‘Nourin No. 2’ early-flowering mutant 1 (N2-ef1) and ‘Tokachi Nagaha’ (TN). Alleles: A, from N2-ef1; B, from TN. Values in parentheses represent the flowering times of the F₃ parent (left number) and number of plants tested in the progeny (right number). The parent–offspring correlation coefficient was 0.911 ($P < 0.001$). DAS, number of days after sowing. ×, mean value.

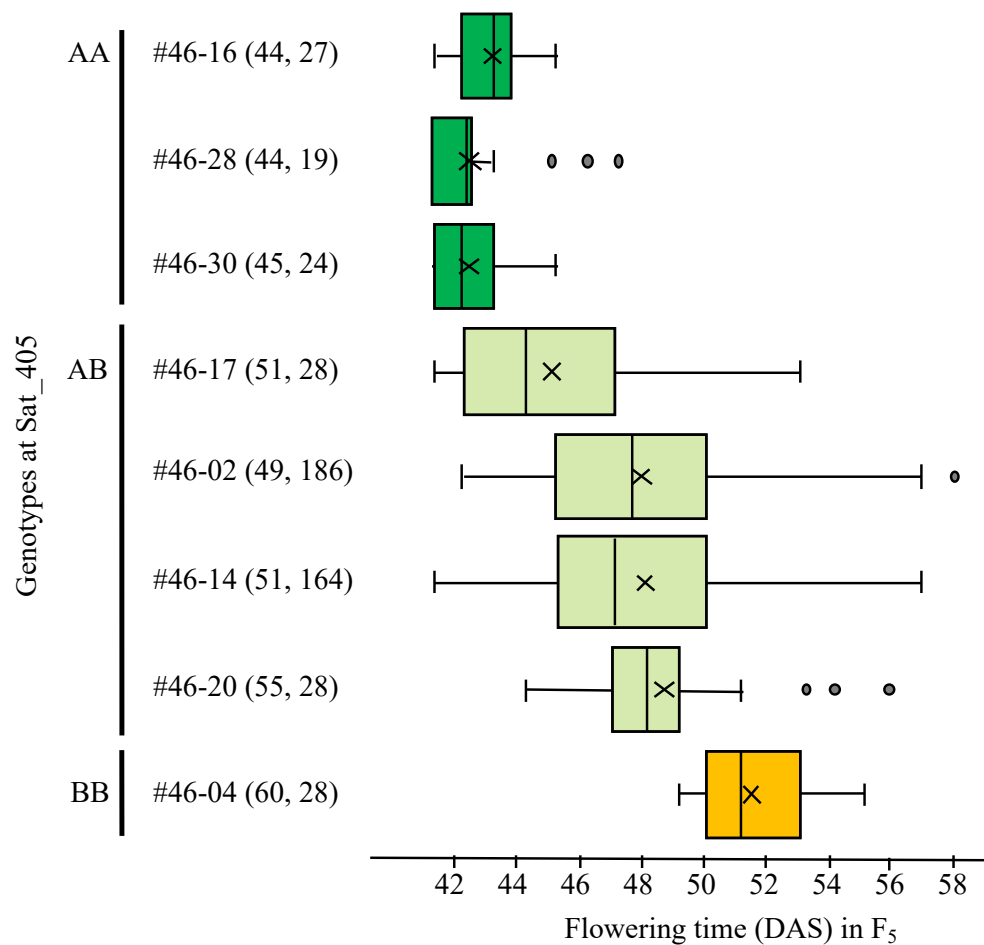


FIGURE 3. Boxplots of segregation of flowering time in the progeny of 8 F_4 plants from family #46 of the cross between ‘Nourin No. 2’ early-flowering mutant 1 (N2-ef1) and ‘Tokachi Nagaha’ (TN). The 8 plants were selected on the basis of the marker genotypes at Sat_405. Alleles: A, from N2-ef1; B, from TN. Values in parentheses represent the flowering time of the F_4 parent (left number) and the number of plants tested in the progeny (right number). The parent–offspring correlation coefficient was 0.921 ($P < 0.001$). DAS, number of days after sowing. ×, mean value.

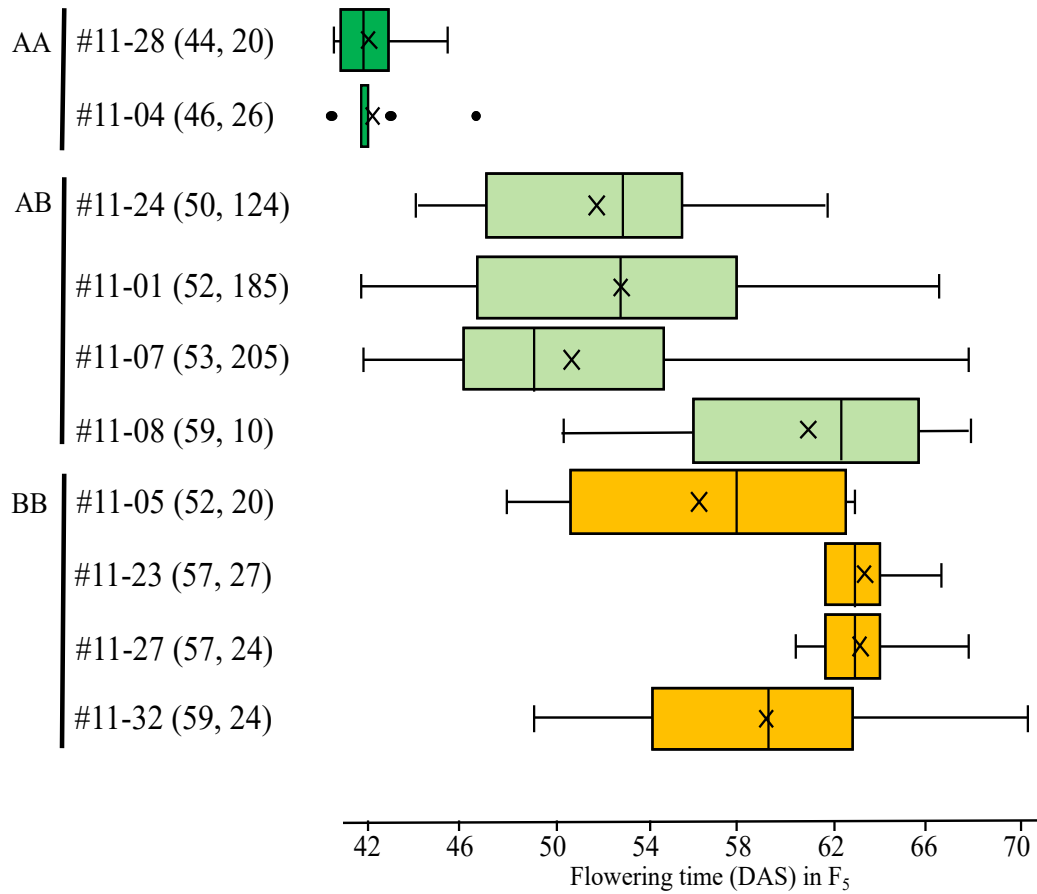


FIGURE 4. Boxplots of segregation of flowering time in the progeny of 10 F₄ plants from family #11 of the cross between ‘Nourin No. 2’ early-flowering mutant 1 (N2-ef1) and ‘Tokachi Nagaha’. The 10 plants were selected on the basis of the marker genotypes at Sat_405. Alleles: A, from N2-ef1; B, from TN. Values in parentheses represent the flowering time of the F₄ parent (left number) and the number of plants tested in the progeny (right number). The parent–offspring correlation coefficient was 0.902 ($P < 0.001$). DAS, number of days after sowing. ×, mean value.

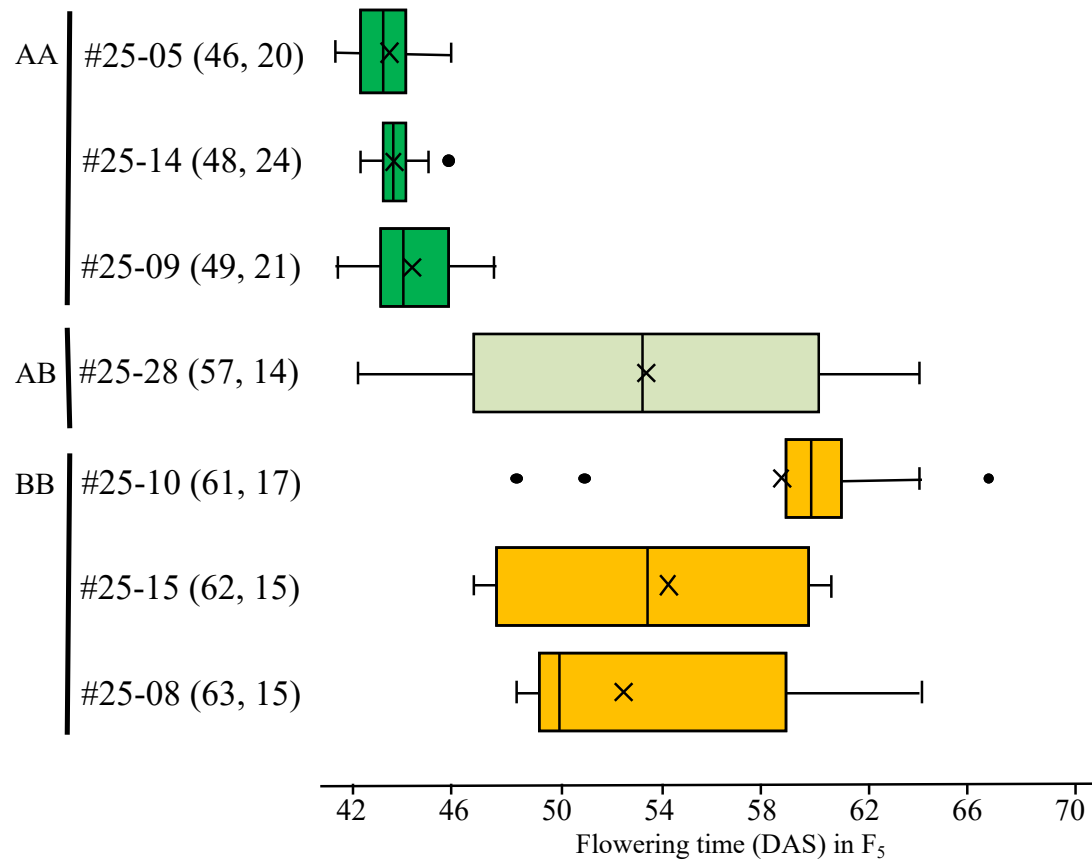


FIGURE 5. Boxplots of segregation of flowering time in the progeny of 7 F_4 plants from family #25 of the cross between ‘Nourin No. 2’ early-flowering mutant 1 (N2-ef1) and ‘Tokachi Nagaha’. The 7 plants were selected on the basis of the marker genotypes at Sat_405. Alleles: A, from N2-ef1; B, from TN. Values in parentheses represent the flowering time of the F_4 parent (left number) and the number of plants tested in the progeny (right number). The parent–offspring correlation coefficient was 0.912 ($P < 0.001$). DAS, number of days after sowing. ×, mean value.

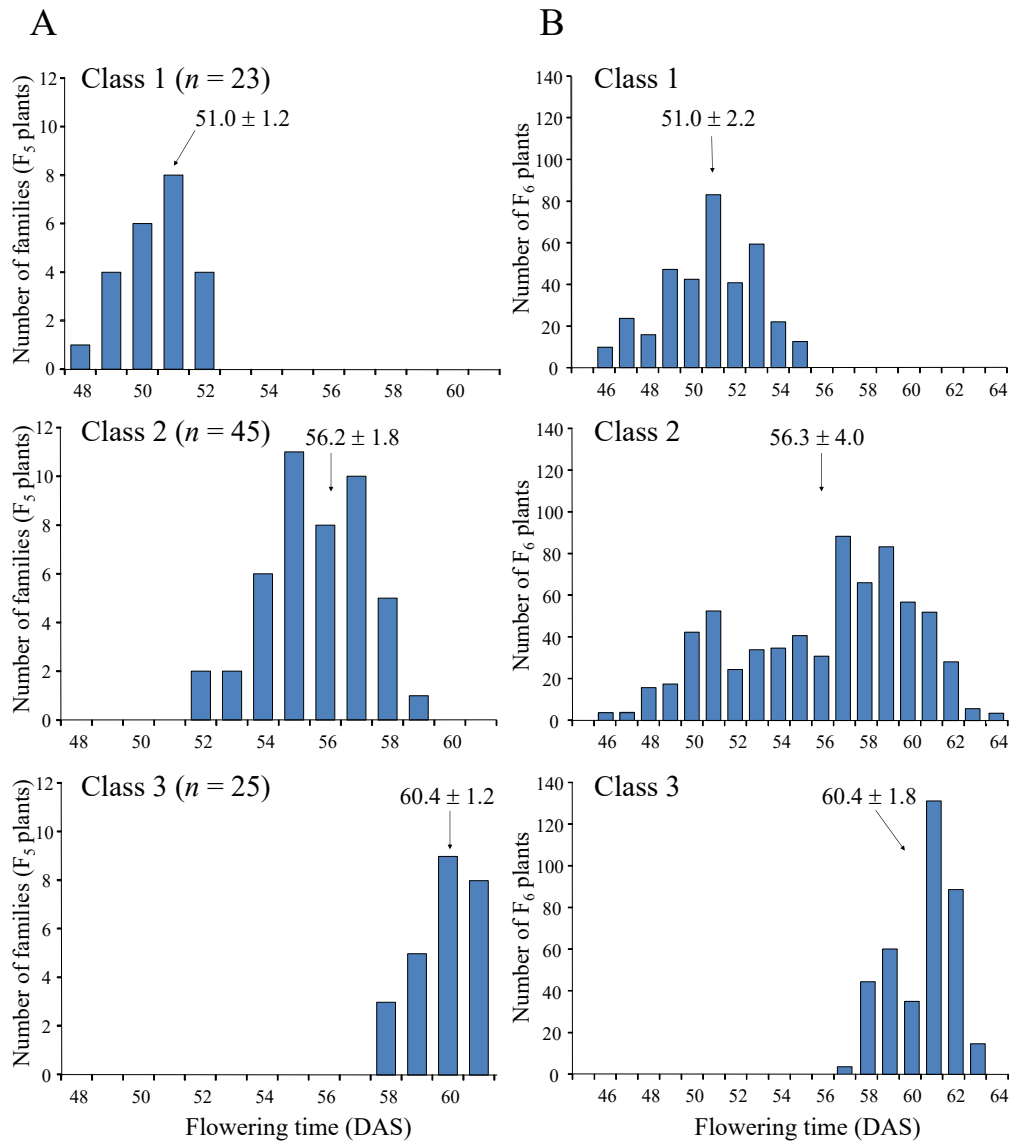


FIGURE 6. Segregation of flowering time in the progeny of F_5 plants of a cross between Nourin No. 2 early flowering mutant 1 (N2-ef1) and ‘Tokachi Nagaha’ (TN). 93 F_5 plants were classified into three classes based on the results of progeny test: Class 1, plants that produced only early-flowering progeny ($n = 23$); Class 2, plants whose progeny showed segregation for flowering time ($n = 45$); Class 3, plants that produced only late-flowering progeny ($n = 25$). (A) average flowering times of F_5 plants, (B) segregations of flowering times in F_6 progeny. DAS, number of days after sowing.

2.3.2 Fine-mapping of the mutant gene *efl*

To delineate the genomic position of *efl*, I analyzed genotypes for of 93 plants in the #46-14 family using additional six SSR markers (Figure 7). All of the plants had the TN allele at BARCSOY_19_0230. I detected nine plants in which recombination occurred between BARCSOY_19_0240 and Sat_405 and constructed a linkage map of 9.7 cM (Figure 7A). The *efl* genotypes estimated by progeny test agreed completely with the genotypes at three completely linked BARCSOY markers (BARCSOY_19_0240 to BARCSOY_19_0310) (Figure 7B).

Additional 500 seeds from plants heterozygous for BARCSOY_19_0240 and BARCSOY_19_0330 were used for fine mapping. Of these, seven seeds showed recombination between the two markers, and were genotyped for three SSRs (BARCSOY_19_0279, BARCSOY_19_304, and BARCSOY_19_310); a linkage map of 5.8 cM was constructed for these five SSR markers (Figure 7A). The seven plants were then grown in the greenhouse during the winter of 2018 and the *efl* genotypes were estimated from the segregation in the progeny in the summer of 2019. A comparison of the graphical genotypes constructed by the SSR markers and the *efl* genotype estimated by the progeny test revealed that the *efl* gene for early-flowering from N2-*efl* was most likely located between BARCSOY_19_0279 and BARCSOY_19_0310 (Figure 7B).

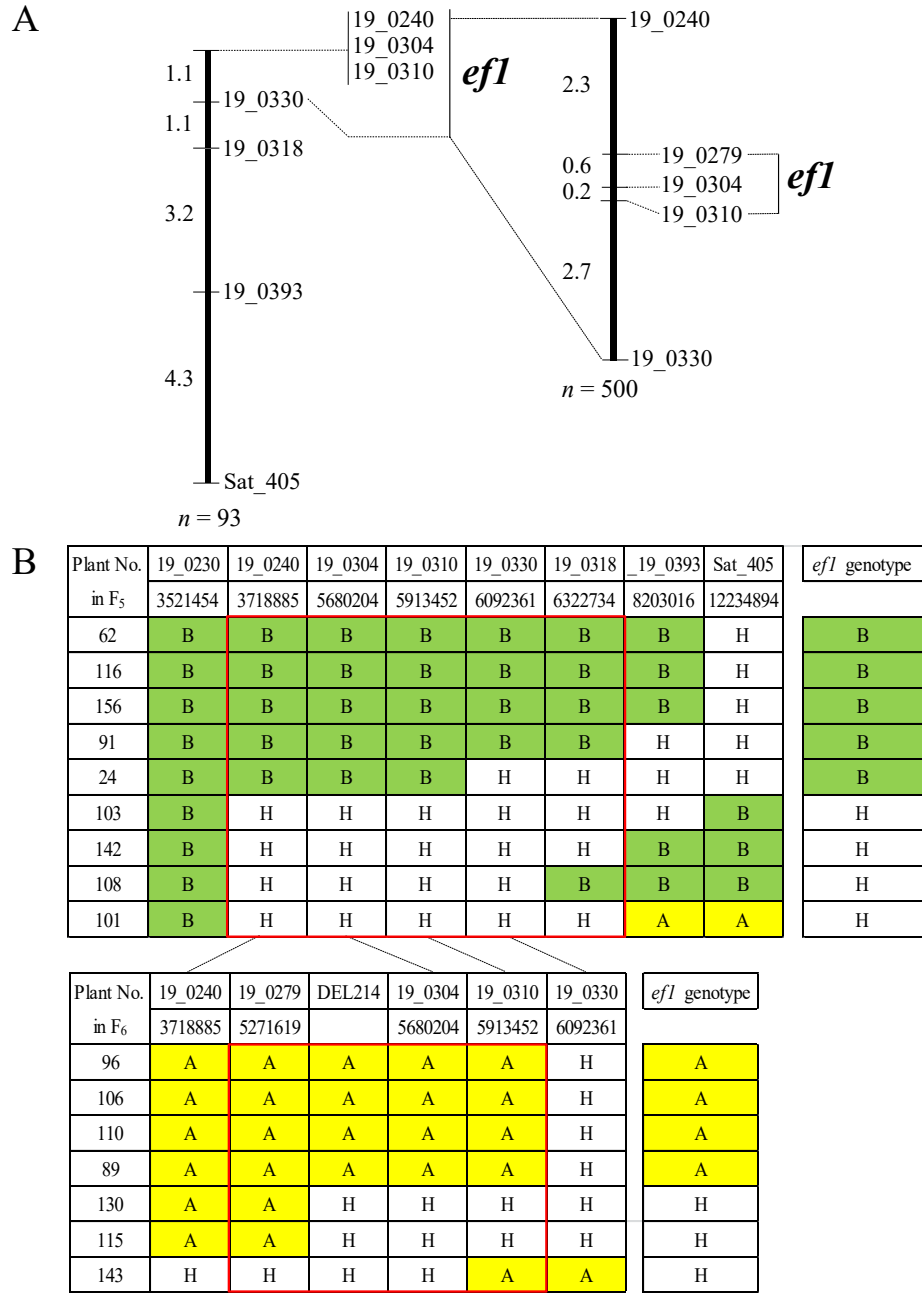


FIGURE 7. Linkage maps and graphical genotypes for the SSR markers flanking a candidate gene for early-flowering (*efl*) from ‘Nourin No. 2’ early-flowering mutant 1. (A) Linkage maps for the SSR markers and *efl*. (B) Graphical genotypes of recombinants among the SSR markers and association with the *efl* genotypes estimated by the progeny test. Numbers under the SSRs are the genomic positions in the Williams 82 reference genome. Colors represent the homozygote for the N2-*efl* allele (yellow), the homozygote for the TN allele (green), and heterozygote (white).

2.3.3 Identification of the molecular basis for *efl* based on resequencing data

The region delineated by fine mapping encompassed a 642-kbp genomic region (5.271 to 5.913 Mbp) of the Williams 82 genome sequence, which contained 32 annotated genes (Glyma.19g037700 to Glyma.19g040800; Table 2). Comparison of resequencing data between N2 and N2-*efl* for the delineated genomic region in the Integrative Genomics Viewer revealed a large deletion of ~214 kbp in N2-*efl* (Figure 8A).

The deletion was confirmed by PCR with primers designed to sandwich it (Figure 8A). As expected, the product amplified with primers F1 and R2 was detected only in N2-*efl*, not in N2 or TN, whereas the product amplified with primers F1/R1 and F2/R2, in which R1 and F2 were designed inside the deleted region, was detected in N2 and TN, but not in N2-*efl* (Figure 8B). Sequence analysis of the product amplified from N2-*efl* confirmed the deletion of 213,570 bp; there was no insertion of other DNA sequence (Figure 8C). By PCR using a set of three primers (F1, R1, and R2 in Figure 8A), I determined the genotypes for the deletion for the seven recombinant plants: all plants had identical genotypes for the deletion (DEL214 marker) and BARCSOY_19_0304 (Figure 7B). Apart from the deletion, no mutation was detected in the remaining genes. Taken together, the results may indicate that the causal DNA polymorphism for the early-flowering phenotype from N2-*efl* is most likely present in the deleted 214 kbp genomic region.

The deleted region contained 11 annotated genes (Table 2): MYB83 (Glyma.19g038100), involved in secondary cell wall biosynthesis (McCarthy et al., 2009; Zhong and Ye, 2011); NADP-dehydrogenase, which is involved in mitochondria (Glyma.19g038200); a protein in the 6-phosphogluconate dehydrogenase family (Glyma.19g038400); UDP-glucosyl transferase 85A4 (Glyma.19g038500); an ATPase E1-E2 type family protein/haloacid dehalogenase-like hydrolase family protein (Glyma.19g038600); glutamine dumper 3 (Glyma.19g038700); an ortholog of *Arabidopsis* CONSTANS (COL2b; Glyma.19g039000); NSP-interacting kinase 1 (Glyma.19g039100); and three proteins with uncharacterized functions (Glyma.19g038300, Glyma.19g038800 and Glyma.19g038900). A survey of the Flowering Interactive Database (FLOR-ID; <http://www.phytosystems.ulg.ac.be/florid/>) storing

306 *Arabidopsis* flowering genes (Bouché et al., 2016) indicated that no annotated genes homologous to *Arabidopsis* flowering genes existed in the deleted region except for Glyma.19g039000. The survey of FLOR-ID further showed that out of the 21 annotated genes in the flanking regions of deletion, Glyma.19g040300 was homologous to a MYB-RELATED PROTEIN 1 (AT5G18240) involved in photoperiod pathway of flowering. However, Glyma.19g040300 in Williams 82 was a dysfunctional gene encoding a truncated protein of 93 amino acids, despite that its homoeolog, Glyma.02g070900, and AT5G18240 had 416 and 402 amino acids, respectively. The resequencing data indicated that N2 and N2-efl possessed the same dysfunctional gene as Williams 82. Taken together, the results obtained strongly suggest that COL2b in the deleted region is involved in the different flowering times between N2-efl and TN/N2; the deletion of COL2b may condition the early-flowering of the mutant.

In addition, the micro-syntenic relationships were examined for the 11 genes in the deletion. There was no distinct micro-syteny between the deleted region and the other genomic regions. All of the genes except for the mitochondrial gene had at least one homolog in different regions of seven chromosomes (Figure 9). The lack of the aberrant phenotype in the mutant lines may be due to the presence of homoeologous copies with redundant functions or the absence of genes that influence the apparent plant phenotypes and growths in the deleted 11 genes.

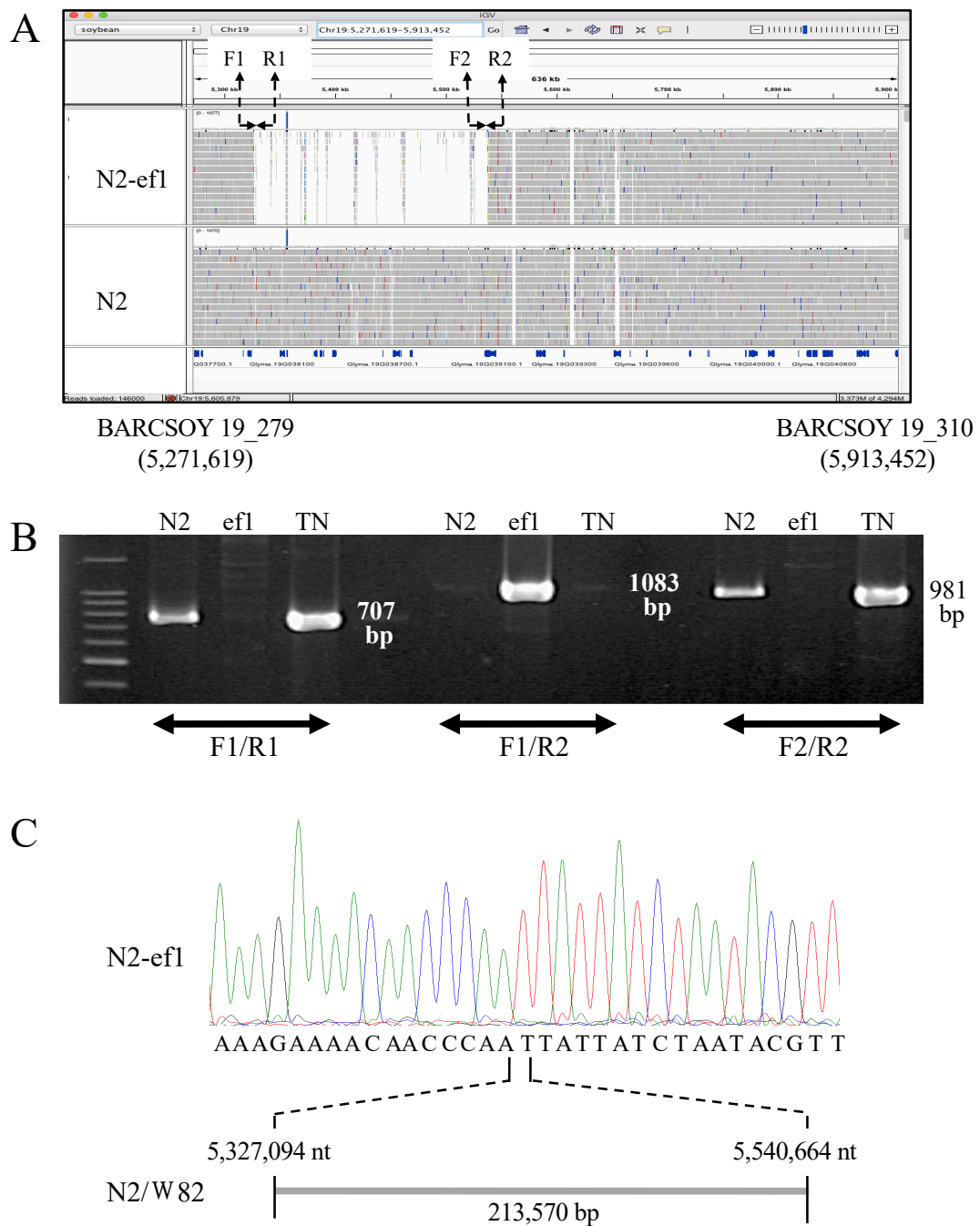


FIGURE 8. Deletion of a 214 kbp genomic region on chromosome 19 as revealed by the resequencing analysis. (A) The deletion was revealed by resequencing analysis of Nourin No. 2 (N2) and its early-flowering mutant (N2-efl). (B) PCR detection of the deletion by using primers (F1, F2, R1, and R2) designed in the boundaries of the deletion. (C) Nucleotide sequence of N2-efl for the boundary region of the deletion. W82, Williams 82 reference genome.

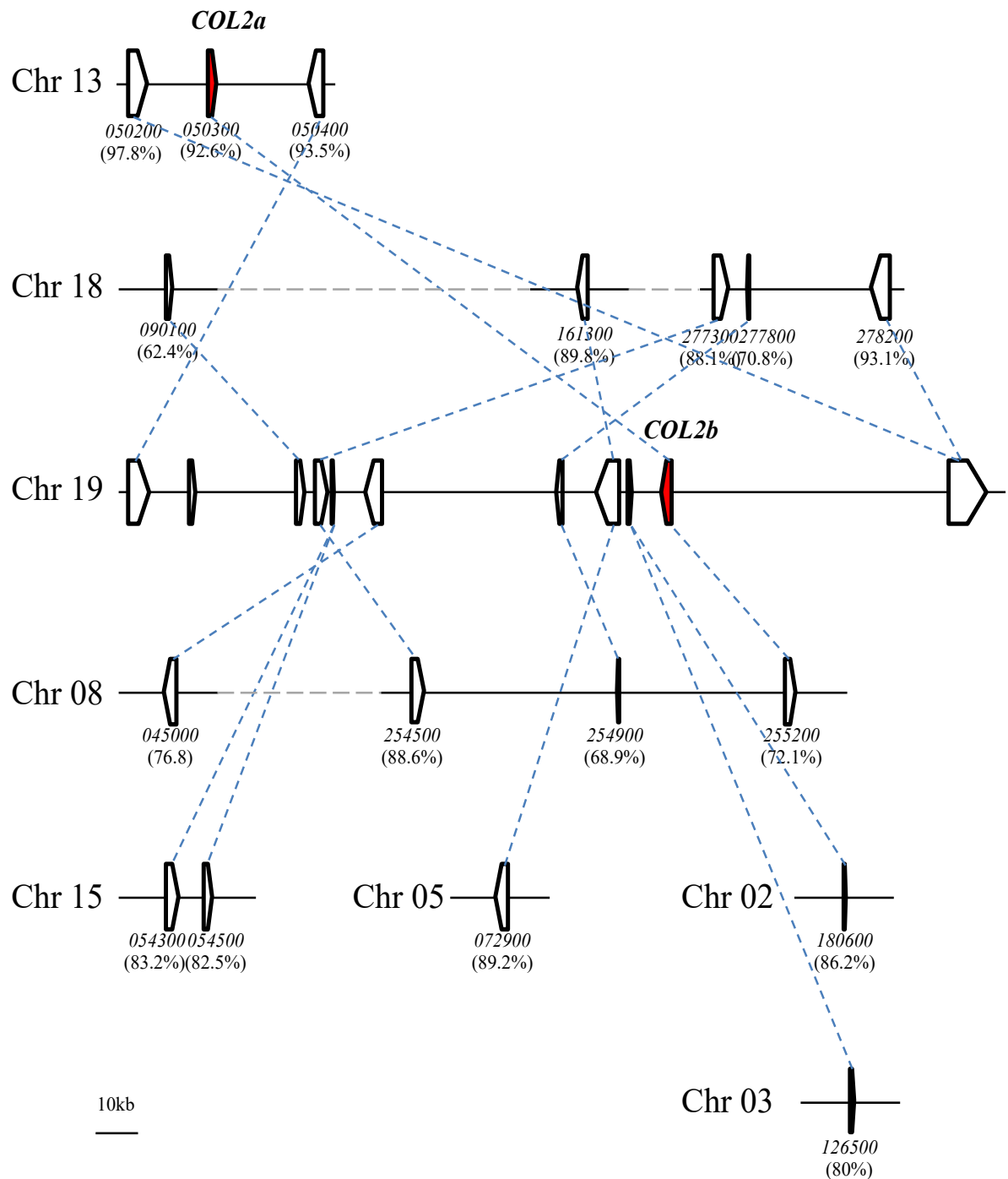


FIGURE 9. Homoeologous relationships of 11 genes in the deleted genomic region in chromosome 19 with their homologs in different chromosomes. There were no distinct homoeologous relationships between the 11 genes and their homologues with high amino acid similarities. Amino acid similarity is presented in the parenthesis.

Table 2. Annotated genes in the genomic region delineated by BARCSOY_19_0279 and BARCSOY_19_0310

Transcript Name	Location	Protein Length (AA)	Top Arabidopsis (TAIR10)	Description
BARCSOYSSR_19_0279 5271619				
Glyma.19G037700	5272862	527114	AT5G11390.1	WPP domain-interacting protein 1
Glyma.19G037800	5279113	5280851	AT3G02340.1	RING/U-box superfamily protein
Glyma.19G037900	5316778	5317846	AT3G0530.1	bZIP42-basic leucine-zipper 42
Glyma.19G038000	5321094	5324327	AT3G02350.1	galacturonosyltransferase 9
Glyma.19G038100	5350324	5355097	AT3G08500.1	myb domain protein 83
Glyma.19G038200	5356194	5357486	NA	NADH dehydrogenase subunit 2 (JN872550.1, <i>Milletia pinnata</i> mitochondrion)
Glyma.19G038300	5381177	5383149	NA	Uncharacterized protein
Glyma.19G038400	5386129	5388865	AT3G02360.1	6-phosphogluconate dehydrogenase family protein
Glyma.19G038500	5389941	5390531	AT1G78270.1	UDP-glucosyl transferase 85A4
Glyma.19G038600	5397711	5401554	AT3G022910.1	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
Glyma.19G038700	5442898	5443608	AT5G57685.1	glutamine dumper 3
Glyma.19G038800	5452151	5457668	NA	hypothetical protein (CP039352, <i>Vigna unguiculata</i> cv. Xiabao 2)
Glyma.19G038900	5459700	5460503	NA	Predicted protein (FN594963, <i>Vitis vinifera</i> cv. PN40024)
Glyma.19G039000	5467686	5469895	AT3G02380.1	CONSTANS-like 2
Glyma.19G039100	5535436	5544281	AT5G16000.1	NSP-interacting kinase 1
Glyma.19G039200	5581743	5589236	NA	hypothetical protein (CP039352, <i>Vigna unguiculata</i> cv. Xiabao 2)
Glyma.19G039300	5606232	5607665	AT1G03940.1	HXXXD-type acyl-transferase family protein
Glyma.19G039400	5652025	5657860	NA	unknown function (BT135446.1, <i>Lotus japonicus</i> clone JCVI-FLJ-3K11 unknown mRNA)
Glyma.19G039500	5666679	5667059	AT5G17450.1	Heavy metal transport/detoxification superfamily protein
BARCSOYSSR_19_0304 5680204				
Glyma.19G039600	5681475	5682707	AT1G03940.1	HXXXD-type acyl-transferase family protein
Glyma.19G039700	5720079	5721875	AT1G03940.1	HXXXD-type acyl-transferase family protein
Glyma.19G039800	5736959	5737278	AT3G55770.6	GATA type zinc finger transcription factor family protein
Glyma.19G039900	5737854	5738513	NA	hypothetical protein (CP039346.1, <i>Vigna unguiculata</i> cv. Xiabao 2)
Glyma.19G040000	5770969	5772171	AT5G38760.1	Late embryogenesis abundant protein family protein
Glyma.19G040100	5773912	5781906	AT3G02460.1	Ypt/Rab-GAP domain of gyp1p superfamily protein
Glyma.19G040200	5791073	5796786	AT3G02450.1	cell division protein ftsH, putative
Glyma.19G040300	5819045	5819405	AT3G04030.1	Homeodomain-like superfamily protein (SF4 - MYB FAMILY TRANSCRIPTION FACTOR-RELATED)
Glyma.19G040400	5826182	5829696	AT4G24400.2	CBL-interacting protein kinase 8
Glyma.19G040500	5831176	5834955	AT1G16710.2	histone acetyltransferase of the CBP family 12
Glyma.19G040600	5840532	5843788	AT1G79000.1	histone acetyltransferase of the CBP family 1
Glyma.19G040700	5843994	5848776	AT1G16710.2	histone acetyltransferase of the CBP family 12
Glyma.19G040800	5871283	5882245	AT5G38600.1	Proline-rich spliceosome-associated (PSP) family protein / zinc knuckle (CCHC-type) family protein
BARCSOYSSR_19_0310 5925831				

11 genes in the deleted region are marked by light blue.

Chapter 3 Characterization of the early-flowering habit in the mutant

3.1 Background and Purposes

The DNA polymorphism responsible for the early flowering *efl* mutant was determined as the 214 kbp deletion in the upper part of chromosome 19 in the previous chapter. Of the 11 annotated genes in the deleted region, *COL2b*, a soybean ortholog of *Arabidopsis CONSTANS*, is supposed to be the best candidate for the early flowering of the mutant lines.

In this chapter, I characterized the flowering behavior and expression profiles of flowering genes of the *efl* mutant, using two sets of near-isogenic lines (NILs).

3.2 Materials and Methods

3.2.1 Evaluation of photo-thermal responses

The NILs for the *efl* mutant and wild-type alleles were developed from two heterozygous F₅ plants (#46-14-32 and #46-14-69). The offspring of an early-flowering F₆ plant that was homozygous for the N2-*efl* allele at the tagging marker and the offspring of a late-flowering plant homozygous for the TN allele were treated as NILs for *efl* and the wild-type alleles, respectively. Two sets of NILs were grown under different combinations of photoperiods and temperatures. Three photoperiod conditions (14, 16 and 18 h) were set in a greenhouse at Hokkaido University in the winter season of 2018 to 2019. Air temperatures in the greenhouse were adjusted to 25°C, with fluctuations from a minimum of 22°C to a maximum of 28°C; lighting was supplied using high intensity discharge lamps (HONDA-T, Panasonic Co, Osaka, Japan) with an average photosynthetically active photon flux density of 120 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and a red-to-far red (R:FR) ratio of 4.5 at 1 m below the light source. Three temperature conditions (18, 25, and 32°C) were set in the growth chambers (KG50, Koito Electric Industries, LTD, Yokohama, Japan); lighting was supplied for 16 h using a combination of fluorescent and incandescent lamps with an average photosynthetically active photon

flux density of $150 \mu\text{mol m}^{-2} \text{s}^{-1}$ and an R:FR ratio of 7.0. The NILs were also grown in the greenhouse and outdoors following different sowing dates (11, 21, and 31 May) in 2019. Four pots for each line, with four plants per pot, were maintained in the greenhouse during the first 10 DAS, and then half of the pots were moved outdoors, while the other half remained in the greenhouse. Average temperatures for the 20 days after the initiation of treatment were 23.7, 22.5, and 25.9°C in the greenhouse and 18.4, 17.5, and 19.8°C in the field in the 11, 21, and 31 May sowings, respectively. Average daylengths for the first 20 days after the treatment ranged from 15 h 10 min to 15 h 20 min in the three sowing treatments (i.e., the difference was negligible). Seeds were directly sown into plastic pots (15 cm in diameter and depth), and then thinned after emergence to four plants per pot. Flowering time of eight plants (two pots) was then examined and recorded individually and expressed as DAS.

3.2.2 RNA extraction and expression analyses

Expression analyses for the two *FT* orthologs, *FT2a* and *FT5a*, and their repressors, *COL2b* and its homologs *COL2a*, and *E1* and *E1Lb*, were performed using fully expanded new leaves of the NILs grown in the growth chambers. Leaves were sampled at Zeitgeber times (ZT) of 3, 12, and 21 h in the second and fourth trifoliate leaf stages. All samples were immediately frozen in liquid nitrogen and stored at -80°C until analysis. Total RNA was isolated from each sample using RNeasy Plant RNA Mini Kit (Invitrogen, Waltham, MA, USA). cDNA was synthesized from total RNA (1 μg) using an oligo(dT)20 primer and random primer cocktail (nonadeoxyribonucleotide mix: pd (N)9, Takara Bio, Otsu, Japan) with M-MLV Reverse Transcriptase (Invitrogen) in a 20- μL volume, according to Khan et al. (2009). The transcript levels were determined by quantitative real-time PCR (qRT-PCR). In brief, each qRT-PCR mixture (20 μL) contained 5 μL of the cDNA synthesis reaction mixture diluted to 1/30 of its original volume, 5 μL of 1.2 μM primer premix, and 10 μL of TB Green Premix ExTaq II (Takara Bio). Expression levels were quantified on a CFX96 Real-Time System (Bio-Rad, Hercules, CA, USA) with PCR cycling conditions of 95°C for 3 min, followed by 40 cycles of 95°C for 10 s, 62°C for 20 s, 72°C for 20 s, and 78°C for 2 s. Values were normalized to 18S ribosomal RNA (18S rRNA). A reaction mixture without reverse transcriptase was also used as a control to confirm the absence of genomic DNA

contamination. The amplification of a single DNA fragment was confirmed by a melting curve analysis and gel electrophoresis of the PCR products. Averages and standard errors of relative expression levels were calculated for three independently synthesized cDNAs. Primers used in the expression analyses are listed in Supplemental Table S3.

3.3 Results and Discussion

3.3.1 Effects of *efl* on flowering

To characterize the effects of *efl* on flowering, I compared flowering time under different photo-thermal conditions between NILs for the *efl* and wild-type alleles.

Under a constant ambient temperature of 25°C, the *efl* lines flowered at almost the same time as the wild-type lines under a 14-h daylength (with no significant difference), but flowered, on average, 2.0 days earlier than the latter under 16 h and 2.2 to 4.3 days earlier under 18 h (Figure 10A). The different alleles also conferred different responses to changes of ambient temperature (Figure 10B). At 25°C under a 16-h daylength, flowering time was almost the same between the NILs (with no significant difference). At 32°C, flowering was promoted slightly in all lines; the *efl* lines flowered 1.9 days earlier than the wild-type lines (Figure 10B). In contrast, at 18°C, the differences in flowering time between NILs were enlarged to 11.5 days and 12.3 days in #46-14-32 and #46-14-69, respectively. The reduction of ambient temperature from 25°C to 18°C greatly delayed flowering in all lines, but the effect differed between the NILs: by 18 to 20 days in the *efl* lines and by 30 to 32 days in the wild-type lines (Figure 10B).

Flowering times of NILs sown at different dates were compared between the greenhouse and outdoor conditions. Average daylengths at 20 DAS were almost the same among the three sowing dates (ca. 15 h 10 min to 15 h 20 min). However, average temperatures during the same periods were 5 to 6°C lower outdoors than in the greenhouse. In the greenhouse, the *efl* lines flowered, on average, 0.9 to 3.5 days earlier than the wild-type lines, whereas they flowered 7.7 to 11.4 days earlier outdoor (Figure 11). The low temperatures outdoors delayed flowering by, on average, 27, 26, and 17

days at the 1st, 2nd, and 3rd sowing, respectively in the wild-type lines and by 18, 19 and 11 days in the *efl* lines. The inhibition of flowering by lower temperatures was thus reduced in plants with the *efl* allele.

Taken together, the results obtained revealed that the mutation moderated the responses of flowering to longer photoperiods and lower ambient temperatures, where the flowering were repressed in the wild-type lines. The effect of mutation was detected in both longer photoperiods and lower temperature, suggesting that the mutated gene was involved in both photoperiod and thermal responses.

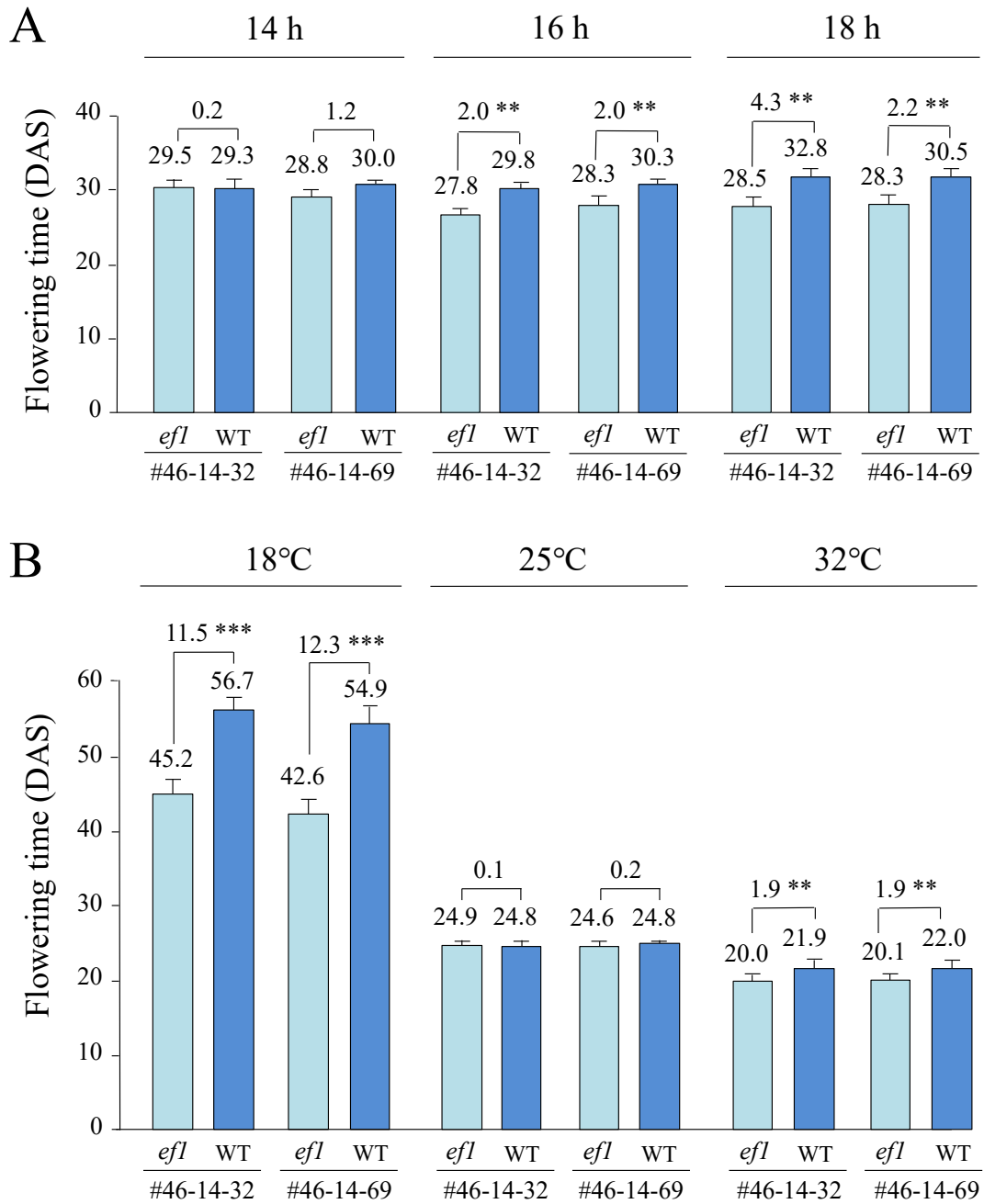


FIGURE 10. Differences in flowering time between two sets of NILs (#46-14-32 and #46-14-69) with the *efl* and wild-type (WT) alleles under different photo-thermal conditions: (A) three photoperiods under a constant ambient temperature of 25°C in the greenhouse, (B) a constant daylength of 16 h under three ambient temperatures in the growth chambers. Differences in average flowering time between NILs are presented above the bars. ** $P < 0.01$; *** $P < 0.001$. DAS, number of days after sowing.

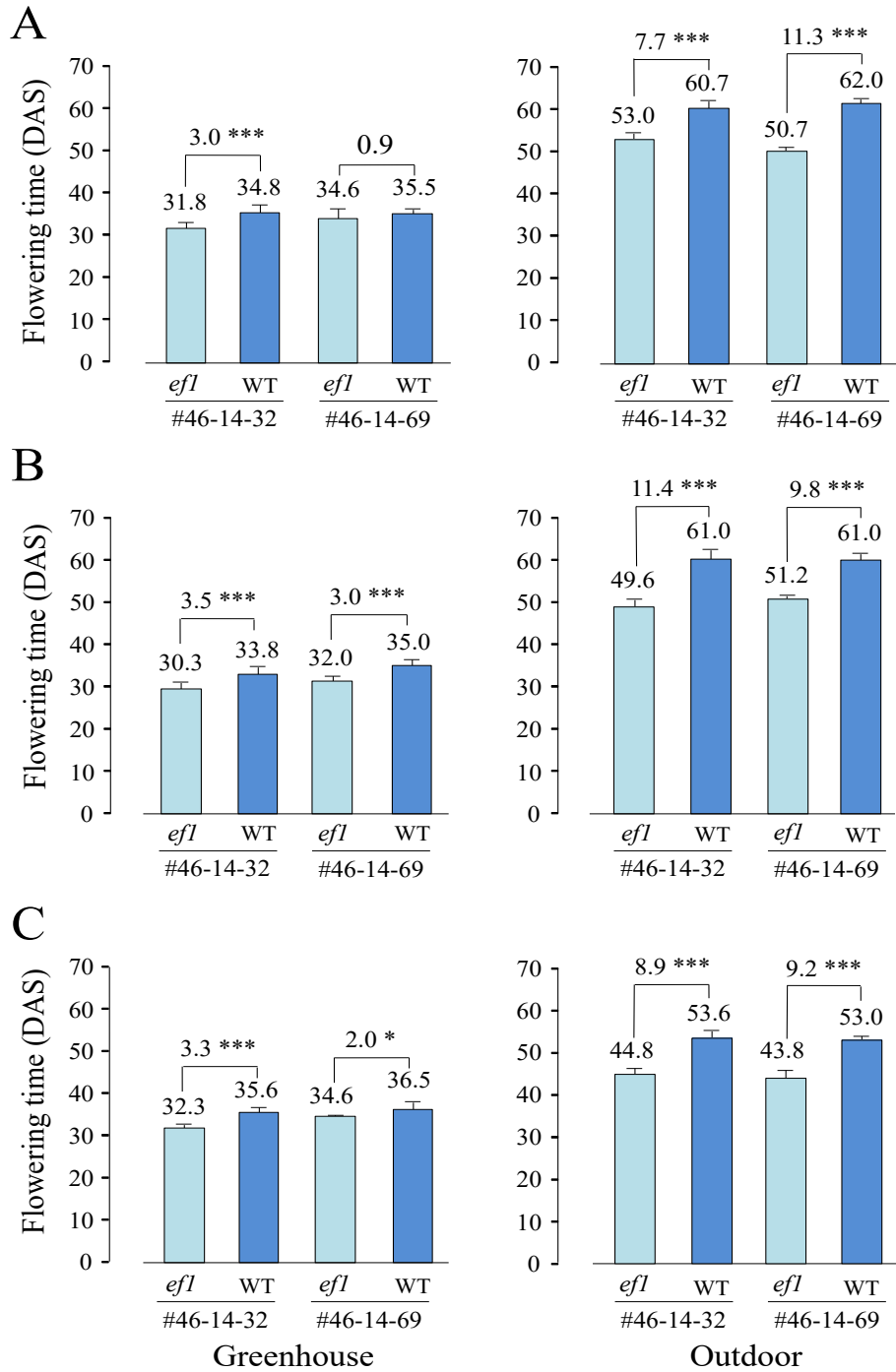


FIGURE 11. Differences in flowering time between two sets of NILs (#46-14-32 and #46-14-69) with the *efl* and wild-type (WT) alleles in the greenhouse and outdoors under natural daylengths at three different sowing times: (A) first sowing (11 May), (B) second sowing (21 May), and (C) third sowing (31 May). Differences in average flowering time between NILs are presented above the bars. * $P < 0.05$; *** $P < 0.001$.

3.3.2 Expression profiles of flowering genes under different thermal conditions

The difference in flowering time between NILs was enlarged in the 18°C condition, compared with the 18 h daylength condition. Transcript abundances of *FT2a* and *FT5a* at the second and fourth trifoliate leaf stages were thus examined in the 18 and 25°C conditions (Figures 12 and 13). At 25°C, there were no consistent differences in *FT2a* or *FT5a* transcripts between the *efl* and wild-type alleles at either growth stage. However, the expression of *FT2a* and *FT5a* was strongly decreased by low temperature at both stages and at the second trifoliate stage, respectively. The decrease caused by low temperature was greater at the second trifoliate leaf stage; the transcript levels at 18°C decreased to approximately 5% and 1%, respectively, of the values at 25°C in #46-14-69 and #46-14-32. Despite the strong repression, the transcript levels at 18°C were significantly higher in the *efl* lines than the wild-type lines at any time points. At the fourth trifoliate leaf stage, the decrease was smaller, particularly in the *efl* lines; for NILs with the *efl* allele, the *FT2a* expression decreased to 10% of the level at 25°C in the #46-14-32 NILs and to 50% of the level at 25°C in the #46-14-69 NILs, and the *FT5a* expression decreased to only 50% of the level at 25°C in the #46-14-32 NILs, versus no decrease in the #46-14-69 NILs. As in the second trifoliate leaf stage, the transcript levels in the fourth trifoliate leaf stage were higher in the *efl* lines than the wild-type lines at 18°C, except for *FT5a* in the 46-14-69 NILs.

Transcript levels of *COL2b/COL2a* and *E1/E1Lb* were examined (Figures 14 and 15). At the second trifoliate leaf stage, the transcript levels for these four genes mostly decreased in response to decreasing temperature, although the reductions were not so marked compared with the *FT2a* and *FT5a* expressions. Furthermore, there was no constant association of transcript abundances between *COL2b/COL2a* or *E1/E1Lb* and the two *FT* genes (Figures 12 ,13 ,14 and 15). In contrast, at the fourth trifoliate leaf stage, low temperature upregulated the *COL2b* expression in the wild-type line of #46-14-69 NIL and shifted the expression peak from 3 ZT at 25°C to 12 ZT at 18°C in that the wild-type line of #46-14-32 NIL. The *COL2a* expression, however, showed no distinct responses to low temperature in both of the *efl* and wild-type lines of two NIL sets. Intriguingly, the *E1* and *E1Lb* expressions at 3 and 12 ZT were upregulated by low temperature in the wild-type line in both sets of NILs; the upregulations were significant

in the #46-14-69 NIL; the effect of cool temperature was thus remarkable in the #46-14-69 NIL.

In summary, the cool temperatures strongly repressed both *FT2a* and *FT5a* expressions at the early growing stage. The repression by low temperatures was released slightly at the later growing stage, but its extent was different between the two *FT* orthologs; the *FT2a* expression still remained repressed. The repression of *FT2a* and *FT5a* expressions at the later growing stage in the wild-type lines by cool temperatures were associated with the upregulation or shift of peak from 3 ZT to 12 ZT in the *COL2b* expression and the upregulation of *E1* and *E1Lb* expressions. The *efl* mutation moderated the repression by cool temperatures to promote flowering in the cool ambient temperatures.

2nd trifoliate leaf stage

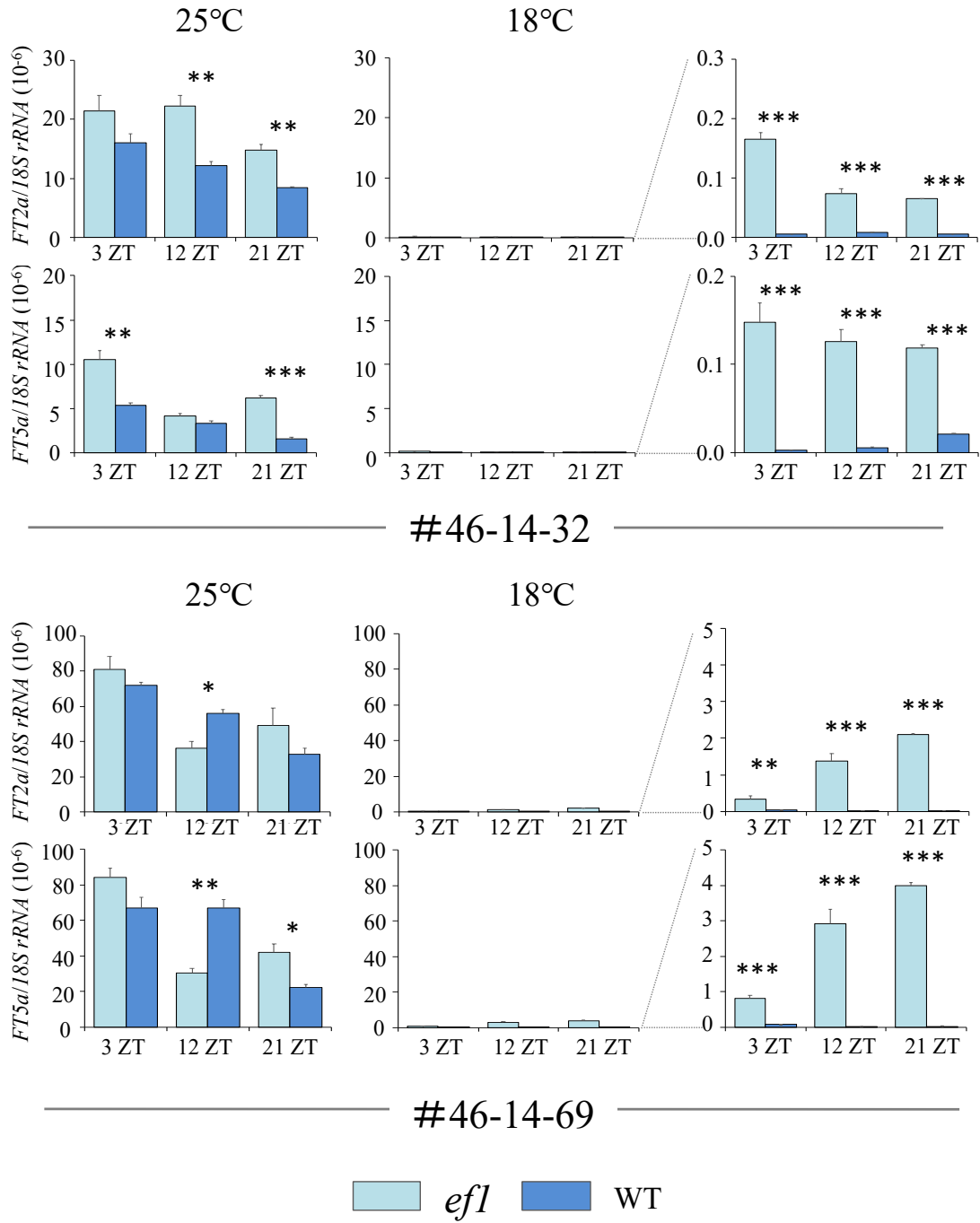


FIGURE 12. Transcript levels of *FT2a* and *FT5a* at the second trifoliate leaf stages in two sets of NILs (#46-14-32 and #46-14-69) with the early-flowering (*efl*) and wild-type (WT) alleles at 18 and 25°C. The enlarged view for the expression levels at 18°C is presented in the right column. ZT values are Zeitgeber times. Error bars are technical repeats ($n = 3$). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

4th trifoliate leaf stage

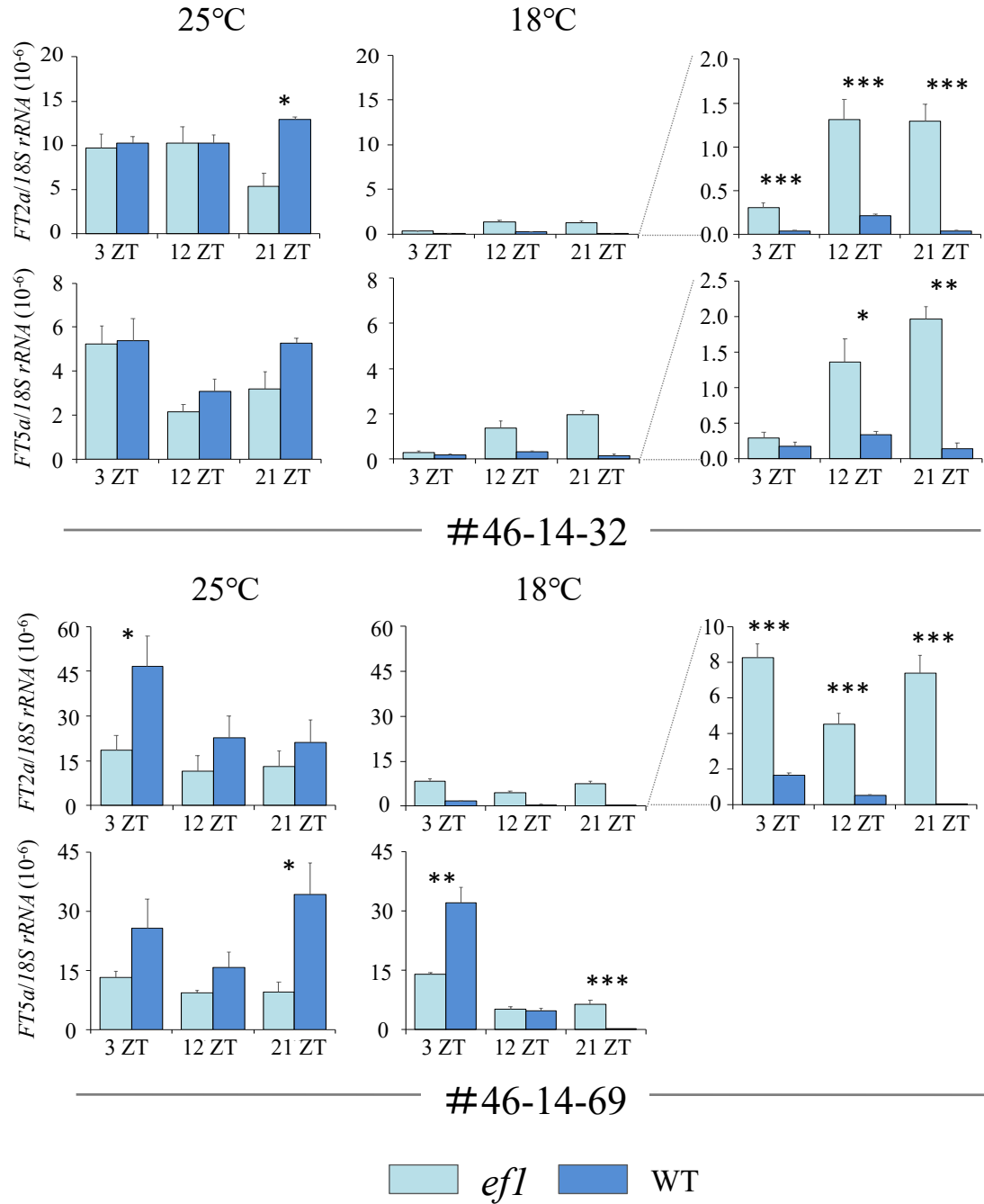
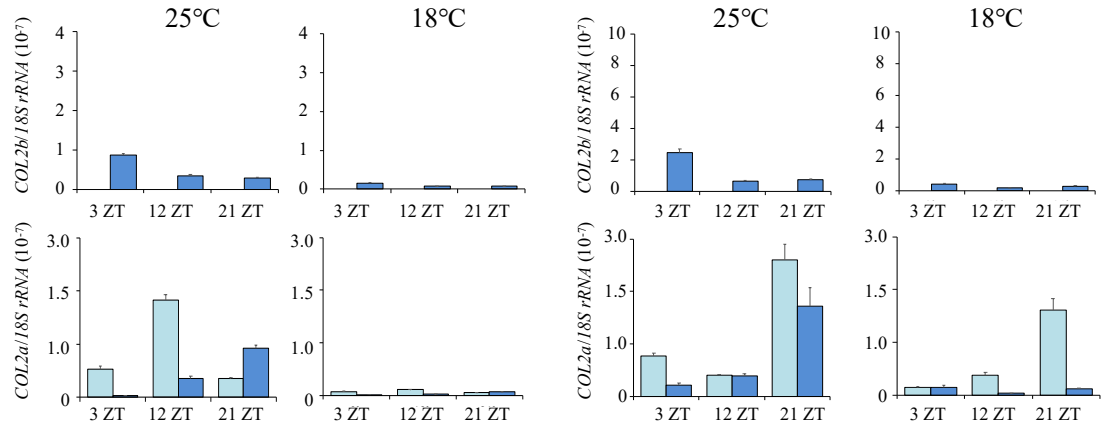


FIGURE 13. Transcript levels of *FT2a* and *FT5a* at the fourth trifoliate leaf stages in two sets of NILs (#46-14-32 and #46-14-69) with the early-flowering (*efl*) and wild-type (WT) alleles at 18 and 25°C. The enlarged view for the expression levels at 18°C is presented in the right column. ZT values are Zeitgeber times. Error bars are technical repeats ($n = 3$). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

2nd trifoliate leaf stage



4th trifoliate leaf stage

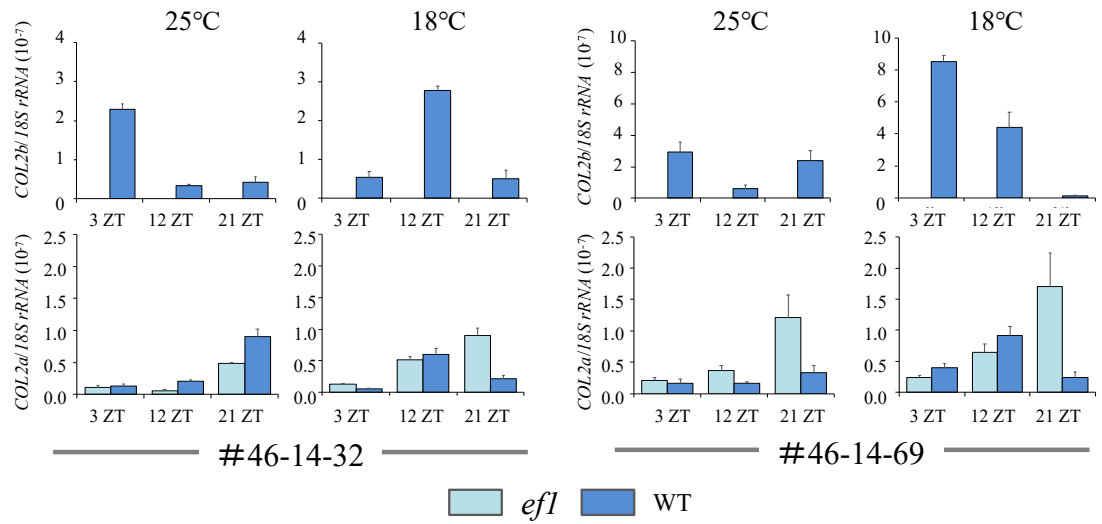
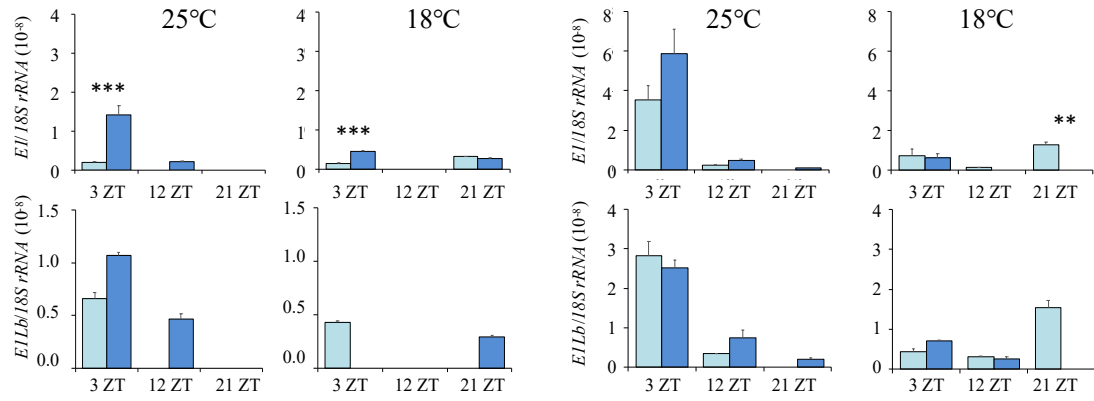


FIGURE 14. Transcript levels of *COL2b* and *COL2a* at the second and fourth trifoliate leaf stages in two sets of NILs (#46-14-32 and #46-14-69) with the early-flowering (*efl*) and wild-type (WT) alleles at 18 and 25°C. ZT values are Zeitgeber times. Error bars are technical repeats ($n = 3$). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

2nd trifoliolate leaf stage



4th trifoliolate leaf stage

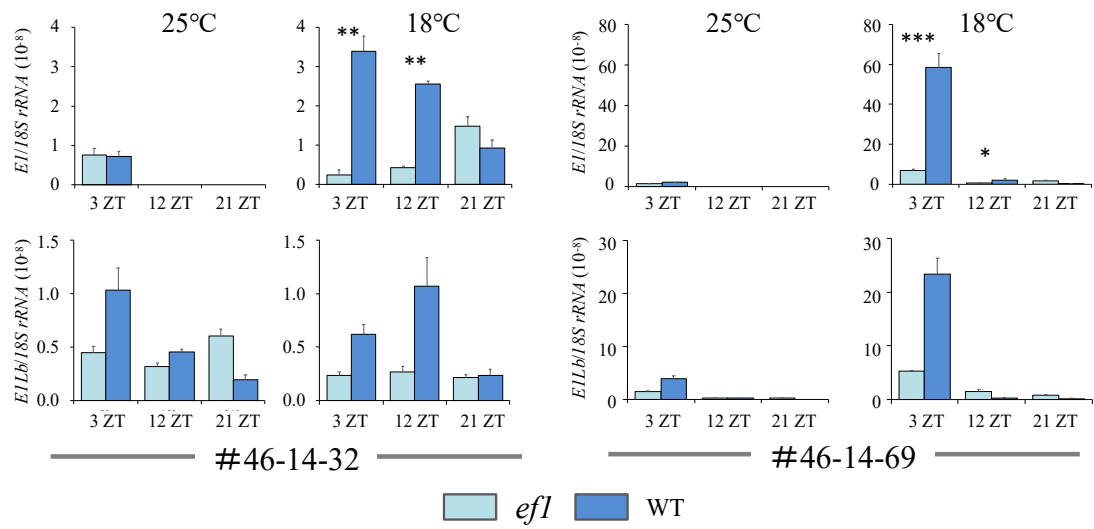


FIGURE 15. Transcript levels of *E1* and *E1Lb* at the second and fourth trifoliolate leaf stages in two sets of NILs (#46-14-32 and #46-14-69) with the early-flowering (*efl*) and wild-type (WT) alleles at 18 and 25°C. ZT values are Zeitgeber times. Error bars are technical repeats ($n = 3$). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Chapter 4 Molecular diversity of *COL* genes

4.1 Background and Purposes

In the previous chapters, I determined the molecular basis for the mutation and characterized the function in the thermal regulation of flowering to chilling temperatures. The results obtained strongly suggest a possible involvement of *COL2b* in the thermal regulation of flowering, particularly the regulation in response to cool weather. No QTLs have been reported in the genomic region harboring *COL2b* (SoyBase, <http://soybase.org>). It is thus interesting to determine whether there is any variation in the coding region of *COL2b* that can be related to the natural variation in flowering time.

As mentioned in Chapter 1, the soybean genome possesses a total of 26 *COL* genes including *COL2b* (Wu et al., 2014). Phylogenic analysis classified these *COL* genes into three clades conserved among flowering plants. Two homoeologous pairs in Clade I, *COL1a/COL1b* and *COL2a/COL2b*, showed the highest sequence similarity to *Arabidopsis CO* (Wu et al., 2014). They control flowering times in soybean and *Arabidopsis*. All of the four copies complement the loss-of-function of *CO* in *Arabidopsis* when overexpressed (Wu et al., 2014). Furthermore, RNA interference and overexpression of these genes promote and delay flowering time in soybean, respectively (Cao et al., 2015; Wu et al., 2019). Therefore, the four *COL* genes function as floral repressors in soybean.

In this chapter, I surveyed the sequence variations of *COL2b* and the other three *COL* genes in 512 cultivated and wild soybean accessions to evaluate their molecular diversities and to determine whether these genes can be the responsible genes for the natural variation of flowering time in soybean.

4.2 Materials and Methods

4.2.1 Resequencing data and SNP calling

In this study, I surveyed the resequencing data deposited in SoyBase (https://soybase.org/data/public/Glycine_max/), which covers three data sets, 17 wild and 14 cultivated soybean genomes from Lam et al. (2010), 106 cultivated soybean genomes from Valliyodan et al. (2016), and 45 wild and 330 cultivated soybean genomes from the United Soybean Board (USB)-funded project. As a whole, the number of accessions tested was 450 in the cultivated soybean and 62 in the wild soybean (Supplemental Table S4).

The single nucleotide polymorphisms (SNP) data for *COL2b* and the three homologs, *COL1a*, *COL1b* and *COL2a*, were filtered by VCFtools (version 0.1.13) (<https://vcftools.github.io/index.html>; Danecek et al., 2011).

4.3 Results and Discussion

Based on the sequences of a total of 512 accessions, the polymorphisms in the *COL* genic region from the translation start site to the end of 3' UTR were determined. In *COL2b*, nine SNPs were detected in the first exon and six SNPs in the intron. In *COL2a*, three SNPs were detected in each of the first and second exons; the other sixteen SNP sites were present in the intron. In *COL1a*, six SNPs were detected in the first exon and thirty-one SNPs in the intron. In *COL1b*, twelve SNPs were detected in the first exon and twenty-three SNPs in the intron (Supplemental Table S6). In this chapter, I will focus on non-synonymous SNPs that led to molecular diversity of *COL* genes (Supplemental Table S5).

In *COL2b*, five of the nine SNPs in the first exon resulted in the amino acid (AA) substitutions (Figure 16). There were no AA substitutions in functionally important domains, B-BOX1, B-BOX2 and CCT domains (Figure 16B). At site 1 (the 13th AA residue), ten wild accessions (SJ17, SJ18, SJ25, SJ27, SJ28, SJ29, SJ31, SJ35, SJ39 and SJ40) and one cultivated accession (USB014) had the AA substitution from

Asparagine (N) of Williams 82 to Serine (S). At site 2 (the 123rd AA residue), five wild accessions (SJ7, SJ10, SJ16, SJ37 and SJ45) and one cultivated accession (USB128) had the AA substitution from Lysine (K) of Williams 82 to Glutamic acid (E). At site 3 (the 146th AA residue), one wild accession (W16) had the AA substitution from Aspartic acid (D) of Williams 82 to Glutamic acid (E). At site 4 (the 166th AA residue), two wild accessions (SJ5 and SJ39) had the AA substitution from Glutamic acid (E) of Williams 82 to Lysine (K). At site 5 (the 167th AA residue), only one wild accession (SJ8) showed the AA substitution from Glutamic acid (E) of Williams 82 to Lysine (K) (Figure 16A). Based on the combinations of AA substitutions, the accessions tested were classified into seven haplotypes; haplotype 1 (Hap 1) for the *COL2b* sequence of Williams 82 reference sequence (*Glycine max Wm82.a2.v1*), haplotype 2 (Hap 2) to haplotype 6 (Hap 6) for the sequences with the non-synonymous SNPs at sites 1 to 5, respectively (Figure 16A). The sequence with SNPs at sites 1 and 4 was designated as haplotype 7 (Hap 7); it was different from Hap 1 in the two AA substitutions. The most predominant haplotype was Hap 1 (95.9%). All haplotypes were observed in the wild soybean. Only two of the 450 cultivated accessions tested possessed variant haplotypes, Hap 2 and Hap 3.

In *COL2a*, three of the six SNPs, two in the first exon and one in the second exon, resulted in the AA substitutions (Figure 17). At site 1 (the 114th AA residue), one wild accession (SJ19) had the AA substitution from Glycine (G) of Williams 82 to Arginine (R). At site 2 (the 167th AA residue), thirteen wild accessions (W1, W2, W3, W4, W5, W7, W9, W11, W12, W13, W14, W15 and W17) and nine cultivated accessions (C1, C2, C12, C17, C24, C30, C33, C34 and C35) had the AA substitution from Lysine (K) of Williams 82 to Glutamic acid (E). At site 3 (the 338th AA residue), one wild accession (SJ3) had the AA substitution from Glutamic acid (E) of Williams 82 to Lysine (K). The accessions were classified into four haplotypes; haplotype 1 (Hap1) for the *COL2a* sequence of Williams 82 reference sequence (*Glycine max Wm82.a2.v1*), haplotypes 2 (Hap 2) to haplotype 4 (Hap 4) for the sequences with the non-synonymous SNPs at sites 1 to 3, respectively (Figure 17A). The most predominant haplotype was Hap 1 (95.3%), and Hap 3 was detected in 22 (13 wild and 9 cultivated) of 512 accessions tested (4.3%).

In *COL1a*, three of the six SNPs in the first exon resulted in the AA substitutions (Figure 18). At site 1 (the 116th AA residue), three wild accessions (SJ27, SJ28 and W05) and 23 cultivated accessions had the AA substitution from Valine (V) of Williams 82 to Isoleucine (I). At site 2 (the 143rd AA residue), five wild accessions (SJ13, SJ15, SJ 22, SJ35 and SJ39) and one cultivated accession (USB182) had the AA substitution from Asparagine (N) of Williams 82 to Serine (S). At site 3 (the 216th AA residue), 25 wild accessions and 252 cultivated accessions (listed in Supplemental Table S4) had the AA substitution from Glutamic acid (E) of Williams 82 to Aspartic acid (D). The accessions were classified into five haplotypes; haplotype 1 (Hap1) for the *COL1a* sequence of Williams 82 reference sequence (*Glycine max Wm82.a2.v1*), haplotypes 2 (Hap 2) to haplotype 4 (Hap 4) for the sequences with the non-synonymous SNPs at sites 1 to 3, respectively. The sequence with SNPs at site 2 and 3 was designated as haplotype 5 (Hap 5); it was different from Hap 1 in the two AA substitutions (Figure 18A). The most predominant haplotype was Hap 4; it was detected in 274 of 512 accessions tested (53.5%). The next was Hap 1 (40.2%).

In *COL1b*, four of the twelve SNPs in the first exon resulted in the AA substitutions (Figure 19). At site 1 (the 6th AA residue), one cultivated (USB-035) accession had an AA substitution from Alanine (A) to Proline (P). At site 2 (the 17th AA residue), one wild accession (SJ8) had an AA substitution from Threonine (T) of Williams 82 to Methionine (M). At site 3 (the 24th AA residue), one wild accession (W13) had an AA substitution from Serine (S) of Williams 82 to Threonine (T). At site 4 (the 174th AA residue), one wild accession (SJ22) and one cultivated accession (USB-035) had the AA substitution from Threonine (T) of Williams 82 to Serine (S) (Figure 19A). The AA substitutions at 17th and 24th AA residues were located in the BBOX1 domain (Figure 19B). The accessions were classified into five haplotypes; haplotype 1 (Hap1) for the *COL1b* sequence of Williams 82 reference sequence (*Glycine max Wm82.a2.v1*), haplotypes 2 (Hap 2) to haplotype 4 (Hap 4) for the sequences with the non-synonymous SNPs at sites 2 to 4, respectively. The sequence with SNPs at sites 1 and 4 was designated as haplotype 5 (Hap 5); it was different from Hap 1 in the two AA substitutions (Figure 19A). The predominant haplotype was Hap 1 (99.2%).

To sum up, most of all accessions shared the same haplotype, Hap 1, with Williams 82 for all *COL* genes except *COL1a*; the frequencies of Hap 1 were over 95.0%. In *COL1a*, there were two major haplotypes, Hap 1 and Hap 4. However, all of the AA substitutions except for two AA substitutions in *COL1b* were located outside of functional domains of CO proteins. Accordingly, a further study is needed to determine whether the two haplotypes, Hap 1 and Hap 4, of *COL1a* have different functions in the control of flowering or not. In *COL1b*, the AA substitutions at 17th and 24th AA residues were located in the BBOX1 domain, suggesting some functional changes (Figure 19B). However, the AA substitutions were detected in only a few wild accessions (0.4% of all 512 accessions). They may not have had any large evolutionary impacts on the control of flowering times in soybean.

Cao et al. (2015) detected two artificially induced mutants for each of *COL1a* and *COL1b* from a mutant population of a Japanese cultivar Fukuyutaka; there were no significant differences in flowering time between the two *colla* mutants and the wild type, while two *collb* mutants flowered significantly earlier than the wild type. *COL1a* and *COL1b* are homoeologous genes sharing 91.4% amino acid identity, which exhibited similar expression patterns to each other in previous works (Fan et al. 2014, Wu et al. 2014). The overexpression of *COL1a* in transgenic lines is confirmed to delay flowering in LD conditions, suggesting that *COL1a* acts as a floral repressor as *COL1b* does (Cao et al., 2015). Accordingly, it may be an unexpected result that the *COL1a* mutants did not show any early flowering phenotypes unlike *COL1b*. Of the two mutants, the *collb*-T314I had the amino acid substitution from threonine to isoleucine at the 314th amino acid residue in the CCT domain, which was conserved in the CO homologs; another mutant, *collb*-P245L, had the amino acid substitution from proline to leucine at the 245th amino acid residue in a commonly conserved region in CO genes (Cao et al., 2015). However, the AA substitutions of two *colla* mutants were located in neither any functionally important domains nor commonly conserved regions (Cao et al., 2015). The AA substitutions detected in *COL1a* might not provide any large changes on the function, because they were located outside the functionally important and/or commonly conserved domains.

In conclusion, the survey of molecular diversity revealed that the extent of the diversity of the peptide sequences was extremely low in *COL2b*; the variant haplotypes were mostly found in the wild accessions. In SoyBase (<http://www.soybase.org>), where the QTLs for flowering time in the literatures published until 2014 are compiled, only one QTL detected in an F₂ population of a cross between the wild accession JP110755 and cultivar ‘Fukuyutaka’ is located on a genomic region of LG F (Chr13) harboring *COL2a*. The absence of QTL near *COL2b* may be consistent with a low level of the naturally occurring genetic variability of the coding sequence as revealed in this chapter. The early-flowering mutant *ef1* may therefore be useful in the molecular genetic study on the thermal regulation of flowering and the breeding for improving the adaptability of cultivars in soybean.

(A) Non-synonymous substitutions and haplotype classification



Position	Number of accessions		Amino acid substitution				
	wild	cultivated	Site1 5469858	Site2 5469529	Site3 5469458	Site4 5469400	Site5 5469397
Haplotype 1 (W82)	43	448	N	K	D	E	E
Haplotype 2	9	1	S	K	D	E	E
Haplotype 3	5	1	N	E	D	E	E
Haplotype 4	1	0	N	K	E	E	E
Haplotype 5	1	0	N	K	D	K	E
Haplotype 6	1	0	N	K	D	E	K
Haplotype 7	1	0	S	K	D	K	E

Glyma.19G039000.1 | Chr19:5467686..5469895 reverse
PTHR31319:SF10 - ZINC FINGER PROTEIN CONSTANS-RELATED

(B) *COL2b* peptide sequence

MLKEGTNNVGG**S****N**TGTTWSRVCDTCLSAPCVLYCHADSAYLCSSCDARV
 HAANRVASRHKRVWVCEACERAPAAFLCKADAASLCSSCDADIHSANPLA
 SRHNRVPILPISGSLFREPEHNH**K**RVEHAFVNEVEEEEEEGVFDEYE**D**VEVA
 ASWLLPHPMKNND**E****E**ENDCGDEGFLFVDEYLDNLVDCCNSCGHNDNQF
 SNVYQHQQNYNTVPQNYVVVPVQVPQHFQPLDFDSSKAGFSYDGSLSQS
 VSVSSMDVGVVPESTVSGISMSHKSPIGTNDLFPPLLMPSHLTPMDREARV
 LRYREKKKTRKFEKKIRYASRKAYAETRPRIKGRFAKRTDVEAEVDQMFS
 T
 KLFNEVGGSIFPTF*

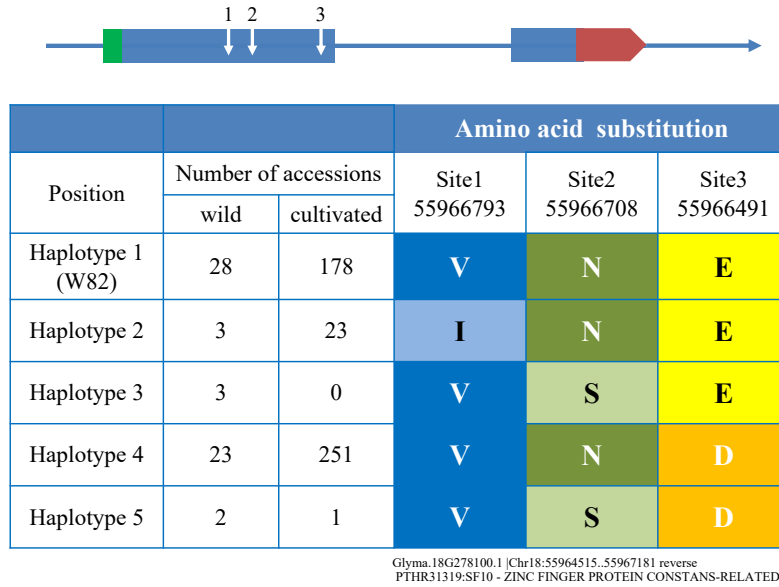
FIGURE 16. Molecular diversity in *COL2b* in 512 soybean accessions. (A) Positions of non-synonymous nucleotide substitutions and haplotype classification. (B) Positions of amino acid substitutions at the *COL2b* peptide sequence (shown in bold and underline). Colored lines represent BBOX1 domain (red), BBOX2 domain (black), and CCT domain (yellow).

Glyma.13G0503001|Chr13:14630394..14632571 forward
PTHR31319-SE10 - ZINC FINGER PROTEIN CONSTANS-RELATED

MLKEGTNNVGGSTGTWSHVCDTCRSAPCVLYCHADSAYLCSSCDARVHA
ANRVASRHERVVWCEACERAPAAFLCKADAASLCSSCDADIHSANPLASR
HHRVPILPISGSLFGEPEHERVYAFVNEVEAESEEEVFDEYDEVEAASWL
LPHPMKNDKIDENGGDKGFLFGDEYFDNLVDNCSCGHNNNQFSNVYDQH
QQNYSNTVPQNYAVVPVQPQHFQPGLDFFDSSKAGFSYDGSLSQSVSVSS
MDVGVLLESTISDISMSHKSPIGTTDLFPPLPMPSHLTTPMDREARVLRyre
KKKTRKFEKKIRYASRKAYAETRPRIKGRFAKRTDVEAEVDQMFTTLFT
EVGGSIPTF*

54

(A) Non-synonymous substitutions and haplotype classification

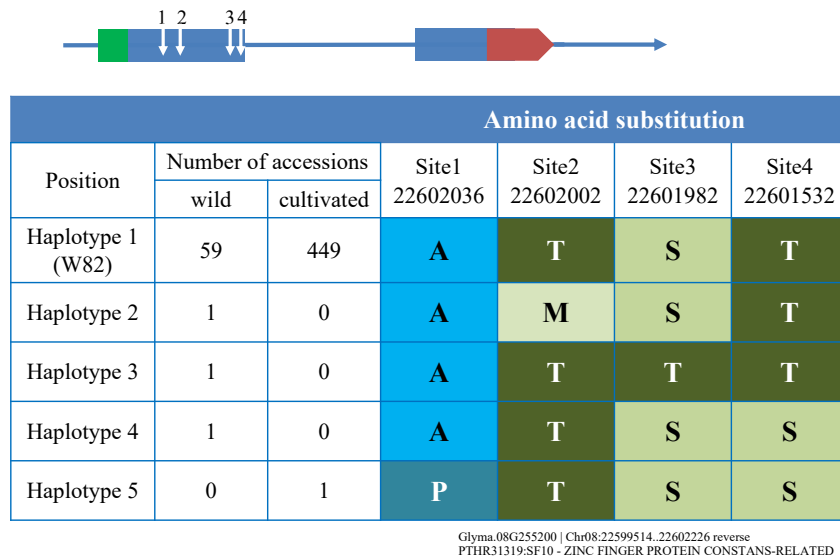


(B) *COL1a* peptide sequence

MLEGQATTPTWPRMCDTCRSVPSTVFCRSHTAFLCATCDTRLHVSLTWH
 ERVWVCEACERAPAAFLCKADAASLCASCDADIHAANPLASRHHRVPIL
 PIAAANNNNNDDDDVAD**V**DDEDETASWLLLNPIKSATVPNTNNNNNNNN
 GFLYNGEVDEYLDLVDNCNSCGDNNHFASAAATTDHYAQHQHFAGVSQ
 KSYAGDSVVPVQQHQHFQLGL**E**FDNSKPAFSYNGSVSVSVSSMDIGV
 VPESPMRDVSIAHTRPPKGTIDLFSGPPIQVPSHFSPMDREARVLRVREKK
 KMRKFEKTIRYASRKAYAETRPRIKGRFAKRTDVEAEVDQMFSSTLITEV
 GYGIVPSF*

FIGURE 18. Molecular diversity in *COL1a* in 512 soybean accessions. (A) Positions of non-synonymous nucleotide substitutions and haplotype classification. (B) Positions of amino acid substitutions at the *COL1a* peptide sequence (shown in bold and underline). Colored lines represent BBOX1 domain (red), BBOX2 domain (black), and CCT domain (yellow).

(A) Non-synonymous substitutions and haplotype classification



(B) *COL1b* peptide sequence

MLDGEATMGTWARMCDTCRSAPSSVFCRAHTAFLCATCDARLHASLTW
 HERVWVCEACERAPAAFLCKADAASLCASCDADIHAANPLASRHHRPILP
 IAAAPGNNDNDNVDDADLDDDDDETASWLLNPNVKSASVPNNNTNNGFS
 YNGEVDEYLDLVDDCDNHHFASVATTTDHYSHQHGHFGVVSHKSYAGDS
 VVPVQHHQHFLGLEFDNSKAAFSYNASVNQSVSVSSMDIGVVPESPMRD
 VSIGHTRTPKGTIDLFSGPPIQVPSHFSPMDREARVRLRYREKKKTRKFEKTIR
 YASRKAYAETRPRIKGRFAKRTDVEAEVDQMFSTTLITEVGYGIVPSF*

FIGURE 19. Molecular diversity in COL1b in 512 soybean accessions. (A) Positions of non-synonymous nucleotide substitutions and haplotype classification. (B) Positions of amino acid substitutions at the COL1b peptide sequence (shown in bold and underline). Colored lines represent BBOX1 domain (red), BBOX2 domain (black), and CCT domain (yellow).

Chapter 5 General discussion

The early flowering of *efl* can be attributed to a deletion in chromosome 19

Flowering is an important developmental process; its timing and pattern influence final yield in seed crops. Molecular dissection of flowering characteristics, such as underlying mechanisms and responses to environmental stimuli, have contributed on the understanding of mechanisms of adaptation and genetic differentiation of crops in a wide range of environments. The early-flowering *efl* mutant focused in this study was selected from the mutant population of a Japanese cultivar Nourin No. 2 generated by ion-beam radiation. The preliminary mapping study suggested that *efl* was a novel mutant in soybean because no QTL for flowering has been reported so far around the SSR markers used for tagging the mutant gene. Furthermore, the marker-assisted analysis for flowering time revealed that the *efl* mutation had the largest effect on flowering time among the genes segregated in the mapping population.

I carried out fine mapping and resequencing analyses and found that the mutation was caused by a deletion of 214 kbp in the genomic positions 5,327,094 to 5,540,664 nucleotide in chromosome 19 in the Williams 82 reference genome (Figures 8). Despite that a total of 11 annotated genes were deleted, the *efl* mutant had no phenotypic aberration in plant architecture or growth. This is a sharp contrast with mutant lines with altered flowering detected by Mikuriya et al. (2017) from the same mutant population; most of the mutant lines had yellowish leaves with reduced chlorophyll contents. The lack of aberrant phenotype in *efl* is most likely attributed to the presence of homoeologous copies of the deleted genes; all of the 11 genes except for a mitochondrial gene had homoeologous copies scattered among seven chromosomes (Figure 9). The paleopolyploid nature of the soybean genome may therefore provide the tolerance against the large deletion as observed in N2-*efl*.

The annotated genes in the deleted region included *COL2b*, a soybean ortholog of *Arabidopsis* *CONSTANS* (Table 2). The other genes in the deleted region encoded proteins involved in cell metabolic pathways, such as the pentose phosphate pathway, glycosylation, ATP biosynthesis, amino acid export, and the mitochondrial electron

transport chain, none of which has been implicated in flowering time control in *Arabidopsis* and the other plant species. The results obtained strongly indicate that the deletion of *COL2b* is the most probable causal factor for the early flowering of the mutant lines.

The *efl* mutation mitigates the repression of flowering by cool temperatures

The *efl* mutant, N2-*efl*, flowered approximately 20 days earlier than its parent, Nourin No. 2, in the natural daylength condition of Sapporo. However, the difference in flowering time was small in Tsukuba, an environment with higher ambient temperatures and shorter daylengths than Sapporo. In order to characterize the flowering characteristics of *efl* mutation, flowering times of NILs for the mutant *efl* and wild-type alleles were compared in various photo-thermal conditions. The *efl* lines flowered, on average, only 2.0 days earlier than the wild-type lines at 25°C and 16-h daylength condition (Figure 10A). The flowering at the 16-h daylength was markedly delayed in both NILs, when the temperature was reduced to 18°C (Figure 10B). However, the effect of cool temperature on flowering time was mitigated in the *efl* lines. The flowering time was delayed by 30 to 32 days in the wild-type lines but only by 18 to 20 days in the *efl* lines; the differences between NILs were enlarged from 2.0 and 3.5 days at 25°C to 11.5 and 12.3 days at 18°C in the two sets of NILs. Increment of ambient temperatures from 25°C to 32°C also influenced flowering times at 16-h daylength condition differentially between the NILs, but the effect was smaller than that of low temperatures (Figure 10B). Therefore, the *efl* mutation may be involved in the thermal regulation of flowering, particularly in responses to cool temperatures, and mainly have a role to moderate the floral repression induced by cool temperatures.

The different flowering times at cool temperatures reflected different transcript abundances particularly for *FT2a* and, to a lesser degree, for *FT5a* between the NILs (Figures 12 and 13). Cooler temperatures greatly decreased the expression of *FT2a* and *FT5a* at the second and forth trifoliate leaf stages. However, the repression was moderated in the NIL with the *efl* allele. In contrast, no large and consistent difference was observed between the NILs in the expression levels of *FT2a* and *FT5a* at 25°C.

The early flowering of *efl* mutation may therefore be attributed to the alleviation of the repression of *FT2a* and *FT5a* expression at cool temperatures.

Involvement of COL2b in the thermal regulation of flowering

Cool temperatures strongly downregulated the *FT2a* and *FT5a* expressions, but their effects differed between the NILs. The expression of *FT2a* and *FT5a* in the *efl* lines that lacked *COL2b* was low but detectable, and correlated with the promotion of flowering at 18°C (Figures 12 and 13). It is therefore tempting to hypothesize that *COL2b* functions as a floral repressor in response to cool temperature. This hypothesis is consistent with the findings of Wu et al. (2019), which demonstrated that the downregulation of *COL2* gene expressions by RNAi resulted in the early flowering. As in the case of *COL1a* and *COL1b*, which are involved in the inhibition of flowering under LD conditions (Cao et al., 2015), *COL2b* may function as a floral repressor, particularly in response to cool temperatures.

In the wild-type lines, the cool temperature upregulated the expression of *COL2b* or shifted the peak from 3 ZT to 12 ZT at the fourth trifoliate leaf stage, but the expression of *COL2b* was slightly decreased in response to cool temperature at the second trifoliate leaf stage; no consistent correlations were detected between the expression levels of *COL2b* and *FT2a* or *FT5a* (Figures 12, 13 and 14). The regulation of *FT2a* and *FT5a* by cool temperatures and the relative role of *COL2b* may vary with the growing stages. In the later growing stage, *COL2b* may be upregulated totally or at least in the late afternoon (12ZT) in response to cool temperature, and in turn repress the *FT2a* and *FT5a* expressions. In the *efl* lines, the lack of *COL2b* may moderate the repression of flowering by cool temperatures.

Cool temperatures also highly upregulated the expression of *E1* in the wild-type lines at the fourth trifoliate leaf stage but had no effect in the NILs with the *efl* allele (Figure 15), suggesting that the *E1* expression is under the control of *COL2b*. Wu et al. (2019) also found a reduction of *E1* peak expression in *COL2*-RNAi transgenic plants, suggesting the control of *E1* expression by *COL2*. Moreover, they analyzed the interaction of *COL2* and *E1*, using an F₂ population from crosses between *COL2*-RNAi

plants and *E1* NILs. In the *e1-as/e1-as* background, *COL2*-RNAi plants flowered earlier than non-RNAi plants, while under the same *COL2*-RNAi background, the *E1/E1* plants showed later flowering than the *e1-as/e1-as* plants. The plants carrying both functional *COL2* and *E1* showed more marked repressive effect than those carrying either of the hypomorph *e1-as* allele or the *COL2*-RNAi transgene, suggesting a synergistic interaction of *COL2* and *E1*. The RNAi approach by Wu et al. (2019) and mutant analysis in this study may therefore indicate that *COL2*, at least *COL2b*, is involved in the induction of *E1* expression. Wu et al. (2019) further found that in the *COL2*-RNAi plants, the *E1* expression was only partly repressed and consequently *E1* retained its function. These findings may indicate that the *E1* expression is controlled by multiple regulatory systems. Additional researches with the overexpression and appropriate genetic materials will be required to confirm whether *COL2b* can really induce the *E1* expression and whether *E1* retains its function in the absence of *COL2b*.

Thermal regulation of flowering by CO in *Arabidopsis* and rice

CO and its homolog Hd1 have a key role on photoperiodic flowering in *Arabidopsis* and rice, respectively. They are also involved in thermal regulation of flowering.

In *Arabidopsis*, high temperature promotes flowering by activating *FT* transcription. Increasing the ambient temperature from the optimum to 27°C accelerates flowering even under non-inductive SD conditions (Balasubramanian et al., 2006; Sanchez-Bermejo et al., 2015). At high temperatures, H2A.Z histone-containing nucleosomes are removed from temperature-sensitive cis-elements in the promoter of *FT* (Kumar and Wigge, 2010; Kumar et al., 2012), enabling the access of specific transcription factors to these cis-elements (Kumar and Wigge, 2010). One of those transcription factors is PIF4, which binds to the *FT* promoter more strongly when plants are grown under SD conditions at 27°C than when they are grown at 22°C (Kumar et al., 2012; Wigge, 2013). The flowering is delayed in a *pif4* null mutant, but not in the wild type under SD conditions at 27°C; the number of rosette leaves at flowering in the wild-type line is the same between 27°C and 22°C under SD conditions, indicating a crucial role of *PIF4* in the thermo-sensory induction of flowering (Kumar et al., 2012). Both CO and PIF4 bind to similar regions of the *FT* promoter (Kumar et al., 2012; Song

et al., 2012) and the expression of *PIF4* can activate *FT* transcription only at times when *CO* is expressed (Fernández et al., 2016). Using co-immunoprecipitation assay, Fernández et al. (2016) found a direct physical interaction between *CO* and *PIF4*, which is important for the activation of *FT* transcription at high temperature under SD conditions. Besides the role as a central component in the photoperiodic pathway, *CO* contributes to the upregulation of *FT* expression at a high temperature under SD conditions through the interaction with *PIF4* (Fernández et al., 2016).

The thermal regulation of flowering in response to ambient temperatures has also been investigated in rice. The reduction of ambient temperature from 27°C to 23°C delays flowering by downregulating expression of *Hd3a* and *RFT1* and of *Ehd1* (Luan et al., 2009; Song et al., 2012b). This repression is caused by increased transcript levels of *Ghd7* at low temperature, particularly under LD conditions; low temperature and LD have a synergistic role in the expression of *Ghd7* (Song et al., 2012b). Unlike the transcription of *COL2b* observed in this study, however, the *Hd1* transcription is unchanged in different temperature conditions, suggesting that *Hd1* is not involved in the temperature flowering regulatory pathway (Luan et al., 2009). Accordingly, *Ghd7* plays a crucial role in controlling flowering in response to photoperiod and ambient temperature in rice (Song et al., 2012b). Under LD conditions, *Hd1* positively regulates the *Ghd7* transcription, and forms a protein complex with *Ghd7* to repress the expression of downstream genes of *Ehd1*, *Hd3a* and *RFT1* (Nemoto et al., 2016; Zhang et al., 2017). *Hd1* therefore controls the *Hd3a* and *RFT1* expression in LD through both the induction of *Ghd7* expression and the interaction with the *Ghd7* protein.

In barley (*Hordeum vulgare*), high ambient temperature (28/24°C, day/night) retards the reproductive development of spring barley that carries dysfunctional alleles of both *PHOTOPERIOD1* (*Ppd-H1*, the barley homolog of *Arabidopsis* *PRR*) and *VRN-H2* (the barley homolog of rice *Ghd7*), relative to the control temperature (20/16°C, day/night) (Ejaz and von Korff, 2017). Both the functional *Ppd-H1* allele and the loss-of-function allele of *HvELF3*, a negative regulator of *Ppd-H1* expression at night (Faure et al., 2012), accelerate flowering and floral development under high temperatures (Ejaz and von Korff, 2017). Two homologs of *Arabidopsis* *CO*, *HvCO1* and *HvCO2*, upregulate *VRN-H2*, resulting in reduced expression of *HvFT1* (the barley

homolog of *FT*), and delay flowering under both LD and SD conditions when ectopically overexpressed in the winter genotype that carries a functional *VRN-H2* allele; however, in the spring genotype that carries a dysfunctional *vrn-H2* allele, *HvCO2* can function as an activator of flowering, depending on the *Ppd-H1* genotype (Mulki and von Korff, 2016).

The findings obtained in these plant species suggest that the stimuli from photoperiod and ambient temperature are integrated by key genes that control the flowering time. Furthermore, the interaction between *Hd1* and *Ghd7* in rice (*HvCO* and *VRN-H2* in barley) may provide an intriguing implication on the interaction between *COL2b* and *E1* in soybean. It is tempting to hypothesize that *COL2b* functions as the repressor in thermal regulation of flowering through not only the upregulation of the *E1* and *EIL* expressions but also the formation of the protein complexes with them, as is the case in rice *Hd1* and *Ghd7*. A further study is needed to confirm whether *COL2b* interacts with *E1* and *EIL* proteins.

Molecular diversity of *COL2b* and its homologs

Neither QTLs nor SNPs associated with time to flowering have been detected in the genomic regions flanking the deletion in the *efl* mutant in previous studies on QTL and GWAS. In order to explore a potential role of *COL2b* in the genetic variation of flowering time, I surveyed the resequence data for 512 cultivated and wild soybean accessions deposited in SoyBase. Most of accessions tested shared the same AA sequence (Haplotype 1) with Williams 82. Only twenty-one accessions (4.1 %), two cultivated and nineteen wild accessions, possessed haplotypes different in at least one amino acid from the former. Furthermore, all of the AA substitutions were located in the outside of any functionally important domains of CO proteins (Figures 16, 17, 18 and 19). Therefore, *COL2b* might have not played important and noticeable roles in the diversification of flowering characteristics in soybean. The early-flowering mutation in N2-*efl* may therefore be useful as a novel genetic resource for better understanding of the molecular mechanisms of flowering and improving the adaptability of soybean cultivars to cool environments.

Gene network for photo-thermal regulation of flowering in soybean

Based on the findings obtained in this study and other researches, I propose a gene network of *COL2b* and other important maturity genes and QTLs that control flowering in response to different environmental cues (Figure 20).

Similar to the functions of *COL1a* and *COL1b* genes in LD, *COL2b* represses flowering particularly under cool temperatures. The repression is accompanied with the downregulation of *FT2a* and *FT5a* expression, which is most likely caused by *COL2b* itself directly or indirectly via the upregulation of *E1* and *E1L* genes. The *E1* expression may therefore be controlled by genes involved in different regulatory pathways; PHYA (E3 and E4)-mediated and PRR3/7 (Tof11 and Tof12)-mediated pathways in LD, ELF3 (J)-mediate pathway in SD, and *COL2b* (*efl*)-mediated pathway in cool temperature (Figure 20).

E2 is the *Arabidopsis GI* homolog (*Gla*) (Watanabe et al., 2011). Taking into account the function of *GI* in *Arabidopsis* and its orthologue in rice, *E2* may be a regulator of the *COL2b* expression, as in *Arabidopsis* and rice. However, the materials used in this study, such as Nourin No.2, Tokachinagaha, and the derived NILs for the *efl* allele, all possess the recessive *e2* allele, suggesting that *E2* is not involved in the upregulation of *COL2b* by cool temperatures observed in this study. There might be unknown gene(s) (X in Figure 20) in the control of the thermal regulation of *COL2b* expression. *Glb* and *Glc*, homoeologs of *E2*, may be probable candidates for the regulator of *COL2b* expression.

E2 has large effects on flowering time across a wide range of latitudes (Jiang et al., 2014; Lu et al., 2015) and a distinct geographical distribution of alleles along latitudes; the *e2* allele is predominant in early-flowering cultivars of higher latitudes (Xu et al., 2013; Wang et al., 2016a). However, the function in the control of flowering still remains unclear. Wang et al. (2016a) found that *E2* could not rescue a late-flowering phenotype in an *Arabidopsis gi-2* mutant and delayed flowering of Col-0 when ectopically expressed, suggesting that *E2* is differentiated functionally from *GI*, which promotes *CO*.

In *Arabidopsis*, *GI* regulates the *miR172* abundance and its target, *TARGET OF EARLY ACTIVATION TAGGED (EAT) 1 (TOE1)*, in response to photoperiod through a unique *CO*-independent pathway; the *miR172* abundance is reduced significantly in the *gi-2* mutant but is unaffected in the *co-2* mutant and does not oscillate with a diurnal rhythm under both LD and SD conditions (Jung et al., 2007; Zhu et al., 2011). Unlike in *Arabidopsis*, however, the expression level of *miR172a* (a member of soybean *miR172s*) has a diurnal rhythm under both LD and SD conditions, although the peak is shifted to later time and the amplitude increases in LD (Wang et al., 2016b). The overexpression of *E2* further does not promote the expression of *miR172a* in LD; it promotes the maturation of *miR172a* through up-regulation of soybean homologues of *DICER-LIKE 1* and *SERRATE*, genes for key enzymes in the miRNA synthesis pathway (Wang et al., 2016b).

Zhao et al. (2015) found that *TOE4a* (Glyma.15g04930), a *TOE1* homolog in soybean, targeted by *miR172a*, is regulated by the maturity genes *E3* and *E4*, and represses flowering-related genes, including *FT2a*, *FT5a*, *API*, and *LFY*. Unlike in *Arabidopsis*, however, *TOE4a* is associated with *COL1a* in soybean; the expression of *COL1a* is upregulated in the *TOE4* overexpression transgenic line, suggesting that the *miR172*-mediated *TOE4a* regulates flowering at least partially dependently on *COLs* (Zhao et al., 2015).

Moreover, Zhao (2015) found no difference in the abundance of mature *miR172a* between ‘Harosoy’ (*e1-as/e2/E3/E4*) and its NIL for *E2* (*e1-as/E2/E3/E4*). Also, there were no differences in the expression levels of *TOE4a* between the NILs for *E2* and *e2* in both ‘Harosoy’ and ‘Bay’ (*E1/E2/E3/E4*) backgrounds. Cao et al. (2016) assumed that the *E2* and *e2* alleles did not have any role on *miR172* maturation or alternatively, an unknown factor(s) might mask the allelic effects. The transcript level of *TOE4a* was highest in ‘Harosoy’ and lowest in its NIL for *e3* and *e4* (*e1-as/e2/e3/e4*), while the transcript level of *miR172* was lowest in ‘Harosoy’ and highest in its NIL for *e3* and *e4* (Zhao et al., 2015). Accordingly, the transcript abundances of *TOE4a* and *miR172a* may rather be controlled by *PHYA* genes, *E3* and *E4*.

Wu et al. (2019) found that the other homolog of *TOE1* designated *TOE1b* (Glyma.19g178200), which possessed a *miR172* target site as well, was expressed higher in Clark (*e1-as/E2/E3/e5*) than its NIL for *E1* (K65-3366; *E1/E2/E3/e5*), suggesting that the functional *E1* allele suppresses *TOE1b*. Overcoming the effect of *E1*, *TOE1b* was upregulated by *e2* in the NIL L66-432 (*E1/e2/E3/e5*) and by *e3* in the NIL L74-441 (*E1/E2/e3/e5*).

Taken together, the function of *E2* in the control of flowering remains still unclear. Using RNAseq analyses, Wu et al. (2019) found that homologs of the circadian clock genes *LHY* and *TIMING OF CAB EXPRESSION 1* (Glyma.17g102200), the clock-controlled *CYCLING DOF FACTOR* (*CDF2*; Glyma.18g260500, *CDF3c*; Glyma.05g025900), *FT2a*, and a homolog of the floral meristem identity gene *AGAMOUS LIKE 8* (*AGL8a*; Glyma.17g081200, *AGL8b*; Glyma.06g205800, *AGL8c*; Glyma.13g052100, *AGL8d*; Glyma.19g034600) were under the control of *E2*. However, *E2* did not appear to have any large effects on the expressions of *COL* genes, unlike in the relationship of *GI* and *CO* in *Arabidopsis* (Mizoguchi et al., 2005) and that of their orthologues in rice (Hayama et al., 2003). Further studies are required to determine whether *E2* functions as an inducer of *COL2b* and how *E2* responds to changes of photoperiods and ambient temperatures.

Concluding remarks

In this study, the DNA polymorphism responsible for the early flowering *efl* mutant was determined as a 214 kbp deletion containing *COL2b*, a soybean ortholog of *Arabidopsis* *CONSTANS*. Characterization of flowering time and expression profile of flowering genes in different photo-thermal conditions using NILs for the *efl* mutant strongly suggested an involvement of *COL2b* in the thermal regulation of flowering, particular in chilling temperatures, and its possible relation with *E1* and *E1L* genes. Further studies are needed to verify the binding of *COL2b* to the promoters of *E1/E1L* genes and the physical interaction of *COL2b* with *E1/E1L* proteins and to determine the function of *E2* and its homologs as probable controllers of *COL2b*, which will provide new insights in an upstream gene network and function of *COL2b* and further in the molecular mechanisms of photoperiodic and thermal regulations of flowering in soybean. Moreover, the lack of functional diversity in the *COL2b* protein in the cultivated soybean germplasm as revealed by a survey on resequence data implies that the early-flowering mutation in N2-*efl* is a useful genetic resource not only for better understanding of the molecular mechanisms of flowering but also for improving the adaptability of soybean cultivars to cool environments. Understanding of floral response to cool temperature will contribute to the development of cultivars adaptive to cool environments in soybean.

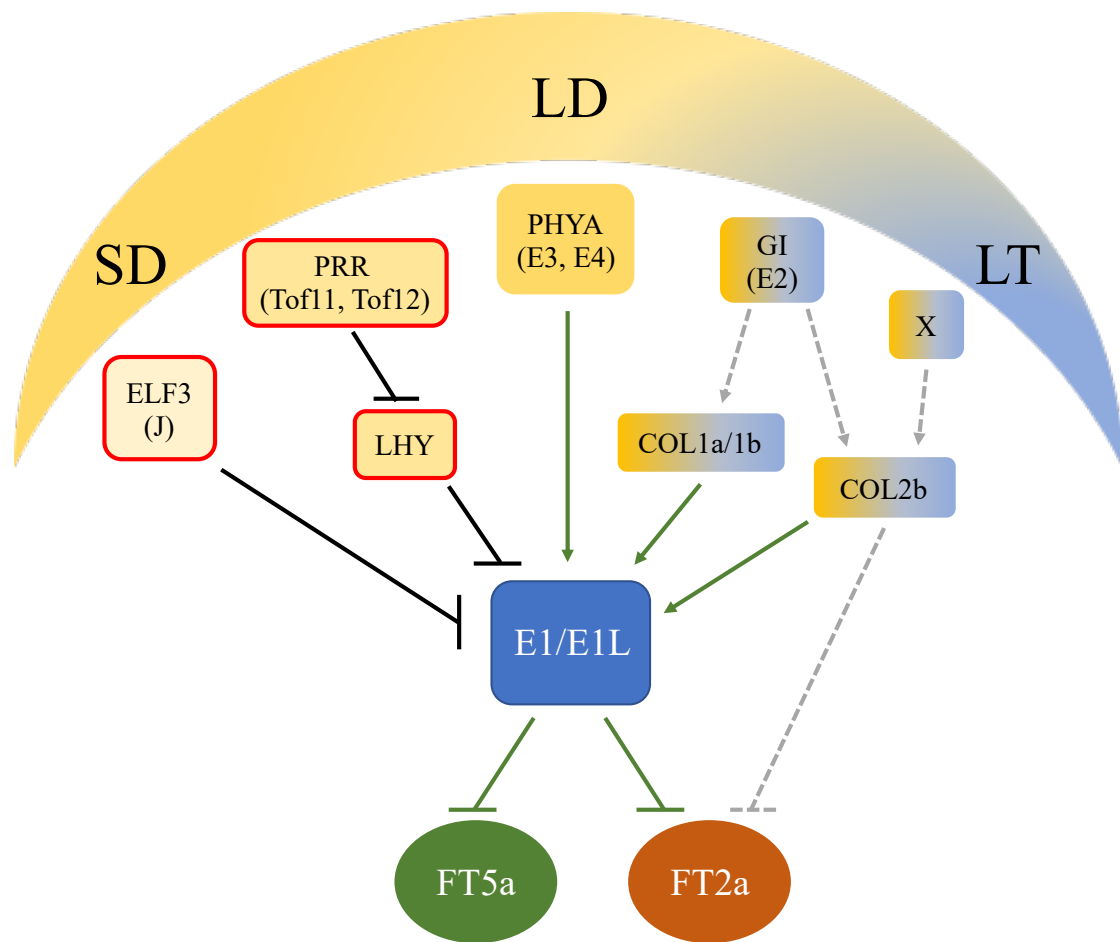


FIGURE 20. Predicted network of COL2b and important maturity genes and QTLs in photo-thermal responses of flowering in soybean.

Various environmental cues, such as SD, LD and cool temperature, control flowering through different regulatory pathways, PHYA (E3 and E4)-mediated pathway, PRR3/7 (Tof11 and Tof12)-mediated pathway, ELF3 (J)-mediated pathway, and COL2b (ef1)-mediated pathway. *COL2b* is one of the genes involved in the control of flowering in cool temperatures. Independently of the *E2* genotype, its expression is induced in cool temperatures by an unknown factor(s), such as X. E2 may also positively contribute to the *COL* expressions, a relationship that has been observed between their orthologues in *Arabidopsis* and rice, but its function should be clarified in a further study. Genes surrounded by red lines are transcription factors which directly associate with the promoter of downstream genes (black lines). Green lines indicate up- or downregulation of downstream genes, as revealed by expression profiling. Dotted lines are predicted interactions. ↓ means promotion and ⊥ means suppression of gene expression.

Chapter 6 Summary

Ambient growth temperature has a large impact on reproductive development in seed crops. The thermal response of flowering in soybean varies with photoperiods and maturity genotypes of cultivars. However, the genetic variation and its underlying molecular mechanism remain barely explored. This study is aiming to determine a molecular genetic basis for an early-flowering soybean mutant generated by the ion-beam mutagenesis and characterize the function of responsible gene in thermal regulation of flowering, particular in response to cool temperature. The results obtained are summarized as follows.

In Chapter 2, the genetic basis for the early-flowering *efl* mutation, which was selected from the mutant population of Nourin No. 2 generated by ion beam mutagenesis, was analyzed in the progeny of a cross with an early-flowering cultivar ‘Tokachi Nagaha’. The segregation analysis with tagging markers revealed that the early flowering was controlled mainly by a single recessive gene designated *efl*. Fine-mapping placed the *efl* gene between BARCSOY_19_0279 and BARCSOY_19_0310, which covered a 642-kbp genomic region, in chromosome 19 and contained 32 annotated genes (Glyma.19g037700 to Glyma.19g040800) in the Williams 82 reference genome sequence. Moreover, a comparison of resequencing data between the mutant and parental lines using the Integrative Genomics Viewer revealed a large deletion of ~214 kbp in the *efl* line, which was confirmed by PCR with primers designed to sandwich it. The deleted region contained 11 annotated genes, of which *COL2b*, a soybean ortholog of *Arabidopsis* *CONSTANS*, is supposed to be the best candidate for the early flowering of the mutant line.

In Chapter 3, the flowering characteristics and expression profiles for flowering genes were evaluated in two sets of near-isogenic lines for the *efl* and wild-type allele. The flowering time was observed in different photo-thermal conditions. The results revealed that the *efl* mutation moderated the responses of flowering to longer photoperiods and lower ambient temperatures, where the flowering was repressed in the wild-type lines. The difference in flowering time between the *efl* and wild-type lines was more marked in cool temperatures than the other photo-thermal conditions.

The expressions of *FT2a* and *FT5a*, important soybean *FLOWERING LOCUS T* orthologs, were strongly repressed at an early growing stage in the cool temperature. The repression was released slightly at the later growing stage, but its extent was different between the two *FT* orthologs. The *FT5a* expression was partially restored to the level in the optimum temperature condition, while the *FT2a* expression still remained repressed. The repression of *FT2a* and *FT5a* expressions in the wild-type lines by cool temperatures were associated with the upregulation or shift of peak from 3 ZT to 12 ZT of the *COL2b* expression and the upregulation of *E1* and *E1Lb* expressions, particularly at the later growing stage. Therefore, the *efl* mutation moderated the repression by cool temperatures to promote flowering in the cool ambient temperatures. The results obtained strongly suggested a possible involvement of *COL2b* in the thermal regulation of flowering, particularly in response to cool weather. Moreover, cool temperatures also highly upregulated the expression of *E1* and its homolog in the wild-type lines at the fourth trifoliate leaf stage but had no effect in the NILs with the *efl* allele, suggesting that the *E1* expression may be under the control of *COL2b*.

In Chapter 4, the sequence variations for *COL2b* and the other three *COL* genes were evaluated for 512 cultivated and wild soybean accessions, using the resequence data deposited in SoyBase. The analyses revealed that the diversity of the peptide sequences was extremely low in *COL2b* and the variant haplotypes were mostly found in the wild accessions. Until now, no QTL and single nucleotide polymorphisms associated with flowering time have been reported in the genome region harboring *COL2b*. The absence of genetic variation associated with *COL2b* may therefore be consistent with a low genetic variability of the coding sequence.

Finally, I discussed a possible involvement of the *COL2b* gene in the thermal regulation of flowering, particular in response to cool temperatures in soybean, and proposed a gene network focusing on the essential flowering regulator *E1* in the photo-thermal regulation of flowering in soybean. The molecular dissection of *efl* mutant demonstrates that *COL2b* play a key role to control the expression of *FT2a* and *FT5a* directly or indirectly through the induction of *E1* expression in response to cool temperature. The *efl* mutant may be used as a promising novel genetic resource in both

understanding the molecular mechanism of photo-thermal responses of flowering and improving the adaptability of soybean cultivars to cool environments.

References

- Andrés, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics*, 13, 627-639. DOI: 10.1038/nrg3291
- Arase, S., Hase, Y., Abe, J., Kasai, M., Yamada, T., Kitamura, K., Narumi, I., Tanaka, A., and Kanazawa, A. (2011). Optimization of ion-beam irradiation for mutagenesis in soybean: effects on plant growth and production of visibly altered mutants. *Plant Biotechnology*, 28, 323-329. DOI:10.5511/plantbiotechnology.11.0111a
- Arya, H., Singh, M. B., and Bhalla, P. L. (2018). Genomic and molecular analysis of conserved and unique features of soybean PIF4. *Scientific Reports*, 8, 1-11. DOI: 10.1038/s41598-018-30043-2
- Balasubramanian, S., Sureshkumar, S., Lempe, J., and Weigel, D. (2006). Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. *PLoS Genetics*, 2, e106. DOI:10.1371/journal.pgen.0020106
- Bernard, R. L. (1971). Two major genes for time of flowering and maturity in soybeans. *Crop Science*, 11, 242-244. DOI: 10.2135/cropsci1971.0011183X001100020022x
- Bonato, E. R., and Vello, N. A. (1999). *E6*, a dominant gene conditioning early flowering and maturity in soybeans. *Genetics and Molecular Biology*, 22, 229-232. DOI: 10.1590/S1415-47571999000200016
- Bouché, F., Lobet, G., Tocquin, P., and Périlleux, C. (2016) FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Research*, 44 (Database issue): D1167–D1171. DOI: 10.1093/nar/gkv1054

Brambilla, V., Gomez-Ariza, J., Cerise, M., and Fornara, F. (2017). The importance of being on time: regulatory networks controlling photoperiodic flowering in cereals. *Frontiers in Plant Science*, 8, 665. DOI: 10.3389/fpls.2017.00665

Buzzell, R. I. (1971). Inheritance of a soybean flowering response to fluorescent-daylength conditions. *Canadian Journal of Genetics and Cytology*, 13, 703-707. DOI: 10.1139/g71-100

Buzzell, R. I., and Voldeng, H. D. (1980). Inheritance of insensitivity to long daylength. *Soybean Genetics Newsletter*, 7, 13.

Cao, D., Li, Y., Lu, S., Wang, J., Nan, H., Li, X., Shi, D., Fang, C., Zhai, H., Yuan, X., Anai, T., Xia, Z., Liu, B., and Kong, F. (2015). *GmCOL1a* and *GmCOL1b* function as flowering repressors in soybean under long-day conditions. *Plant and Cell Physiology*, 56, 2409-2422. DOI: 10.1093/pcp/pcv152

Cao, D., Takeshima, R., Zhao, C., Liu, B., Abe, J., and Kong, F. (2017). Molecular mechanisms of flowering under long days and stem growth habit in soybean. *Journal of Experimental Botany*, 68, 1873-1884. DOI: 10.1093/jxb/erw394

Carroll, B. J., McNeil, D. L., and Gresshoff, P. M. (1985). Isolation and properties of soybean [*Glycine max* (L.) Merr.] mutants that nodulate in the presence of high nitrate concentrations. *Proceedings of the National Academy of Sciences, USA* 82, 4162-4166. DOI: 10.1073/pnas.82.12.4162

Cheng, L., Wang, Y., Zhang, C., Wu, C., Xu, J., Zhu, H., Leng, J., Bai, Y., Guan, R., Hou, W., Zhang, L., and Han, T. (2011). Genetic analysis and QTL detection of reproductive period and post-flowering photoperiod responses in soybean. *Theoretical and Applied Genetics*, 123, 421-429. DOI: 10.1007/s00122-011-1594-8

Cober, E. R., Tanner, J. W., and Voldeng, H. D. (1996). Soybean photoperiod-sensitivity loci respond differentially to light quality. *Crop Science*, 36, 606-610. DOI: 10.2135/cropsci1996.0011183X003600030014x

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

Cober, E. R., Stewart, D. W., and Voldeng, H. D. (2001). Photoperiod and temperature responses in early-maturing, near-isogenic soybean lines. *Crop Science*, 41, 721-727. DOI:10.2135/cropsci2001.413721x

Cober, E. R., Molnar, S. J., Charette, M., and Voldeng, H. D. (2010). A new locus for early maturity in soybean. *Crop Science*, 50, 524-527. DOI: 10.2135/cropsci2009.04.0174

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., and Coupland, G. (2007). FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science*, 316, 1030-1033. DOI: 10.1126/science.1141752

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, H. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., and Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156-2158. DOI: 10.1093/bioinformatics/btr330

Dissanayaka, A., Rodriguez, T. O., Di, S., Yan, F., Githiri, S. M., Rodas, F. R., Abe, J., and Takahashi, R. (2016). Quantitative trait locus mapping of soybean maturity gene *E5*. *Breeding Science*, 60, 407-415 DOI: 10.1270/jsbbs.15160

Doi, K., Izawa, T., Fuse, T., Yamanouchi, U., Kubo, T., Shimatani, Z., Yano, M., and Yoshimura, A. (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 and Development*, 18, 926-936. DOI: 10.1101/gad.1189604

Doyle, J.J. and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12, 13-15.

Ejaz, M., and von Korff, M. (2017). The genetic control of reproductive development under high ambient temperature. *Plant Physiology*, 173, 294-306. DOI:10.1104/pp.16.01275

Fan, C., Hu, R., Zhang, X., Wang, X., Zhang, W., Zhang, Q., Ma, J., and Fu, Y. F. (2014). Conserved CO-FT regulons contribute to the photoperiod flowering control in soybean. *BMC Plant Biology*, 14, 9. DOI: 10.1186/1471-2229-14-9

Faure, S., Turner, A. S., Gruszka, D., Christodoulou, V., Davis, S. J., von Korff, M., and Laurie, D. A. (2012). Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. *Proceedings of the National Academy of Sciences, USA* 109, 8328-8333. DOI:10.1073/pnas.1120496109

Fehr, W. R., Welke, G. A., Hammond, E. G., and Cianzio, S. R. (1991). Inheritance of reduced palmitic acid content in seed oil of soybean. *Crop Science*, 31, 88-89. DOI: 10.2135/cropsci1991.0011183X003100010022x

Fernández, V., Takahashi, Y., Le Gourrierc, J., and Coupland, G. (2016). Photoperiodic and thermosensory pathways interact through CONSTANS to promote flowering at

high temperature under short days. *The Plant Journal*, 86, 426-440. DOI:10.1111/tpj.13183

Fujii, T., Ikenaga, M., and Lyman, J. T. (1966). Radiation effects on *Arabidopsis thaliana*—II killing and mutagenic efficiencies of heavy ionizing particles. *Radiation Botany*, 6, 297-306. DOI: 10.1016/S0033-7560(66)80041-8

Fujii, T., Ikenaga, M., and Lyman, J. T. (1967). Killing and mutagenic efficiencies of heavy ionizing particles in *Arabidopsis thaliana*. *Nature*, 213, 175-176. DOI: 10.1038/213175b0

Funatsuki, H., Kawaguchi, K., Matsuba, S., Sato, Y., and Ishimoto, M. (2005). Mapping of QTL associated with chilling tolerance during reproductive growth in soybean. *Theoretical and Applied Genetics*, 111, 851-861. DOI: 10.1007/s00122-005-0007-2

Githiri, S. M., Yang, D., Khan, N. A., Xu, D., and Takahashi, R. (2007). QTL analysis of low temperature-induced browning in soybean seed coats. *Journal of Heredity*, 98, 360-366. DOI: 10.1093/jhered/esm042

Hadley, P., Roberts, E. H., Summerfield, R. J., and Minchin, F. R. (1984). Effects of temperature and photoperiod on flowering in soya bean [*Glycine max* (L.) Merrill]: a quantitative model. *Annals of Botany*, 53, 669-681. DOI: 10.1093/oxfordjournals.aob.a086732

Hajika, M., Igita, K., and Kitamura, K. (1991). A line lacking all the seed lipoxygenase isozymes in soybean [*Glycine max* (L.) Merrill] induced by gamma-ray irradiation. *Japanese Journal of Breeding*, 41, 507-509. DOI: 10.1270/jsbbs1951.41.507

Hayama, R., Yokoi, S., Tamaki, S., Yano, M., and Shimamoto, K. (2003). Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature*, 422, 719-722. DOI: 10.1038/nature01549

Hayashi, M., Harada, K., Fujiwara, T., and Kitamura, K. (1998). Characterization of a 7S globulin-deficient mutant of soybean (*Glycine max* (L.) Merrill). *Molecular and General Genetics MGG*, 258, 208-214. DOI: 10.1007/s004380050724

Hirano, T., Kazama, Y., Ohbu, S., Shirakawa, Y., Liu, Y., Kambara, T., Fukunishi, N., and Abe, T. (2012). Molecular nature of mutations induced by high-LET irradiation with argon and carbon ions in *Arabidopsis thaliana*. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 735, 19-31. DOI: 10.1016/j.mrfmmm.2012.04.010

Imaizumi, T., Schultz, T. F., Harmon, F. G., Ho, L. A., and Kay, S. A. (2005). FKF1 F-box protein mediates cyclic degradation of a repressor of *CONSTANS* in *Arabidopsis*. *Science*, 309, 293-297. DOI: 10.1126/science.1110586

Jaeger, K. E., and Wigge, P. A. (2007). FT protein acts as a long-range signal in *Arabidopsis*. *Current Biology*, 17, 1050-1054. DOI: 10.1016/j.cub.2007.05.008

Jiang, B., Nan, H., Gao, Y., Tang, L., Yue, Y., Lu, S., Ma, L., Cao, D., Sun, S., Wang, J., Wu, C., Yuan, X., Hou, W., Kong, F., Han, T., and Liu, B. (2014). Allelic combinations of soybean maturity loci *E1*, *E2*, *E3* and *E4* result in diversity of maturity and adaptation to different latitudes. *PLoS one*, 9, e106042. DOI: 10.1371/journal.pone.0106042

Jiang, B., Zhang, S., Song, W., Khan, M. A. A., Sun, S., Zhang, C., Wu, T., Wu, C., and Han, T. (2019). Natural variations of *FT* family genes in soybean varieties covering a wide range of maturity groups. *BMC Genomics*, 20, 230. DOI:10.1186/s12864-019-5577-5

Jung, J. H., Seo, Y. H., Seo, P. J., Reyes, J. L., Yun, J., Chua, N. H., and Park, C. M. (2007). The GIGANTEA-regulated *microRNA172* mediates photoperiodic flowering independent of *CONSTANS* in *Arabidopsis*. *The Plant Cell*, 19, 2736-2748. DOI: 10.1105/tpc.107.054528

Kamiya, M., and Kiguchi, T. (2003). Rapid DNA extraction method from soybean seeds. *Breeding Science*, 53, 277-279. DOI:10.1270/jsbbs.53.277

Kaufmann, K., Smaczniak, C., and Angenent, G. (2009). The transcription machinery underlying flower formation. *NPC Highlights*, 2009, 5-7.

Khan, N. A., Takahashi, R., Abe, J., and Komatsu, S. (2009) Identification of cleistogamy associated proteins in floral buds of near-isogenic lines of soybean using proteome analysis. *Peptides*, 30, 2095-2102. DOI:10.1016/j.peptides.2009.08.012

King, G. J. (2015). Crop epigenetics and the molecular hardware of genotype× environment interactions. *Frontiers in Plant Science*, 6, 968. DOI: 10.3389/fpls.2015.00968

Kitamura, K. (1991). Spontaneous and induced mutations of seed proteins in soybean (*Glycine max* L. Merrill). *Gamma Field Symposia*, 30, 61-69.

Komiya, R., Ikegami, A., Tamaki, S., Yokoi, S., and Shimamoto, K. (2008). *Hd3a* and *RFT1* are essential for flowering in rice. *Development*, 135, 767-774. DOI: 10.1242/dev.008631

Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B. M., Watanabe, S., Yamada, T., Harada, K., and Abe, J. (2010). Two coordinately regulated homologs of *FLOWERING LOCUS T*

are involved in the control of photoperiodic flowering in soybean. *Plant Physiology*, 154, 1220-1231. DOI: 10.1104/pp.110.160796

Kong, F., Nan, H., Cao, D., Li, Y., Wu, F., Wang, J., Lu, S., Yuan, X., Cober, E. R., Abe, J., and Liu, B. (2014). A new dominant gene *E9* conditions early flowering and maturity in soybean. *Crop Science*, 54, 2529-2535. DOI:10.2135/cropsci2014.03.0228

Kumagai, E., and Sameshima, R. (2014). Genotypic differences in soybean yield responses to increasing temperature in a cool climate are related to maturity group. *Agricultural and Forest Meteorology*, 198, 265-272. DOI: 10.1016/j.agrformet.2014.08.016

Kumar, S. V., and Wigge, P. A. (2010). H2A. Z-containing nucleosomes mediate the thermosensory response in *Arabidopsis*. *Cell*, 140, 136-147. DOI: 10.1016/j.cell.2009.11.006

Kumar, S. V., Lucyshyn, D., Jaeger, K. E., Alós, E., Alvey, E., Harberd, N. P., and Wigge, P. A. (2012). Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature*, 484, 242. DOI: 10.1038/nature10928

Lam, H. M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F., Li, M., He, W., Qin, N., Wang, B., Li, J., Jian, M., Wang, J., Shao, G., Wang, J., Sun, S. S., and Zhang, G. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nature Genetics*, 42, 1053. DOI: 10.1038/ng.715

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Bowtie: An ultrafast memory-efficient short read aligner. *Genome Biology*, 10, R25. DOI: 10.1186/gb-2009-10-3-r25

Li, Y., Dong, Y., Wu, H., Hu, B., Zhai, H., Yang, J., and Xia, Z. (2019a). Positional cloning of the flowering time QTL *qFT12-1* reveals the link between the clock related *PRR* homolog with photoperiodic response in soybeans. *Frontiers in Plant Science*, 10, 1303. DOI: 10.3389/fpls.2019.01303

Li, M., Liu, Y., Tao, Y., Xu, C., Li, X., Zhang, X., Han, Y., Yang, X., Sun, J., Li, W., Li, D., Zhao, X., and Zhao, L. (2019b). Identification of genetic loci and candidate genes related to soybean flowering through genome wide association study. *BMC Genomics*, 20, 987. DOI: 10.1186/s12864-019-6324-7

Lin, H. X., Yamamoto, T., Sasaki, T., and Yano, M. (2000). Characterization and detection of epistatic interactions of 3 QTLs, *Hd1*, *Hd2*, and *Hd3*, controlling heading date in rice using nearly isogenic lines. *Theoretical and Applied Genetics*, 101, 1021-1028. DOI: 10.1007/s001220051576

Liu, B., Fujita, T., Yan, Z. H., Sakamoto, S., Xu, D., and Abe, J. (2007). QTL mapping of domestication-related traits in soybean (*Glycine max*). *Annals of Botany*, 100, 1027-1038. DOI: 10.1093/aob/mcm149

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

Liu, B., and Abe, J. (2010). QTL mapping for photoperiod insensitivity of a Japanese soybean landrace Sakamotowase. *Journal of Heredity*, 101, 251-256. DOI: 10.1093/jhered/esp113

Lu, S., Li, Y., Wang, J., Srinives, P., Nan, H., Cao, D., Wang, Y., Li, J., Li, X., Fang, C., Shi, X., Yuan, X., Wantanabe, S., Feng, X., Liu, B., Abe, J., and Kong, F. (2015). QTL

mapping for flowering time in different latitude in soybean. *Euphytica*, 206, 725-736. DOI: 10.1007/s10681-015-1501-5

Lu, S., Zhao, X., Hu, Y., Liu, S., Nan, H., Li, X., Fang, C., Cao, D., Shi, X., Kong, L., Su, T., Zhang, F., Li, S., Wang, Z., Yuan, X., Cober, E. R., Weller, J. L., Liu, B., Hou, X., Tian, Z., and Kong, F. (2017). Natural variation at the soybean *J* locus improves adaptation to the tropics and enhances yield. *Nature Genetics*, 49, 773. DOI:10.1038/ng.3819

Lu, S., Dong, L., Fang, C., Liu, S., Kong, L., Cheng, Q., Chen, L., Su, T., Nan, H., Zhang, D., Zhang, L., Wang, Z., Yang, Y., Yu, D., Liu, X., Yang, Q., Lin, X., Tang, Y., Zhao, X., Yang, X., Tian, C., Xie, Q., Yuan, X., Tian, Z., Liu, B., Weller, J. L., and Kong, F. (2020). Stepwise selection on homeologous PRR genes controlling flowering and maturity during soybean domestication. *Nature Genetics*, 52, 428-436. DOI: 10.1038/s41588-020-0604-7

Luan, W., Chen, H., Fu, Y., Si, H., Peng, W., Song, S., Liu, W., Hu, G., Sun, Z., Xie, D., and Sun, C. (2009). The effect of the crosstalk between photoperiod and temperature on the heading-date in rice. *PLoS One*, 4, e5891. DOI:10.1371/journal.pone.0005891

Lü, J., Suo, H., Yi, R., Ma, Q., and Nian, H. (2015). Glyma11g13220, a homolog of the vernalization pathway gene *VERNALIZATION 1* from soybean [*Glycine max* (L.) Merr.], promotes flowering in *Arabidopsis thaliana*. *BMC Plant Biology*, 15, 232. DOI: 10.1186/s12870-015-0602-6

Lyu, J., Cai, Z., Li, Y., Suo, H., Yi, R., Zhang, S., and Nian, H. (2020). The floral repressor *GmFLC-like* is involved in regulating flowering time mediated by low temperature in soybean. *International Journal of Molecular Sciences*, 21, 1322. DOI: 10.3390/ijms21041322

Malosetti, M., van Eeuwijk, F. A., Boer, M. P., Casas, A. M., Elía, M., Moralejo, M., Bhat, P. R., Ramsay, L., and Molina-Cano, J. L. (2011). Gene and QTL detection in a three-way barley cross under selection by a mixed model with kinship information using SNPs. *Theoretical and Applied Genetics*, 122, 1605-1616. DOI: 10.1007/s00122-011-1558-z

Mao, T., Li, J., Wen, Z., Wu, T., Wu, C., Wang, D. and Han, T. (2017). Association mapping of loci controlling genetic and environmental interaction of soybean flowering time under various photo-thermal conditions. *BMC Genomics*, 18, 415. DOI:10.1186/s12864-017-3778-3

McBlain, B. A., and Bernard, R. L. (1987). A new gene affecting the time of flowering and maturity in soybeans. *Journal of Heredity*, 78, 160-162. DOI: 10.1093/oxfordjournals.jhered.a110349

McCarthy, R. L., Zhong, R., and Ye, Z. H. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant and Cell Physiology*, 50, 1950-1964. DOI:10.1093/pcp/pcp139

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20, 1297-1303. DOI:10.1101/gr.107524.110

Mei, M., Deng, H., Lu, Y., Zhuang, C., Liu, Z., Qiu, Q., Qiu, Y., and Yang, T. C. (1994). Mutagenic effects of heavy ion radiation in plants. *Advances in Space Research*, 14, 363-372. DOI: 10.1016/0273-1177(94)90489-8

Mikuriya, S., Kasai, M., Nakashima, K., Natasia, Hase, Y., Yamada, T., Abe, J., and Kanazawa, A. (2017). Frequent generation of mutants with coincidental changes in

multiple traits via ion-beam irradiation in soybean. *Genes and Genetic Systems*, 92, 153-161. DOI:10.1266/ggs.16-00034

Mizoguchi, T., Wright, L., Fujiwara, S., Cremer, F., Lee, K., Onouchi, H., Mouradov, A., Fowler, S., Kamada, H., Putterill, J., and Coupland, G. (2005). Distinct roles of *GIGANTEA* in promoting flowering and regulating circadian rhythms in *Arabidopsis*. *The Plant Cell*, 17, 2255-2270. DOI: 10.1105/tpc.105.033464

Molnar, S. J., Rai, S., Charette, M., and Cober, E. R. (2003). Simple sequence repeat (SSR) markers linked to *E1*, *E3*, *E4*, and *E7* maturity genes in soybean. *Genome*, 46, 1024-1036. DOI: 10.1139/g03-079

Mulki, M. A., and von Korff, M. (2016). *CONSTANS* controls floral repression by up-regulating *VERNALIZATION2* (*VRN-H2*) in barley. *Plant Physiology*, 170, 325-337. DOI:10.1104/pp.15.01350

Nemoto, Y., Nonoue, Y., Yano, M., and Izawa, T. (2016). *Hdl*, a *CONSTANS* ortholog in rice, functions as an *Ehd1* repressor through interaction with monocot-specific CCT-domain protein Ghd7. *The Plant Journal*, 86, 221-233. DOI: 10.1111/tpj.13168

Ogiso-Tanaka, E., Shimizu, T., Hajika, M., Kaga, A., and Ishimoto, M. (2019). Highly multiplexed AmpliSeq technology identifies novel variation of flowering time-related genes in soybean (*Glycine max*). *DNA Research*, 26, 243-260. DOI: 10.1093/dnares/dsz005

Pooprompan, P., Wasee, S., Toojinda, T., Abe, J., Chanprame, S., and Srinives, P. (2006). Molecular marker analysis of days to flowering in vegetable soybean (*Glycine max* (L.) Merrill). *Kasetsart Journal*, 40, 573-581.

Rawlings, J. E., Hanway, D. G., and Gardner, C. O. (1958). Variation in Quantitative Characters of Soybeans After Seed Irradiation. *Agronomy Journal*, 50, 524-528. DOI: 10.2134/agronj1958.00021962005000090009x

Ray, J. D., Hinson, K., Mankono, J., and Malo, M. F. (1995). Genetic control of a long-juvenile trait in soybean. *Crop Science*, 35, 1001-1006. DOI:10.2135/cropsci1995.0011183X003500040012x

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., and Getz, G. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29, 24-26. DOI: 10.1038/nbt.1754

Rosenzweig, V. E., Aksyonova, E. A., Milash, S. B., Goloenko, D. V., and Davydenko, O. G. (2008). Prospects of exploiting of photoperiod sensitivity gene *E7* in early soybean breeding and revealing of its sources with SSR-markers. *Soybean Genetics Newsletter*, 35, 1-7.

Saindon, G., Beversdorf, W. D., and Voldeng, H. D. (1989). Adjustment of the soybean phenology using the *E4* locus. *Crop Science*, 29, 1361-1365. DOI: 10.2135/cropsci1989.0011183X002900060006x

Samach, A., Onouchi, H., Gold, S. E., Ditta, G. S., Schwarz-Sommer, Z., Yanofsky, M. F., and Coupland, G. (2000). Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. *Science*, 288, 1613-1616. DOI: 10.1126/science.288.5471.1613

Samanfar, B., Molnar, S. J., Charette, M., Schoenrock, A., Dehne, F., Golshani, A., Belzile, F., and Cober, E. R. (2017). Mapping and identification of a potential candidate gene for a novel maturity locus, *E10*, in soybean. *Theoretical and Applied Genetics*, 130, 377-390. DOI: 10.1007/s00122-016-2819-7

Sanchez-Bermejo, E., Zhu, W., Tasset, C., Eimer, H., Sureshkumar, S., Singh, R., Sundaramoorthi, V., Colling, L., and Balasubramanian, S. (2015). Genetic architecture of natural variation in thermal responses of *Arabidopsis*. *Plant Physiology*, 169, 647-659. DOI: 10.1104/pp.15.00942

Sawa, M., Nusinow, D. A., Kay, S. A., and Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in *Arabidopsis*. *Science*, 318, 261-265. DOI: 10.1126/science.1146994

Sayama, T., Hwang, T., Komatsu, K., Takada, Y., Takahashi, M., Kato, S., Sasama, H., Higashi, A., Nakamoto, Y., Funatsuki, H., and Ishimoto, M. (2011). Development and application of a whole-genome simple sequence repeat panel for high-throughput genotyping in soybean. *DNA Research*, 18, 107-115. DOI: 10.1093/dnares/dsr003

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., Song, Q., Thelen, J. J., Cheng, J., Xu, D., Hellsten, U., May, G. D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M. K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X., Shinozaki, K., Nguyen, H. T., Wing, R. A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R. C., and Jackson, S. A. (2010). Genome sequence of the palaeopolyploid soybean. *Nature*, 463, 178. DOI:10.1038/nature08670

Sebastian, S. A., Fader, G. M., Ulrich, J. F., Forney, D. R., and Chaleff, R. S. (1989). Semidominant soybean mutation for resistance to sulfonylurea herbicides. *Crop Science*, 29, 1403-1408. DOI: 10.2135/cropsci1989.0011183X002900060014x

Shikazono, N., Suzuki, C., Kitamura, S., Watanabe, H., Tano, S., and Tanaka, A. (2005). Analysis of mutations induced by carbon ions in *Arabidopsis thaliana*. *Journal of Experimental Botany*, 56, 587-596. DOI: 10.1093/jxb/eri047

Shrestha, R., Gómez-Ariza, J., Brambilla, V., and Fornara, F. (2014). Molecular control of seasonal flowering in rice, *Arabidopsis* and temperate cereals. *Annals of Botany*, 114, 1445-1458. DOI: 10.1093/aob/mcu032

Song, Q., Jia, G., Zhu, Y., Grant, D., Nelson, R. T., Hwang, E. Y., Hyten, D. L., and Cregan, P. B. (2010). Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. *Crop Science*, 50, 1950-1960. DOI:10.2135/cropsci2009.10.0607

Song, Y. H., Smith, R. W., To, B. J., Millar, A. J., and Imaizumi, T. (2012a). FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science*, 336, 1045-1049. DOI: 10.1126/science.1219644

Song, Y., Gao, Z., and Luan, W. (2012b). Interaction between temperature and photoperiod in regulation of flowering time in rice. *Science China Life Sciences*, 55, 241-249. DOI:10.1007/s11427-012-4300-4

Song, Y. H., Shim, J. S., Kinmonth-Schultz, H. A., and Imaizumi, T. (2015). Photoperiodic flowering: time measurement mechanisms in leaves. *Annual Review of Plant Biology*, 66, 441-464. DOI: 10.1146/annurev-arplant-043014-115555

Sun, F., Xu, M., Park, C., Dwiyanti, MS., Nagano, A. J., Zhu, J., Watanabe, S., Kong, F., Liu, B., Yamada, T., and Abe, J. (2019) Characterization and quantitative trait locus mapping of late-flowering from a Thai soybean cultivar in a photoperiod-insensitive genetic background. *PLoS One*, 14: e0226116. DOI: 10.1371/journal.pone.0226116

Takagi, Y., ABM, M., Yanagita, T., Matsueda, T., and Murayama, A. (1990). Linolenic acid content in soybean improved by X-ray irradiation. *Agricultural and Biological Chemistry*, 54, 1735-1738. DOI: 10.1271/bbb1961.54.1735

Takahashi, K., Banba, H., Kikuchi, A., Ito, M., and Nakamura, S. (1994). An Induced Mutant Line Lacking the α -subunit of β -conglycinin in Soybean [*Glycine max* (L.) Merrill]. *Japanese Journal of Breeding*, 44, 65-66. DOI: 10.1270/jsbbs1951.44.65

Takeshima, R., Hayashi, T., Zhu, J., Zhao, C., Xu, M., Yamaguchi, N., Sayama, T., Ishimoto, M., Kong, L., Shi, X., Liu, B., Tina, Z., Yamada, T., and Abe, J. (2016). A soybean quantitative trait locus that promotes flowering under long days is identified as *FT5a*, a *FLOWERING LOCUS T* ortholog. *Journal of Experimental Botany*, 67, 5247-5258. DOI: 10.1093/jxb/erw283

Tanaka, A., Shikazono, N., and Hase, Y. (2010). Studies on biological effects of ion beams on lethality, molecular nature of mutation, mutation rate, and spectrum of mutation phenotype for mutation breeding in higher plants. *Journal of Radiation Research*, 51, 223-233. DOI: 10.1269/jrr.09143

Thomas, J. M., Boote, K. J., Pan, D., and Allen, L. H. (2010). Elevated temperature delays onset of reproductive growth and reduces seed growth rate of soybean. *Journal of Agro Crop Science*, 1, 19-32.

Tsubokura, Y., Watanabe, S., Xia, Z., Kanamori, H., Yamagata, H., Kaga, A., Katayose, Y., Abe, J., Ishimoto, M., and Harada, K. (2014). Natural variation in the genes responsible for maturity loci *E1*, *E2*, *E3* and *E4* in soybean. *Annals of Botany*, 113, 429-441. DOI: 10.1093/aob/mct269

Tsuji, H., Taoka, K. I., and Shimamoto, K. (2011). Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. *Current Opinion in Plant Biology*, 14, 45-52. DOI: 10.1016/j.pbi.2010.08.016

Upadhyay, A. P., Ellis, R. H., Summerfield, R. J., Roberts, E. H., and Qi, A. (1994). Characterization of photothermal flowering responses in maturity isolines of soybean [*Glycine max* (L.) Merrill] cv. Clark. *Annals of Botany*, 74, 87-96. DOI: 10.1093/aob/74.1.87

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science*, 303, 1003-1006. DOI: 10.1126/science.1091761

Valliyodan, B., Qiu, D., Patil, G., Zeng, P., Huang, J., Dai, L., Chen, C., Li, Y., Joshi, T., Song, L., Vuong, T. D., Musket, T. A., Xu, D., Shannon, J. G., Chen, S., Liu, X., and Nguyen, H. T. (2016). Landscape of genomic diversity and trait discovery in soybean. *Scientific Reports*, 6, 23598. DOI: 10.1038/srep23598

Verbyla, A. P., Cavanagh, C. R., and Verbyla, K. L. (2014). Whole-genome analysis of multienvironment or multitrait QTL in MAGIC. *G3: Genes, Genomes, Genetics*, 4, 1569-1584. DOI: 10.1007/s00122-014-2337-4

Wang, Y., Gu, Y., Gao, H., Qiu, L., Chang, R., Chen, S., and He, C. (2016a). Molecular and geographic evolutionary support for the essential role of *GIGANTEA* in soybean domestication of flowering time. *BMC Evolutionary Biology*, 16, 79. DOI: 10.1186/s12862-016-0653-9

Wang, T., Sun, M. Y., Wang, X. S., Li, W. B., and Li, Y. G. (2016b). Over-expression of *GmGla*-regulated soybean *miR172a* confers early flowering in transgenic *Arabidopsis*

thaliana. *International journal of molecular sciences*, 17, 645. DOI: 10.3390/ijms17050645

Wang, F., Nan, H., Chen, L., Fang, C., Zhang, H., Su, T., Li, S., Cheng, Q., Dong, L., Liu, B., Kong, F., and Lu, S. (2019). A new dominant locus, *E11*, controls early flowering time and maturity in soybean. *Molecular Breeding*, 39, 70. DOI: 10.1007/s11032-019-0978-3

Wang, L., Sun, S., Wu, T., Liu, L., Sun, X., Cai, Y., Li, J., Jia, H., Yuan, S., Chen, L., Jiang, B., Wu, C., Hou, W., and Han, T. (2020). Natural variation and CRISPR/Cas9-mediated mutation in *GmPRR37* affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnology Journal*, 1-13. DOI: 10.1111/pbi.13346

Watanabe, S., Tajuddin, T., Yamanaka, N., Hayashi, M., and Harada, K. (2004). Analysis of QTLs for reproductive development and seed quality traits in soybean using recombinant inbred lines. *Breeding Science*, 54, 399-407. 10.1270/jsbbs.54.399

Watanabe, S., Hideshima, R., Xia, Z., Tsubokura, Y., Sato, S., Nakamoto, Y., Yamanaka, N., Takahashi, R., Ishimoto, M., Anai, T., Tabata, S., and Harada, K. (2009). Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics*, 182, 1251-1262. DOI: 10.1534/genetics.108.098772

Watanabe, S., Xia, Z., Hideshima, R., Tsubokura, Y., Sato, S., Yamanaka, N., Takahashi, R., Anai, T., Tabata, S., Kitamura, K., and Harada, K. (2011). A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics*, 188, 395-407. DOI: 10.1534/genetics.110.125062

Watanabe, S., Harada, K., and Abe, J. (2012). Genetic and molecular bases of photoperiod responses of flowering in soybean. *Breeding Science*, 61, 531-543. DOI: 10.1270/jsbbs.61.531

Weng, J., Xie, C., Hao, Z., Wang, J., Liu, C., Li, M., Zhang, D., Bai, L., Zhang, S., and LI, X. (2011). Genome-wide association study identifies candidate genes that affect plant height in Chinese elite maize (*Zea mays* L.) inbred lines. *PLoS One*, 6, e29229. DOI: 10.1371/journal.pone.0029229

Weng, X., Wang, L., Wang, J., Hu, Y., Du, H., Xu, C., Xiong, Y., LI, X., Xiao, J., and Zhang, Q. (2014). *Grain number, plant height, and heading date7* is a central regulator of growth, development, and stress response. *Plant Physiology*, 164, 735-747. DOI: 10.1104/pp.113.231308

Wigge, P. A. (2011). *FT*, a mobile developmental signal in plants. *Current Biology*, 21, R374-R378. DOI: 10.1016/j.cub.2011.03.038

Wigge, P. A. (2013). Ambient temperature signaling in plants. *Current Opinion in Plant Biology*, 16, 661-666. DOI: 10.1016/j.pbi.2013.08.004

Williams, J. H., and Hanway, D. G. (1961). Genetic variation in oil and protein content of soybeans induced by seed irradiation. *Crop Science*, 1, 34-36. DOI: 10.2135/cropsci1961.0011183X000100010011x

Wu, F., and Hanzawa, Y. (2014). Photoperiodic control of flowering in plants. In *Handbook of Plant and Crop Physiology*, CRC Press, 150-167.

Wu, F., Price, B. W., Haider, W., Seufferheld, G., Nelson, R., and Hanzawa, Y. (2014). Functional and evolutionary characterization of the *CONSTANS* gene family in short-

day photoperiodic flowering in soybean. *PLoS One*, 9, e85754. DOI: 10.1371/journal.pone.0085754

Wu, F., Kang, X. K., Wang, M., Haider, W., Price, W. B., Hajek, B., and Hanzawa, Y. (2019). Transcriptome-enabled network inference revealed the *GmCOL1* feed-forward loop and its roles in photoperiodic flowering of soybean. *Frontiers in Plant Science*, 10, 1221. DOI: 10.3389/fpls.2019.01221

Xia, Z., Watanabe, S., Yamada, T., Tsubokura, Y., Nakashima, H., and Zhai, H., Anai, T., Sato, S., Yamazaki, T., Lü, S., Wu, H., Tabata, S., and Harada, K. (2012). Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proceedings of the National Academy of Sciences*, USA 109, E2155-E2164. DOI: 10.1073/pnas.1117982109

Xu, M., Xu, Z., Liu, B., Kong, F., Tsubokura, Y., Watanabe, S., Xia, Z., Harada, K., Kanazawa, A., Yamada, T., and Abe, J. (2013). Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. *BMC Plant Biology*, 13, 91. DOI:10.1186/1471-2229-13-91

Xu, M., Yamagishi, N., Zhao, C., Takeshima, R., Kasai, M., Watanabe, S., Kanazawa, A., Yoshikawa, N., Liu, B., Yamada, T., and Abe, J. (2015). The soybean-specific maturity gene *E1* family of floral repressors controls night-break responses through down-regulation of *FLOWERING LOCUS T* orthologs. *Plant Physiology*, 168, 1735-1746. DOI: 10.1104/pp.15.00763

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., Zhou, H., Yu, S., Xu, C., LI, X., and Zhang, Q. (2008). Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nature Genetics*, 40, 761. DOI: 10.1038/ng.143

Yamaguchi, N., Sayama, T., Sasama, H., Yamazaki, H., Miyoshi, T., Tanaka, Y., and Ishimoto, M. (2014). Mapping of quantitative trait loci associated with terminal raceme length in soybean. *Crop Science*, 54, 2461-2468. DOI: 10.2135/cropsci2014.03.0226

Yamanaka, N., Watanabe, S., Toda, K., Hayashi, M., Fuchigami, H., Takahashi, R., and Harada, K. (2005). Fine mapping of the *FT1* locus for soybean flowering time using a residual heterozygous line derived from a recombinant inbred line. *Theoretical and Applied Genetics*, 110, 634-639. DOI: 10.1007/s00122-004-1886-3

Yang, G., Zhai, H., Wu, H., Zhang, X., Lu, S., and Wang, Y., Li, Y., Hu, B., Wang, L., Wen, Z., Wang, D., Wang, S., Harada, K., Xia, Z., and Xie, F. (2017). QTL effects and epistatic interaction for flowering time and branch number in a soybean mapping population of Japanese x Chinese cultivars. *Journal of Integrative Agriculture*, 16, 1900-1912. DOI: 10.1016/S2095-3119(16)61539-3

Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., Baba, T., Yamamoto, K., Umehara, Y., Nagamura, Y., Sasaki, T. (2000). *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *The Plant Cell*, 12, 2473-2483. DOI: 10.1105/tpc.12.12.2473

Yue, Y., Liu, N., Jiang, B., Li, M., Wang, H., Jiang, Z., Pan, H., Xia, Q., Ma, Q., Han, T., and Nian, H. (2017). A single nucleotide deletion in *J* encoding GmELF3 confers long juvenility and is associated with adaption of tropic soybean. *Molecular Plant*, 10, 656-658. DOI: 10.1016/j.molp.2016.12.004

Zatybekov, A., Abugalieva, S., Didorenko, S., Gerasimova, Y., Sidorik, I., Anuarbek, S., and Turuspekoy, Y. (2017). GWAS of agronomic traits in soybean collection included in breeding pool in Kazakhstan. *BMC Plant Biology*, 17, 179. DOI: 10.1186/s12870-017-1125-0

Zhang, Y., Zhao, L., Li, H., Gao, Y., Li, Y., Wu, X., Teng, W., Han, Y., Zhao, X., and Li, W. (2013). *GmGBPI*, a homolog of human ski interacting protein in soybean, regulates flowering and stress tolerance in *Arabidopsis*. *BMC Plant Biology*, 13, 21. DOI: 10.1186/1471-2229-13-21

Zhang, J., Song, Q., Cregan, P. B., Nelson, R. L., Wang, X., Wu, J., and Jiang, G. L. (2015). Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics*, 16, 217. DOI: 10.1186/s12864-015-1441-4

Zhang, Z., Hu, W., Shen, G., Liu, H., Hu, Y., Zhou, X., Liu, T., and Xing, Y. (2017). Alternative functions of Hd1 in repressing or promoting heading are determined by Ghd7 status under long-day conditions. *Scientific Reports*, 7, 1-11. DOI: 10.2135/cropsci1971.0011183X001100020022x

Zhang, J., Li, X., Lin, H., and Chong, K. (2019). Crop improvement through temperature resilience. *Annual Review of Plant Biology*, 70, 753-780. DOI: 10.1146/annurev-arplant-050718-100016

Zhao, J., Bradfield, J. P., Zhang, H., Sleiman, P. M., Kim, C. E., Glessner, J. T., Deliard, S., Thomas, K. A., Fracklton, E. C., Li, M., Chiavacci, R. M., Berkowitz, R. I., Hakonarson, H., and Grant, S. F. A. (2011). Role of BMI-associated loci identified in GWAS meta-analyses in the context of common childhood obesity in European Americans. *Obesity*, 19, 2436-2439. DOI: 10.1038/oby.2011.237

Zhao, L., Wang, Z., Lu, Q., Wang, P., Li, Y., Lü, Q., Song, X., Li, D., Gu, Y., Liu, L., and Li, W. (2013). Overexpression of a *GmGBPI* ortholog of soybean enhances the responses to flowering, stem elongation and heat tolerance in transgenic tobaccos. *Plant Molecular Biology*, 82, 279-299. DOI: 10.1007/s11103-013-0062-z

Zhao, X. (2015) The function of *miR172* and its targets in the regulation of photoperiod-mediated flowering in soybean. PhD thesis, University of Chinese Academy of Sciences.

Zhao, X., Cao, D., Huang, Z., Wang, J., Lu, S., Xu, Y., Liu, B., Kong, F., and Yuan, X. (2015). Dual functions of *GmTOE4a* in the regulation of photoperiod-mediated flowering and plant morphology in soybean. *Plant Molecular Biology*, 88, 343-355. DOI: 10.1007/s11103-015-0322-1

Zhao, C., Takeshima, R., Zhu, J., Xu, M., Sato, M., Watanabe, S., Kanazawa, A., Liu, B., Kong, F., Yamada, T., and Abe, J. (2016). A recessive allele for delayed flowering at the soybean maturity locus *E9* is a leaky allele of *FT2a*, a *FLOWERING LOCUS T* ortholog. *BMC Plant Biology*, 16, 20. DOI: 10.1186/s12870-016-0704-9

Zhong, R., and Ye, Z. H. (2011). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant and Cell Physiology*, 53, 368-380. DOI: 10.1093/pcp/pcr185

Zhu, Q. H., and Helliwell, C. A. (2011). Regulation of flowering time and floral patterning by *miR172*. *Journal of Experimental Botany*, 62, 487-495. DOI: 10.1093/jxb/erq295

Zhu, J., Takeshima, R., Harigai, K., Xu, M., Kong, F., Liu, B., Kanazawa, A., Yamada, T., and Abe, J. (2019). Loss of function of the *El-Like-b* gene associates with early flowering under long-day conditions in soybean. *Frontiers in Plant Science*, 9, 1867. DOI: 10.3389/fpls.2018.01867

Appendix

Supplemental Table S1. SSR markers used in the association test.

No. of markers	Molecular linkage group	Chr	SSR marker							
1	D1a	6	Sat_413	Satt184	CSSR390	Sat_353	Satt370	Satt436	Gm01.SSR3.1	Sat_160
2	D1b	6	Sat_227	Gm02.SSR1.2	Satt634	Gm02.SSR3.2	Satt141	Satt703	Satt459	Sat_415
3	N	5	Sct_195	Sat_186	Sat_208	Sat_033	Satt521	Satt237	Sat_306	CSSR126
4	C1	8	Satt565	Sat_337	GMES0780	AW277661	Sat_311	Gm04.SSR2.2	AI794821	Gm04.SSR3.3
5	A1	4	Sat_137	Sat_344	Gm05.SSR1.2	Sat_384	Satt717	Satt545	Satt211	Satt599
6	C2	7	AW734043	Sat_130	Satt281	Sat_153	Gm06.SSR1.1	Satt277	Satt307	Satt371
7	M	4	Sat_391	Satt150	Satt540	Satt536	Sct_147	Satt551	Satt346	Satt336
8	A2	8	Satt589	Sat_162	Satt424	Sat_115	Satt233	Satt329	Satt409	Satt378
9	K	3	CSSR438	GMES5734	Satt178	Satt417	Satt559	Satt499	Sat_352	Satt588
10	O	3	Gm10.SSR1.1	Satt653	Satt345	Satt478	Satt592	Sat_038	Satt153	CSSR141

11	B1	4	BE806308	Sat_272	Satt426	Satt197	Satt519	Satt583	Satt359	CSSR18
12	H	8	Satt666	GMES0816	Satt568	GMES4012	Sat_206	Satt302	Sat_218	Satt442
13	F	7	Satt348	Sat_390	CSSR535	Satt663	Satt510	Sat_197	Satt657	Sat_417
14	B2	5	Satt577	Sat_342	Satt467	Satt416	Satt318	Sat_424	Satt560	Gm14.SSR1.2
15	E	4	Satt213	Satt411	Satt212	CSSR59	Sat_124	Sat_380	CSSR63	Sat_381
16	J	4	Satt249	Satt285	GMES5027	Satt414	Sat_361	Satt011	Satt244	Sat_144
17	D2	5	Sat_333	Sct_192	Sat_284	CSSR172	Sat_292	Satt488	GMES4770	Satt386
18	G	8	Satt163	Sat_163	Sat_315	Satt394	Satt199	Satt288	Satt472	Sat_372
19	L	6	Sat_408	Sat_301	Sat_405	Sat_150	Satt156	Sat_113	Sat_184	CSSR548
20	I	6	GMES1405	Satt419	Sat_219	Satt354	Gm20.SSR1.2	Satt292	Sat_420	CSSR366

Red, polymorphic markers used in the association test.

Primer information; Sayama et al. (2011)

Supplemental Table S2. Primers for SSR markers used in fine mapping

SSR marker	Primer sequence (5' - 3')
BARCSOYSSR_19_0230	F AAAATTTTAAAAAGTAGCACGGTTA
	R TGGGAGATTCTTTTTCTACCTGA
BARCSOYSSR_19_0240	F TGAACCATTGGAAAAGGGAG
	R CGATGGCATTGCAGTCTTTA
BARCSOYSSR_19_0279	F CTCTCGTAGGAGGGACCAAG
	R TCAACCGCAAATATTTTATGTCA
BARCSOYSSR_19_0304	F GCGTTTTGCAATTCAAGTCA
	R GAGTTGACCCAACAGTAAGTGAAA
BARCSOYSSR_19_0310	F AATCAATTGGGGAACAGCAT
	R CACTTAGGTGGTATTTGCCCA
BARCSOYSSR_19_0330	F AGAAAATGCTGCTTTGCTAGA
	R CGGAAGAACTTCCCATTTC
BARCSOYSSR_19_0318	F CCCCATGAGTATGAAAGGAAA
	R GGTTTGTGAGGAGTTGTTGC
BARCSOYSSR_19_0393	F CGAATTGAAATAGGCTGAATGA
	R GCAGATATGCTTTTGATCCCA

Supplemental Table S3. Primers used in qRT-PCR

Gene		Primer sequence (5' - 3')	Accession/Glyma number	Reference
<i>18s</i>	F	TGATTAACAGGGACAGTCGG	X02623	Khan et al. (2009) ¹⁾
<i>rRNA</i>	R	ACGGTATCTGATCGTCTTCG		
<i>FT2a</i>	F	GGATTGCCAGTTGCTGCTGT	Glyma.16g150700	Kong et al. (2010) ²⁾
	R	GAGTGTGGGAGATTGCCAAT		
<i>FT5a</i>	F	GATTCTTGGATGCCTTACTCC	Glyma.16g044100	Kong et al. (2010) ²⁾
	R	GGCATGCTCTAGCATTGCAA		
<i>COL2a</i>	F	GGAACCAGAGCATGAACGCGTGTA	Glyma.13g050300	This study
	R	AAACCCTTATCACCACCATTCTCGTC		
<i>COL2b</i>	F	GAGTTGTACCCGAATCAACAGTAAGTGG	Glyma.19g039000	This study
	R	GAGGCATATCTTATTTTCTTCTCGAAC		
<i>COL1a</i>	F	CAGATGTTCTCGACAACACT	Glyma.18g278100	This study
	R	CGAGATATGATTGACTTGAC		
<i>COL1b</i>	F	CATCAAAACACCACTCTGAC	Glyma.08g255200	This study
	R	GCGTCAGCCTTGCAGAGGAA		
<i>E1</i>	F	CACTCAAATTAAGCCCTTTCA	Glyma.06g207800	Xia et al. (2012) ³⁾
	R	TTCATCTCCTCTTCATTTTTGTTG		
<i>E1Lb</i>	F	GTGTAAACACTCAAAGTCCTT	Glyma.04g143300	Xu et al. (2015) ⁴⁾
	R	CTCCTCTTCATTTTTGTTGCTGC		
<i>E2</i>	F	GTTGCAACCCCACTACAGCCT	Glyma.10g221500	This study
<i>(Gla)</i>	R	GAGGCAGAGCCAAAGCCTATG		
<i>Glb</i>	F	GTTGCAACCCCGCTACAACCA	Glyma.20g170000	This study
	R	TCGAGGCAGAGCCAATGCCA		

1) Khan et al. (2009). Peptides, 30, 2095-2102.

2) Kong et al. (2010). Plant Physiology, 154, 1220-1231.

3) Xia et al. (2012). Proc. Natl. Acad. Sci. U.S.A., 10, :E2155-2164.

4) Xu et al. (2015). Plant Physiology, 168, 1735-1746.

Supplemental Table S4. The list of 512 accessions used in SNP calling for COL genes and their haplotypes.

Sequence ID	Data source	PI number	Description	Common Name	Origin/Country	COL1a	COL1b	COL2a	COL2b
HN001	#1	PI518664	cultivated	Hutcheson	U.S.	-	-	-	-
HN002	#1	PI548402	cultivated	Peking	China	-	-	-	-
HN003	#1	PI89772	cultivated	7193	China	-	-	-	-
HN004	#1	PI90763	cultivated	7570	China	-	-	-	-
HN005	#1	PI404166	cultivated	Krasnoarmejskaja	China	-	-	-	-
HN006	#1	PI407788A	cultivated	ORD 8113	Korea, South	Haplotype4	-	-	-
HN007	#1	PI424298	cultivated	KAS 300-10	Korea, South	Haplotype4	-	-	-
HN008	#1	PI437655	cultivated	Er-huan-jan	China	-	-	-	-
HN009	#1	PI495017C	cultivated	(Beijing da qing don)	China	Haplotype4	-	-	-
HN010	#1	PI468915	cultivated		China	-	-	-	-
HN011	#1	PI507354	cultivated	Tokei 421	Japan	-	-	-	-
HN012	#1	PI567305	cultivated	Hei dou zi	China	-	-	-	-
HN013	#1	S05-11482	cultivated		U.S.	Haplotype4	-	-	-
HN014	#1	PI548667	cultivated	Essex	U.S.	Haplotype4	-	-	-
HN015	#1	PI437654	cultivated	Er-hej-jan	China	-	-	-	-
HN016	#1	PI567387	cultivated	Huang huai dou	China	-	-	-	-

HN017B	#1	PI437725	cultivated	Te-zu-gan	China	-	-	-	-
HN018	#1	PI437690	cultivated	Pin-din-guan	China	-	-	-	-
HN020	#1	PI88788	cultivated	5913	China	Haplotype4	-	-	-
HN021	#1	PI209332	cultivated	No. 4	Japan	-	-	-	-
HN022	#1	PI404198B	cultivated	(Sun huan do)	China	-	-	-	-
HN023	#1	PI424608A	cultivated	KAS 681-21	Korea, South	Haplotype4	-	-	-
HN024	#1	PI548316	cultivated	Cloud	China	Haplotype4	-	-	-
HN025	#1	PI567516C	cultivated	(Ba yue zha)	China	Haplotype4	-	-	-
HN026	#1	PI612611	cultivated	Browngilgun	Korea, North	Haplotype4	-	-	-
HN027	#1	S10-11227	cultivated		U.S.	-	-	-	-
HN028	#1	PI572239	cultivated	Holladay	U.S.	Haplotype4	-	-	-
HN029	#1	IA3023	cultivated		U.S.	Haplotype4	-	-	-
HN030	#1	PI598124	cultivated	Maverick	U.S.	Haplotype4	-	-	-
HN031	#1	PI79691-4	cultivated		China	-	-	-	-
HN032	#1	PI86006	cultivated	Kiio Shokuzu	Japan	Haplotype4	-	-	-
HN033	#1	PI87617	cultivated	Miyongaikon	Korea, North	Haplotype4	-	-	-
HN034	#1	PI87631-1	cultivated	Kindaizu	Japan	-	-	-	-
HN035	#1	PI196175	cultivated	Yu tae	Korea, South	Haplotype4	-	-	-
HN037	#1	PI200508	cultivated	Natsu Daizu	Japan	-	-	-	-
HN038	#1	PI248515	cultivated	White Hilum Iwata Variety No. 2	Japan	Haplotype4	-	-	-
HN039	#1	PI366121	cultivated		Japan	Haplotype4	-	-	-
HN040	#1	PI378702	cultivated		Japan	Haplotype4	-	-	-

HN041	#1	PI398593	cultivated	KAS 390-4	Korea, South	Haplotype4	-	-	-
HN042	#1	PI398595	cultivated	KAS 390-5	Korea, South	-	-	-	-
HN043	#1	PI398610	cultivated	KAS 390-8	Korea, South	Haplotype4	-	-	-
HN044	#1	PI398614	cultivated	KAS 390-9	Korea, South	Haplotype4	-	-	-
HN045	#1	PI407162	cultivated	K1-D	Korea, South	Haplotype4	-	-	-
HN046	#1	PI407184	cultivated	K29	Korea, South	Haplotype4	-	-	-
HN047	#1	PI407729	cultivated		China	Haplotype4	-	-	-
HN048	#1	PI407965	cultivated	KAERI 504-4	Korea, South	Haplotype4	-	-	-
HN049	#1	PI408105A	cultivated	KAS 633-19	Korea, South	Haplotype4	-	-	-
HN050	#1	PI416937	cultivated	Houjaku Kuwazu	Japan	Haplotype4	-	-	-
HN051	#1	PI424078	cultivated	74077	Korea, South	Haplotype4	-	-	-
HN052	#1	PI424079	cultivated	74079	Korea, South	Haplotype4	-	-	-
HN053	#1	PI424088	cultivated	74088	Korea, South	Haplotype4	-	-	-
HN054	#1	PI437169B	cultivated	(VNIISC-4)	Russian Federation	Haplotype4	-	-	-
HN055	#1	PI437679	cultivated	Nan-cou	China	Haplotype4	-	-	-
HN056	#1	PI437863A	cultivated	DV-2841	China	Haplotype4	-	-	-
HN057	#1	PI438258	cultivated	VIR 4714	China	Haplotype4	-	-	-
HN058	#1	PI458515	cultivated	Tie Zhugan	China	-	-	-	-
HN059	#1	PI464920B	cultivated	(Jin Dou 33)	China	-	-	-	-
HN060	#1	PI467312	cultivated	Cha-mo-shi-dou	China	-	-	-	-
HN061	#1	PI471938	cultivated	197	Nepal	Haplotype4	-	-	-
HN062	#1	PI475783B	cultivated	(Tsing 2)	China	-	-	-	-
HN063	#1	PI483463	cultivated		China	-	-	-	-

HN064	#1	PI518751	cultivated	NS-20	Former Serbia and Montenegro	Haplotype4	-	-	-
HN065	#1	PI542044	cultivated	Kunitz	U.S.	-	-	-	-
HN066	#1	PI547862	cultivated	L83-570	U.S.	-	-	-	-
HN067	#1	PI548317	cultivated	Columbia	China	-	-	-	-
HN068	#1	PI548349	cultivated	Ilsoy	Korea, North	Haplotype4	-	-	-
HN069	#1	PI548415	cultivated	Sooty	China	Haplotype4	-	-	-
HN070	#1	PI548511	cultivated	Beeson 80	U.S.	Haplotype4	-	-	-
HN071	#1	PI548657	cultivated	Jackson	U.S.	-	-	-	-
HN072	#1	PI549031	cultivated		China	Haplotype4	-	-	-
HN073	#1	PI552538	cultivated	Dunbar	U.S.	-	-	-	-
HN074	#1	PI561271	cultivated	Pei xian da quing dou	China	Haplotype4	-	-	-
HN075	#1	PI567230	cultivated	WJK-China-23	China	-	-	-	-
HN076	#1	PI567336B	cultivated	(Lao hei dou)	China	Haplotype4	-	-	-
HN077	#1	PI567343	cultivated	Ma huang dou	China	Haplotype4	-	-	-
HN078	#1	PI567354	cultivated	You huang dou	China	-	-	-	-
HN079	#1	PI567357	cultivated	Du jia qiao huang dou	China	Haplotype4	-	-	-
HN080	#1	PI567383	cultivated	Da ke huang dou	China	-	-	-	-
HN081	#1	PI567519	cultivated	Bai hua chi	China	Haplotype4	-	-	-
HN082	#1	PI567611	cultivated	Ba yue zha	China	Haplotype4	-	-	-

HN083	#1	PI567651	cultivated	Shang cai er cao ping ding shi	China	-	-	-	-
HN084	#1	PI567690	cultivated	Fu yang (7)	China	Haplotype2	-	-	-
HN085	#1	PI567719	cultivated	Fu yang (43)	China	Haplotype2	-	-	-
HN086	#1	PI567731	cultivated	Fu yang (56)	China	Haplotype4	-	-	-
HN087	#1	PI591539	cultivated	L91-8558	U.S.	-	-	-	-
HN088	#1	PI593258	cultivated	Macon	U.S.	-	-	-	-
HN089	#1	PI594012	cultivated	Heuksatangdu	Korea, South	Haplotype4	-	-	-
HN090	#1	PI594512A	cultivated	Bian zi jiang se dou	China	-	-	-	-
HN091	#1	PI594599	cultivated	Chang de chun hei dou	China	Haplotype4	-	-	-
HN092	#1	PI597387	cultivated	Pana	U.S.	Haplotype4	-	-	-
HN093	#1	PI603154	cultivated	GL 2622 /96	Korea, North	Haplotype4	-	-	-
HN094	#1	PI603170	cultivated	GL 2683 /96	Korea, North	Haplotype4	-	-	-
HN095	#1	PI603175	cultivated	GL 2688 /96	Korea, North	Haplotype4	-	-	-
HN096	#1	PI603176A	cultivated		Korea, North	Haplotype4	-	-	-
HN097	#1	PI603497	cultivated	Hua dou	China	Haplotype4	-	-	-
HN098	#1	PI605869A	cultivated	Sample 140	Vietnam	Haplotype4	-	-	-
HN099	#1	PI639740	cultivated	LD00-3309	U.S.	Haplotype4	-	-	-
HN100	#1	PI647086	cultivated	N8001	U.S.	Haplotype4	-	-	-
HN101	#1	PI658519	cultivated	LD00-2817P	U.S.	Haplotype4	-	-	-
HN102	#1	S07-5049	cultivated		U.S.	Haplotype4	-	-	-
HN103	#1	V71-370	cultivated		U.S.	-	-	-	-

HN104	#1	FC31721	cultivated			Haplotype4	-	-	-
HN105	#1	PI438471	cultivated	Fiskeby III	Sweden	Haplotype4	-	-	-
HN106	#1	PI417091	cultivated	Kuro mame	Japan	-	-	-	-
HN107	#1	PI417015	cultivated	Kawanagare (Iwate)	Japan	-	-	-	-
USB-001	#1	PI548656	cultivated	Lee	U.S.	Haplotype4	-	-	-
USB-002	#1	FC33243	cultivated	Anderson		Haplotype4	-	-	-
USB-003	#1	PI81785	cultivated	Chusei Hadaka	Japan	Haplotype4	-	-	-
USB-004	#1	PI166105	cultivated	Bhart	India	Haplotype4	-	-	-
USB-005	#1	PI180501	cultivated	Strain No. 18	Germany	Haplotype4	-	-	-
USB-006	#1	PI240664	cultivated	Bilomi No. 3	Philippines	Haplotype4	-	-	-
USB-007	#1	PI266806C	cultivated	No. 4	China	Haplotype4	-	-	-
USB-009	#1	PI360957	cultivated	Karafuto No. 1	Japan	Haplotype4	-	-	-
USB-010	#1	PI361087	cultivated	Medias 23	Romania	-	-	-	-
USB-011	#1	PI361093	cultivated	Novosadska Br. 1	Serbia	-	-	-	-
USB-012	#1	PI437265D	cultivated	(Dobruzanca D)	Moldova	Haplotype4	-	-	-
USB-013	#1	PI438335	cultivated	SAO 196-C	Algeria	-	-	-	-
USB-014	#1	PI468908	cultivated		China	Haplotype4	-	-	Haplotype2
USB-015	#1	PI497964A	cultivated	I.C. 9461	India	Haplotype4	-	-	-
USB-017	#1	PI507180	cultivated	Rikuu 21	Japan	-	-	-	-
USB-018	#1	PI518668	cultivated	TN 4-86	U.S.	-	-	-	-
USB-020	#1	PI603458A	cultivated	Shui dou	China	-	-	-	-
USB-022	#1	PI548193	cultivated	T201	U.S.	-	-	-	-
USB-024	#1	PI548325	cultivated	Flambeau	Russian Federation	Haplotype4	-	-	-

USB-025	#1	PI548360	cultivated	Korean	Korea, North	-	-	-	-
USB-026	#1	PI548364	cultivated	Macoupin	United States	-	-	-	-
USB-028	#1	PI548447	cultivated	Cherokee	China	-	-	-	-
USB-029	#1	PI548452	cultivated	Dixie	Korea, North	Haplotype4	-	-	-
USB-031	#1	PI548490	cultivated	Tanner	Taiwan	-	-	-	-
USB-032	#1	PI548520	cultivated	Preston	U.S.	-	-	-	-
USB-033	#1	PI548561	cultivated	Hodgson	U.S.	-	-	-	-
USB-035	#1	PI549018	cultivated	ZYD 3939	China	Haplotype4	Haplotype5	-	-
USB-036	#1	PI549026	cultivated	Gao li huang	China	Haplotype2	-	-	-
USB-037	#1	PI549041A	cultivated	ZYD 2709	China	-	-	-	-
USB-038	#1	PI559932	cultivated	Manokin	U.S.	-	-	-	-
USB-039	#1	PI567171	cultivated	Hei he No. 1	China	Haplotype4	-	-	-
USB-040	#1	PI567426	cultivated	Bai huang dou	China	-	-	-	-
USB-041	#1	PI567558	cultivated	Liu shi ri jin huang da dou	China	Haplotype4	-	-	-
USB-042	#1	PI567604A	cultivated	Xin huang dou	China	-	-	-	-
USB-043	#1	PI578309	cultivated	I-64	Nepal	-	-	-	-
USB-044	#1	PI578495	cultivated	Jin dou No. 4	China	-	-	-	-
USB-045	#1	PI592523	cultivated	Glacier	U.S.	-	-	-	-
USB-046	#1	PI592954	cultivated	Nin zhen No. 1	China	Haplotype2	-	-	-
USB-047	#1	PI594307	cultivated	Tsurusengoku	Japan	Haplotype4	-	-	-
USB-048	#1	PI597476	cultivated	Deogyukong	South Korea	Haplotype4	-	-	-

USB-049	#1	PI603426G	cultivated	Ben di yuan huang dou	China	-	-	-	-
USB-050	#1	PI603442	cultivated	Ke qi xiao hei dou	China	-	-	-	-
USB-059	#1	FC29333	cultivated	Laredo	Unknown	-	-	-	-
USB-060	#1	FC31697	cultivated		Costa Rica	-	-	-	-
USB-061	#1	PI54591	cultivated	No. 31	China	-	-	-	-
USB-062	#1	PI54608-1	cultivated	No. 48	China	-	-	-	-
USB-063	#1	PI54614	cultivated	No. 54	China	Haplotype4	-	-	-
USB-064	#1	PI54615-1	cultivated	No. 55	China	Haplotype4	-	-	-
USB-065	#1	PI58955	cultivated	Common Yellow Variety	China	Haplotype4	-	-	-
USB-066	#1	PI62203	cultivated	937	China	-	-	-	-
USB-067	#1	PI68521-1	cultivated	205	China	-	-	-	-
USB-068	#1	PI68604-1	cultivated	285	China	-	-	-	-
USB-069	#1	PI68732-1	cultivated	204	China	Haplotype4	-	-	-
USB-070	#1	PI70080	cultivated	6908	China	Haplotype4	-	-	-
USB-071	#1	PI70466-3	cultivated	7336	China	Haplotype4	-	-	-
USB-072	#1	PI71465	cultivated	No. 33	China	-	-	-	-
USB-073	#1	PI80837	cultivated	Mejiro	Japan	-	-	-	-
USB-074	#1	PI81041	cultivated	Kuro Daizu	Japan	Haplotype4	-	-	-
USB-075	#1	PI83881	cultivated	Orukon	Korea, North	Haplotype4	-	-	-
USB-076	#1	PI83942	cultivated	Kuro churyu	Korea, South	-	-	-	-
USB-077	#1	PI84631	cultivated	S-56	Korea, South	-	-	-	-

USB-078	#1	PI84637	cultivated	S-62	Korea, South	Haplotype2	-	-	-
USB-079	#1	PI84656	cultivated	S-81	Korea, South	Haplotype4	-	-	-
USB-080	#1	PI84946-2	cultivated	(Kandokon)	Korea, South	Haplotype4	-	-	-
USB-081	#1	PI84973	cultivated	Takiya	Japan	Haplotype4	-	-	-
USB-084	#1	PI86904	cultivated	Fukota	Korea, South	-	-	-	-
USB-085	#1	PI86972-2	cultivated	(Pakute)	Korea, South	Haplotype4	-	-	-
USB-086	#1	PI87620	cultivated	Kuromeshoryu	Korea, North	Haplotype4	-	-	-
USB-087	#1	PI88313	cultivated	5702	China	-	-	-	-
USB-088	#1	PI88468	cultivated	Iganzu	China	Haplotype4	-	-	-
USB-089	#1	PI89005-5	cultivated	5950	China	-	-	-	-
USB-090	#1	PI89775	cultivated	7221	China	Haplotype4	-	-	-
USB-091	#1	PI90479P	cultivated	7413	China	-	-	-	-
USB-093	#1	PI91100-3	cultivated	6554	China	Haplotype4	-	-	-
USB-094	#1	PI91159-4	cultivated	6614	China	Haplotype4	-	-	-
USB-095	#1	PI91160	cultivated	6615	China	Haplotype4	-	-	-
USB-096	#1	PI92651	cultivated	7846	China	-	-	-	-
USB-097	#1	PI94159-3	cultivated	(Kiizaya)	Japan	-	-	-	-
USB-098	#1	PI95860	cultivated	155	Korea, South	Haplotype4	-	-	-
USB-100	#1	PI103088	cultivated	Ming Chuan	China	Haplotype4	-	-	-
USB-101	#1	PI123440	cultivated	No. 2	Myanmar	-	-	-	-
USB-102	#1	PI153231	cultivated	B-63	Unknown	Haplotype4	-	-	-
USB-104	#1	PI154189	cultivated	No. 57	Netherland	Haplotype4	-	-	-
USB-106	#1	PI159925	cultivated	Glycine H	Peru	Haplotype2	-	-	-

USB-107	#1	PI165563	cultivated	Bhart	India	-	-	-	-
USB-108	#1	PI165675	cultivated	Nanking 332	China	Haplotype4	-	-	-
USB-109	#1	PI171428	cultivated	Large Yellow Soybean	China	-	-	-	-
USB-110	#1	PI171451	cultivated	Kosamame	Japan	-	-	-	-
USB-112	#1	PI179935	cultivated	Bhart	India	Haplotype4	-	-	-
USB-113	#1	PI189873	cultivated	Miko Saumon	France	Haplotype4	-	-	-
USB-115	#1	PI209333	cultivated	No. 3	Japan	Haplotype4	-	-	-
USB-116	#1	PI209334	cultivated	No. 9	Japan	-	-	-	-
USB-117	#1	PI232992	cultivated	Kono-Kuradaizu	Japan	Haplotype4	-	-	-
USB-119	#1	PI253661B	cultivated	No. 12	China	-	-	-	-
USB-120	#1	PI261272C	cultivated	(Tua Nao)	Thailand	Haplotype4	-	-	-
USB-121	#1	PI274453	cultivated		Japan	-	-	-	-
USB-122	#1	PI291294	cultivated		China	Haplotype4	-	-	-
USB-123	#1	PI291309D	cultivated		China	-	-	-	-
USB-124	#1	PI291310C	cultivated		China	Haplotype4	-	-	-
USB-125	#1	PI297505	cultivated	Czi ti No. 5	China	Haplotype4	-	-	-
USB-126	#1	PI297520	cultivated	Iregi Universal	Hungary	Haplotype4	-	-	-
USB-127	#1	PI324924	cultivated	Rhosa	South Africa	Haplotype4	-	-	-
USB-128	#1	PI342434	cultivated		Japan	Haplotype2	-	-	Haplotype3
USB-129	#1	PI342619A	cultivated		Russian Federation	Haplotype4	-	-	-
USB-131	#1	PI361066B	cultivated	(F. 56-17)	Romania	Haplotype4	-	-	-
USB-132	#1	PI361070	cultivated	Faur	Romania	Haplotype4	-	-	-

USB-133	#1	PI361080	cultivated	Kormovaia 15	Russian Federation	-	-	-	-
USB-134	#1	PI372403B	cultivated	(Caloria)	Austria	Haplotype4	-	-	-
USB-135	#1	PI372418	cultivated	Novosadska Br. 4	Serbia	Haplotype4	-	-	-
USB-136	#1	PI374207	cultivated	N-9	India	-	-	-	-
USB-139	#1	PI378658	cultivated	Dnepropetrovsk 12	Ukraine	-	-	-	-
USB-140	#1	PI378663	cultivated	Habarovskaja II	Russian Federation	-	-	-	-
USB-141	#1	PI378680E	cultivated	(VNIIMK 9186)	Russian Federation	Haplotype4	-	-	-
USB-143	#1	PI379618	cultivated	TC 1	Taiwan	Haplotype4	-	-	-
USB-144	#1	PI391577	cultivated	Ch'a ye sheng tou	China	-	-	-	-
USB-145	#1	PI391583	cultivated	Jilin No. 10	China	Haplotype4	-	-	-
USB-146	#1	PI398633	cultivated	KAS 390-17-2	Korea, South	Haplotype4	-	-	-
USB-147	#1	PI398965	cultivated	KLS 628-1	Korea, South	Haplotype4	-	-	-
USB-148	#1	PI404187	cultivated	Suj nii hun mao ju	China	Haplotype4	-	-	-
USB-149	#1	PI407701	cultivated	Hei long No. 3	China	Haplotype4	-	-	-
USB-150	#1	PI407708A	cultivated	Feng shou No. 10	China	Haplotype4	-	-	-
USB-151	#1	PI407742	cultivated	16	China	Haplotype4	-	-	-
USB-152	#1	PI416751	cultivated	A-B(D)	Japan	Haplotype4	-	-	-
USB-153	#1	PI416838	cultivated	Choutan	Japan	Haplotype4	-	-	-
USB-154	#1	PI417215	cultivated	Ooita Aki Daizu 2	Japan	-	-	-	-
USB-155	#1	PI417242	cultivated	Pekin dai seitou	China	-	-	-	-
USB-156	#1	PI417345B	cultivated	(Shou outou)	China	-	-	-	-
USB-157	#1	PI417381	cultivated	Tenpoku shirome	Japan	Haplotype4	-	-	-
USB-158	#1	PI417479	cultivated	Yougetsu	Japan	-	-	-	-

USB-159	#1	PI417500	cultivated	Escura A	Brazil	Haplotype4	-	-	-
USB-160	#1	PI417529	cultivated	A38	Germany	-	-	-	-
USB-161	#1	PI417581	cultivated	H-060072	United States	Haplotype4	-	-	-
USB-162	#1	PI423926	cultivated	Tousan 72	Japan	Haplotype4	-	-	-
USB-163	#1	PI424038B	cultivated	74053	Korea, South	Haplotype4	-	-	-
USB-166	#1	PI424195A	cultivated	ISZ-3	Hungary	Haplotype4	-	-	-
USB-167	#1	PI430595	cultivated	58-161	China	Haplotype2	-	-	-
USB-168	#1	PI436684	cultivated	Tie-feng 8	China	Haplotype4	-	-	-
USB-169	#1	PI437110A	cultivated	VIR 244	Russian Federation	-	-	-	-
USB-170	#1	PI437112A	cultivated	VIR 249	Russian Federation	-	-	-	-
USB-171	#1	PI437127A	cultivated	Imeretinscaja	Georgia	Haplotype4	-	-	-
USB-173	#1	PI437160	cultivated	Krasnodarscaja 13	Russian Federation	-	-	-	-
USB-174	#1	PI437165A	cultivated	Toncostebel'naja 27	Russian Federation	-	-	-	-
USB-175	#1	PI437240	cultivated	CSchi 1069	Moldova	Haplotype4	-	-	-
USB-177	#1	PI437376A	cultivated	Ussurijscaja 308	Russian Federation	-	-	-	-
USB-178	#1	PI437485	cultivated	VIR 1048	Russian Federation	-	-	-	-
USB-179	#1	PI437500A	cultivated	VIR 3810	Russian Federation	-	-	-	-
USB-180	#1	PI437505	cultivated	VIR 3853	Russian Federation	Haplotype4	-	-	-
USB-182	#1	PI437662	cultivated	Gun'-tszu-lin' 658	China	Haplotype5	-	-	-
USB-183	#1	PI437685D	cultivated	(Phun-zhun)	China	-	-	-	-
USB-184	#1	PI437695A	cultivated	S-185	China	-	-	-	-
USB-185	#1	PI437776	cultivated	VIR 1302	China	Haplotype4	-	-	-
USB-186	#1	PI437788A	cultivated	VIR 3018	China	-	-	-	-

USB-187	#1	PI437793	cultivated	VIR 3024	China	Haplotype4	-	-	-
USB-188	#1	PI437814A	cultivated	An'da	China	Haplotype4	-	-	-
USB-189	#1	PI437838	cultivated	DV-254	Russian Federation	-	-	-	-
USB-192	#1	PI437991B	cultivated	VIR 1657	China	Haplotype4	-	-	-
USB-193	#1	PI438019B	cultivated	VIR 1883	China	Haplotype4	-	-	-
USB-195	#1	PI438083	cultivated	VIR 2506	China	Haplotype4	-	-	-
USB-196	#1	PI438112B	cultivated	VIR 2623	China	-	-	-	-
USB-198	#1	PI438230A	cultivated	VIR 4521	China	Haplotype4	-	-	-
USB-199	#1	PI438239B	cultivated	VIR 4536	China	Haplotype4	-	-	-
USB-201	#1	PI438309	cultivated	VIR 3017	China	-	-	-	-
USB-202	#1	PI438323	cultivated	Grignon 53-F-3	France	Haplotype4	-	-	-
USB-203	#1	PI438336	cultivated	Sao 208	Algeria	-	-	-	-
USB-204	#1	PI438347	cultivated	35S.277	Australia	Haplotype2	-	-	-
USB-206	#1	PI438496B	cultivated	(Peking)	United States	Haplotype2	-	-	-
USB-207	#1	PI438496C	cultivated	(Peking)	United States	Haplotype4	-	-	-
USB-208	#1	PI438500	cultivated	Virginia	United States	Haplotype4	-	-	-
USB-209	#1	PI445824A	cultivated	Wolfsthaler	Germany	Haplotype4	-	-	-
USB-210	#1	PI458505	cultivated	Da Bai mei	China	Haplotype4	-	-	-
USB-212	#1	PI458510	cultivated	Ji Ti No. 1	China	Haplotype4	-	-	-
USB-213	#1	PI464896	cultivated	Jou Nong No. 5	China	-	-	-	-
USB-214	#1	PI464912	cultivated	Dan Dou 1	China	-	-	-	-
USB-215	#1	PI464923	cultivated	Tie Fen 16	China	Haplotype4	-	-	-
USB-216	#1	PI467347	cultivated	Zi-hua-cuo-zi	China	Haplotype2	-	-	-

USB-217	#1	PI468408B	cultivated	(Qi Huang No. 1)	China	Haplotype4	-	-	-
USB-218	#1	PI475820	cultivated		China	-	-	-	-
USB-220	#1	PI476352B	cultivated	(Colnon)	Kyrgyzstan	Haplotype4	-	-	-
USB-221	#1	PI479735	cultivated	Silihuang	China	Haplotype4	-	-	-
USB-222	#1	PI490766	cultivated	Dawudou	China	-	-	-	-
USB-223	#1	PI495020	cultivated	Xu dou 2	China	-	-	-	-
USB-224	#1	PI497953	cultivated	I.C. 192	India	Haplotype4	-	-	-
USB-226	#1	PI497967	cultivated	PLSO 96	India	-	-	-	-
USB-227	#1	PI504288	cultivated	S	Japan	-	-	-	-
USB-228	#1	PI506933	cultivated	Kouiku 1	Japan	Haplotype4	-	-	-
USB-229	#1	PI506942	cultivated	Koushurei 235	Japan	Haplotype4	-	-	-
USB-230	#1	PI507017	cultivated	Madara ooha tsuru mame	Japan	-	-	-	-
USB-232	#1	PI507088	cultivated	Nattou Kotsubu	Japan	Haplotype4	-	-	-
USB-233	#1	PI507293B	cultivated	(Shoukin ou)	Japan	Haplotype4	-	-	-
USB-234	#1	PI507458	cultivated	Tousan kei BL 521	Japan	Haplotype4	-	-	-
USB-235	#1	PI507467	cultivated	Tousan kei F 764	Japan	Haplotype4	-	-	-
USB-236	#1	PI507471	cultivated	Tousan kei na 16	Japan	Haplotype4	-	-	-
USB-238	#1	PI507480	cultivated	Tousan kei YL 24	Japan	Haplotype4	-	-	-
USB-240	#1	PI514671	cultivated	Feng shou No. 7	China	Haplotype4	-	-	-
USB-241	#1	PI518727	cultivated	Ju huang	China	Haplotype2	-	-	-
USB-242	#1	PI532463B	cultivated	(He bei No. 1)	China	Haplotype4	-	-	-
USB-245	#1	PI538386A	cultivated	1886	China	Haplotype4	-	-	-

USB-246	#1	PI540552	cultivated	Hoyt	U.S.	Haplotype4	-	-	-
USB-248	#1	PI542972	cultivated	H7190	U.S.	-	-	-	-
USB-249	#1	PI548162	cultivated	T48	U.S.	-	-	-	-
USB-250	#1	PI548169	cultivated	T117	U.S.	-	-	-	-
USB-251	#1	PI548178	cultivated	T145	U.S.	Haplotype4	-	-	-
USB-252	#1	PI548198	cultivated	T209	U.S.	-	-	-	-
USB-253	#1	PI548200	cultivated	T211H	U.S.	Haplotype4	-	-	-
USB-254	#1	PI548256	cultivated	T279	U.S.	Haplotype4	-	-	-
USB-255	#1	PI548336	cultivated	Habaro	Russian Federation	Haplotype4	-	-	-
USB-256	#1	PI548356	cultivated	Kanro	Korea, North	-	-	-	-
USB-257	#1	PI548427	cultivated	Wilson	China	Haplotype4	-	-	-
USB-258	#1	PI548383	cultivated	Mansoy	China	Haplotype4	-	-	-
USB-259	#1	PI548400	cultivated	Patoka	China	Haplotype2	-	-	-
USB-260	#1	PI548411	cultivated	Seneca	China	Haplotype4	-	-	-
USB-261	#1	PI548359	cultivated	Kingwa	China	Haplotype4	-	-	-
USB-262	#1	PI548473	cultivated	Monetta	China	Haplotype4	-	-	-
USB-263	#1	PI548474	cultivated	Nanda	Korea, North	Haplotype4	-	-	-
USB-264	#1	PI548479	cultivated	Otootan	Taiwan	Haplotype4	-	-	-
USB-265	#1	PI548521	cultivated	BSR 201	U.S.	Haplotype4	-	-	-
USB-266	#1	PI548524	cultivated	Weber	U.S.	-	-	-	-
USB-268	#1	PI548571	cultivated	Harlon	Canada	Haplotype4	-	-	-
USB-269	#1	PI548572	cultivated	Harly	Canada	-	-	-	-
USB-270	#1	PI548582	cultivated	McCall	U.S.	-	-	-	-

USB-271	#1	PI548619	cultivated	Sparks	U.S.	-	-	-	-
USB-272	#1	PI548633	cultivated	Wye	U.S.	-	-	-	-
USB-274	#1	PI548696	cultivated	Dortchsoy 67	U.S.	-	-	-	-
USB-276	#1	PI548978	cultivated	Gail	U.S.	-	-	-	-
USB-277	#1	PI549017	cultivated	ZYD 3938	China	-	-	-	-
USB-278	#1	PI549021A	cultivated	Na hei dou	China	Haplotype4	-	-	-
USB-279	#1	PI549028	cultivated	Feng da li	China	Haplotype4	-	-	-
USB-281	#1	PI549040	cultivated	ZYD 2704	China	Haplotype4	-	-	-
USB-282	#1	PI556511	cultivated	A3127	U.S.	Haplotype4	-	-	-
USB-283	#1	PI561318A	cultivated	Hui nan bai hua xiao hei dou	China	-	-	-	-
USB-284	#1	PI561371	cultivated	Fen dou 15	China	-	-	-	-
USB-285	#1	PI561387	cultivated	Kosuzu	Japan	Haplotype4	-	-	-
USB-286	#1	PI561389B	cultivated	(Okura Natto)	Japan	-	-	-	-
USB-287	#1	PI561701	cultivated	G88-20092	U.S.	Haplotype4	-	-	-
USB-290	#1	PI567074B	cultivated	MARIF 2672	Indonesia	-	-	-	-
USB-291	#1	PI567173	cultivated	Hei he 51	China	Haplotype4	-	-	-
USB-293	#1	PI567225	cultivated	Kisinevskaja 90	Moldova	-	-	-	-
USB-294	#1	PI567226	cultivated	Har'kovskaja Zernoukosnaja	Russian Federation	Haplotype4	-	-	-
USB-295	#1	PI567231	cultivated	WJK-China-46	China	-	-	-	-
USB-296	#1	PI567238	cultivated	W6 6210	China	-	-	-	-

USB-297	#1	PI567262A	cultivated	Similar to: Gu tian type	China	Haplotype2	-	-	-
USB-298	#1	PI567307	cultivated	Hei huang dou	China	Haplotype4	-	-	-
USB-301	#1	PI567346	cultivated	Niu mao huang dou	China	-	-	-	-
USB-302	#1	PI567352A	cultivated	Yang yan qing dou	China	Haplotype4	-	-	-
USB-303	#1	PI567353	cultivated	Yang yan ren dou	China	-	-	-	-
USB-304	#1	PI567361	cultivated	Lu fang huang dou	China	-	-	-	-
USB-306	#1	PI567407	cultivated	Xiao dou	China	-	-	-	-
USB-307	#1	PI567408	cultivated	Xiao jin huang	China	-	-	-	-
USB-308	#1	PI567410B	cultivated	(Yang huang dou)	China	Haplotype4	-	-	-
USB-309	#1	PI567415A	cultivated	Bai da huang dou	China	-	-	-	-
USB-310	#1	PI567416	cultivated	Bai dou	China	Haplotype4	-	-	-
USB-311	#1	PI567418A	cultivated	Bai hei dou	China	-	-	-	-
USB-312	#1	PI567428	cultivated	Bai ji yao	China	Haplotype4	-	-	-
USB-313	#1	PI567435B	cultivated	(Hei hei dou)	China	-	-	-	-
USB-314	#1	PI567439	cultivated	Hong jia huang dou	China	-	-	-	-
USB-315	#1	PI567488A	cultivated	Di liu huang dou No. 2	China	-	-	-	-
USB-316	#1	PI567489A	cultivated	Er da li huang dou	China	-	-	-	-
USB-317	#1	PI567525	cultivated	Cao qing huang dou	China	Haplotype2	-	-	-
USB-318	#1	PI567532	cultivated	Dai ye xiao huang dou	China	Haplotype4	-	-	-
USB-319	#1	PI567548	cultivated	Hua li hu zi	China	-	-	-	-

USB-320	#1	PI567576	cultivated	Ping ding huang	China	-	-	-	-
USB-322	#1	PI567675	cultivated	Yu cheng xiao tie jiao huang	China	Haplotype4	-	-	-
USB-323	#1	PI567685	cultivated	Zhong mou tie jiao er cao	China	-	-	-	-
USB-324	#1	PI567698A	cultivated	Fu yang (17)	China	Haplotype4	-	-	-
USB-325	#1	PI567726	cultivated	Fu yang (50)	China	-	-	-	-
USB-326	#1	PI567746	cultivated	Pei xian da bai jiao	China	-	-	-	-
USB-327	#1	PI567780B	cultivated	(Tong shan zheng ji dou)	China	Haplotype4	-	-	-
USB-328	#1	PI567782	cultivated	OAC Dorado	Canada	Haplotype4	-	-	-
USB-330	#1	PI574477	cultivated	Fen dou 31	China	Haplotype4	-	-	-
USB-331	#1	PI578375B	cultivated	(Aan tu dang di hei dou)	China	-	-	-	-
USB-332	#1	PI578412	cultivated	Gong jiao 6308-1	China	Haplotype4	-	-	-
USB-333	#1	PI578499A	cultivated	Lu yue bai	China	Haplotype4	-	-	-
USB-334	#1	PI578493	cultivated	Huang bao zhu	China	Haplotype4	-	-	-
USB-335	#1	PI578503	cultivated	Tie jia si li huang	China	Haplotype4	-	-	-
USB-336	#1	PI578504	cultivated	Xiang dou No. 3	China	Haplotype4	-	-	-
USB-338	#1	PI587588A	cultivated	Tai xing niu mao huang yi	China	-	-	-	-
USB-339	#1	PI587588B	cultivated	(Tai xing niu mao huang yi)	China	-	-	-	-

USB-341	#1	PI587712B	cultivated	(E dou No. 1)	China	Haplotype4	-	-	-
USB-342	#1	PI587804	cultivated	Jing 789	China	Haplotype4	-	-	-
USB-343	#1	PI587811A	cultivated	ZDD005777	China	Haplotype2	-	-	-
USB-345	#1	PI592937	cultivated	Jin dou 14	China	-	-	-	-
USB-346	#1	PI592940	cultivated	Jin dou 17	China	-	-	-	-
USB-347	#1	PI592952	cultivated	Zheng 77249	China	Haplotype4	-	-	-
USB-348	#1	PI592960	cultivated	Dong nong 38	China	-	-	-	-
USB-351	#1	PI593953	cultivated	Sui nong No. 10	China	Haplotype4	-	-	-
USB-353	#1	PI594170B	cultivated	(Geden shirazu)	Japan	-	-	-	-
USB-354	#1	PI594456A	cultivated	Xiao jin huang	China	-	-	-	-
USB-357	#1	PI594880	cultivated	Song zi dou	China	Haplotype4	-	-	-
USB-358	#1	PI594922	cultivated	Graham	U.S.	Haplotype4	-	-	-
USB-361	#1	PI597464	cultivated	Zhe chun No. 3	China	Haplotype4	-	-	-
USB-362	#1	PI597471A	cultivated	Zheye 29	China	-	-	-	-
USB-363	#1	PI597478B	cultivated	(Paldalkong)	Korea, South	Haplotype4	-	-	-
USB-364	#1	PI598358	cultivated	TN 5-95	U.S.	-	-	-	-
USB-366	#1	PI602502B	cultivated	(Xiong yue xiao huang dou)	China	Haplotype4	-	-	-
USB-367	#1	PI602993	cultivated	Pi xian ruan tiao zhi	China	-	-	-	-
USB-368	#1	PI603162	cultivated	GL 2631 /96	Korea, North	Haplotype4	-	-	-
USB-369	#1	PI603290	cultivated	Zao shu 18	China	Haplotype4	-	-	-
USB-370	#1	PI603345	cultivated	ZDD00403	China	Haplotype4	-	-	-
USB-371	#1	PI603389	cultivated	Huang ke	China	Haplotype4	-	-	-

USB-372	#1	PI603397	cultivated	Hei qi huang da dou	China	Haplotype4	-	-	-
USB-373	#1	PI603399	cultivated	Xiao bai qi	China	Haplotype4	-	-	-
USB-374	#1	PI603463	cultivated	Dong jie No. 1	China	Haplotype4	-	-	-
USB-375	#1	PI603488	cultivated	ZDD19294	China	Haplotype4	-	-	-
USB-376	#1	PI603492	cultivated	Qi hei dou	China	Haplotype4	-	-	-
USB-377	#1	PI603494	cultivated	Hai dou zi	China	-	-	-	-
USB-378	#1	PI603495B	cultivated	(Hong mi lan dou zi)	China	Haplotype4	-	-	-
USB-381	#1	PI603526	cultivated	Hei you dou	China	-	-	-	-
USB-382	#1	PI603549	cultivated	Mei dou	China	Haplotype4	-	-	-
USB-383	#1	PI603555	cultivated	Hua da hei dou	China	Haplotype4	-	-	-
USB-384	#1	PI603556	cultivated	ZDD08563	China	-	-	-	-
USB-385	#1	PI603559	cultivated	ZDD08590	China	-	-	-	-
USB-386	#1	PI603675	cultivated	Huai yin gua dou jia	China	-	-	-	-
USB-387	#1	PI603698J	cultivated	(Dan yang shui bai dou)	China	Haplotype4	-	-	-
USB-388	#1	PI603722	cultivated	Nan chong ba yue huang	China	Haplotype2	-	-	-
USB-389	#1	PI605765B	cultivated	(Ninh minh)	Vietnam	-	-	-	-
USB-390	#1	PI606374	cultivated	Cao bang 8	Vietnam	Haplotype2	-	-	-
USB-392	#1	PI612730	cultivated	Zhong huong No. 10	China	Haplotype4	-	-	-
USB-393	#1	PI612754	cultivated	ZY 645	China	Haplotype4	-	-	-
USB-395	#1	PI628812	cultivated	MG/BR-46 (Conquista)	Brazil	Haplotype4	-	-	-

USB-396	#1	PI628913	cultivated	BR-30	Brazil	Haplotype4	-	-	-
USB-398	#1	PI628963	cultivated	La Suprema	Brazil	Haplotype2	-	-	-
USB-399	#1	PI631123	cultivated	TW98-1	U.S.	-	-	-	-
USB-400	#1	PI632418	cultivated	Tara	U.S.	-	-	-	-
USB-401	#1	PI632650	cultivated	DT 22	Vietnam	Haplotype4	-	-	-
USB-402	#1	PI639528B	cultivated	VIR 233	Russian Federation	-	-	-	-
USB-403	#1	PI639543	cultivated	VIR 3715	Russian Federation	Haplotype4	-	-	-
USB-404	#1	PI639550E	cultivated	(KSHI 713)	Moldova	Haplotype4	-	-	-
USB-405	#1	PI639559B	cultivated	(VYTKA 2)	Ukraine	-	-	-	-
USB-407	#1	PI639570	cultivated	VIR 7010	Philippines	-	-	-	-
USB-409	#1	PI322692	cultivated	Max C.P1159A8	Australia	Haplotype4	-	-	-
USB-411	#1	PI567788	cultivated	Bienville	U.S.	Haplotype4	-	-	-
C01	#2		cultivated		Advanced line; Shandong, China	-	-	Haplotype3	-
C02	#2		cultivated		Advanced line; Liaoning, China	-	-	Haplotype3	-
C08	#2		cultivated		Advanced line; USA	-	-	-	-
C12	#2		cultivated		Advanced line; Shanxi, China	-	-	Haplotype3	-
C14	#2		cultivated		Advanced line; Brazil	-	-	-	-
C16	#2		cultivated		Taiwan	Haplotype2	-	-	-
C17	#2		cultivated		Landrace; Sichuan, China	-	-	Haplotype3	-

C19	#2		cultivated		Landrace; Jilin, China	-	-	-	-
C24	#2		cultivated		Advanced line; Jiangxi, China	Haplotype2	-	Haplotype3	-
C27	#2		cultivated		Advanced line; Hebei, China	Haplotype2	-	-	-
C30	#2		cultivated		Advanced line; Henan, China	-	-	Haplotype3	-
C33	#2		cultivated		Advanced line; Heilongjiang, China	-	-	Haplotype3	-
C34	#2		cultivated		Landrace; Guangxi, China	-	-	Haplotype3	-
C35	#2		cultivated		Landrace; Guangdong, China	Haplotype2	-	Haplotype3	-
SJ-01	#1	PI378684A	wild		Japan	Haplotype4	-	-	-
SJ-02	#1	PI378696A	wild		Japan	Haplotype4	-	-	-
SJ-03	#1	PI378698	wild		Japan	-	-	-	-
SJ-04	#1	PI407083	wild	RB 1072	Japan	Haplotype4	-	-	-
SJ-05	#1	PI407270	wild	K45-B	Korea, South	Haplotype4	-	-	Haplotype5
SJ-06	#1	PI407281	wild	K103-C	Korea, South	Haplotype4	-	-	-
SJ-07	#1	PI407286	wild	Tsuru Mame	Japan	-	Haplotype2	-	Haplotype3
SJ-08	#1	PI407303	wild		China	-	-	-	Haplotype6
SJ-09	#1	PI407307	wild		China	Haplotype4	-	-	-
SJ-10	#1	PI407308	wild		Korea, South	Haplotype4	-	-	Haplotype3

SJ-11	#1	PI407315	wild		Korea, South	Haplotype4	-	-	-
SJ-12	#1	PI424008A	wild	74006	Korea, South	-	-	-	-
SJ-13	#1	PI468918	wild		China	Haplotype3	-	Haplotype2	-
SJ-14	#1	PI507609	wild	NIAR 040023	Japan	Haplotype4	-	-	-
SJ-15	#1	PI507619B	wild	NIAR 040016	Japan	Haplotype3	-	-	-
SJ-16	#1	PI507667	wild	NIAR 090024	Japan	Haplotype4	-	-	Haplotype3
SJ-17	#1	PI507752	wild	VIR 8448	Russian Federation	-	-	-	Haplotype2
SJ-18	#1	PI507830B	wild	VIR 9044	Russian Federation	-	-	-	Haplotype2
SJ-19	#1	PI532450	wild	GD50344	China	Haplotype4	-	Haplotype4	-
SJ-20	#1	PI549047	wild	Mrs. Wong's Picnic #1	China	Haplotype4	-	-	-
SJ-21	#1	PI549048	wild		China	-	-	-	-
SJ-22	#1	PI562558	wild	KE12	Korea, South	Haplotype5	Haplotype4	-	-
SJ-23	#1	PI562568	wild	KF24	Korea, South	-	-	-	-
SJ-24	#1	PI597458B	wild		China	Haplotype4	-	-	-
SJ-25	#1	PI597459C	wild	ZYD3234	China	-	-	-	Haplotype2
SJ-26	#1	PI597471D	wild	(Zheyue 29)	China	Haplotype4	-	-	-
SJ-27	#1	PI639586	wild	KB 14	Russian Federation	Haplotype2	-	-	Haplotype2
SJ-28	#1	PI639590A	wild	KZ 659	Russian Federation	Haplotype2	-	-	Haplotype2
SJ-29	#1	PI639590B	wild	(KZ 659)	Russian Federation	-	-	-	Haplotype2
SJ-30	#1	PI639621	wild	KT 156	Russian Federation	Haplotype4	-	-	-
SJ-31	#1	PI507847	wild	L 598	Russian Federation	Haplotype4	-	-	Haplotype2
SJ-32	#1	PI424014	wild	74012	Korea, South	Haplotype4	-	-	-

SJ-33	#1	PI424107A	wild	74106	Korea, South	-	-	-	-
SJ-34	#1	PI407175	wild	K4-B	Korea, South	-	-	-	-
SJ-35	#1	PI407288	wild		China	Haplotype3	-	-	Haplotype2
SJ-36	#1	PI507615	wild	NIAR 040012	Japan	Haplotype4	-	-	-
SJ-37	#1	PI562557	wild	KE10	Korea, South	-	-	-	Haplotype3
SJ-38	#1	PI378699A	wild		Japan	Haplotype4	-	-	-
SJ-39	#1	PI407262	wild	K42-A	Korea, South	Haplotype5	-	-	Haplotype7
SJ-40	#1	PI562534	wild	KA5	Korea, South	Haplotype4	-	-	Haplotype2
SJ-41	#1	PI407183	wild	K8-B	Korea, South	-	-	-	-
SJ-42	#1	PI407038	wild	RB 1072	Japan	Haplotype4	-	-	-
SJ-43	#1	PI407318A	wild		Korea, South	-	-	-	-
SJ-44	#1	PI407229	wild	K47-A	Korea, South	-	-	-	-
SJ-45	#1	PI507638	wild	NIAR 060010	Japan	Haplotype4	-	-	Haplotype3
W01	#2		wild		Beijing area, China	-	-	Haplotype3	-
W02	#2		wild		Liaoning, China	-	-	Haplotype3	-
W03	#2		wild		Inner Mongolia, China	-	-	Haplotype3	-
W04	#2		wild		Henan, China	-	-	Haplotype3	-
W05	#2		wild		Henan, China	Haplotype2	-	Haplotype3	-
W06	#2		wild		Heilongjian, China	-	-	-	-
W07	#2		wild		Liaoning, China	-	-	Haplotype3	-
W08	#2		wild		Heilongjian, China	-	-	-	-
W09	#2		wild		Liaoning, China	-	-	Haplotype3	-
W10	#2		wild		Heilongjian, China	-	-	-	-

W11	#2	wild	Shanxi, China	-	-	Haplotype3	-
W12	#2	wild	Anhui, China	-	-	Haplotype3	-
W13	#2	wild	Inner Mongolia, China	-	Haplotype3	Haplotype3	-
W14	#2	wild	Inner Mongolia, China	-	-	Haplotype3	-
W15	#2	wild	Henan, China	-	-	Haplotype3	-
W16	#2	wild	Heilongjian, China	-	-	-	Haplotype4
W17	#2	wild	Liaoning, China	-	-	Haplotype3	-

Supplemental Table S5. Positions of non-synonymous SNPs in haplotypes of *COL* gene.

COL2b

Non-synonymous SNP based	SNP position				
	5469858	5469529	5469458	5469400	5469397
Haplotype 1	A	A	C	G	G
Haplotype 2	G	A	C	G	G
Haplotype 3	A	G	C	G	G
Haplotype 4	A	A	A	G	G
Haplotype 5	A	A	C	A	G
Haplotype 6	A	A	C	G	A
Haplotype 7	G	A	C	A	G

COL2a

Non-synonymous SNP based	SNP position		
	14630733	14630892	14632456
Haplotype 1	G	A	G
Haplotype 2	C	A	G
Haplotype 3	G	G	G
Haplotype 4	G	A	A

COL1a

Non-synonymous SNP based	SNP position		
	55966793	55966708	55966491
Haplotype 1	G	A	G
Haplotype 2	A	A	G
Haplotype 3	G	G	G
Haplotype 4	G	A	C
Haplotype 5	G	G	C

COL1b

Non-synonymous SNP based	SNP position			
	22602036	22602002	22601982	22601532
Haplotype 1	G	C	T	A
Haplotype 2	C	C	T	A
Haplotype 3	G	T	T	A
Haplotype 4	G	C	A	A
Haplotype 5	G	C	T	T

Supplemental Table S6. Positions of SNPs of four *COLs* gene in exon (red) and intron (black).

Gene	Chr	Position	W82	SNP	Haplotype
<i>COL2b</i>	Chr19	5468378	A	G	/
<i>COL2b</i>	Chr19	5468407	T	A	/
<i>COL2b</i>	Chr19	5468450	A	C	/
<i>COL2b</i>	Chr19	5468847	A	G	/
<i>COL2b</i>	Chr19	5469091	T	C	/
<i>COL2b</i>	Chr19	5469136	A	C	/
<i>COL2b</i>	Chr19	5469397	C	T	6
<i>COL2b</i>	Chr19	5469400	C	T	5&7
<i>COL2b</i>	Chr19	5469458	G	T	4
<i>COL2b</i>	Chr19	5469529	T	C	3
<i>COL2b</i>	Chr19	5469590	G	A	/
<i>COL2b</i>	Chr19	5469779	T	A	/
<i>COL2b</i>	Chr19	5469788	T	C	/
<i>COL2b</i>	Chr19	5469812	C	T	/
<i>COL2b</i>	Chr19	5469858	T	C	2&7
<i>COL2a</i>	Chr13	14630432	C	T	/
<i>COL2a</i>	Chr13	14630733	G	C	2
<i>COL2a</i>	Chr13	14630892	A	G	3
<i>COL2a</i>	Chr13	14631202	T	C	/
<i>COL2a</i>	Chr13	14631208	T	C	/
<i>COL2a</i>	Chr13	14631216	T	G	/
<i>COL2a</i>	Chr13	14631255	A	G	/
<i>COL2a</i>	Chr13	14631285	G	A	/
<i>COL2a</i>	Chr13	14631372	G	A	/
<i>COL2a</i>	Chr13	14631416	C	T	/
<i>COL2a</i>	Chr13	14631466	C	A	/
<i>COL2a</i>	Chr13	14631515	G	T,A	/
<i>COL2a</i>	Chr13	14631543	A	T	/
<i>COL2a</i>	Chr13	14631612	C	T	/
<i>COL2a</i>	Chr13	14631855	T	A	/
<i>COL2a</i>	Chr13	14631926	C	A	/
<i>COL2a</i>	Chr13	14632023	T	A	/
<i>COL2a</i>	Chr13	14632070	T	C	/

<i>COL2a</i>	Chr13	14632076	A	G	/
<i>COL2a</i>	Chr13	14632269	T	C	/
<i>COL2a</i>	Chr13	14632440	A	T	/
<i>COL2a</i>	Chr13	14632456	G	A	4
<i>COL1a</i>	Chr18	55964548	A	T	/
<i>COL1a</i>	Chr18	55965212	T	C	/
<i>COL1a</i>	Chr18	55965223	A	G	/
<i>COL1a</i>	Chr18	55965234	T	C	/
<i>COL1a</i>	Chr18	55965313	T	C	/
<i>COL1a</i>	Chr18	55965330	A	G	/
<i>COL1a</i>	Chr18	55965339	A	G	/
<i>COL1a</i>	Chr18	55965422	C	T	/
<i>COL1a</i>	Chr18	55965452	C	A	/
<i>COL1a</i>	Chr18	55965571	T	G	/
<i>COL1a</i>	Chr18	55965606	C	A	/
<i>COL1a</i>	Chr18	55965857	A	G	/
<i>COL1a</i>	Chr18	55965869	C	G	/
<i>COL1a</i>	Chr18	55965876	A	C	/
<i>COL1a</i>	Chr18	55965945	A	C	/
<i>COL1a</i>	Chr18	55965972	C	T	/
<i>COL1a</i>	Chr18	55966007	G	C	/
<i>COL1a</i>	Chr18	55966097	T	C	/
<i>COL1a</i>	Chr18	55966128	A	T	/
<i>COL1a</i>	Chr18	55966142	T	C	/
<i>COL1a</i>	Chr18	55966177	G	A	/
<i>COL1a</i>	Chr18	55966193	A	G	/
<i>COL1a</i>	Chr18	55966235	T	A	/
<i>COL1a</i>	Chr18	55966240	T	C	/
<i>COL1a</i>	Chr18	55966252	A	G	/
<i>COL1a</i>	Chr18	55966261	C	G	/
<i>COL1a</i>	Chr18	55966276	G	A	/
<i>COL1a</i>	Chr18	55966308	T	C	/
<i>COL1a</i>	Chr18	55966332	A	G	/
<i>COL1a</i>	Chr18	55966366	G	C	/
<i>COL1a</i>	Chr18	55966393	G	A	/
<i>COL1a</i>	Chr18	55966491	C	G	4&5
<i>COL1a</i>	Chr18	55966533	A	G	/

<i>COL1a</i>	Chr18	55966708	T	C	3&5
<i>COL1a</i>	Chr18	55966793	C	T	2
<i>COL1a</i>	Chr18	55967004	A	C	/
<i>COL1a</i>	Chr18	55967030	T	A	/
					/
<i>COL1b</i>	Chr08	22599531	T	A	/
<i>COL1b</i>	Chr08	22599532	T	A	/
<i>COL1b</i>	Chr08	22599573	C	T	/
<i>COL1b</i>	Chr08	22599713	T	C	/
<i>COL1b</i>	Chr08	22600197	G	A	/
<i>COL1b</i>	Chr08	22600256	C	T	/
<i>COL1b</i>	Chr08	22600331	C	A	/
<i>COL1b</i>	Chr08	22600335	C	G	/
<i>COL1b</i>	Chr08	22600513	A	T	/
<i>COL1b</i>	Chr08	22600564	C	T	/
<i>COL1b</i>	Chr08	22600651	T	G	/
<i>COL1b</i>	Chr08	22600849	A	G	/
<i>COL1b</i>	Chr08	22600891	G	A	/
<i>COL1b</i>	Chr08	22600947	G	A	/
<i>COL1b</i>	Chr08	22601121	T	C	/
<i>COL1b</i>	Chr08	22601140	A	G	/
<i>COL1b</i>	Chr08	22601185	A	T	/
<i>COL1b</i>	Chr08	22601205	A	G	/
<i>COL1b</i>	Chr08	22601258	G	A,T	/
<i>COL1b</i>	Chr08	22601303	G	A	/
<i>COL1b</i>	Chr08	22601317	T	G	/
<i>COL1b</i>	Chr08	22601365	A	G	/
<i>COL1b</i>	Chr08	22601503	T	C	/
<i>COL1b</i>	Chr08	22601527	C	A	/
<i>COL1b</i>	Chr08	22601532	T	A	5
<i>COL1b</i>	Chr08	22601614	C	A	/
<i>COL1b</i>	Chr08	22601721	T	C	/
<i>COL1b</i>	Chr08	22601770	A	G	/
<i>COL1b</i>	Chr08	22601887	G	A	/
<i>COL1b</i>	Chr08	22601977	C	T	/
<i>COL1b</i>	Chr08	22601982	A	T	4
<i>COL1b</i>	Chr08	22602002	G	A	3
<i>COL1b</i>	Chr08	22602036	C	G	2

<i>COL1b</i>	Chr08	22602063	A	G	/
<i>COL1b</i>	Chr08	22602149	G	A	/

NOTE:

Part of the results presented in this thesis has been published in *Frontiers in Plant Physiology*.

Zhang, J., Xu, M., Dwiyanti, M. S., Watanabe, S., Yamada, T., Hase, Y., Kanazawa, A., Sayama, T., Ishimoto, M., Liu, B., and Abe, J. (2020). A soybean deletion mutant that moderates the repression of flowering by cool temperatures. *Frontiers in Plant Science*, 11, 429. DOI: 10.3389/fpls.2020.00429

Acknowledgements

Above all, I would like to express my sincere appreciation to my supervisor Professor Jun Abe for the constant encouragement and continuous support of my doctoral study. During the preparation of this thesis, he has spent much time reading through each draft and provided me with inspiring advice. Without his patient instruction, insightful criticism and expert guidance, my doctoral thesis could not have reached its present form.

I acknowledge Professor Akira Kanazawa, Professor Yuji Kishima, Lecturer Tetsuya Yamada and Assistant Professor Maria Stefanie Dwiyantri for their numerous valuable comments and suggestions for my doctoral study and thesis. The knowledge that I gained from them will be of everlasting significance to my future research.

I also appreciate Professor Baohui Liu (School of Life Science, Guangzhou University), Lecturer Satoshi Watanabe (Faculty of Agriculture, Saga University), Dr. Yoshihiro Hase (Takasaki Advanced Radiation Research Institute, National Institutes for Quantum and Radiological Science and Technology), Dr. Takashi Sayama (Western Region Agricultural Research Center, National Agriculture and Food Research Organization), Dr. Masao Ishimoto (Institute of Crop Science, National Agriculture and Food Research Organization) for their constructive suggestions and cooperation to my study.

I am really grateful to all members of Plant Genetics and Evolution laboratory. Special thanks for Dr. Meilan Xu, Dr. Maria Stefanie Dwiyantri, Dr. Jianghui Zhu, Dr. Fei Sun, Dr. Cheolwoo Park, Mr. Shohei Kusumoto and Miss. Ying Qi for their advices and experimental help. Finally, I would like to thank my parents for their financial support and encouragement during my study life.