
Divergence of flowering genes in soybean

MOON YOUNG KIM¹, JIN HEE SHIN¹, YANG JAE KANG¹, SANG REA SHIM¹ and SUK-HA LEE^{1,2*}

¹Department of Plant Science and Research Institute for Agriculture and Life Sciences, ²Plant Genomics and Breeding Institute, Seoul National University, Seoul, 151-921, Korea

*Corresponding author (Fax, +82-2-873-2056; Email, sukhalee@snu.ac.kr)

Soybean genome sequences were blasted with *Arabidopsis thaliana* regulatory genes involved in photoperiod-dependent flowering. This approach enabled the identification of 118 genes involved in the flowering pathway. Two genome sequences of cultivated (Williams 82) and wild (IT182932) soybeans were employed to survey functional DNA variations in the flowering-related homologs. Forty genes exhibiting nonsynonymous substitutions between *G. max* and *G. soja* were catalogued. In addition, 22 genes were found to co-localize with QTLs for six traits including flowering time, first flower, pod maturity, beginning of pod, reproductive period, and seed filling period. Among the genes overlapping the QTL regions, two *LHY/CCAI* genes, *GI* and *SFR6* contained amino acid changes. The recently duplicated sequence regions of the soybean genome were used as additional criteria for the speculation of the putative function of the homologs. Two duplicated regions showed redundancy of both flowering-related genes and QTLs. ID 12398025, which contains the homeologous regions between chr 7 and chr 16, was redundant for the *LHY/CCAI* and *SPAI* homologs and the QTLs. Retaining of the *CRYI* gene and the pod maturity QTLs were observed in the duplicated region of ID 23546507 on chr 4 and chr 6. Functional DNA variation of the *LHY/CCAI* gene (Glyma07g05410) was present in a counterpart of the duplicated region on chr 7, while the gene (Glyma16g01980) present in the other portion of the duplicated region on chr 16 did not show a functional sequence change. The gene list catalogued in this study provides primary insight for understanding the regulation of flowering time and maturity in soybean.

[Kim MY, Shin JH, Kang YJ, Shim SR and Lee S-H 2012 Divergence of flowering genes in soybean. *J. Biosci.* **37** 857–870]
DOI 10.1007/s12038-012-9252-0

1. Introduction

Flowering refers to the plant transition from the vegetative stage to the reproductive stage. Plants have developed different mechanisms for controlling the timing of flowering in order to maximize their reproductive success, and examples of these mechanisms include photoperiod response and vernalization (Kim *et al.* 2009a). From an agricultural perspective, the control of flowering time is critical for grain yield and the production of dry matter in crops (Cockram *et al.* 2007). Thus, understanding the molecular mechanisms of flowering and identification of relevant genes should enable more efficient plant breeding. For example, the introduction of genes involved in the control of early flowering genes may permit the cultivation of some crops in short seasoned areas and late flowering genes may lead to a longer vegetative growth period for vegetative crops (Roux *et al.* 2006).

Furthermore, synchronization of the flowering time of two genotypes may enable the crossing of two genotypes that do not naturally bloom at the same time.

The evolution of flowering time is a key factor in the domestication and adaptation to new environments (Fuller 2007). Domestication is reported to cause physiological changes including changes in photoperiod sensitivity and synchronized flowering (Doebley *et al.* 2006). To date, several flowering genes and their causative changes related to crop domestication have been identified. For example, the vernalization (*Vrn*) and photoperiod (*Ppd*) genes were identified and found to be involved in the domestication and adaptation of wheat and barley (Cockram *et al.* 2007). Recently, an integrative candidate gene strategy was conducted to identify loci related to the domestication of sunflower (*Helianthus annuus*) (Blackman *et al.* 2011). Homologs of the flowering genes in model species were characterized in

Keywords. Flowering time; functional DNA variation; genome duplication; photoperiod; QTL; soybean

sunflower. These data along with previously identified QTLs and sequence differences were used to narrow down the genes involved in domestication (Blackman *et al.* 2011).

Soybean (*Glycine max*), which flowers in response to a short photoperiod, is classified as a short-day plant. Wild soybean (*G. soja*) is geographically distributed around a latitude of 30° and is exclusively found in East Asia including the Korean Peninsula, Japan, China and the Russian Far East (Hymowitz 1970). In contrast, modern soybean is cultivated from the equator up to high-latitude regions of 50° or more (Liu *et al.* 2008). The broad adaptability of cultivated soybean results from the selection of genotypes harbouring beneficial alleles in a large number of genes and QTLs involved in the flowering pathway, as well as introgression of these alleles into other genotypes. In soybean, numerous loci involved in flowering and maturity have been reported. Eight *E* loci have been detected by classical methods: *E1* and *E2* (Bernard 1971); *E3* (Buzzel 1971); *E4* (Buzzel and Voldeng 1980); *E5* (McBlain and Bernard 1987); *E6* (Bonato and Vello 1999); *E7* (Cober and Voldeng 2001); and *E8* (Cober *et al.* 2010). These *E* loci display differential sensitivity to light quality and photoperiod (Cober *et al.* 1996a, b). Upon isolating the functional genes underlying the *E2*, *E3*, and *E4* loci, they were found to be genes related to the light-dependent photoperiodic pathway. The genes for *E3* and *E4* encode light-absorbing photoreceptors, *GmPhyA3* and *GmPhyA2*, respectively (Liu *et al.* 2008; Watanabe *et al.* 2009). The gene for *E2* encodes a homolog of the circadian clock-controlled *GIGANTEA* (*GI*) from *Arabidopsis thaliana* (Watanabe *et al.* 2011). The *GI* homologs have been identified in several crops, including *Pisum sativum* (*LATE BLOOMER1*) and *Oryza sativa* (*Os-GIGANTEA*), and found to be functionally conserved (Hecht *et al.* 2007; Itoh *et al.* 2010; Izawa *et al.* 2011). The gene corresponding to the *Dt1* locus, which controls soybean growth habit, was identified as a homolog of *Arabidopsis terminal flower 1* (*TFL1*) (Liu *et al.* 2010; Tian *et al.* 2010). In comparison to cereal crops including rice (*O. sativa*), wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*), soybean contains a small number of identified genes related to the flowering pathway. For the cereal crops, a number of reviews describe the molecular pathways that control the seasonal flowering response (Cockram *et al.* 2007; Greenup *et al.* 2009; Izawa *et al.* 2003).

The recent release of the genome sequences of cultivated and wild soybean has allowed for comparative molecular genetics at the sequence level (Schmutz *et al.* 2010, Kim *et al.* 2010). Key genes and distinct pathways involved in flowering have been identified through many *Arabidopsis* studies. This information has enabled researchers to expand these insights to the investigation of flowering in other plant species (Jung and Muller 2009; Roux *et al.* 2006; Jarillo and Pineiro 2011; Yant *et al.* 2009). In this review, the

soybean homologs are compiled and compared to the *Arabidopsis* genes involved in light-dependent flowering. Specific focus is given to the co-localization of select genes with previously identified flowering time QTLs and functional DNA variations between cultivated and wild soybeans. Furthermore, the functional conservation of flowering QTLs in recently duplicated regions is discussed.

2. Soybean homologs of the flowering-related genes of *Arabidopsis thaliana*

In an attempt to identify soybean homologs, 20 *Arabidopsis* genes reportedly related to the photoperiodic flowering pathway were selected. The protein sequences of these *Arabidopsis* genes were subjected to BLAST analysis using the soybean whole genome sequences (<http://www.phytozome.com/soybean>). *Arabidopsis* flowering-related genes were recognized with TAIR10 functional descriptions. The peptide sequences were retrieved from TAIR10_pep_20110103_representative_gene_model and downloaded from ftp://ftp.arabidopsis.org/home/tair/Proteins/TAIR10_protein_lists/. The peptide sequences were blasted against *Glycine max* peptide sequences, downloaded from [phytozome.net](http://www.phytozome.net) with an e-value cutoff of 1e-50. *Arabidopsis* sequences that were highly similar (between 36% and 89%) were matched to nearly the same soybean peptides. Thus, to reduce the redundancy of soybean gene hits, similar *Arabidopsis* genes, such as *PHYA/B/C*, were merged into single categories. As a paleopolyploid, soybean has experienced two rounds of whole genome duplications (Shoemaker *et al.* 1996; Schlueter *et al.* 2004; Schmutz *et al.* 2010). This resulted in the identification of a large number of soybean homologs (table 1). Among the selected soybean genes, previously published genes were identified by searching two public databases (<http://www.soybase.org> and <http://www.ncbi.nlm.nih.gov>) and are discussed below.

Phytochromes (PHY) are red (R) light- and far-red (FR) light-absorbing photoreceptors present in the leaves of most plants (Chen *et al.* 2004). R absorption changes phytochrome into a biologically active form (Pfr), while FR absorption converts it into an inactive form (Pf). Mutations in the phytochrome genes cause altered flowering phenotypes in *Arabidopsis* and rice (*Oryza sativa*) (Reed *et al.* 1994; Takano *et al.* 2005). Eight soybean genes orthologous to *Arabidopsis PHYA/PHYB/PHYC* were detected. Three of these have been previously reported as *GmPhyA1*, *GmPhyA2* and *GmPhyA3* (Liu *et al.* 2008; Watanabe *et al.* 2009). The map-based cloning of the flowering time QTL identified four *GmPhyA3* alleles, and earlier flowering and maturity phenotypes are caused by mutations in the *GmPhyA3* gene (Watanabe *et al.* 2009). *GmPhyA2* and *GmPhyA3* are encoded by the previously reported flowering and maturity *E4* and *E3* loci, respectively (Liu *et al.* 2008; Watanabe *et al.*

Table 1. Soybean homologs of flowering time genes

Category	<i>Arabidopsis</i>	<i>Arabidopsis</i> locus ID	Soybean	Soybean gene model ID	Published soybean gene
Photoreceptor	<i>PHY A/B/C</i>	AT1G09570	<i>PHYA</i>	Glyma20g22160	<i>GmPhyA2, E4</i> , Liu <i>et al.</i> 2008
		AT2G18790	<i>PHYA</i>	Glyma10g28170	<i>GmPhyA1</i> , Liu <i>et al.</i> 2008
		AT5G35840	<i>PHYA</i>	Glyma19g41210	<i>GmPhyA3</i> , Watanabe <i>et al.</i> 2009
			<i>PHYB</i>	Glyma09g03990	
			<i>PHYB</i>	Glyma15g14980	
			<i>PHYA</i>	Glyma03g38620	
			<i>PHYE</i>	Glyma09g11600	
			<i>PHYE</i>	Glyma15g23400	
		<i>CRY 1/2</i>	AT4G08920	<i>CRY1</i>	Glyma13g01810
	AT1G04400		<i>CRY1</i>	Glyma14g35020	
			<i>CRY1</i>	Glyma04g11010	<i>GmCRY1a</i> , Zhang <i>et al.</i> 2008
			<i>CRY1</i>	Glyma06g10830	
			<i>CRY2</i>	Glyma20g35220	
			<i>CRY2</i>	Glyma10g32390	<i>CmCRY2a</i> , Zhang <i>et al.</i> 2008
	Circadian clock mediator	<i>TOC1</i>	AT5G61380	<i>TOC1</i>	Glyma04g33110
			<i>TOC1</i>	Glyma06g21120	<i>GmTOC1</i> , Liu <i>et al.</i> 2009, Hudson 2010
			<i>TOC1</i>	Glyma17g11040	
			<i>TOC1</i>	Glyma05g00880	
<i>GI</i>		AT1G22770	<i>GI</i>	Glyma10g36600	<i>GmGla</i> , Watanabe <i>et al.</i> 2011
			<i>GI</i>	Glyma09g07240	
			<i>GI</i>	Glyma20g30980	
<i>LHY/CCA1</i>		AT1G01060	<i>LHY</i>	Glyma19g45030	<i>GmLHY-like</i> , Hudson 2010
		AT2G46830	<i>LHY</i>	Glyma16g01980	
			<i>LHY</i>	Glyma03g42260	<i>GmLCL2</i> , Liu <i>et al.</i> 2009
<i>ELF3</i>		AT2G25930	<i>LHY</i>	Glyma07g05410	<i>GmCCA1a</i> , Hudson 2010
			<i>ELF3</i>	Glyma04g05280	
			<i>ELF3</i>	Glyma17g34980	
<i>CO</i>		AT5G15840	<i>ELF3</i>	Glyma14g10530	
			<i>COL2</i>	Glyma13g07030	
			<i>CO</i>	Glyma19g05170	
			<i>COL2</i>	Glyma08g28370	
			<i>COL2</i>	Glyma18g51320	
			<i>COL5</i>	Glyma13g01290	
			<i>COL5</i>	Glyma17g07420	
			<i>COL4</i>	Glyma06g06300	
	<i>COL4</i>		Glyma04g06240		
	<i>CDF1/FKF1</i>		AT5G62430	<i>CDF3</i>	Glyma06g20950
AT1G68050		<i>FKF1</i>	Glyma05g34530		
		<i>FKF1</i>	Glyma08g05130		
		<i>ZTL</i>	Glyma13g00860		
<i>SPA1/COPI</i>	AT2G46340 AT2G32950	<i>ZTL</i>	Glyma09g06220		
		<i>ZTL</i>	Glyma17g06950		
		<i>ZTL</i>	Glyma15g17480		
		<i>SPA2</i>	Glyma08g02490		
		<i>SPA1</i>	Glyma07g06420		
	<i>SPA1</i>	Glyma16g03030			

Table 1. (continued)

Category	<i>Arabidopsis</i>	<i>Arabidopsis</i> locus ID	Soybean	Soybean gene model ID	Published soybean gene		
Floral pathway integrators			<i>SPA2</i>	Glyma11g02110			
			<i>SPA2</i>	Glyma05g37070			
			<i>SPA2</i>	Glyma01g43360			
			<i>SPA3</i>	Glyma12g35320			
			<i>SPA3</i>	Glyma12g25240			
			<i>SPA3</i>	Glyma06g37080			
			<i>SPA3</i>	Glyma13g35190			
			<i>COPI</i>	Glyma02g43540			
			<i>COPI</i>	Glyma14g05430			
			<i>LOV1</i>	AT2G02450	<i>NAC035</i>	Glyma07g40140	
					<i>NAC035</i>	Glyma17g00650	
					<i>FEZ</i>	Glyma02g11900	
			<i>RFI2</i>	AT2G47700		Glyma20g38050	
				Glyma10g29230			
				Glyma19g42100			
	<i>SFR6</i>	AT4G04920	<i>SFR6</i>	Glyma13g31480			
			<i>SFR6</i>	Glyma13g24970			
			<i>SFR6</i>	Glyma15g07830			
	<i>TEMI</i>	AT1G25560	<i>TEMI</i>	Glyma01g22260			
			<i>TEMI</i>	Glyma02g11060			
			<i>TEMI</i>	Glyma20g32730			
			<i>TEMI</i>	Glyma10g34760			
		<i>FT/TSF</i>	AT1G65480	<i>FT</i>	Glyma16g26660	<i>GmFT2a</i> , Kong et al. 2010	
			AT4G20370	<i>FT</i>	Glyma16g26690	<i>GmFT2b</i> , Kong et al. 2010	
				<i>FT</i>	Glyma19g28390	<i>GmFT3b</i> , Kong et al. 2010	
				<i>FT</i>	Glyma16g04840	<i>GmFT3a</i> , Kong et al. 2010	
				<i>FT</i>	Glyma08g47820	<i>GmFT6</i> , Kong et al. 2010	
			<i>FT</i>	Glyma16g04830	<i>GmFT5a</i> , Kong et al. 2010		
			<i>FT</i>	Glyma19g28400	<i>GmFT5b</i> , Kong et al. 2010		
			<i>TSF</i>	Glyma18g53690	<i>GmFT1b</i> , Kong et al. 2010		
			<i>TSF</i>	Glyma18g53680	<i>GmFT1a</i> , Kong et al. 2010		
			<i>FT</i>	Glyma08g47810	<i>GmFT4</i> , Kong et al. 2010		
	<i>SOCI</i>	AT2G45660	<i>AGL20</i>	Glyma18g45780			
			<i>AGL20</i>	Glyma09g40230			
			<i>AGL19</i>	Glyma05g03660			
			<i>AGL14</i>	Glyma05g03660			
			<i>AGL42</i>	Glyma20g29300			
	<i>SVP/AGL24</i>	AT2G22540	<i>SVP</i>	Glyma01g02880			
		AT4G24540	<i>SVP</i>	Glyma02g04710			
			<i>SVP</i>	Glyma06g10020			
Floral meristem identity genes	<i>AGL18</i>	AT3G57390	<i>AGL18</i>	Glyma02g33040			
	<i>API</i>	AT1G69120	<i>API</i>	Glyma16g13070			
			<i>API</i>	Glyma08g36380			
			<i>API</i>	Glyma01g08150			
			<i>API</i>	Glyma02g13420			
			<i>AGL8</i>	Glyma06g22650			
			<i>AGL8</i>	Glyma17g08890			

Table 1. (continued)

Category	<i>Arabidopsis</i>	<i>Arabidopsis</i> locus ID	Soybean	Soybean gene model ID	Published soybean gene
<i>TFL1</i>		AT5G03840	<i>AGL8</i>	Glyma05g07380	GmTFL1, Tian <i>et al.</i> 2010/ GmTFL1b, Liu <i>et al.</i> 2010 GmTFL1a, Liu <i>et al.</i> 2010, Tian <i>et al.</i> 2010 <i>GmBFT</i> , Jian <i>et al.</i> 2008
			<i>AGL8</i>	Glyma08g27680	
			<i>TFL1</i>	Glyma19g37890	
			<i>TFL1</i>	Glyma03g35250	
			<i>ATC</i>	Glyma10g08340	
			<i>ATC</i>	Glyma13g22030	
			<i>TFL1</i>	Glyma16g32080	
			<i>TFL1</i>	Glyma09g26550	
			<i>ATC</i>	Glyma13g39360	
			<i>LFY</i>		
<i>LFY</i>	Glyma04g37900				
<i>AP2/TOE2/TOE3/SMZ</i>		AT4G36920 AT5G60120 AT5G67180 AT3G54990	<i>AP2</i>	Glyma01g39520	
			<i>AP2</i>	Glyma17g18640	
			<i>AP2</i>	Glyma05g18170	
			<i>AP2</i>	Glyma11g05720	
			<i>RAP2.7</i>	Glyma19g36200	
			<i>RAP2.7</i>	Glyma15g04930	
			<i>RAP2.7</i>	Glyma13g40470	
			<i>TOE2</i>	Glyma13g40470	
			<i>RAP2.7</i>	Glyma03g33470	
			<i>RAP2.7</i>	Glyma11g15650	
<i>RAP2.7</i>	Glyma12g07800				
<i>RAP2.7</i>	Glyma02g09600				

2009). The presence of the multiple phytochrome genes, along with various alleles exhibiting altered phenotypes, is thought to be related to the adaptation to various environments. Cryptochromes (CRY), blue light-sensitive photoreceptors, are related to circadian clock rhythm through the stabilization of a key photoperiodic gene *Constans* (*CO*) (Liu *et al.* 2008). *Arabidopsis* contains two *CRY* genes, whereas soybean contains seven *CRY* genes. The expression of a soybean homolog, *GmCRY1a* (Glyma04g11010), showed a circadian rhythm pattern (Zhang *et al.* 2008).

TIMING OF CAB EXPRESSION1 (*TOC1*) and *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*) play a pivotal role in the *Arabidopsis* circadian clock (Ding *et al.* 2007). Of the four *TOC1* orthologs identified in soybean, one of them (*GmTOC1*, Glyma06g21120) had been previously cloned (Liu *et al.* 2009). Of the three *CCA1* orthologs identified, Glyma03g42260 had been cloned and designated *GmLCL2* (Liu *et al.* 2009). Both *GmTOC1* and *GmLCL2* exhibited a circadian expression pattern, indicating that their functions were conserved between *Arabidopsis* and soybean (Liu *et al.* 2009). *LATE ELONGATED HYPOCOTYL* (*LHY*) and *CCA1*

repress flowering in short-day (SD) and long-day (LD) conditions, but promote flowering under continuous light conditions. *Arabidopsis* *EARLY FLOWERING3* (*ELF3*) is correlated to photoperiod and photomorphogenesis sensitivity (Zagotta *et al.* 1996). Three soybean *ELF3* orthologs were detected. *Arabidopsis* *CONSTANS* (*CO*) protein regulates flowering time by activating *FLOWERING LOCUS T* (*FT*) (Jarillo and Pineiro 2011). Eight *CO* orthologs were present in soybean. *CYCLING DOF FACTOR 1* (*CDF1*) delays flowering by repressing *CO* transcription in *Arabidopsis* (Fornara *et al.* 2009). One *CDF1* ortholog was identified in soybean. Six orthologs of *FLAVIN-BINDING KELCH REPEAT F-BOX PROTEIN1* (*FKF1*), responsible for regulating *CO* transcription and controlling the stability of *CDF1* (Imai-zumi *et al.* 2005), were identified. *SUPPRESSOR OF PHYA-105 1* (*SPA 1*) is a negative regulator involved in the phyA-specific signalling pathway in a light-dependent manner (Hoecker *et al.* 1998). The *SPA1* and *CONSTITUTIVE PHOTOMORPHOGENIC1* (*COP1*) complex promotes the degradation of *CO* protein in the dark (Jarillo and Pineiro 2011). *Long vegetative phase 1* (*LOV1*) was isolated from a

late-flowering *Arabidopsis* mutant. *LOV1* controls flowering time by negatively, regulating *CO* expression and is related to the cold response (Yoo et al. 2007). *RED AND FAR-RED INSENSITIVE 2 (RFI2)* functions downstream of *PHYA* and *PHYB*. As a result, the mutation impairs several red and far-red light-mediated responses (Chen and Ni 2006). The *sensitive to freezing 6 (srf6)* mutant of *Arabidopsis* exhibits reduced expression of *CCA1*, *TOC1*, and *GIGANTEA* and displays a late-flowering phenotype under long days (Knight et al. 2008).

FLOWERING LOCUS T (FT) is required for flowering and widely conserved among plant species. *FT*, a target of *CO*, is expressed in the leaves and translocated to the shoot apical meristem (Corbesier et al. 2007). The expression of *FT* in *Arabidopsis* and its homolog *Heading date 3a (Hd3a)* in rice was up-regulated in leaves under LD and SD conditions, respectively (Corbesier et al. 2007; Tamaki et al. 2007). *TWIN SISTER OF FT (TSF)* is an *FT* homolog and a direct target of *CO*. Extensive sequence variation in the *TSF* locus was detected in various *Arabidopsis* accessions and this may confer variation in flowering time (Yamaguchi et al. 2005). In soybean, 10 *FT/TSF* homologs were identified. These homologs were previously reported to be located in five different homeologous regions (Kong et al. 2010). The large number of *FT* genes was produced by tandem duplications and whole genome duplication in soybean (Kong et al. 2010). Among these genes, *GmFT2a* and *GmFT5a* were found to be controlled by the *PHYA*-mediated photoperiod response (Kong et al. 2010). *GIGANTEA (GI)* is located in the nucleus and functions upstream of *CO* and *FT* (Fowler et al. 1999). *GI* binds to *FKF1*. This leads to a degradation of *CDF1* by the *FKF1-GI* complex, leading to the repression of *CO* expression and promotion of flowering (Sawa et al. 2007). Three *GI* genes were found in soybean, and one (*GmGIa*, *Glyma10g36600*) was found to be encoded by the previously reported flowering and maturity locus *E2* (Watanabe et al. 2011). A missense mutation found in the *e2* allele caused elevation of soybean *FT (GmFT2a)* mRNA. This indicates that soybean *GI* suppresses *FT* expression (Watanabe et al. 2011). *FT* moves to the shoot apical meristem and activates the expression of *SUPPRESSION OF OVEREXPRESSION OF CO1 (SOC1)*. *SOC1* is crucial to the promotion of flowering by *CO* (Yoo et al. 2005).

AGAMOUS-LIKE 18 (AGL18) encodes a MADS-box-containing protein and represses the floral transition in *Arabidopsis* (Adamezyk et al. 2007). *APETALA1 (API)* and *LEAFY (LFY)* are flower meristem genes that interact with each other (Yant et al. 2009). *AGAMOUS-LIKE 24 (AGL24)* is repressed by *API* in emerging flower primordia (Liu et al. 2007). *TERMINAL FLOWER 1 (TFL1)* is an ortholog of the *Antirrhinum CENTRORADIALIS (CEN)*. *TFL1* functions as a floral repressor by preventing the expression of *LFY* and *API* (Bradley et al. 1996; Liu et al. 2010). The functional

gene underlying the soybean determinate stem (*Dt1*) locus was homologous to *Arabidopsis TFL* and designated *GmTFL1 (Glyma19g37890)* (Liu et al. 2010; Tian et al. 2010). The expression patterns of *Glyma06g17170* and *Glyma03g35250* have been reported (Jian et al. 2008; Liu et al. 2010). Genetic diversity among the minicore collection and the allelic variation at the *GmTFL1* locus reveal the effect of artificial selection on the growth habit (Tian et al. 2010). *APETALA 2 (AP2)* is related to various aspects of plant development including flowering and seed mass control (Jofuku et al. 2005; Ohto et al. 2005). *TARGET OF EARLY ACTIVATION TAGGED 3 (EAT3)* and *SCHLAFMUTZE (SMZ)* are *AP2*-like transcription factors involved in flowering repression (Aukerman and Sakai 2003; Mathieu et al. 2009). *TEMPRANILLO 1 (TEMI)* directly binds to *FT* and abolishes the *FT* expression (Castillejo and Pelaz 2008).

3. Functional DNA variations of flowering-related gene homologs in cultivated and wild soybeans

The genetic diversity in cultivated crops has been decreased by domestication. This has resulted in nucleotide substitutions and phenotypic differences between wild and cultivated crops (Tanksley and McCouch 1997). A genetic bottleneck created by domestication was reported in soybean, and 81% of the rare alleles have been lost (Hyten et al. 2006). Genome resequencing of wild soybean (*G. soja* var. IT182932) revealed 2.5 Mb of nucleotide substitutions and 406 kb of insertions/deletions relative to *G. max* (Williams 82) (Kim et al. 2010). Changes in amino acids can affect protein function and have biological effects. The phenotypic change from indeterminate to determinate growth habit is a good illustration that shows the potential effects of nucleotide substitutions in soybean (Tian et al. 2010). Additionally, wild soybean (*G. soja*) is different from cultivated soybean (*G. max*) in terms of flowering time and growth habit. The wild soybean IT182932, collected from the middle part of Korean Peninsula, shows very late flowering. Williams 82 is a maturity group III cultivar. The difference in days to flower between two genotypes is almost 1 month in South Korea (127° 2' E longitude, 37° 6' N latitude). Thus, functional DNA variations that cause amino acid changes in flowering-related soybean homologs were investigated in *G. max* (Williams 82) and *G. soja* (IT182932) (table 2). Among the 126 flowering-related homologs, 40 genes exhibited nonsynonymous substitutions between Williams 82 and IT182932. The number of amino acid changes varied from one to six. Nearly half of the photoreceptors identified in this study were found to have SNPs between Williams 82 and IT182932. New SNPs not previously reported were identified between IT182932 and Williams 82 at *GmPhyA2 (Glyma20g22160)* encoded by the *E4* locus and at *GmGIa*

Table 2. Functional DNA variations in the flowering-related genes of cultivated (Williams 82) and wild (IT182932) soybeans

Soybean	Soybean gene ID	Nucleotide substitution*	Amino acid change*
<i>PHYA</i>	Glyma20g22160	452 T/C	151 L/S,
<i>PHYB</i>	Glyma09g03990	1967 A/G 2132 T/A	656 H/R, 711 F/Y
<i>PHYB</i>	Glyma15g14980	47 T/C, 48 T/C 1468 A/G, 2402 C/T 2563 C/T	16 V/A, 16 V/A, 490 R/G, 801 A/V, 855 H/Y
<i>PHYE</i>	Glyma15g23400	9231 T/A, 145 T/C 1055 C/A, 1213 G/A 1934 C/T 3221 G/A	31 L/Q, 49 S/P, 352 S/Y, 405 G/S, 645 A/V 1074 R/H
<i>CRY1</i>	Glyma14g35020	1468 G/T 1566 A/T	490 A/S, 522 E/D
<i>CRY2</i>	Glyma20g35220	1556 T/C 1558 G/T 1579 T/G	519 V/A, 520 V/F, 527 S/A
<i>CRY2</i>	Glyma02g00830	80 A/T, 486 G/C 1495 C/A	27 K/M, 162 M/I, 499 H/N
<i>TOC1</i>	Glyma06g21120	131 T/C 1376 T/G	44 L/S, 459 I/S
<i>TOC1</i>	Glyma17g11040	1047 T/C	349 I/T
<i>TOC1</i>	Glyma05g00880	518 T/C	173 L/S
<i>GI</i>	Glyma10g36600	658 A/G	220 I/V
<i>LHY</i>	Glyma19g45030	2023 T/A	675 C/S
<i>LHY</i>	Glyma07g05410	256 T/G 1481 G/T 1863 C/G	86 S/A, 494 G/V, 621 D/E
<i>ELF3</i>	Glyma04g05280	217 A/G, 518 A/G	73 R/G, 173 Q/R
<i>ELF3</i>	Glyma14g10530	29 C/T	10 S/L
<i>CO</i>	Glyma19g05170	367 A/G	123 K/E
<i>COL2</i>	Glyma18g51320	648 G/C	216 E/D
<i>COL5</i>	Glyma13g01290	1090 A/T	364 T/S
<i>COL4</i>	Glyma04g06240	602 C/T	201 S/L
<i>SPA2</i>	Glyma05g37070	400 A/C, 404 T/C, 647 A/G, 694 G/T 1194 A/C	134 K/Q, 135 I/T, 216 Q/R, 232 A/S, 398 K/N
<i>SPA2</i>	Glyma01g43360	899 G/C	300 C/S
<i>SPA3</i>	Glyma12g35320	2327 A/G	776 H/R
<i>SPA3</i>	Glyma12g25240	1382 A/T	461 Y/F
<i>COP1</i>	Glyma02g43540	977 G/A	326 S/N
<i>SFR6</i>	Glyma13g24970	1417 G/A	473 V/I
<i>TEM1</i>	Glyma01g22260	19 C/A, 400 T/G	7 L/M, 134 S/A
<i>TEM1</i>	Glyma20g32730	89 T/G, 551 A/C	30 L/R, 184 Q/P
<i>AGL19</i>	Glyma05g03660	652 A/G	218 T/A
<i>AGL18</i>	Glyma02g33040	405 A/C	135 E/D
<i>API</i>	Glyma01g08150	614 G/C	205 R/P
<i>API</i>	Glyma02g13420	398 T/A	133 V/D
<i>AGL8</i>	Glyma05g07380	667 T/A	223 C/S
<i>TFL1</i>	Glyma10g08340	17 C/A	6 T/K
<i>ATC</i>	Glyma13g39360	332 T/C	111 L/P
<i>AP2</i>	Glyma17g18640	65 G/A	22 C/Y
<i>AP2</i>	Glyma11g05720	407 C/A	136 T/K
<i>RAP2.7</i>	Glyma19g36200	139 G/T	47 A/S
<i>RAP2.7</i>	Glyma15g04930	1330 A/T	444 I/F
<i>RAP2.7</i>	Glyma03g33470	112 T/C, 316 C/T	38 Y/H, 106 P/S
<i>RAP2.7</i>	Glyma02g09600	317 T/C	106 V/A

*The number prior to nucleotide/amino acid change indicates the position of the variation on the corresponding gene.

(Glyma10g36600) encoded by the *E2* locus. Loss-of-function alleles at the *E4* locus (*GmPhyA2*) has been reported to promote photoperiod insensitivity and to result in early flowering under LD condition during early

summer (Liu *et al.* 2008). A null mutant harbouring premature stop codon in *GmGla* (Glyma10g36600) showed an earlier flowering phenotype than its wild genotype (Watanabe *et al.* 2011; Shin and Lee 2012). Compared to the

wild soybean IT182932, amino acid changes in *GmPhyA2* and *GmGla* may promote flowering in Williams 82. Interestingly, no SNPs were found between Williams 82 and IT182932 at the 10 *FT* homologs, key components of the flowering signal and highly conserved genes. Additionally, homologs for five *Arabidopsis* genes (*CDF/FKF*, *LOV*, *RF12*, *SVP/AGL24*, and *LFY*) did not have SNPs between Williams 82 and IT182932. This information provides a good source for an investigation into the different flowering patterns of wild and domesticated soybean. Further study at the population level is necessary to uncover the exact mechanism.

4. Co-localization of the flowering genes with QTLs

To speculate concerning the function of the identified soybean homologs to the *Arabidopsis* genes involved in the light-dependent flowering pathway, the genes were investigated to determine whether they co-localized with the previously reported QTLs for flowering time and maturity. The QTLs associated with flowering time were retrieved from the Soybase Web site (<http://www.soybase.org/dlpages/index.php>). This included 36 loci for six traits (flowering time, first flower, pod maturity, beginning of pod, reproductive period and seed filling period) (figure 1). The chromosomal distribution of these QTLs is shown in figure 1. Several QTL clusters associated with flowering time were observed on chromosome (chr) 6, 7, and 19. On chr 6, 10 QTLs for pod maturity, first flower, reproductive period and flowering time were clustered in the middle and end of the chromosome. On the long arm of chr 7, 10 QTLs for five traits exception to flowering time were clustered. There was also a cluster of 13 QTLs in the distal arm of chr 19. Highly clustered QTL regions may be the result of pleiotropic genes involved in several flowering traits or a cluster of genes working as a functional unit. In *Arabidopsis*, rearrangement within MAP2-5, a MADS-box multigene family cluster, is suspected to be involved in the variation of ecologically important phenotypes like flowering time (Caicedo et al. 2009; Rosloski et al. 2010). The pleiotropic effects of flowering-related genes have been previously reported. In *Arabidopsis*, *FLOWERING LOCUS C* was reported to regulate both flowering time and seed germination (Chiang et al. 2009). In rice, *OsEF3*, the homolog of *Arabidopsis* *ELF3*, regulates heading date, root development and 1000-grain weight. These functions are all different from the gene function of *Arabidopsis* *ELF3*. The *Ghd8* QTL plays a role in regulating grain productivity, plant height and heading date (Fu et al. 2009; Yan et al. 2011). These pleiotropic effects of QTLs may result from an upstream key regulator of a specific developmental pathway, similar to the manner in which *Ghd8* influences the traits of post-vegetative stages.

To determine the co-localization of the flowering time-related genes with the QTLs, a 2 Mb region flanking the marker linked to the QTL was determined. The genetic distance of QTLs is highly variable due to the recombination frequency being affected by population size. The QTL physical locations are seldom determined, even if the linked marker sequences and their genomic positions are known. Therefore, this study used a relatively wide 4 Mb span surrounding the marker linked to the QTL. This enabled the consideration of as many potential genes as possible. The physical location of the markers linked to the QTLs were downloaded from Soybase (<http://www.soybase.org/dlpages/index.php>), with only SSR and RFLP marker locations available (Grant et al. 2010).

Among 114 genes related to the flowering pathway, 22 genes were located within the QTL regions controlling flowering and maturity (table 3). These genes may be highly associated with the corresponding QTL traits. Notably, the genes in the highly clustered QTL regions of chr 6, 7 and 19 are candidates of key regulators for flowering time-associated traits (table 3).

Two genes homologous to *SVP* and *CRY1* and the gene homolog of *AGL18* co-localized with the *Podmat13-3* and *Podmat13-4* QTLs for pod maturity on chr 6, respectively. Two genes, *Glyma07g05410* (*LHY/CCA1*) and *Glyma07g06420* (*SPA1/COP1*), were within the regions surrounding eight QTLs for first flower, pod maturity, and reproductive period on chr 7. On chr 19, the nine QTL regions for first flower, pod maturity, reproductive period, and flowering time harbored five genes homologous to *AP2/TOE3/SMZ*, *FT/TSF/TFL1*, *PHY A/B/C*, *RF12* and *LHY/CCA1*. The homologs of *LHY/CCA1* were simultaneously detected in the QTL regions on chr 7 and 19. This result indicates that the soybean *LHY/CCA1* homologs, *Glyma07g05410* and *Glyma19g45030*, might be involved in regulating flowering time. Moreover, the discovery of clusters of flowering-related gene homologs in the QTL cluster region of chr 19 may support the hypothesis that gene clusters of functional units may be involved in the control of flowering time.

In individual QTL regions associated with flowering, homologs of flowering genes were detected. The homologs of *AGL18* and *CRY 1/2* existed in the pod maturity QTL regions on chr 1 and 4, respectively. These homologs were also detected in the pod maturity QTL on chr 6. These genes may participate in the regulation of pod maturity. In the region containing the *Podmat9-1* and *Reprod3-1* QTLs, the homologs of *LHY/CCA1* and *SPA1/COP1* were detected on chr 16 and also in relation to the QTLs for the common traits of pod maturity and reproductive period on chr 7. This observation indicates that the *LHY/CCA1* and *SPA1/COP1* homologs may influence pod maturity and reproductive period. A *CCA1* homolog was reportedly expressed in soybean

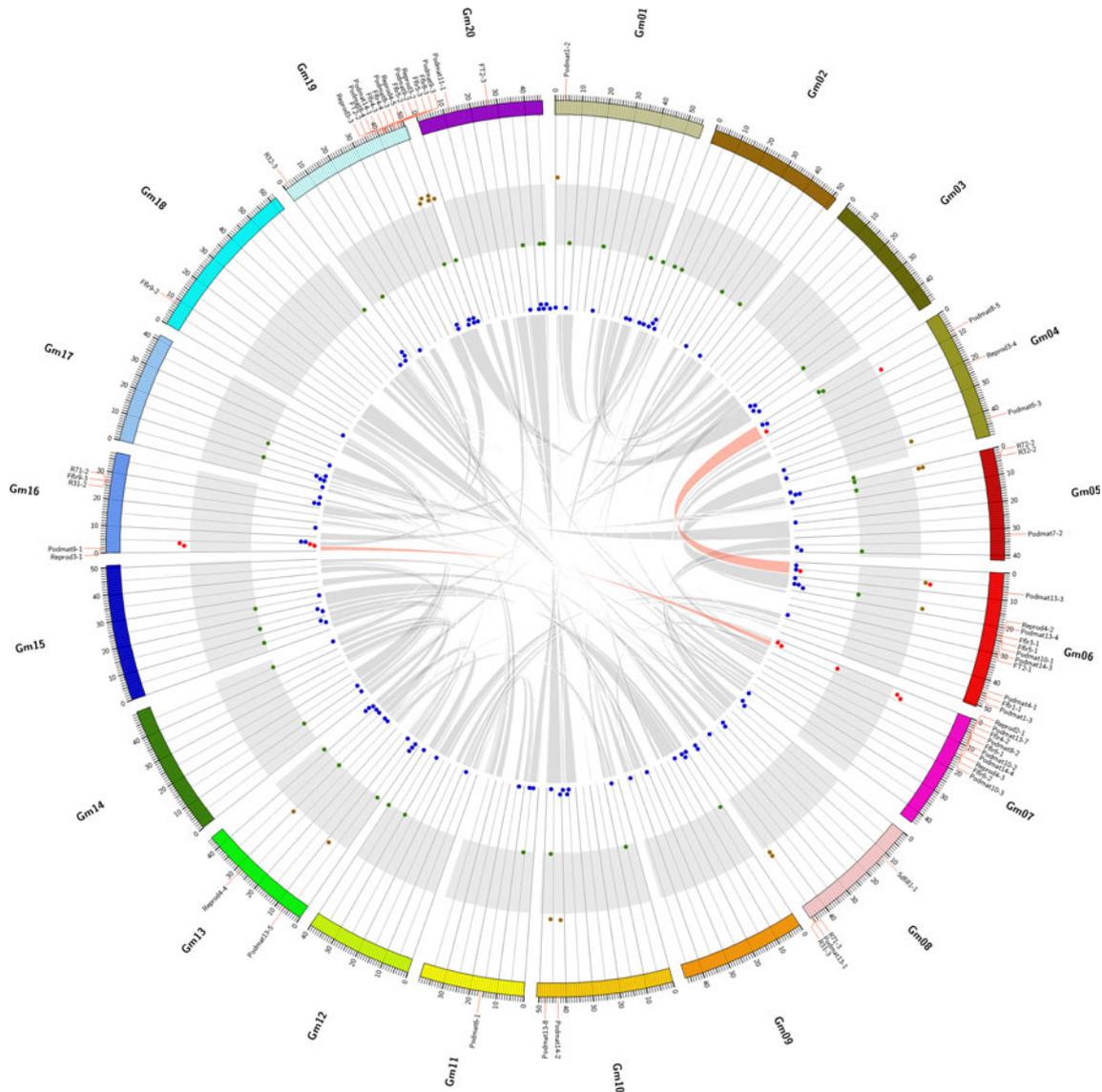


Figure 1. Summary of the chromosomal distribution of the soybean homologs of the *Arabidopsis* flowering-related genes and the QTLs for flowering time and maturity. From the inside to the outside, the circles indicate recent-duplicated genomic regions, 114 gene homologs, 40 genes with nonsynonymous SNPs between cultivated (Williams 82) and wild (IT182932) soybeans, 22 genes co-localizing with the QTLs, and QTLs associated with flowering time and maturity, respectively. Among 118 genes, one and two genes, represented by red dots, were conserved in two recently-duplicated genomic regions (reddish pink) on chr 4 vs chr 6 and chr 7 vs chr 16, respectively, and these regions contained the QTLs for flowering time and maturity simultaneously. Thirty-six loci representing six traits (flowering time, first flower, pod maturity, pod initiation, reproductive period, and seed filling period) were displayed. The figure was created using the circular genome data visualization software Circos (<http://circos.ca/>).

seeds with circadian rhythm. This gene, having a predicted function of protein synthesis, fatty acid metabolism and photosynthesis, was also expressed in soybean seeds with circadian rhythm oscillation (Hudson 2010). These previous report indicate that the circadian rhythm may affect the pod maturation of seed development and support the putative role of the *CCA1* homolog near the pod maturity QTL in this study.

Of the 22 genes overlapping the QTL chromosomal locations, four genes were found to encode protein sequence differences between cultivated (Williams 82) and wild (IT182932) soybeans (tables 2 and 3). *Glyma07g5410* and *Glyma19g45030*, homologs of *LHY/CCA1*, contained three and one amino acid changes, respectively. These two genes were surrounded by the QTL clusters on chr 7 and chr 19. This finding indicates excellent candidates for the regulation

Table 3. The list of the genes co-localized with the QTLs for flowering time and maturity

Chr	QTL	Position (bp)*	Flowering time-related gene ID	Annotation
Gm01	Podmat1–2	3772148...3772781	Glyma01g02880	SVP
Gm04	Podmat8–5	7819345...819499	Glyma04g11010	CRY1/2
Gm04	Podmat6–3	42916785...42917215	Glyma04g37900	LFY
Gm05	R32–2, R72–2	3442359...3442620	Glyma05g03660	SOC1
Gm06	Podmat13–3	7320610...7320826	Glyma06g10020 Glyma06g10830	SVP CRY1
Gm06	Podmat13–4	20018743...20018949	Glyma06g22650	AGL18
Gm07	Fflr4–2, Podmat13–7, Podmat8–2, Reprod2–1	2414416...2414616	Glyma07g05410	LHY/CCA1
Gm07	Fflr6–1, Podmat10–2, Podmat14–4, Reprod4–3	4510428...4510539	Glyma07g05410 Glyma07g06420	LHY/CCA1 SPA1/COP1
Gm08	Podmat13–1, R31–3, R71–3	46365992...46366258	Glyma08g47810 Glyma08g47820	FT/TSF/TFL1 FT/TSF/TFL1
Gm10	Podmat14–2	42983781...42984021	Glyma10g34760 Glyma10g36600	TEM1 GI
Gm13	Podmat13–5	6747258...6747869	Glyma13g07030	CO
Gm13	Reprod4–4	28185673...28185985	Glyma13g24970	SFR6
Gm16	Podmat9–1 Reprod3–1	1631658...1631885	Glyma16g01980 Glyma16g03030	LHY/CCA1 SPA1/COP1
Gm19	Fflr4–3 Fflr4–4, Podmat8–4 Podmat8–3 Reprod4–5 Fflr5–2, Podmat9–2 Reprod3–2 FT2–2 Fflr5–3, Podmat9–3, Reprod3–3,	42110146...42110419 42415414...42415665 42835184...42835381 43523447...43523702 46620622...46621239 47049000...47049225 49191231...49191478	Glyma19g36200 Glyma19g37890 Glyma19g41210 Glyma19g42100 Glyma19g45030	AP2/TOE3/SMZ FT/TSF/TFL1 PHY A/B/C RFI2 LHY/CCA1

*The genes were located within 4-Mb genomic regions surrounding the marker linked to the QTLs associated with flowering time and maturity.

of the flowering pathway in soybean. LHY, CCA1 and TOC1 are all key genes of circadian clock (Ding *et al.* 2007). In *Arabidopsis*, TOC1 regulates the floral transition in a LHY/CCA1-dependent manner and LHY/CCA1 functions upstream of TOC1 in regulating a photomorphogenic process. Mutants at these loci exhibit early flowering than wild types, resulting from circadian defect (Ding *et al.* 2007; Mizoguchi *et al.* 2002; Strayer *et al.* 2000). Thus, amino acid changes of LHY/CCA1 are supposed to be associated with

difference in flowering time between Williams 82 and IT182932. A nonsynonymous SNP (A/G) causing Ile→Val in the *GI* gene homolog (Glyma10g36600) was found co-localized with the *Podmat14–2* QTL. In addition, the *SFR6* gene (Glyma13g24970) overlapped with the chromosomal region of the *Redprod 4–4* QTL and was observed to carry the amino acid change of Val→Ile via the DNA variation of G→A. This discovery is supportive of *GI* and *SFR6* genes as candidates involved in the soybean flowering pathway.

Table 4. Flowering-related genes and QTLs conserved in recently duplicated genomic regions

Recently duplicated region ID	Chr	Position (bp)	Gene ID	QTL	Chr	Position (bp)	Gene ID	QTL
12398025	Gm07	2782652.. 6242275	Glyma07g05410 (LHY/CCA1) Glyma07g06420 (SPA1/COPI)	Reprod2-1	Gm16	247496.. 3305627	Glyma16g01980 (LHY/CCA1) Glyma16g03030 ((SPA1/COPI))	Podmat9-1
				Reprod4-3				Reprod3-1
				Fflr4-2				
				Fflr6-1				
				Podmat10-2				
				Podmat13-7				
				Podmat14-4				
				Podmat8-2				
23546507	Gm04	7268.. 998668	Glyma04g11010 (CRY1)	Podmat8-5	Gm06	24815.. 8483404	Glyma06g10830 (CRY1)	Podmat13-3

5. Genetic redundancy of the flowering genes and QTLs in recently duplicated genomic regions

To obtain additional bioinformatic support for the prediction of putative functions for the identified soybean flowering-related gene homologs, the genomic regions that typically show conserved gene contents were investigated. Of the recently duplicated sequence regions in the soybean genome (downloaded from Phytozome), the regions retaining the genes involved in the flowering pathway and also the QTLs for flowering time and maturity were selected. Two recently duplicated regions were found to exhibit redundancy of both flowering-related genes and the QTLs; their identity numbers are ID 12398025 and ID 23546507 (table 4). ID 12398025 contains the homeologous regions between chr 7 and chr 16. These regions are redundant for the *LHY/CCA1* and *SPA1* homologs and the QTLs controlling pod maturity, first flower and reproductive period. Retaining of the *CRY1* gene and the pod maturity QTLs also occurred in the duplicated region of ID 23546507 on chr 4 and chr 6. Interestingly, functional DNA variations of the *LHY/CCA1* gene (*Glyma07g05410*) were present in a counterpart of the chr 7-duplicated region (table 2). In contrast, the gene (*Glyma16g01980*) in the other portion of the duplicated region on chr 16 did not show any functional change at the sequence level. This functional sequence variation between two paralogs may influence functional diversification of the corresponding gene. The findings presented here indicate that the conserved *LHY/CCA1*, *SPA1* and *CRY1* genes in the homeologous regions on chrs 7 vs 16 and chrs 4 vs 6 may influence reproductive development including pod maturity.

The conservation of multiple QTLs across homeologous regions has been observed for other traits in soybean. This includes soybean cyst nematode resistance, corn earworm resistance, seed protein content, seed size, yield and oil

content (Kim *et al.* 2009b; Shin *et al.* 2008). Also, a homeologous tetrad in the genomic region surrounding the bacterial leaf pustule resistance gene (*rxp*) was identified and generated by two rounds of duplication, and harboured the conserved QTLs for disease resistance and seed size (Kim *et al.* 2009b). QTL mapping of flowering time at three different latitudes indicated that the flowering time QTLs in the duplicated regions have maintained ancestral gene function for the control of flowering time. These QTLs have diverged in an environment-specific manner (Liu *et al.* 2011). A paralog (*GmPhyA2*) of the *phytochrome A* genes, encoded by the *E4* locus for photoperiod insensitivity on chr 20, shares the homologous regions with chr 10 (Liu *et al.* 2008). Together with *GmPhyA3* being responsible for the *E3* locus controlling maturity on chr 19, genetic redundancy of the *phytochrome A* gene was due to complex genome constitution by genome duplication. Variations in these genes may contribute to the diverse flowering response and maturity time in soybean (Liu *et al.* 2008; Watanabe *et al.* 2009).

In *Brassica* species, intensive studies have revealed relationships between genes or genome duplication events and the functional conservation/diversification of flowering time. It is thought that these multiple flowering time QTLs may represent copies of a single ancestral gene, quite possibly a homolog of *A. thaliana* *CONSTANS (CO)* (Axelsson *et al.* 2001). The polyploid *Brassica* genome contains four copies of *FLOWERING LOCUS C (FLC)* that correspond to several flowering time loci on different linkage groups (Kole *et al.* 2001; Schranz *et al.* 2002). These *FLC* paralogs may perform similar functions in the regulation of flowering time, and multiple genes may have an additive effect (Schranz *et al.* 2002). In addition, three *FLOWERING LOCUS T* paralogs are associated with two major QTL clusters for flowering time in *B. napus*. These clusters modulate functional differences in the flowering time between winter and spring cultivars of oilseed *Brassica* (Wang *et al.* 2009).

6. Conclusion

As the immense quantity of information on genome sequence and gene contents in crops, as well as model plants, accumulates and computational tools are rapidly developed, new avenues become available to decipher the genetic elements involved in the control of complex traits. These include traits such as domestication syndrome and disease resistance. Furthermore, great efforts have attempted to integrate current biological and genetic knowledge with genomic information. A good example is the Soybase Web site. This Web site permitted a new approach for the identification of the most likely candidate genes of target traits using synthesized information from both top-down and bottom-up methods (Blackman *et al.* 2011).

In this review, soybean genes homologous to genes in *A. thaliana* involved in the flowering pathway were examined. Functional nucleotide variations in these genes between cultivated (Williams 82) and wild (IT182932) soybeans were investigated, and the subset co-localizing with the relevant QTLs constituted a group of candidates controlling flowering time. As additional criterion, conservation of the genes and QTLs in recently duplicated genomic regions was applied to develop evidence in support of the candidacy of these genes. The regulation of the timing of transition from the vegetative to the reproductive stage is a major goal for plant breeding. This regulation could enable the development of novel varieties that are better adapted to challenging environments and climate conditions. Even though the gene regulatory network controlling flowering time is well described in model plants, limited studies using a candidate gene approach have been conducted in soybean. A few genes involved in flowering and maturity have been previously isolated in soybean (Lie *et al.* 2008; Watanabe *et al.* 2009, 2011). Thus, the gene list catalogued in this text provides a primary insight into understanding the regulation of flowering time and maturity in soybean. This information should help in the determination of target gene as the starting point for molecular biological research. Also, this analytical process can be applied to identify the candidates of a trait of interest under various contexts. This provides the potential for rapidly gaining new insights into challenging biological pathways.

Acknowledgements

This research was supported by a grant from the Next Generation BioGreen 21 Program (Code No. PJ00811701), Rural Development Administration, the Republic of Korea.

References

Adamczyk BJ, Lehti-Shiu MD and Fernandez DE 2007 The MADS domain factors *AGL15* and *AGL18* act redundantly as repressors of the floral transition in *Arabidopsis*. *Plant J.* **50** 1007–1019

- Aukerman MJ and Sakai H 2003 Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. *Plant Cell* **15** 2730–2741
- Axelsson T, Shavorskaya O and Lagercrantz U 2001 Multiple flowering time QTLs within several *Brassica* species could be the result of duplicated copies of one ancestral gene. *Genome* **44** 856–864
- Bernard RL 1971 Two major genes for time of flowering and maturity in soybeans. *Crop Sci.* **11** 242–244
- Blackman BK, Rasmussen DA, Strasburg JL, Raduski AR, Burke JM, Knapp SJ, Michaels SD and Rieseberg LH 2011 Contributions of flowering time genes to sunflower domestication and improvement. *Genetics* **187** 271–287
- Bonato ER and Vello NA 1999 *E6*, a dominant gene conditioning early flowering and maturity in soybeans. *Genet. Mol. Biol.* **22** 229–232
- Bradley D, Carpenter R, Copsey L, Vincent C, Rothstein S and Coen E 1996 Control of inflorescence architecture in *Antirrhinum*. *Nature* **379** 791–797
- Buzzel RI 1971 Inheritance of a soybean flowering response to fluorescent-daylength conditions. *Can. J. Genet. Cytol.* **13** 703–707
- Buzzel RI and Voldeng HD 1980 Inheritance of insensitivity to long day length. *Soybean Genet. Newsl.* **7** 26–29
- Caicedo AL, Richards C, Ehrenreich IM and Purugganan MD 2009 Complex rearrangements lead to novel chimeric gene fusion polymorphisms at the *Arabidopsis thaliana* *MAF2–5* flowering time gene cluster. *Mol. Biol. Evol.* **26** 699–711
- Castillejo C and Pelaz S 2008 The balance between *CONSTANS* and *TEMPRANILLO* activities determines *FT* expression to trigger flowering. *Curr. Biol.* **18** 1338–1343
- Chen M and Ni M 2006 *RED AND FAR-RED INSENSITIVE 2*, a RING-domain zinc finger protein, mediates phytochrome-controlled seedling deetiolation responses. *Plant Physiol.* **140** 457–465
- Chen M, Chory J and Fankhauser C 2004 Light signal transduction in higher plants. *Annu. Rev. Genet.* **38** 87–117
- Chiang GCK, Barua D, Kramer EM, Amasino RM and Donohue K 2009 Major flowering time gene, *FLOWERING LOCUS C*, regulates seed germination in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **106** 11661–11666
- Cober ER and Voldeng HD 2001 A new soybean maturity and photoperiod-sensitivity locus linked to *E1* and *T*. *Crop Sci.* **41** 698–701
- Cober ER, Molnar SJ, Charette M and Voldeng HD 2010 A new locus for early maturity in soybean. *Crop Sci.* **50** 524–527
- Cober ER, Tanner JW and Voldeng HD 1996a Genetic control of photoperiod response in early-maturing, near-isogenic soybean lines. *Crop Sci.* **36** 601–605
- Cober ER, Tanner JW and Voldeng HD 1996b Soybean photoperiod-sensitivity loci respond differentially to light quality. *Crop Sci.* **36** 606–610
- Cockram J, Jones H, Leigh FJ, O'Sullivan D, Powell W, Laurie DA and Greenland AJ 2007 Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. *J. Exp. Bot.* **58** 1231–1244
- 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

- Ding Z, Doyle MR, Amasino RM and Davis SJ 2007 A complex genetic interaction between *Arabidopsis thaliana* *TOC1* and *CCA1/LHY* in driving the circadian clock and in output regulation. *Genetics* **176** 1501–1510
- Doebley JF, Gaut BS and Smith BD 2006 The molecular genetics of crop domestication. *Cell* **127** 1309–1321
- Fornara F, Panigrahi KC, Gissot L, Sauerbrunn N, Ruhl M, Jarillo JA and Coupland G 2009 *Arabidopsis* DOF transcription factors act redundantly to reduce *CONSTANS* expression and are essential for a photoperiodic flowering response. *Dev. Cell* **17** 75–86
- Fowler S, Lee K, Onouchi H, Samach A, Richardson K, Coupland G and Putterill J 1999 *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. *EMBO J.* **18** 4679–4688
- Fu C, Yang X O, Chen X, Chen W, Ma Y, Hu J and Li S 2009 *OsEF3*, a homologous gene of *Arabidopsis* *ELF3*, has pleiotropic effects in rice. *Plant Biol.* **11** 751–757
- Fuller DQ 2007 Contrasting patterns in crop domestication and domestication rates: Recent archaeobotanical insights from the old world. *Annal. Bot.* **100** 903–924
- Grant D, Nelson RT, Cannon SB and Shoemaker RC 2010 SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* **38** D843–846
- Greenup A, Peacock WJ, Dennis ES and Trevaskis B 2009 The molecular biology of seasonal flowering-responses in *Arabidopsis* and the cereals. *Ann. Bot.* **103** 1165–1172
- Hecht V, Knowles CL, Vander Schoor JK, Liew LC, Jones SE, Lambert MJ and Weller JL 2007 Pea *LATE BLOOMER1* is a *GIGANTEA* ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. *Plant Physiol.* **144** 648–661
- Hoecker U, Xu Y and Quail PH 1998 *SPA1*: a new genetic locus involved in phytochrome A-specific signal transduction. *Plant Cell* **10** 19–33
- Hudson KA 2010 The circadian clock-controlled transcriptome of developing soybean seeds. *The Plant Genome* **3** 1–11
- Hymowitz T 1970 On the domestication of the soybean. *Economic Bot.* **24** 408–421
- Hyten DL, Song QJ, Zhu YL, Choi IY, Nelson RL, Costa JM, Specht JE, Shoemaker RC and Cregan PB 2006 Impacts of genetic bottlenecks on soybean genome diversity. *Proc. Natl. Acad. Sci. USA* **103** 16666–16671
- Imaizumi T, Schultz TF, Harmon FG, Ho LA and Kay SA 2005 *FKF1* F-box protein mediates cyclic degradation of a repressor of *CONSTANS* in *Arabidopsis*. *Science* **309** 293–297
- Itoh H, Nonoue Y, Yano M and Izawa T 2010 A pair of floral regulators sets critical day length for *Hd3a* florigen expression in rice. *Nature Genet.* **42** 635–638
- Izawa T, Mihara M, Suzuki Y, Gupta M, Itoh H, Nagano AJ, Motoyama R, Sawada Y, Yano M, Hirai MY, Makino A and Nagamura Y 2011 *Os-GIGANTEA* confers robust diurnal rhythms on the global transcriptome of rice in the field. *Plant Cell* **23** 1741–1755
- Izawa T, Takahashi Y and Yano M 2003 Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. *Curr. Opin. Plant Biol.* **6** 113–120
- Jarillo JA and Pineiro M 2011 Timing is everything in plant development. The central role of floral repressors. *Plant Sci.* **181** 364–378
- Jian B, Liu B, Bi YR, Hou WS, Wu CX and Han TF 2008 Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Mol. Biol.* **9** 59
- Jofuku KD, Omidyar PK, Gee Z and Okamoto JK 2005 Control of seed mass and seed yield by the floral homeotic gene *APE-TALA2*. *Proc. Natl. Acad. Sci. USA* **102** 3117–3122
- Jung C and Muller AE 2009 Flowering time control and applications in plant breeding. *Trends Plant Sci.* **14** 563–573
- Kim DH, Doyle MR, Sung S and Amasino RM 2009 Vernalization: winter and the timing of flowering in plants. *Annu. Rev. Cell Dev. Biol.* **25** 277–299
- Kim KD, Shin JH, Van K, Kim DH and Lee SH 2009 Dynamic rearrangements determine genome organization and useful traits in soybean. *Plant Physiol.* **151** 1066–1076
- Kim MY, Lee S, Van K, Kim TH, Jeong SC, Choi IY, Kim DS, Lee YS, *et al.* 2010 Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. *Proc. Natl. Acad. Sci. USA* **107** 22032–22037
- Knight H, Thomson AJ and McWatters HG 2008 *SENSITIVE TO FREEZING6* integrates cellular and environmental inputs to the plant circadian clock. *Plant Physiol.* **148** 293–303
- Kole C, Quijada P, Michaels SD, Amasino RM and Osborn TC 2001 Evidence for homology of flowering-time genes *VFR2* from *Brassica rapa* and *FLC* from *Arabidopsis thaliana*. *Theor. Appl. Genet.* **102** 425–430
- Kong F, Liu B, Xia Z, Sato S, Kim BM, Watanabe S, Yamada T, Tabata S, Kanazawa A, 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 Physiol.* **154** 1220–1231
- 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
- Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, Xia Z, Nagamatsu A, Arai M, *et al.* 2010 The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis* *TERMINAL FLOWER1*. *Plant Physiol.* **153** 198–210
- Liu C, Zhou J, Bracha-Drori K, Yalovsky S, Ito T and Yu H 2007 Specification of *Arabidopsis* floral meristem identity by repression of flowering time genes. *Development* **134** 1901–1910
- Liu H, Wang HG, Gao PF, Xu JH, Xu TD, Wang JS, Wang BL, Lin CT and Fu YF 2009 Analysis of clock gene homologs using unifoliolates as target organs in soybean (*Glycine max*). *J. Plant Physiol.* **166** 278–289
- Liu W, Kim MY, Kang YJ, Van K, Lee YH, Srinives P, Yuan DL, Lee S-H 2011 QTL identification of flowering at three different latitudes reveals homeologous genomic regions that control flowering in soybean. *Theor. Appl. Genet.* **123** 545–553
- Mathieu J, Yant LJ, Murdter F, Kuttner F and Schmid M 2009 Repression of flowering by the miR172 target SMZ. *PLoS Biol.* **7** e1000148
- McBlain BA and Bernard RL 1987 A new gene affecting the time of flowering and maturity in soybean. *J. Hered.* **78** 160–162

- Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song HR, Carré IA and Couplang G 2002 LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in *Arabidopsis*. *Dev. Cell* **2** 629–641
- Ohto MA, Fischer RL, Goldberg RB, Nakamura K and Harada JJ 2005 Control of seed mass by *APETALA2*. *Proc. Natl. Acad. Sci. USA* **102** 3123–3128
- Reed JW, Nagatani A, Elich TD, Fagan M and Chory J 1994 Phytochrome-A and Phytochrome-B have overlapping but distinct functions in *Arabidopsis* development. *Plant Physiol.* **104** 1139–1149
- Rosloski SM, Jali SS, Balasubramanian S, Weigel D and Grbic V 2010 Natural diversity in flowering responses of *Arabidopsis thaliana* caused by variation in a tandem gene array. *Genetics* **186** 263–276
- Roux F, Touzet P, Cuguen J and Corre V L 2006 How to be early flowering: an evolutionary perspective. *Trends Plant Sci.* **11** 375–381
- Sawa M, Nusinow DA, Kay SA and Imaizumi T 2007 *FKF1* and *GIGANTEA* complex formation is required for day-length measurement in *Arabidopsis*. *Science* **318** 261–265
- Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ and Shoemaker RC 2004 Mining EST databases to resolve evolutionary events in major crop species. *Genome* **47** 868–876
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, et al. 2010 Genome sequence of the palaeopolyploid soybean. *Nature* **463** 178–183
- Schranz ME, Quijada P, Sung SB, Lukens L, Amasino R and Osborn TC 2002 Characterization and effects of the replicated flowering time gene *FLC* in *Brassica rapa*. *Genetics* **162** 1457–1468
- Shin JH, Van K, Kim DH, Kim KD, Jang YE, Choi BS, Kim MY and Lee SH 2008 The lipoxigenase gene family: a genomic fossil of shared polyploidy between *Glycine max* and *Medicago truncatula*. *BMC Plant Biol.* **8** 133
- Shin JH and Lee S-H 2012 Molecular markers for the E2 and E3 genes controlling flowering and maturity in soybean. *Mol. Breed.* DOI 10.1007/s11032-012-9743-6
- Shoemaker RC, Polzin K, Labate J, Specht J, Brummer EC, Olson T, Young N, Concibido V, et al. 1996 Genome duplication in soybean (*Glycine* subgenus *soja*). *Genetics* **144** 329–338
- Strayer C Oyama T, Schultz TF, Raman R, Somers DE, Más P, Panda S, Kreps JA and Kay SA 2000 Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* **4** 768–771
- Takano M, Inagaki N, Xie XZ, Yuzurihara N, Hihara F, Ishizuka T, Yano M, Nishimura M, et al. 2005 Distinct and cooperative functions of phytochromes A, B, and C in the control of deetiolation and flowering in rice. *Plant Cell* **17** 3311–3325
- Tamaki S, Matsuo S, Wong HL, Yokoi S and Shimamoto K 2007 Hd3a protein is a mobile flowering signal in rice. *Science* **316** 1033–1036
- Tanksley SD and McCouch SR 1997 Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* **277** 1063–1066
- Tian Z, Wang X, Lee R, Li Y, Specht JE, Nelson RL, McClean PE, Qiu L and Ma J 2010 Artificial selection for determinate growth habit in soybean. *Proc. Natl. Acad. Sci. USA* **107** 8563–8568
- Wang J, Long Y, Wu BD, Liu J, Jiang CC, Shi L, Zhao JW, King GH and Meng JL 2009 The evolution of *Brassica napus* *FLOWERING LOCUST* paralogues in the context of inverted chromosomal duplication blocks. *BMC Evol. Biol.* **9** 271–284
- 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
- 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 aresidual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics* **188** 395–407
- Yamaguchi A, Kobayashi Y, Goto K, Abe M and Araki T 2005 *TWIN SISTER OF FT (TSF)* acts as a floral pathway integrator redundantly with *FT*. *Plant Cell Physiol.* **46** 1175–1189
- Yan WH, Wang P, Chen HX, Zhou HJ, Li QP, Wang CR, Ding ZH, Zhang YS, Yu SB, Xing YZ and Zhang QF 2011 A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol. Plant* **4** 319–330
- Yant L, Mathieu J and Schmid M 2009 Just say no: floral repressors help *Arabidopsis* bide the time. *Curr. Opin. Plant Biol.* **12** 580–586
- Yoo SK, Chung KS, Kim J, Lee JH, Hong SM, Yoo SJ, Yoo SY, Lee JS and Ahn JH 2005 *CONSTANS* activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to promote flowering in *Arabidopsis*. *Plant Physiol.* **39** 770–778
- Yoo SY, Kim Y, Kim SY, Lee JS and Ahn JH 2007 Control of flowering time and cold response by a NAC-domain protein in *Arabidopsis*. *PLoS One* **2** e642
- Zagotta MT, Hicks KA, Jacobs CI, Young JC, Hangarter RP and Meeks-Wagner DR 1996 The *Arabidopsis* *ELF3* gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. *Plant J.* **10** 691–702
- Zhang Q, Li H, Li R, Hu R, Fan C, Chen F, Wang Z, Liu X, Fu Y and Lin C 2008 Association of the circadian rhythmic expression of *GmCRY1a* with a latitudinal cline in photoperiodic flowering of soybean. *Proc. Natl. Acad. Sci. USA* **105** 21028–21033