# Supporting Information

Zhang *et al*. 10.1073/pnas.0810585105

## SI Materials and Methods

**Plant Materials and Growth Conditions.** Eleven soybean cultivars (*Glycine max*) and 328 wild soybean (*G. soja*) accessions examined in this study were obtained from Soybean Germplasm Resources and Molecular Genetics, Chinese Academy of Agriculture Sciences, Beijing, China. An elite cultivar KN18 (Kennong 18) was used to clone *GmCRY* genes and to study GmCRY gene expression. The cultivars used in this study are: Guizao 2 (GZ2), Fudou 1 (FD1), Zhechun 3 (ZC3), Zhongdou 31 (ZD31), Zhonghuang 13 (ZH13), Jidou 12 (JD12), Tiefeng 31 (TF31), Changnong 13 (CN13), Kennong 18 (KN18), Suinong 14 (SN14), and Heihe 27 (HH27).

For the common garden experiment, seeds were sowed on May 9, 2007, grown in a field near Beijing (≈40°N, ≈116°E). The earliest-flowering accession started to flower on June 16 (day length, 15 h), whereas the latest-flowering accession started to flower on October 1 (day length, 11 h 50 min), so the vegetative growth and floral transition of soybean in this experiment were completed in LD photoperiods. The temperature during this period was between 12 °C and 37.3 °C.

For experiments conducted in the growth room with defined day length (LD, 18 hL/6 hD, SD, 8h L/16 hD) and temperature (25–28 °C), the cool white fluorescent lights (TLD 18W/54, Philips) lights were used. The experiments using blue light (436 nm) and red light (658 nm) were performed in the Blue-LED or Red-LED growth chambers (Percival Scientific Inc.). Fluence rates were measured using a Li250 quantum photometer (Li-Cor). For the measurement of soybean flowering time, the days to flowering and the numbers of trifoliate leaves were scored when the petals of the first flower were visible.

**Gene and mRNA Analyses.** Soybean genes encoding cryptochromes were analyzed using the soybean EST and genome sequence database online (http://www.phytozome.net/soybean and http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi). The full-length *GmCRY1a* and *GmCRY2a* cDNA were cloned from the cultivar KN18, using the 5′-RACE and 3′-RACE cloning kits according to the manufacturer's instruction (Invitrogen). The cDNA sequences of *GmCRY1a* (accession, DQ401046) and *GmCRY2a* (accession: DQ40104712) were deposited in the GenBank. The mRNA expression was analyzed by qPCR as described (1), using a Thermal Cycler (Applied Biosystems). Twelve- to 20-day-old soybean seedlings were used to study RNA (and protein) expression. The primers used to detect GmCRY genes and the internal control tubulin gene are:

*GmCRY1*F : ATGGTGGCTCAAGAACAGTTTGG
*GmCRY1*R : GAGTGTCAGTGGATCGCTTAGTG
*GmCRY2*F: AAGGCTCCATGTGCTACTATTTCAG
*GmCRY2*R: GCG TAG GCG GTT CAT TCT TGG
*GmTUB*-F: TCTTGGACAACGAAGCCATCT
*GmTUB*-R: TGGTGAGGGACGAAATGATCT
*GmCRY1aF-1486:* GAGAATGGAACTGAGGAAGGACT
*GmCRY1aR-2043*: CCCAGTCTGAGGTAGCTGCCTCC
*GmCRY2aF-1486*: TCTGAACCAAGGGATGAAGTTGT
*GmCRY2aR-1902*: CATAGCTCCATCTTTGCTTGAAC.

**Immunological Analyses.** Antibodies were prepared against GmCRY1a and GmCRY2a, because they were the first cloned (Fig. S1*C*). Immunoblot analyses of samples from *E. coli* (Fig. S1 *E* and *F*) and Arabidopsis expressing the respective GmCRY proteins (Fig. S3*A*) confirmed that the anti-GmCRY1a antibody did not cross-react with GmCRY2a, whereas the anti-GmCRY2a antibody did not cross-react with GmCRY1a.

Immunostain was as described (2), using the anti-GmCRY antibodies and rhodamine Red-x conjugated goat anti-rabbit IgG (1:200 dilution) (Jackson ImmunoResearch Lab). Images were captured and analyzed using an Olympus BX-51 microscope with U-MWU2 and U-MWG2 filter with the Olympus DP Controller software, and processed using Image J and Adobe Photoshop 7.0 software.

Immunoblots were performed using the ECL (enhance chemiluminance) method as described previously (1). The signals of the cryptochrome band detected in immunoblots were digitized and quantified using lumi-imager (Chemi DOC-It Imaging System, UVP) with Vison WorksLS software. For MG-132 treatment the unifoliolate of 12-day-old etiolated soybean seedlings were collected and excised into 2- to 5-mm-long sections under weak red light and incubated in 50 $\mu$M MG132 (EMD Chemicals) dissolved in 0.1% DMSO (from a fresh stock of 50 mM in DMSO), or in 0.1% DMSO mock control, in the dark for 5 h at room temperature. Explants were then exposed to blue light (20 $\mu$mol/m$^2$/s) for the indicated durations.

**Data Analyses.** Because neither anti-GmCRY1a nor anti-GmCRY2a can be easily stripped off from an immunoblot, all of the immunoblot was probed only once. The relative loading of samples were normalized by the band signal of a non-specific band (NS) shown on the same immunoblot. To normalize the variations in signal strengths of samples from different immunoblots resulting from variable times of the ECL reaction and x-ray film exposure, we included an "ECL control" sample in each gel. The ECL control sample was an aliquot of the same sample prepared at noon from the cultivar KN18 grown in LD photoperiods, the same amount of this protein sample was loaded in one lane of each SDS/PAGE gel, and the respective cryptochrome band signal of the ECL control sample was scored as the ECL control signal. The relative abundance of cryptochrome in each sample was calculated by the two-way normalization with the formula $GmCRY^X/(NS^X)(GmCRY^{ECL})$, in which $GmCRY^X$ is the band signal of GmCRY of the sample tested, $NS^X$ is the individual loading control (NS band signal) of the respective lane, and $GmCRY^{ECL}$ is the ECL control signal (GmCRY band signal of the ECL control sample) in the respective gel.

The relative abundance of cryptochromes at noon was estimated by the formula $GmCRY1a\ (N/\Sigma_{MNE}) = GmCRY1aN/[(GmCRY1aM+GmCRY1aN+GmCRY1aE)]$, in which GmCRY1aM, GmCRY1aN, and GmCRY1aE are the relative abundance of GmCRY1a in samples collected in the morning (0.5 h after light on), noon (middle of the light phase), or evening (0.5 h before light off), respectively, from plants grown in LD or SD as indicated.

One-way ANOVA was used to analyze variances of flowering times, GmCRY expressions, and latitudes of cultivars. Pearson correlation and linear regression were calculated between variances of flowering time, latitude, and GmCRY expression, using the software SAS (version 8.0, SAS Institute) and Excel (version 2003, Microsoft).

**Transgenic Analyses.** Transgenic Arabidopsis were prepared and analyzed as described (2). GFP-GmCRY constructs, prepared as described in the vector pEGAD (1), were used for stable transformation in Arabidopsis and transient transformation in soybean. GFP fluorescent images were observed using a confocal fluorescent microscopy (Leica TCS SP2). Images were

analyzed using the Leica Software, and processed using Photoshop 7.0 (Adobe Systems). At least 15 transgenic lines (35S::GFP-GmCRY1a/*cry2*) showed both the expression of GFP-GmCRY1a and accelerated flowering. But none of the multiple transgenic lines (35S::GFP-GmCRY1a/*cry2*) that expressed GFP-GmCRY2a showed altered flowering time (Fig. 1, and data not shown). For transient soybean transformation, adult leaves were infiltrated with *Agrobacterium* containing *pEGAD-GmCRY1a*, *pEGAD-GmCRY2a,* or the control plasmid pEGAD. *Agrobacteria* grown on plate (LB, 50 mg/L, kanamycin, 50 mg/L gentamycin, 100 mg/L rifampicin) for 2 days were collected and resuspended in 5% sucrose to 0.5 $OD_{600}$. The cell suspension was pressure-infiltrated into unifoliolate leaves of plants grown in LD at the third trifoliate stage. After infiltration, the plants were left to grow until flower. Seven days after infiltration, tissue samples were collected from infected areas of leaves for qPCR analysis to access the transgene expression. In one experiment, we attempted to increase transformation efficiency by including a detergent (0.6% Silwet L-77) in the solution used for leaf infiltration. However, the detergent concentration was apparently too high to result in viable transformed cell. Because no *GmCRY1a* mRNA was detected in the infected plants transformed using the same *Agrobacteria* cells in this experiment, its result served as another negative control to access the effect of GmCRY1a on flowering time of soybean (Fig. S3 *G* and *H*).


1. Yu X, *et al.* (2007) Derepression of the NC80 motif is critical for the photoactivation of Arabidopsis CRY2. *Proc Natl Acad Sci USA* 104:7289–7294.
2. Yu X, *et al.* (2007) Arabidopsis Cryptochrome 2 Completes Its Posttranslational Life Cycle in the Nucleus. *Plant Cell* 19:3146–3156.
3. Wroblewski T, Tomczak A, Michelmore R (2005) Optimization of Agrobacterium-mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. *Plant Biotechnol J* 3:259–273.
4. Izawa T (2007) Adaptation of flowering-time by natural and artificial selection in Arabidopsis and rice. *J Exp Bot* 58:3091–3097.
5. Morse W, Cartter J, Williams L (1947) Soybeans: culture and varieties. *US Dep Agric Farmers' Bull* No. 1520:1–38.
6. Boerma HR, Specht JE (2004) (American Society of Agronomy, Inc., Madison, WI).

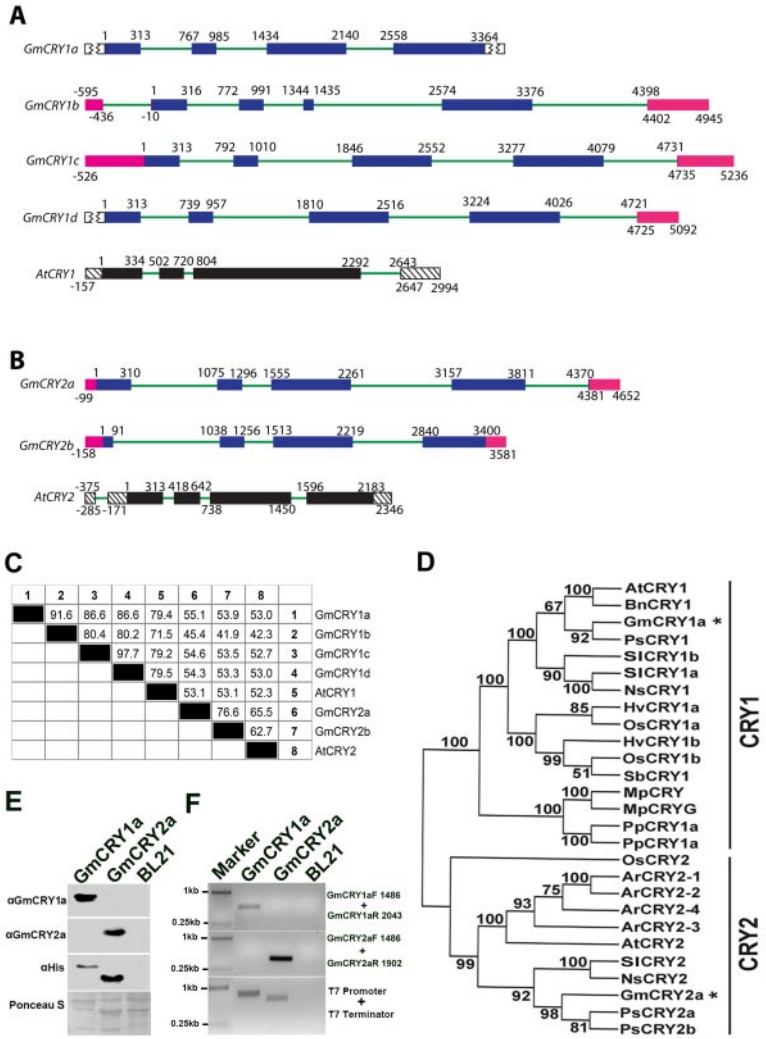


**Fig. S1.** Schematic diagram depicting the structures of the soybean cryptochrome genes. (*A* and *B*) Gene structure comparisons of the 4 soybean *GmCRY1* genes and the Arabidopsis *AtCRY1* gene (*A*), or the two soybean *GmCRY2* genes and the Arabidopsis *AtCRY2* gene (*B*), according to the preliminary soybean genome annotation database (http://www.phytozome.net/soybean). The numbers on the top of the schematic diagram indicate positions of the DNA sequence corresponding to the ORF of the respective gene, with the ATG codon starting at position 1. The pink-colored boxes of *GmCRYs* and dashed boxes of *AtCRYs* represent UTR. The blue boxes of *GmCRYs* and black boxes of *AtCRYs* represent exons. The green lines represent introns. The saw-edged white boxes represent a hypothetic UTR that are not predicted by the current annotation online. (*C* and *D*) A pair-wise amino acid sequence similarity analysis (*C*) and a phylogenetic analysis (*D*) of GmCRYs and AtCRYs apoproteins. The percentage identities of the amino acid sequence between different cryptochromes are shown (*C*). The phylogenetic tree was constructed using the Neighbor-Joining method (Mega 3.1), and the bootstrap values (1,000 replicates) are indicated (*D*). The gene accessions are: *ArCRY2–1*(AB092681.1), *ArCRY2–2* (AB092682.1), *ArCRY2–3* (AB092683.1), *ArCRY2–4* (AB092684.1), *AtCRY1* (At1g04400), *AtCRY2* (At4g08920), *GmCRY1a* (DQ401046), *GmCRY2a* (DQ40104712), *BnCRY1*(AJ344565), *HvCRY1a* (DQ201150), *HvCRY1b* (DQ201153), *SlCRY1a* (AF130423.1), *SlCRY1b* (AF545572.1), *SlCRY2* (AF130426.1), *MpCRY* (AB126657.1), *MpCRYG* (AB126658.1), *NsCRY1* (DQ231576.1), *NsCRY1* (DQ231577.1), *OsCRY1a* (AB073546.1), *OsCRY1b* (AB073547.2), *OsCRY2* (AB098568.1), *PpCRY1a* (AB027528.2), *PpCRY1b* (AB060693.1), *PsCRY1* (AY161310.1), *PsCRY2a* (AY508972.1), *PsCRY2b* (AY161312.1), and *SbCRY1* (AF545572.1). Ar, *Armoracia rusticana*; At, *Arabidopsis thaliana*; Bn, *Brassica napus*; Hv, *Hordeum vulgare*; Gm, *Glycine max*; Mp, *Marchantia polymorpha*; Ns, *Nicotiana sylvestris*; Os, *Oryza sativa*; Pp, *Physcomitrella patens*; Ps, *Pisum sativum*; Sb, *Sorghum bicolor*; Sl, *Solanum lycopersicon*. (*E*) Immunoblot analysis to verify the specificity of anti-GmCRY1a and anti-GmCRY2a antibodies. Immunoblots of samples prepared from *E. coli* strains expressing 6-His tagged GmCRY1aC or GmCRY2aC as indicated were probed with anti-GmCRY2aC, (αGmCRY1a) or anti-GmCRY2aC (αGmCRY2a), or anti-His (αHis) antibodies. (*F*) PCR reactions, using the primer pairs indicate on the right, to confirm that the *E. coli* strains used in (*E*) indeed contained the respective GmCRY1a or GmCRY2a coding sequences (see *SI Text*).


**Fig. S2.** Amino acid sequence alignment of Arabidopsis and soybean cryptochromes The alignment was conducted using ClustarW (MegAlign, DNAStar). Identical residues are highlighted by black boxes. Pink frame indicate the residues of the DAS domain.

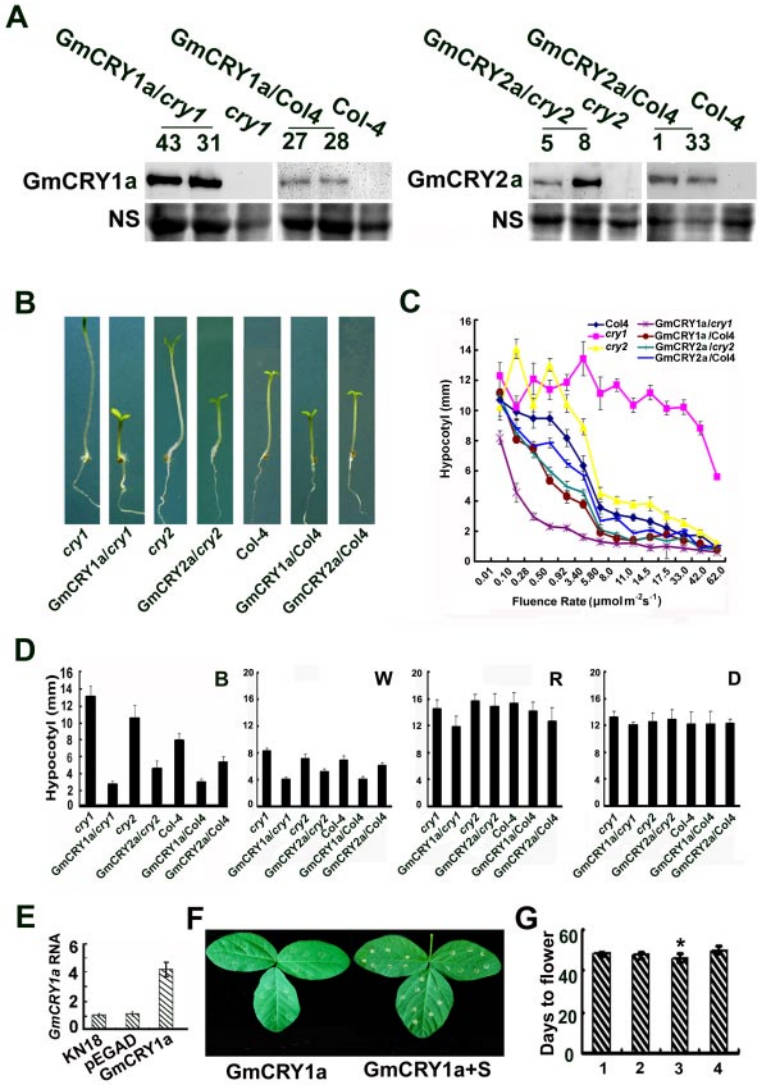


**Fig. S3.** GmCRY1a and GmCRY2a mediate blue light inhibition of hypocotyl elongation, but only GmCRY1a promotes floral initiation. (*A*) Immunoblot showing GmCRY1a or GmCRY2a expression in transgenic Arabidopsis lines of the indicated genotypes. Protein extracts of transgenic Arabidopsis expressing *35S*::*GFP-GmCRY1a* or *35S*::*GFP-GmCRY2a* were analyzed by immunoblot probed with anti-GmCRY1a (GmCRY1a, left panels) or anti-GmCRY2a (GmCRY1a, right panels), respectively. NS, non-specific band indicating relative loading. (*B*) Seedlings of the indicated genotype grown in continuous blue light (2 $\mu$molm$^{-2}$s$^{-1}$). (*C*) A fluence rate response showing hypocotyl lengths of seedlings grown under indicated fluence rates. (*D*) Hypocotyl lengths of 3-day-old seedlings of indicated genotype grown in blue (B, 2 $\mu$mol/m$^2$/s), white (W, 8 $\mu$mol/m$^2$/s), red light (R, 5 $\mu$mol/m$^2$/s), or in the dark (D). (*E*) Transgenic plants expressing *35S*::*GFP-GmCRY1a* in the *cry1* mutant background (three independent lines, 3, 31, and 43) showed accelerated flowering in SD photoperiods. Left, 50-day-old plants; right, flowering time of the indicated lines. (*F*) qPCR results showing the level of *GmCRY1a* mRNA in the control soybean plants (KN18), and plants transiently expressing *35S*::*GFP* (pEGAD) or *35S*::*GFP-GmCRY1a*. (*G*) Soybean trifoliate leaves 7 days after infection with the Agrobacterium containing plasmid expressing *35S*::*GFP-GmCRY1a* in the presence (GmCRY1a+S) or absence (GmCRY1a) of the detergent Silwet. (*H*) Flowering time of soybean plants not transformed (1) transformed with *35S*::*GFP*, (2) or *35S*::*GFP-GmCRY1a*, (3–4) with (4) or without (1–3) Silwet. Means of days to flower and standard deviations ($n > 20$) are shown (asterisk: $P <$ 0.05%). Note: Because of technical difficulties in preparing transgenic soybean plants, we adapted an *in planta* transient transformation assay to test the activity of GmCRY1a and GmCRY2a in soybean (3). In this experiment, adult leaves of soybean grown in LD were infected with *Agrobacteria* cells harboring the Ti plasmid that encodes the GFP-GmCRY1a fusion protein under the control of the 35S promoter (Fig. 1*F*). After *Agrobacteria* infection, plants were left to grow until flowering. As expected, the *GFP-GmCRY1a* mRNA was detected in leaves infected with *Agrobacteria* containing the GFP-GmCRY1a plasmid, but not in the control plants uninfected or infected with *Agrobacteria* containing the ''empty'' vector (Fig. 1*F*). Compared with the controls, plants expressing *GFP-GmCRY1a* showed a mild but statistically significant acceleration of flowering (Fig. 1*H*). In contrast, transient expression of *GFP-GmCRY2a* resulted in the relatively high level of *GmCRY2a* mRNA expression but no acceleration of flowering (data not shown). Another group of control plants, infected with *Agrobacteria* containing the GFP-GmCRY1a plasmid in the presence of detergent (0.6% Silwet L-77), also showed no accelerated flowering, apparently due to the lack of GFP-GmCRY1a expression in the infected cells damaged by the detergent (Fig. 1*F*, GmCRY1a+S, see the white spots of dead tissue).

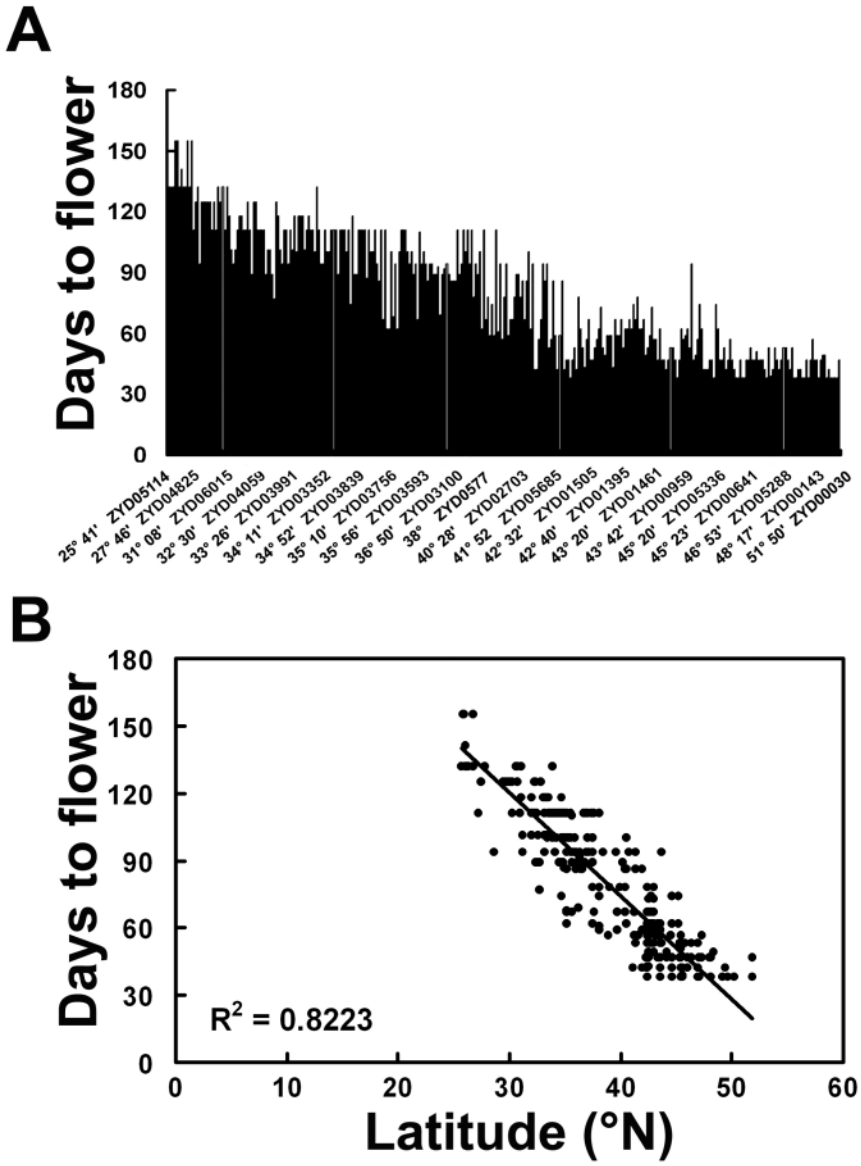


**Fig. S4.** A latitudinal cline in flowering time of 328 accessions of the wild soybean (*Glycine soja*). (*A*) Flowering time, presented as ''Days to Flower'' of 328 different wild soybean accessions grown under natural LD photoperiods as described (see Method and Materials). The symbols of accessions (arbitrarily assigned by Excel from 328 accessions) and latitude of the site of collection are indicated. (*B*) Correlation of flowering time and latitude ($R^2 = 0.8218$, $P < 0.00001$). Note: We analyzed flowering time of 328 accessions of *G soja*, which were originally collected at different locations in China with the latitudes ranging from approximately 25°N to 50°N. Plants were sowed in the spring of 2007, grown in a field near Beijing (≈40°N, ≈116°E) in the natural LD photoperiod (see Materials and Methods), and the flowering time was measured in the summer and fall of 2007 (Fig. S4). The results of this common garden experiment demonstrated a strong latitudinal cline of photoperiodic flowering in the wild soybean accessions. Similar to that observed in the domesticated soybean cultivars (Fig. 4), the wild soybean collected from higher latitude flowered earlier than that collected from lower latitudes ($R^2 = 0.812$, $P < 0.0001$) when they were grown under the natural LD photoperiod condition. Therefore, domestication does not seem to positively contribute to the strong latitudinal cline in photoperiodic flowering of cultivated soybean. To the contrary, domestication may have negatively contributed to the latitudinal cline in flowering time of the soybean cultivars ($R^2 = 0.739$ in *G. max*), comparing to its wild relative ($R^2 = 0.822$ in *G. soja*). This seems consistent with the reduced photoperiodic sensitivity being a ''symptom'' of the ''domestication syndrome'' (4). It is intriguing that soybean is the only major crop commonly cultivated by matching ''maturity groups'' with the latitude of the site of cultivation (5, 6). Whether this phenomenon was due to relatively short domestication history of soybean in comparison to other major crops remains to be further investigated.

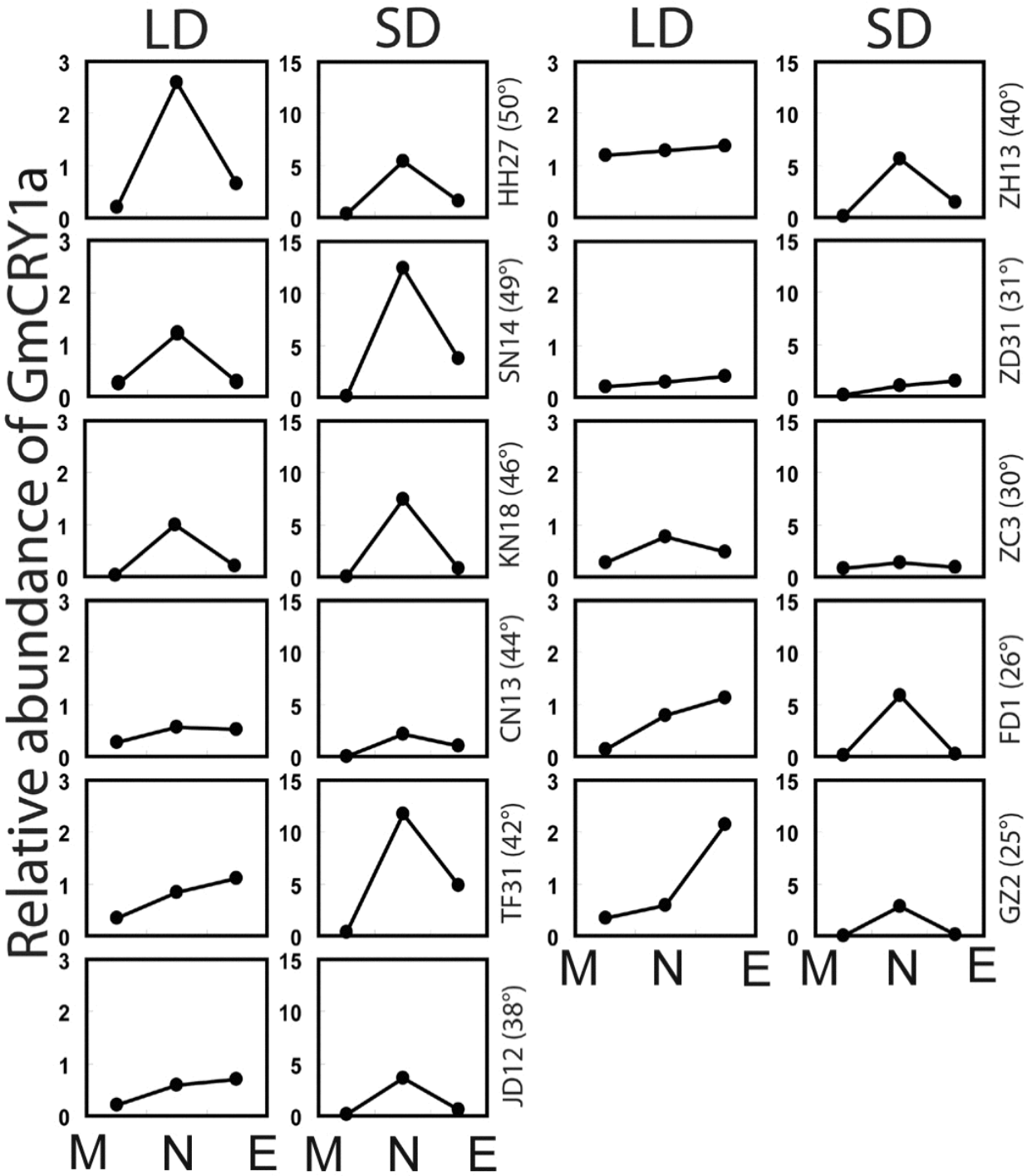


**Fig. S5.** A comparison of different patterns of rhythmic expression of the GmCRY1a protein in different soybean cultivars. The same data shown in Fig. 5 *A* and *B* are plotted here to show the oscillation of GmCRY1a expression level in each cultivar. Note the peak of GmCRY1a expression shown may not be the highest level of expression, because samples of only three time points were examined in this experiment (see Fig. 3 for comparison).

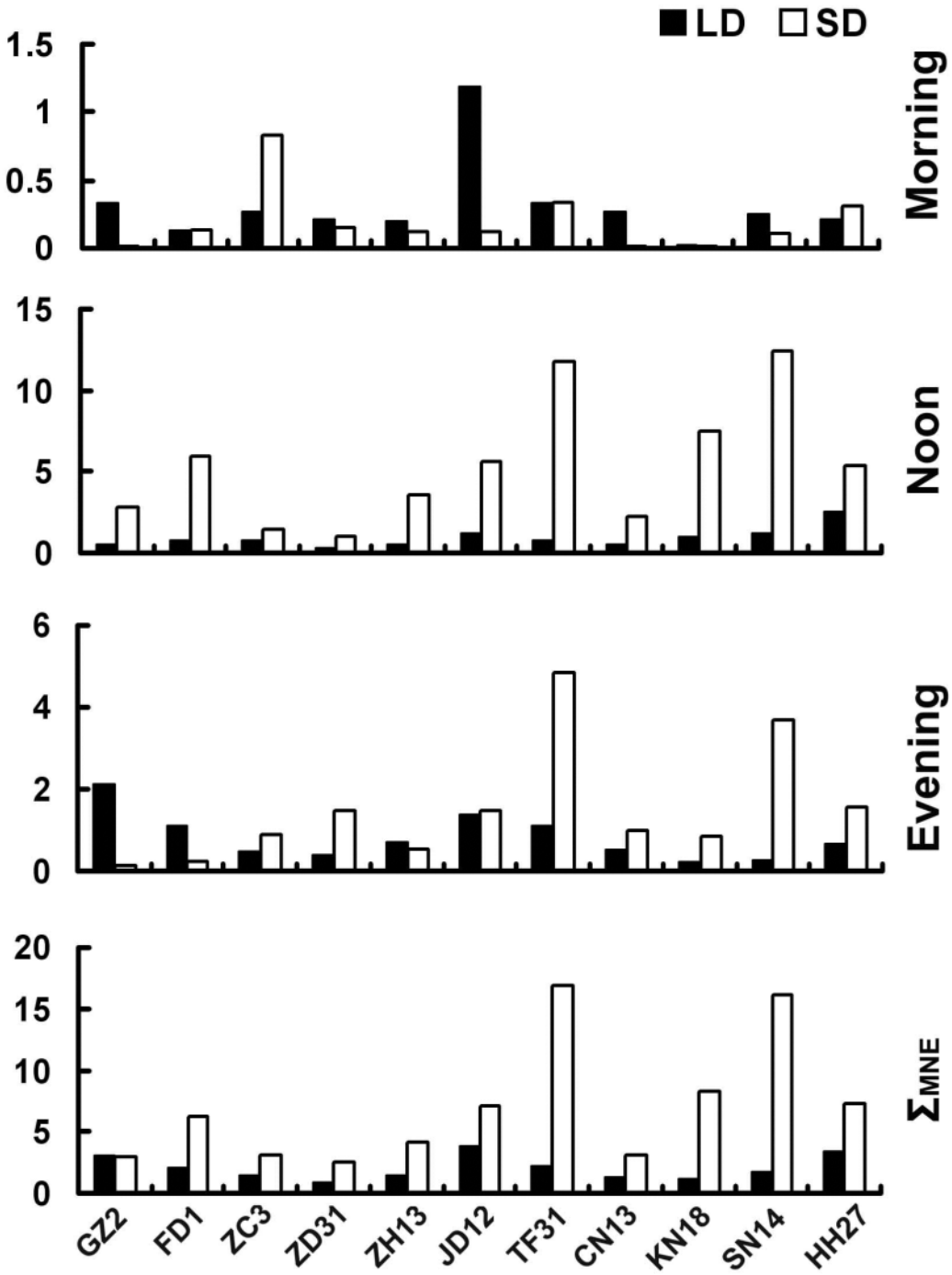


**Fig. S6.** A comparison of relative abundance of the GmCRY1a protein at different time of the day in different cultivars grown in LD or SD. The same data shown in Fig. 5 *A* and *B* are plotted here to show that the level of GmCRY1a protein is generally higher in SD than in LD. $\sum_{MNE}$, sum of GmCRY protein abundance in the morning, noon, and evening.

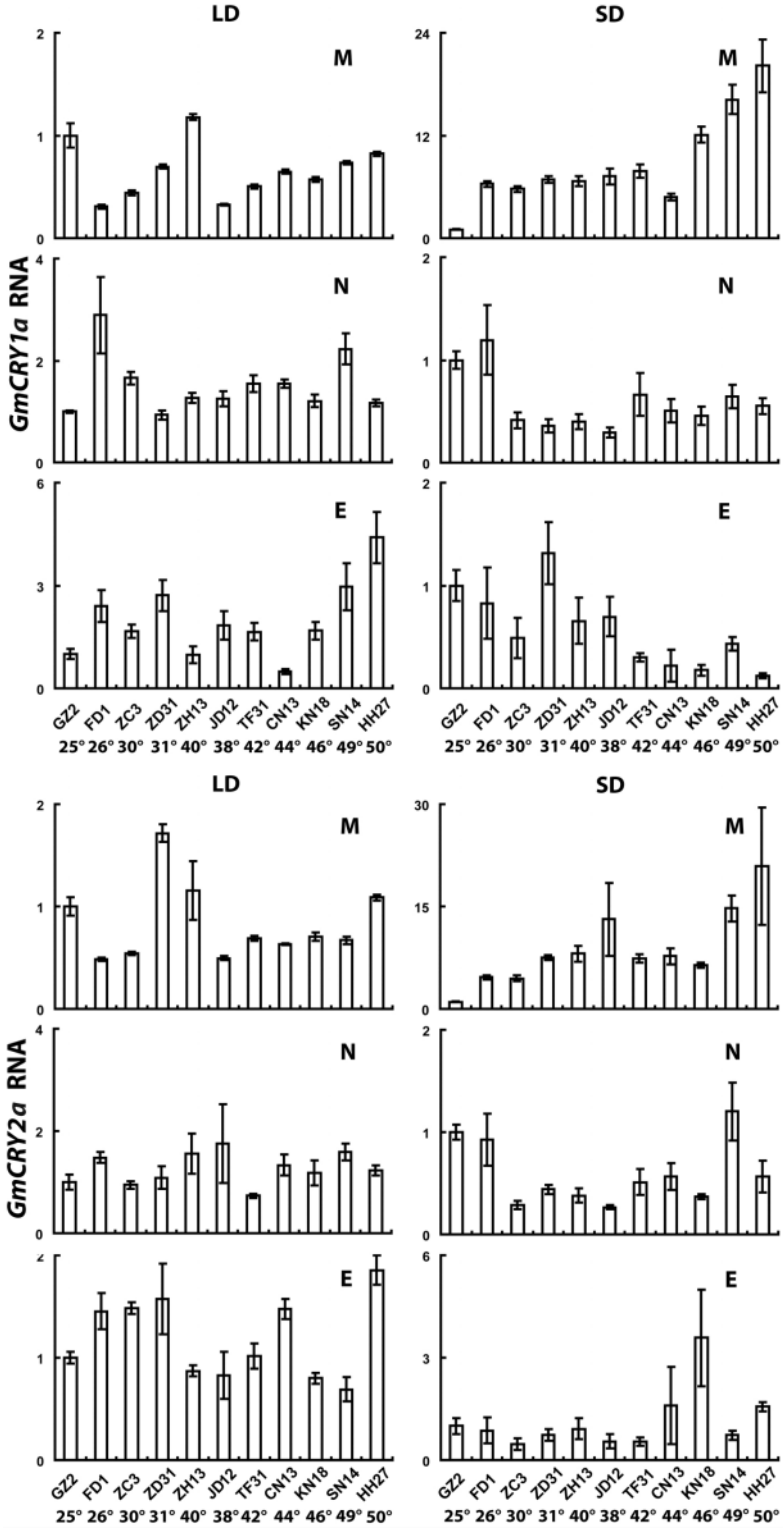


**Fig. S7.** A comparison of relative abundance of the *GmCRY1a* mRNA at different time of the day and in different cultivars grown in LD or SD. Results of the qPCR experiment showing the relative level of *GmCERY1a* and *GmCRY2a* mRNA in samples of different time, photoperiod, and cultivars indicated.