Identification and Molecular Characterization of FKF1 and GI Homologous Genes in Soybean

Fang Li¹, Xiaomei Zhang¹, Ruibo Hu², Faqiang Wu¹, Jinhua Ma¹, Ying Meng¹, Yongfu Fu¹*

¹ MOA Key Lab of Soybean Biology (Beijing), National Key Facility of Crop Gene Resource and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Haidian District, Beijing, China, ²CAS Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Energy Genetics, Qingdao Institute of BioEnergy and BioProcess Technology, Chinese Academy of Sciences, Qingdao, Shandong, China

Abstract

In Arabidopsis, FKF1 (FLAVIN BINDING, KELCH REPEAT, F-BOX1) and GI (GIgANTEA) play important roles in flowering pathway through regulating daytime CO (CONSTANS) expression, and such a function is conserved across plants studied. But related reports are limited for soybean. In this study, we cloned FKF1 and GI homologos in soybean, and named as GmFKF1, GmFKF2, GmGI1, GmGI2, and GmGI3, respectively. GmFKF1 had two alternative splicing forms, GmGI1a and GmGI1b. GmFKF1 transcripts were diurnally regulated, with a peak at zeitgeber time 12 (ZT12) in long days and at ZT10 in short days. The diurnal phases between GmGIs transcript levels greatly differed. GmGI2 expression was regulated by both the circadian clock and photoperiod. But the rhythmic phases of GmGI1 and GmGI3 expression levels were mainly conferred by long days. GmFKFs shared similar spatio-temporal expression profiles with GmGIs in all of the tissue/organisms in different developmental stages in both LD and SD. Both GmFKF and GmGI proteins were targeted to the nucleus. Yeast two hybrid assays showed GmFKF1/GmFKF2 interacted with GmGI1/GmGI2/GmCDF1 (CYCLING DOF FACTOR CDF1 homolog in soybean); and the LOV (Light, Oxygen, or Voltage) domain in GmFKF1/GmFKF2 played an important role in these interactions. N-terminus of GmGI2 was sufficient to mediate its interaction with GmCDF1. Interestingly, N-terminus not full of GmGI3 interacted with GmFKF1/GmFKF2/GmCDF1. Ectopic over-expression of the GmFKF1 or GmFKF2 in Arabidopsis enhanced flowering in SD. Collectively, GmFKF and GmGI homologs in soybean had conserved functional domains at DNA sequence level, but specific characters at function level with their homologs in other plants.

Citation: Li F, Zhang X, Hu R, Wu F, Ma J, et al. (2013) Identification and Molecular Characterization of FKF1 and GI Homologous Genes in Soybean. PLoS ONE 8(11): e79036. doi:10.1371/journal.pone.0079036

Editor: Meng-xiang Sun, Wuhan University, China

Received August 5, 2013; Accepted September 26, 2013; Published November 13, 2013

Copyright: © 2013 Li et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was partly supported by 973 Program (2010CB125906), Transgenic program (2011ZX08009-001 and 2011ZX08004-005), 863 program (2013AA102602), and the National Natural Science Founds (31000681). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: fufu19cn@163.com

Introduction

An internal time-keeping mechanism or oscillator known as the circadian clock has been found in most organisms from cyanobacteria, plants to humans [1]. Many physiological and biological behaviors of plants are conferred by the circadian clock, including photosynthesis, leaf movements, hormone production, metabolic activities, growth and development, fitness, and the transition to flowering [2-6]. Of which, triggering flowering at the appropriate time is vital for plants to successfully maximize metabolic activities, growth and development, fitness, and the transition to flowering [2-6]. Of which, triggering flowering at the appropriate time is vital for plants to successfully maximize reproduction [7], therefore photoperiodic pathway is an important mechanism of flowering [8,9].

The photoperiod pathway is controlled by the circadian clock, which is regulated by different components. FLAVIN BINDING, KELCH REPEAT, F-BOX1 (FKF1) and GIgANTEA (GI) genes are controlled by the circadian clock. In turn, FKF1 and GI can mediate the stability of some key clock proteins. FKF1 has three functional domains: the LOV domain, the F-box motif, and the Kelch repeats, all of which are highly conserved in two F-box proteins ZEITLUPE (ZTL) and LOV KELCH PROTEIN2 (LKP2) [10-12]. FKF1 interacts with GI through the LOV domain to form a complex in a blue-light dependent manner in the late afternoon under LD conditions [13]. Both FKF1 and GI can physically interact with CYCLING DOF FACTOR 1 (CDF1) [13,14] and result in degradation of CDF1 by ubiquitin–proteasome system [15]. In this process, the F-box motif is involved in formation of the SCF complex, whereas the Kelch repeats are responsible for substrate protein recognition [16,17]. In addition, the Kelch repeats of FKF1 can interact with CDF1 [14]. The CDF1 protein is a transcription repressor of CONSTANS (CO) by directly binding to the DoF binding site in the CO promoter. Under LD conditions, the timing of FKF1 and GI expression is in phase and sufficient FKF1-GI complex is formed to activate CO transcription during the day. Meanwhile, the CO protein is stabilized by light at the end of the day in LD [18]. Such a daytime CO expression triggers the expression of the floral integrators FLOWERING LOCUS T (FT) and TWIN SISTER OF FT (TSF), which are known as florigens [19-21], leading to floral initiation [19,22-25]. In contrast, under SD conditions the expression of FKF1 and GI are out of phase and small amount of FKF1-GI complex is formed in light, causing a low abundance of CO mRNA during the day [13]. It has been reported FKF1 and GI regulate flowering time besides through the CO/FT module [10,26-28]. FKF1 may regulate FT expression independent of CO through the same FKF1-mediated CDF proteins degradation mechanism. FKF1 is also involved in the stabilization of CO
proteins in the long-day afternoon and increasing the expression of FT [15,29]. Similarly, GI can activate FT expression either through directly binding FT promoter or via miR172 [30,31].

Compared with the photoperiodic flowering regulation, the FKF1 protein has a subtle role in regulation of the circadian clock and may contribute to the ubiquitin-dependent degradation of TIMING OF CAB EXPRESSION 1 (TOC1) and PSEUDO RESPONSE REGULATORS 5 (PRR5) [1,32]. On the contrary, GI is required for maintaining normal circadian system. Mutations in the Arabidopsis thaliana GI gene cause alteration of circadian rhythms in the clock-associated genes, CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY). The circadian period of the clock-controlled gene cab2::luc is altered in the gi mutant [33]. GI stabilizes ZTL proteins though...
forming a complex with ZTL and prevents TOC1 and PRR5 from ZTL-dependent degradation in the afternoon [34,55]. In addition, GI plays multiple roles in plant growth and development [36–42]. So far, the information on the isolation and function analysis of FKF1 in other plant species is limited [43,44]. However, GI is highly conserved in seed plants, such as Oryza sativa [45], Hordeum vulgare [46], Triticum aestivum [47], Zea mays [48], Pisum sativum [49], Allium cepa [49], Brachypodium distachyon [50], Pharbitis nil [51] and Lemma gibba [52].

Soybean (Glycine max), a typical short-day plant, is one of the important oil and protein crops in the world. Normal flowering is important for soybean to get yield maximization at a given ecological conditions. Recently, a series of quantitative trait loci and major genes controlling flowering and maturity (E1 to E8 and Dt1 gene, corresponding to Ardiloides TERMINAL FLOWER 1) homologs, and the functions of those in Arabidopsis, Triticum aestivum, Oryza sativa, Zea mays, Glycine max, Allium cepa, Brachypodium distachyon, Pharbitis nil, and Lemma gibba were evaluated through over-expression analysis.

**Results**

**Homologs of FKF1 and GI in Soybean Genome**

The coding sequences of FKF1 and GI genes homologs in soybean genome were respectively isolated by BLAST search against the soybean database (http://www.phytozome.org) using Arabidopsis FKF1 and GI sequences as the query, and then cloned from the soybean cultivar Kennong 18 (Table S1 in File S1). The soybean genome contained two FKF1 homologs (hereafter GmFKF1 and GmFKF2, Gm for Glycine max), and three GI homologs (hereafter GmGI1, GmGI2, and GmGI3). Interestingly, alternative splicing occurred in the 11th exon of GmGI1, resulting in two different versions (namely GmGI1a and GmGI1b). In the same way, one of CYCLING DOF FACTOR 1 (CDF1) homologs (namely GmCDF1) was also obtained. FKF1 and GI homologs in soybean genome shared comparable exon sizes and similar gene structures with those in Arabidopsis, but the intron size of soybean genes were considerably larger (Figure 1A, B), indicating more complex regulation for soybean FKF1 and GI homologs. GmFKF1 and GmFKF2 shared 95% of amino acids identity with each other. Both GmFKF1 and GmFKF2 shared high peptide identities with AtFKF1 (83.5% and 81.4%, respectively), and included all the known functional domains in AtFKF1: the LOV domain, F-box motif and Kelch repeats (Figure S1 in File S1). The LOV domains from GmFKF1 and GmFKF2 shared high similarity with that from AtFKF1, AtPHOT1 and AtPHOT2 in the secondary structure elements [69] (Figure S2 in File S1). Similarly, the protein sequences of GmGI1a, GmGI1b, GmGI2, and GmGI3 were highly conserved with approximately 76% of amino acids identities with AtGI. GmGI1 shared peptide identity of 94%–97% with GmGI3, and 80%–82% with GmGI2. The GmGIs proteins also contained multiple transmembrane domains.

**Figure 2. Circadian rhythms of FKF1 homologs in soybean.** GmFKF1 (A) and GmFKF2 (B) gene expression under different light regimes. LD, 16 hr light/8 hr dark; SD, 8 hr light/16 hr dark; LL, constant light; DD, constant dark. Soybean gene ACT11 was used as a control for normalization. doi:10.1371/journal.pone.0079036.g002
and nuclear localization signals in similar positions as in the GIs from Arabidopsis, Triticum and Brachypodium (Figure S3 in File S1).

To examine the evolutionary relationships of FKF1 or GI homologs from several plant species, phylogenetic trees were constructed using MEGA4.0 software (Figure 1C, D). Both trees clearly divided into two major clades, one corresponding to the monocots, and the other to the dicots. The GmFKFs and GmGIs proteins were clustered within the latter clade, and within the legume sub-clade together with FKF1 and GI homologs from Medicago truncatula and Pisum sativum. Interestingly, GmGI1 and GmGI3 were more closely to the Medicago truncatula and Pisum sativum GI homologs than to GmGI2, and GmGI2 had much shorter introns than GmGI1 and GmGI3 (Figure 1B), suggesting that GmGI1 and GmGI3 diverged from GmGI2 before soybean speciation.

Figure 3. Circadian rhythms of GmGIs transcript levels. GmGI1α (A), GmGI1β (B), GmGI2 (C) and GmGI3 (D) gene expression under different light regimes. LD, 16 hr light/8 hr dark; SD, 8 hr light/16 hr dark; LL, constant light; DD, constant dark. Soybean gene ACT11 was used as a control for normalization. doi:10.1371/journal.pone.0079036.g003
Diurnal Rhythms of FKF1 and GI Transcripts in Soybean

In Arabidopsis, the transcription of FKF1 and GI is both controlled by the circadian clock and photoperiod [13]. To determine whether the expression of GmFKFs and GmGIs fluctuate diurnally, RT-qPCR was performed using unifoliolate leaves under both LD and SD conditions. The level of GmFKF1 and GmFKF2 transcripts shared nearly the same patterns under both light regimes and showed clear circadian rhythms, with a peak at Zeitgeber time 12 h (ZT12) in LD and at ZT10 in SD, respectively (Figure 2). Such patterns were consistent with those observed in Arabidopsis and Nicotiana [10,26,70], and different from that of AaFKF1 (onion FKF1), whose transcript accumulation varied greatly according to the photoperiodicity [43].

To determine whether the diurnal rhythms of GmFKF1 and GmFKF2 transcripts were stable under free running conditions, the seedlings grown for one week in LD or SD were respectively transferred to continuous light (LL) and continuous dark (DD) conditions. The diurnal rhythms of GmFKF1 and GmFKF2 expression levels in LD continued cycling under LL and DD conditions and peaked 2 hr earlier in LD-LL (Figure 2). Differently, the peaking values were significantly higher on the first cycle of LD-LL and LD-DD conditions. Similarly, GmFKF1 and GmFKF2 transcripts still maintained circadian rhythms in SD-LL, but showed a longer period of 2 hr than in SD. In SD-DD, the expression levels of GmFKF1 and GmFKF2 transcript levels apparently increased. Of these, GmGI1 showed diurnal rhythm and a shorter period, while GmGI1β and GmGI3 hardly showed circadian rhythms. In SD-DD, the expression levels of GmGI1 and GmGI3 reduced to nearly be undetectable (Figure 3A, B, D). Thus, the circadian expressions of GmFKFs and GmGIs were diurnally regulated, and the expression patterns of GmGI2 and GmGI3 were in phase under both LD and SD conditions (Figure 2, 3). Although GmGI1 and GmGI2 as well as GmGI3 shared high peptide identity each other, the diurnal expression patterns of them apparently differed.

Spatial and Temporal Expression Profiles of GmFKFs and GmGIs

To better understand GmFKFs and GmGIs functions, spatio-temporal expression patterns of GmFKFs and GmGIs were systematically performed using RT-qPCR. GmFKFs and GmGIs shared similarly tissue/organ-specific expression patterns and were detected in a variety of tissues/organs under both light regimes.
Moreover, \textit{GmFKF}s and \textit{GmGI}s showed higher expression levels in most organs in SD than in LD (Figure 4, 5). Under LD conditions, \textit{GmFKF} and \textit{GmGI} transcripts had the highest level in the 2nd trifoliolates and floral buds at flowering. Under SD conditions, \textit{GmFKF}s and \textit{GmGI}1 showed the highest expression levels in roots at unifoliolate opening and in leaves at flowering (Figure 4, 5A, B). However, \textit{GmGI}2 and \textit{GmGI}3 had the highest level in roots at both vegetative and reproductive phases (Figure 5C, D). It has been reported that \textit{AtFKF1} mRNA was detected throughout the plant, with the highest level in leaves [10] and that \textit{AtGI} had higher expression level in the inflorescence apices, young flowers, and young siliques [27]. Therefore, \textit{GmFKF}s and \textit{GmGI}s displayed different tissue/organ expression patterns as day-length changed.

The expression patterns of \textit{GmFKF}s and \textit{GmGI}s during the developmental progress were also carried out. In this case, \textit{GmFKF}s and \textit{GmGI}2 shared similar expression patterns in both LD and SD, with very low expression levels in most of leaves in LD. Short days increased \textit{GmFKF}s and \textit{GmGI}2 expression levels in all leaves and the levels slightly decreased over time (Figure 6, 7C). Contrary to \textit{GmGI}2, \textit{GmGI}1 had high transcript levels in LD and low levels in SD at all development stages (Figure 7A, B). The transcript level of \textit{GmGI}3 was very low in most of leaves regardless of day-length (Figure 7D). So \textit{GmGI}s had totally different day-length responses in their developmental expression patterns. In \textit{Arabidopsis}, \textit{FKF1} and \textit{GI} expression levels increase slightly over time during development, consistent with their role in floral promotion.

Figure 5. Expression profiles of \textit{GmGI}s in various tissues/organs. The expression of \textit{GmGI}1 (A), \textit{GmGI}2 (B), \textit{GmGI} (C) and \textit{GmGI}3 (D) were investigated in LD (left panel) and SD (right panel). The samples were collected as Figure 4. The soybean \textit{UKN1} gene was used as the reference gene. doi:10.1371/journal.pone.0079036.g005
The results suggested GmGI2 and GmFKFs may involve flowering regulation under SD conditions.

Nuclear Localization of the GmFKF and GmGI Proteins
To investigate the subcellular localizations of the FKF1 and GI proteins, a yellow fluorescent protein (YFP)-coding sequence was fused to the C-terminus of GmFKF and GmGI genes driven by a 35S promoter. The fusion constructs of GmFKF1/2-YFP or GmGI1a/2/3-YFP was co-transformed with the nuclear marker gene CFP-AHL22, a positive control [71], into Arabidopsis mesophyll protoplasts. The results showed that both FKF1 and GI proteins were mainly targeted to the nucleus (Figure 8), similar to the localization patterns of AtFKF1 and AtGI [72,73]. The nuclear localization of GmGI proteins were consistent with the presence of NLS-like (nuclear localization signals) sequences in the middle of GmGI proteins (Figure S3 in File S1), which were also present in GI homolog proteins in other plants [47,50,73].

Interactions between GmFKF and GmGI Proteins
In Arabidopsis, FKF1 interacts with GI in a blue light dependent manner under LD conditions, and subsequently degrades the CO repressor CDF1 and induces CO expression [13,14]. To explore the possible roles of GmFKFs and GmGIs, a series of yeast two-hybrid assays were performed. Both GmFKF1 and GmFKF2 interacted with GmGI1α and GmGI2 in yeast, and the LOV domains from GmFKF1 and GmFKF2 proteins were sufficient to interact with GmGI1α and GmGI2 (Figure 9A), similar to the AtFKF1 LOV domain [13]. Meanwhile, GmFKF1 and GmFKF2 had an interaction with GmCDF1, but no interactions were observed between GmFKF1/GmFKF2 LOV domain and GmCDF1, suggesting the importance of C-terminus Kelch repeats for this interaction (Figure 9A) [14]. In addition, GmGI2 also interacted with GmCDF1 (Figure 9A). Contrary to the AtGI N-terminus [13], the N-terminus of GmGI1 and GmGI2 interacted with GmFKF1/GmFKF2 LOV domain but not the full length of GmFKF1 and GmFKF2 (Figure 9B). Interestingly, GmGI3 N-terminus, not GmGI3 full protein, interacted with full proteins as well as LOV domains of GmFKF1 and GmFKF2 (Figure 9A, B). GmGI2 and GmGI3 N termini also interacted with GmCDF1 (Figure 9B). As expected, the C-terminus of GmGIs did not interact with GmFKF and GmFKF LOV domains (Figure S4 in File S1). It has been reported the phototropin LOV1 form a stable homo-dimer in vitro regardless of light conditions [69], so did the GmFKF1/GmFKF2 LOV domains in our test (Figure 10A). Moreover, the interaction was also occurred between the GmFKF1 LOV domain and the GmFKF2 LOV domain (Figure 10A). In Arabidopsis, AtGI forms a tetramer in solution [74]. Similarly, the homo- or hetero-dimer were also observed among GmGIs and the N terminus played important roles in these interactions (Figure 10B).

Potential Roles of GmFKF1 and GmFKF2 in Flowering Regulation
The functions of GmFKF1 and GmFKF2 in the regulation of flowering were carried out by overexpressing these genes in Arabidopsis. GmFKF1 and GmFKF2 were respectively over-expressed
Figure 7. Expression profiles of GmGIs during development. GmGI1α (A), GmGI1β (B), GmGI2 (C) and GmGI3 (D) expression levels were performed in leaves. The samples and developmental stages in LD (left panel) and SD (right panel) were the same as Figure 6. The soybean UKN1 gene was used as the normalization transcripts.

doi:10.1371/journal.pone.0079036.g007
Figure 8. Nuclear localization of GmFKF1, GmFKF2, GmGI1α, GmGI1β, GmGI2 and GmGI3 proteins in Arabidopsis protoplasts. The vector indicated the empty vector as negative control; The CFP was for CFP-AHL22, a nuclear marker (Xiao et al., 2009); the YFP was for YFP fluorescence; the red signal was due to auto-fluorescence of chloroplasts; the last panel showed superimposition of the former three panels. doi:10.1371/journal.pone.0079036.g008
in *Arabidopsis* Col-0 background and the transgenic lines (T2 generation) were grown under both LD and SD conditions. Neither the flowering time nor the rosette leaf number of the transgenic lines over-expressing *GmFKF1* or *GmFKF2* had significantly difference from that of the control plants in LD (Figure 11A). However, the transgenic plants flowered much earlier than control plants in SD (Figure 11B). The results indicated that *FKF1* may promote flowering, specifically under SD conditions.

Most of genes in soybean are present in multiple copies resulting from three duplication events occurring in the evolutionary course [75]. To determine whether function redundancy were present between *GmFKF1* and *GmFKF2*, the double transgenic plants (T1 generation) were obtained. Again, there was little difference with respect to flowering time or rosette leaf number between the double transgenic plants and the wild type in LD (Figure 11C). The results indicated that *FKF1* may promote flowering, specifically under SD conditions.

Discussion

The homolog sequences of *FKF1* and *GI* in soybean have been reported [44,56,68], but the biochemical features and molecular functions of *FKF1* and *GI* proteins has not systematically studied. In this study, we obtained the soybean *FKF1* and *GI* homologs by homology-based cloning method. Phylogenetic analysis showed high amino acid sequence identities presented between *GmFKFs* and *AtFKF1* [10,26] and between *GmGIs* and other *GI* proteins [10,26,45,47,49–51] (Figure S1 in File S1, and Figure S3 in File S1). In addition, *GmGI2* probably evolved earlier than *GmGI1* and *GmGI3* (Figure 1D). The similar nuclear localizations of *GmFKF* and *GmGI* proteins to their homologs in other plants indicated their conserved functions [13,72].

The expression of *GmFKFs* and *GmGI2* were regulated by both the clock and light (Figure 2, 3C), as that of *FKF1* and *GI* in *Arabidopsis* and other plant species [10,27,33,43,46,47,49–51]. However, expression divergence among *GmGIs* was obvious. Compared with *GmGI2*, the circadian rhythms of *GmGI1* and *GmGI3* transcripts were greatly affected by day-length and long days largely contributed to the rhythmic phases of *GmGI1* and *GmGI3* (Figure 1D). The similar nuclear localizations of *GmFKF* and *GmGI* proteins to their homologs in other plants indicated their conserved functions [13,72].

The expression of *GmFKFs* and *GmGI2* were regulated by both the clock and light (Figure 2, 3C), as that of *FKF1* and *GI* in *Arabidopsis* and other plant species [10,27,33,43,46,47,49–51]. However, expression divergence among *GmGIs* was obvious. Compared with *GmGI2*, the circadian rhythms of *GmGI1* and *GmGI3* transcripts were greatly affected by day-length and long days largely contributed to the rhythmic phases of *GmGI1* and *GmGI3* (Figure 1D). The similar nuclear localizations of *GmFKF* and *GmGI* proteins to their homologs in other plants indicated their conserved functions [13,72].

The expression of *GmFKFs* and *GmGI2* were regulated by both the clock and light (Figure 2, 3C), as that of *FKF1* and *GI* in *Arabidopsis* and other plant species [10,27,33,43,46,47,49–51]. However, expression divergence among *GmGIs* was obvious. Compared with *GmGI2*, the circadian rhythms of *GmGI1* and *GmGI3* transcripts were greatly affected by day-length and long days largely contributed to the rhythmic phases of *GmGI1* and *GmGI3* (Figure 1D). The similar nuclear localizations of *GmFKF* and *GmGI* proteins to their homologs in other plants indicated their conserved functions [13,72].

The expression of *GmFKFs* and *GmGI2* were regulated by both the clock and light (Figure 2, 3C), as that of *FKF1* and *GI* in *Arabidopsis* and other plant species [10,27,33,43,46,47,49–51]. However, expression divergence among *GmGIs* was obvious. Compared with *GmGI2*, the circadian rhythms of *GmGI1* and *GmGI3* transcripts were greatly affected by day-length and long days largely contributed to the rhythmic phases of *GmGI1* and *GmGI3* (Figure 1D). The similar nuclear localizations of *GmFKF* and *GmGI* proteins to their homologs in other plants indicated their conserved functions [13,72].

The expression of *GmFKFs* and *GmGI2* were regulated by both the clock and light (Figure 2, 3C), as that of *FKF1* and *GI* in *Arabidopsis* and other plant species [10,27,33,43,46,47,49–51]. However, expression divergence among *GmGIs* was obvious. Compared with *GmGI2*, the circadian rhythms of *GmGI1* and *GmGI3* transcripts were greatly affected by day-length and long days largely contributed to the rhythmic phases of *GmGI1* and *GmGI3* (Figure 1D). The similar nuclear localizations of *GmFKF* and *GmGI* proteins to their homologs in other plants indicated their conserved functions [13,72].
second trifoliolates and the floral meristem in LD, but in roots in SD (Figure 4, 5). Contrast to GmGI2, the level of GmGI1 transcripts was much higher in LD than in SD, and GmGI3 lowly expressed under both LD and SD conditions (Figure 7). Therefore, developmental cues had effects on the expression of GmGI and GmFKF (Figure 6, 7). However, different from AtFKF1 and AtGI, which increase their expression along with development [10,27], GmFKFs and GmGI2 expression levels reduced slightly in SD over time (Figure 6, 7C).

The synchronous expression of GmFKFs and GmGIs was conserved and obvious, which displayed in the same time point (circadian rhythms and peaks), the same location (leaves), both vegetative and reproductive stages. And the same subcellular localization (the nucleus) implied interactions probably occurred between GmFKFs and GmGIs. Such interactions were confirmed by yeast two hybrid assays and independent of light (Figure 9), while the FKF1-GI complex formation in Arabidopsis is in a blue light dependent manner [13]. The subfunctionalization of GmFKFs and GmGIs also happened on interactions between different copies of these genes. For example, GmGI1b and GmGI3 transcripts peaked earlier than GmFKF1 and GmFKF2 in both light regimes. Consistent with this, neither GmFKF1 nor GmFKF2 interacted with GmGI1b and GmGI3, but did with GmGI1a and GmGI2. In addition, both GmFKF1 and GmFKF2 interacted with GmCDF1, so did GmGI2 (Figure 9A). The LOV domain of GmFKFs may be important for intermediating the interaction with their partners, supported by data from Arabidopsis FKF1 [13,14,26,76–78]. In soybean, GmFKF1/GmFKF2 LOV domains interacted with GmGI1a and GmGI2 (Figure 9A). Meanwhile, C-terminus Kelch repeats of GmFKFs may be related to an interaction with GmCDF1 (Figure 9A). In Arabidopsis, sufficient levels of the FKF1–GI complex in LD are required for the proper induction of CO [13,26,27]. Therefore, the interactions between GmFKFs and GmGIs may confer to the promotion of flowering in soybean.

In Arabidopsis, the fkh1 mutant flowers much later than wild type in long days, whereas it flowers much normally in short days [10]. On the contrary, both GmFKF1 and GmFKF2 promoted soybean flowering in SD, not in LD (Figure 11A, B). However, the promotion effect of GmFKF1 or GmFKF2 was inhibited when GmFKF1 and GmFKF2 were co-overexpressed (Figure 11D), suggesting GmFKF1 and GmFKF2 antagonized each other on the regulation of flowering. Provided that GmFKF1 and GmFKF2 shared the synchronous expression, the fine tune between GmFKF1 and GmFKF2 was important for soybean flowering regulation. However, overexpression of GmGIs had no significant difference in flowering time and 35S::GmGI1a and 35S::GmGI1b could not rescue gi-1 mutant phenotype (data not shown), inferring soybean GmGIs had a species-specific flowering activity and the FKF1-GI–CDF1–CO-FT pathway was unique in soybean. Additionally, the transgenic plants over-expressing both GmFKF2 and GmGI1a produced more rosette leaves and exhibited more vigorous growth and senescence retardation compared to the wild type plants in LD (Figure S5 in File S1), suggesting their functions on vegetative growth. GmGI3, which may be a more ancestral gene than GmGI1 and GmGI2, may had a subtle function in flowering regulation, because the full protein did not interact with GmFKF1/2, but its N-terminus did interact with GmFKF1/2 and GmCDF1 (Figure 9). Consistent with this, it has been reported the truncated protein (735aa) of GmGia (corresponding to GmGI3 in this study) showed a significantly earlier flowering phenotype than the wild type under natural day length conditions [56]. Undoubtedly, more information in situ not ex situ regarding the biochemical functions of
the soybean FKF1 and GI genes and their mutants in soybean needed to elucidate their bona fide functions.

Materials and Methods

Plant Materials and Growth Conditions

Ectopic expressing materials were Arabidopsis thaliana wild-type Col-0 stored in our lab. Seeds were pretreated at 4°C for 3 days in the dark after treated with sodium hypochlorite, then transferred to a growth chamber 22°C under long days (LD, 16 h light/8 h dark) or short days (SD, 8 h light/16 h dark). Light intensity was approximately 100 μmol·m⁻²·s⁻¹ provided by white fluorescent illumination. Kennong 18 (KN18), a soybean cultivar, was used in this study. Seedlings were grown in the phytotron under long-day (LD, 16 h light/8 h dark) or short-day (SD, 8 h light/16 h dark) conditions. The temperature was 26°C. The light source was from cool-white fluorescent illumination. The samples used for analysis of tissue-organ expression patterns were harvested as described by Hu et al [79]. The samples during development (from fully expanded unifoliolate stage until to flowering onset) comprised unifoliolate leaves and different trifoliolate leaves as well as flower buds. The samples for spatio-temporal expression analysis were harvested at ZT12 in LD and at ZT8 in SD. The fully expanded unifoliolate leaves were harvested as photoperiod samples at 2 hr intervals over 48 hr when seedlings were exposed to either long-day or short-day conditions, then the seedlings were transferred to either constant light (LL) or constant darkness (DD) conditions and the samples were continuously collected as in LD or SD. All samples were immediately frozen in liquid nitrogen and stored at -80°C until required.

Isolation and Phylogenetic Analysis of Soybean FKF1 and GI Homologs

The Arabidopsis FKF1 and GI amino acid sequences have been used as baits to search for potential homologous genes in the soybean database (Phytozome 4.0), and homologs sequences were acquired based on the E value = 0 and then used to conduct
BLASTP in Arabidopsis database [http://www.arabidopsis.org/Blast/index.jsp] for preliminary screening. Finally, two FKF1 and three GI homologs were determined based on further analysis in Softberry database [http://www.softberry.com/beryl.pluralm]. Gene-specific primers were designed at both ends of gene non-coding/coding regions and used to conduct nested or semi-nested PCR employed KN10 cDNA as the template. All primers used were listed in Table S2 in File S1. The PCR products were introduced into the pEASY-T1 vector (TransGen Biotech, CHN) after purification and several independent clones were sequenced for real sequences.

Phylogenetic analysis of FKF1 and GI homologs were conducted, respectively. Predicted and published protein sequences were obtained from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and then selected in Pfam database based on functional domains annotation. Multiple alignments were conducted using ClustalX. Phylogenetic neighbor-joining trees were constructed with MEGA 4.0 based on the full-length amino acid sequences. Bootstrap analysis was performed estimate nodal support based on 1,000 replicates.

**RNA Isolation, cDNA Preparation and Real-time Quantitative RT-PCR**

Total RNA was extracted using TRIzol reagent (Invitrogen, CA, USA) according to the manufacturer’s recommendations. Contaminating genomic DNA was removed with the RNA-free DNAseI kit (Invitrogen, CA, USA). RNA purity was controlled by determining the ratio of A260/A280 ranging between 1.8 and 2.0 and A260/A230 greater than 2.0. First-strand cDNA was synthesized using 5 μg purified RNA with the M-MLV reverse transcriptase kit (Invitrogen, CA, USA) and oligo-dT primers, according to the manufacturer’s protocol. Then the cDNA product was diluted 1:20 prior to use.

RT-qPCR reaction was performed as previously described [79]. Each reaction was performed in three technical replicates and the Ct value was exported using the StepOne Software v2.0 (ABI Applied). The blank controls were included with H2O as the template for each reaction. The soybean UK1 gene was used as reference control for tissue-organ and developmental genes analysis, while the gene ACT11 was used as the normalization transcripts for diurnal rhythms analysis [79]. Each reaction was performed in three technical replicates and reciprocal hybrids were determined by alignment of amino acid sequences between GmGIs and AtGIs [13]. The bait and prey pairs were co-transformed into the yeast strain AH109 (MATa, trp1, leu2) and the positive colonies were screened by growing on the SD dropout medium (−Trp/−Leu/−His/−Ade). Negative controls and self-activation experiments were also investigated. Transformation was conducted according to the manufacturer’s protocol (Clontech). The experiment was performed in three biological replicates and reciprocal hybrids were included.

**Ectopic Expression in Arabidopsis**

The coding sequences of GmFKFs and GmGIs genes were introduced into the expression vectors pLeela (Basta-resistance) and pGBW2 (Kanamycin and Hygromycin resistance) by using gateway method, then individually or collectively transformed into Arabidopsis wild-type Col-0 plants using the floral dipping method [82], which were infected with Agrobacterium tumefaciens strain PV3101 MP90RK. The transgenic plants were screened using 50 mg L\(^{-1}\) glufosinate ammonium or 50 mg L\(^{-1}\) hygromycin or both. At least three independent transgenic lines were used to measure the flowering time and rosette leaf number. The transgenic lines and the control plants were grown under the same conditions in growth chamber.

**Supporting Information**

**File S1** Table S1 A list of FKFIs and GIs as well as CDFI homolog genes in soybean. Table S2 A list of primers for gene cloning and RT-qPCR. **Figure S1** Amino acid sequence alignment of GmFKFs (Glycine max) and AtFKF1 (Arabidopsis thaliana). Identical and similar amino acids were indicated in black-shaded and grey-shaded, respectively. LOV domain, F-box motif, and Kelch repeats were highlighted in yellow fonts, green fonts and pink fonts, respectively. **Figure S2** Comparison of conserved amino acid sites among several LOV domains. These LOV domains were from Glycine max FKF1 and FKF2, Arabidopsis thaliana FKF1, PHOT1_LOV1, PHOT1_LOV2, PHOT2_LOV1 and PHOT2_LOV2. The conserved cysteines were highlighted in black background. The loop regions of GmFKF1/GmFKF2/AtFKF1 LOV were in white fonts and green-shaded. The predicted secondary structure elements of GI proteins from several species. These GI proteins included Glycine max GI (GmGIs), Arabidopsis thaliana GI (AtGI), Triticum aestivum L. GI (TaGIs) and Brachypodium distachyon GI (BdGI). The black and grey regions represented amino acid positions these LOV domains started and ended. **Figure S3** Alignment of conserved amino acid sites in GI proteins from several species. These GI proteins included Glycine max GI (GmGIs), Arabidopsis thaliana GI (AtGI), Triticum aestivum L. GI (TaGIs) and Brachypodium distachyon GI (BdGI). The black and grey regions represented identical and similar amino acids, respectively. The red region represented transmembrane domains and the blue region represented NLS-like (nuclear localization signals) motifs. **Figure S4** Interactions between GmGI C-terminal and GmFKFs or GmFKF (LF) truncated proteins in yeast. GmFKFs (LF) included the LOV domain and F-box motif of
GmFKF1 and GmFKF2. GmGI C-terminal were fused to the GAL4 activation domain (Prey). GmFKFs and GmFKFs (L) were fused to the GAL4 DNA binding domain (Bait). The empty vector only contained the activation domain (AD). -LW, synthetic dropout (SD) yeast growth medium lacking leucine and tryptophan; SDNH lacking leucine, tryptophan, histidine; and adenine. Figure S5 Phenotypes of GmGI1A and GmFKF2 co-overexpressing lines in Arabidopsis Col under LD conditions. The double transgenic plants exhibited more vigorous growth and senescence retardation and produced more rosette leaves (A). Comparison of the rosette leaf number (B) and the days to flowering (C) between the double transgenic plants and the control plants. 

Author Contributions

Conceived and designed the experiments: YF. Performed the experiments: FL. NZ. RH. FW. YM. Analyzed the data: FL. YF. Contributed reagents/materials/analysis tools: FL. Wrote the paper: FL.

References

1. Baudry A, Ito S, Song YH, Strait AA, Kiba T, et al. (2010) F-box proteins FK1 and LKP2 act in concert with ZEL1/TOC1 to regulate Arabidopsis clock progression. Plant Cell 22: 606–622.
2. Ouyang Y, Anderson CR, Kondo T, Golden SS, Johnson CH (1998) Resonating circadian clocks enhance fitness in cyanobacteria. Proc Natl Acad Sci U S A 95: 8660–8664.
3. Green RM, Tingay S, Wang ZY, Tobin EM (2002) Circadian rhythms confer a higher level of fitness to Arabidopsis plants. Plant Physiol 129: 576–584.
4. Dodd AN, Salathia N, Hall A, Kevei E, Toth R, et al. (2005) Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science 309: 630–633.
5. Hamer SL (2009) The circadian system in higher plants. Annu Rev Plant Biol 60: 357–377.
6. Song YH, Ito S, Imaizumi T (2010) Similarities in the circadian clock and photoperiodism in plants. Curr Opin Plant Biol 13: 594–603.
7. Perezio M, Juarez JA (2010) Ubiquination in the control of photoperiodic flowering. Plant Sci 198: 98–109.
8. Fornara F, de Montaigu A, Coupland G (2010) SnapShot: Control of flowering in Arabidopsis. Cell 141: 550, 550 e551–552.
9. Andrews F, Coupland G (2012) The genetic basis of flowering responses to seasonal cues. Nat Rev Genet 13: 627–639.
10. Nelson DC, Lasswell J, Rieg LE, Cohen MA, Bartel B (2000) FK1, a clock-controlled gene that regulates the transition to flowering in Arabidopsis. Cell 101: 331–341.
11. Somers DE, Schultz TF, Mihmamow M, Kay SA (2000) ZEL1/TOC1 encodes a novel clock-associated PAS protein from Arabidopsis. Cell 101: 319–329.
12. Schultz TF, Kiyosue T, Yanovsky M, Wara M, Kay SA (2001) A role for LKP2 in the circadian clock of Arabidopsis. Plant Cell 13: 2659–2670.
13. Sawa M, Nusinovic D, Kay SA, Imaizumi T (2007) FK1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science 318: 261–265.
14. Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA (2005) FK1/F-box protein mediates cyclical degradation of a repressor of CONSTANS in Arabidopsis. Science 309: 293–297.
15. Fornara F, Panigrahi KC, Gisost L, Sauerbunn N, Ruhl M, et al. (2009) Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. Dev Cell 17: 75–86.
16. Yasuhara M, Mitsui S, Hiranó H, Takahara R, Tokioka Y, et al. (2004) Identification of ASK and clock-associated proteins as molecular partners of LKP2 (LOV kelch protein 2) in Arabidopsis. J Exp Bot 55: 2015–2027.
17. Andrade MA, Gonzalez-Guzman M, Serrano R, Rodriguez PL (2001) A family of F-box proteins. Plant Mol Biol 46: 603–614.
18. Koyama Y, Weigel D (2007) Move on up, it's time for change–mobile signals controlling photoperiod-dependent flowering. Genes Dev 21: 2371–2384.
19. Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, et al. (2007) FT protein mediates long-distance signaling in floral induction of Arabidopsis. J Exp Bot 58: 23073–23083.
20. Sawa M, Kay SA (2011) GIGANTEA directly activates flowering Locus T in Arabidopsis thaliana. Proc Natl Acad Sci U S A 108: 11698–11703.
21. Jung JH, Seo YH, Seo PJ, Reyes JL, Yun J, et al. (2007) The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in Arabidopsis. Plant Cell 19: 2736–2748.
22. Wang L, Fujisawa S, Somers DE (2010) PR5 regulates phosphorylation, nuclear import and subnuclear localization of TOC1 in the Arabidopsis circadian clock. EMBO J 29: 1905–1915.
23. Park DH, Somers DE, Kim YS, Cho YH, Lim HK, et al. (1999) Control of circadian rhythms and photoperiodic flowering by the Arabidopsis GIGANTEA gene. Science 283: 1579–1582.
24. Kim WY, Fujisawa S, Sub SS, Kim J, Kim Y, et al. (2007) ZEL1/TOC1 is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature 449: 356–360.
25. Fujisawa S, Wang L, Han L, Sub SS, Salome PA, et al. (2008) Post-translational regulation of the Arabidopsis circadian clock through selective proteolysis and phosphorylation of pseudo-response regulator proteins. J Biol Chem 283: 23073–23083.
26. Fowler S, Thomashow MF (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CFY cold response pathway. Plant Cell 14: 1675–1690.
27. Caso S, Ye M, Jiang S (2005) Involvement of GIGANTEA gene in the regulation of the cold stress response in Arabidopsis. Plant Cell Rep 24: 683–690.
28. Falsot J, Amin R, Gover A, Ori N, Samach A (2006) Novel roles for GIGANTEA revealed under environmental conditions that modify its expression in Arabidopsis and Medicago truncatula. Planta 224: 1255–1260.
29. Sun S, Atkinson RM (2006) Molecular genetic studies of the memory of vernalisation. J Exp Bot 57: 3369–3377.
30. Penfield S, Hall A (2009) A role for multiple circadian clock genes in the response to signals that break seed dormancy in Arabidopsis. Plant Physiol 151: 1722–1732.
31. Dalchau N, Back SJ, Briggs HM, Robertson FC, Dodd AN, et al. (2011) The circadian oscillator gene GIGANTEA mediates a long-term response of the Arabidopsis thaliana circadian clock to sucrose. Proc Natl Acad Sci U S A 108: 5104–5109.
32. Edwards J, Martin AP, Andrussias F, Ollifer CE, Patrick JW, et al. (2010) GIGANTEA is a component of a regulatory pathway determining wall ingrowth deposition in phloem parenchyma transfer cells of Arabidopsis thaliana. Plant J 63: 651–661.
33. Taylor A, Masahia Aj, Thomas B (2010) Conservation of Arabidopsis thaliana photoperiodic flowering time genes in onion (Allium cepa L.). Plant Cell Physiol 51: 1638–1647.
34. Tiwari D, Kumudini S, Dinkins RD (2010) Expression of flowering-time genes in soybean E1 near-isogenic lines under short and long day conditions. Planta 231: 951–963.
35. Yamaya R, Yokoi S, Tamaki S, Yano M, Shimamoto K (2003) Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature 422: 719–722.
36. van der Put DP, Griffiths S, Christodoulou V, Laurie DA (2005) Characterisation of a barley (Hordeum vulgare L.) homologue of the Arabidopsis flowering time regulator GIGANTEA. Theor Appl Genet 110: 925–931.
37. Zhao XY, Liu MS, Li JR, Guan CM, Zhang XS (2005) The wheat TaG1, involved in photoperiodic flowering, encodes an Arabidopsis GI ortholog. Plant Mol Biol 58: 53–64.
63. Tian Z, Wang X, Lee R, Li Y, Specht JE, et al. (2010) Artificial selection for circadian clock-associated gene gianteal affects maize developmental transition. Plant Cell Environ 36: 1379–1390.

62. Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, et al. (2010) The soybean stem growth habit gene Dt1 is an ortholog of Arabidopsis TERMINAL FLOWER1 and suppresses one-shot induction of flowering in Phbahsil, a typical short-day plant. Plant Cell Physiol 52: 635–650.

61. Sun H, Jia Z, Cao D, Jiang B, Wu C, et al. (2011) GmFT2a, a soybean homolog of the Arabidopsis FLOWERING LOCUS T, is involved in flowering transition and suppresses one-shot induction of flowering in Phbahsil, a typical short-day plant. Plant Physiol 154: 1220–1231.

60. Hong SY, Lee S, Seo PJ, Yang MS, Park CM (2010) Identification and molecular characterization of a Brachypodium distachyon GIGANTEA gene: functional conservation in monocot and dicot plants. Plant Mol Biol 72: 485–497.

59. Wu FQ, Zhang XM, Li DM, Fu YF (2011) Ectopic expression reveals a functional clock-controlled gene in soybean (Glycine max). J Plant Physiol 168: 1952–1963.

58. Watanabe S, Harada K, Abe J (2012) Genetic and molecular bases of soybean maturity and flowering. Bred Sci 61: 531–543.

57. Liu H, Wang H, Gao P, Xu J, Xu T, et al. (2009) Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. Proc Natl Acad Sci U S A 106: E2155–E2164.

56. Watanabe S, Hideshima R, Xia Z, Tsubokura Y, Sato S, et al. (2011) A map-based cloning strategy employing a residual heterozygous line reveals that the soybean maturity locus E3 is involved in soybean maturity and flowering. Genetics 188: 395–407.

55. Liu H, Wang H, Gao P, Xu J, Xu T, et al. (2009) Analysis of clock gene homology using unifoliolates as target organs in soybean (Glycine max). J Plant Physiol 166: 278–289.

54. Abe J, Xu DH, Suzuki Y, Kanazawa A, Shimamoto Y (2003) Soybean GIGANTEA, a circadian clock-related gene, is involved in flowering transition and suppresses one-shot induction of flowering in Pharbitis nil, a typical short-day plant. Plant Cell Physiol 153: 198–210.

53. Watanabe S, Harada K, Abe J (2012) Genetic and molecular bases of soybean maturity and flowering. Bred Sci 61: 531–543.

52. Serikawa M, Miwa K, Kondo T, Oyama T (2008) Functional conservation of clock-related genes in flowering plants: overexpression and RNA interference analyses of the circadian rhythm in the monocotyledon Lemna gibba. Plant Physiol 146: 473–484.

51. Higuchi Y, Sage-Ono K, Sasaki R, Ohtsuki N, Hoshino A, et al. (2011) Identification and functional analysis of the circadian rhythm clock-related genes in soybean. J Biosci 37: 857–870.

50. Hong SY, Lee S, Seo PJ, Yang MS, Park CM (2010) Identification and characterization of clock-related genes in flowering plants: overexpression and RNA interference analyses of the circadian rhythm in the monocotyledon Lemna gibba. Plant Physiol 146: 473–484.

49. Hecht V, Knowles CL, Vander Schoor JK, Liew LC, Jones SE, et al. (2007) Pea late bloomer1 is a GIGANTEA ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. Plant Physiol 144: 648–661.

48. Bendix C, Mendoza JM, Stanley DN, Meeley R, Harmon FG (2013) The circadian clock-associated gene gianteal affects maize developmental transition. Plant Cell Environ 36: 1379–1390.

47. Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, et al. (2010) The soybean stem growth habit gene Dt1 is an ortholog of Arabidopsis TERMINAL FLOWER1. Plant Physiol 153: 198–210.

46. Tian Z, Wang X, Lee R, Li Y, Specht JE, et al. (2010) Artificial selection for circadian clock-associated gene gianteal affects maize developmental transition. Plant Cell Environ 36: 1379–1390.