nature genetics

Natural variation at the soybean *J* locus improves adaptation to the tropics and enhances yield

Sijia Lu[1,2,8], Xiaohui Zhao[1,2,8], Yilong Hu[3,4,8], Shulin Liu[4,5,8], Haiyang Nan[1,8], Xiaoming Li[1,4], Chao Fang[1,4], Dong Cao[1,2], Xinyi Shi[1], Lingping Kong[1,4], Tong Su[1,4], Fengge Zhang[1,4], Shichen Li[1,4], Zheng Wang[5], Xiaohui Yuan[1], Elroy R Cober[6], James L Weller[7], Baohui Liu[1,2], Xingliang Hou[3], Zhixi Tian[5] & Fanjiang Kong[1,2]

Soybean is a major legume crop originating in temperate regions, and photoperiod responsiveness is a key factor in its latitudinal adaptation. Varieties from temperate regions introduced to lower latitudes mature early and have extremely low grain yields. Introduction of the long-juvenile (LJ) trait extends the vegetative phase and improves yield under short-day conditions, thereby enabling expansion of cultivation in tropical regions. Here we report the cloning and characterization of *J*, the major classical locus conferring the LJ trait, and identify *J* as the ortholog of *Arabidopsis thaliana EARLY FLOWERING 3* (*ELF3*). *J* depends genetically on the legume-specific flowering repressor *E1*, and J protein physically associates with the *E1* promoter to downregulate its transcription, relieving repression of two important *FLOWERING LOCUS T* (*FT*) genes and promoting flowering under short days. Our findings identify an important new component in flowering-time control in soybean and provide new insight into soybean adaptation to tropical regions.

Soybean (*Glycine max* (L.) Merr.) is one of the most economically important plant oil and protein crops and provides more than a quarter of the world's protein for food and animal feed[1]. Soybean is a short-day plant (SDP) and is highly sensitive to photoperiod. As in many photoperiodic species, its yield in a given latitudinal zone depends critically on an appropriate degree of photoperiod sensitivity[2,3]. This is reflected in the fact that traditional soybean cultivars are adapted to a relatively narrow range of latitudes[4]. Cultivated soybean was domesticated from its wild relative (*Glycine soja* Sieb. & Zucc.) 5,000 years ago[5,6] in temperate regions of China between 32° and 40° N (ref. 7). The subsequent expansion of cultivation worldwide saw soybean introduced to Korea, Japan, and South and Southeast Asia approximately 2,000 years ago, to North America in 1765, and to South America at the beginning of the last century[8]. Incorporation of new genetic resources has enabled the gradual extension of commercial soybean cultivation toward higher latitudes. However, until relatively recently, commercial cultivation at latitudes below 22° was impeded because of the poor adaptation of existing elite temperate cultivars, which flower early and produce extremely low grain yields when grown under the short photoperiods inherent in these tropical regions[3,9,10].

Identification of the LJ trait in the 1970s was a key breakthrough in overcoming this barrier[9–11]. LJ soybean varieties flower much later than temperate varieties under inductive short-day conditions[11] and also show delayed maturity, leading to improvements in plant height, node number, lodging degree, grain yield, and other important agronomic characteristics in the field at low latitudes[12–14]. The LJ trait was first introduced into soybean cultivars in central–western Brazil and subsequently enabled expansion of soybean production to regions below 15° latitude and even as far as the equator[3,9]. Largely on the basis of this innovation, Brazil has become the world's second largest soybean producer, despite the fact that 40% of its total soybean growing area is located below 24° S and thus outside the ancestral soybean adaptation zone[3,9,10].

While much attention has been given to the mechanisms of adaptation that have enabled efficient production at higher latitudes[2,15], less is known about the genetic basis for expansion into tropical regions. However, an important component of the LJ trait in North and South American cultivars is reported to be contributed by a major locus, *J* (refs. 11,12,16,17). Although the importance of *J* for low-latitude adaptation has been known for several decades[11,16], its molecular characterization has not yet been achieved. Here we identified the *J* gene through map-based cloning, examined the distribution of functional variants, and characterized the role and interactions of *J* in the soybean flowering pathway.

[1]Key Laboratory of Soybean Molecular Design Breeding, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin, China. [2]School of Life Sciences, Guangzhou University, Guangzhou, China. [3]Key Laboratory of South China Agricultural Plant Molecular Analysis and Genetic Improvement, and Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China. [4]University of the Chinese Academy of Sciences, Beijing, China. [5]State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [6]Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, Central Experimental Farm, Ottawa, Ontario, Canada. [7]School of Plant Science, University of Tasmania, Hobart, Tasmania, Australia. [8]These authors contributed equally to this work. Correspondence should be addressed to F.K. (kongfj@iga.ac.cn), Z.T. (zxtian@genetics.ac.cn), X.H. (houxl@scib.ac.cn) or B.L. (liubh@iga.ac.cn).

Received 10 October 2016; accepted 24 February 2017; published online 20 March 2017; doi:10.1038/ng.3819

RESULTS

Genetic control of the LJ trait

As a representative LJ accession, we initially selected PI 159925, a Peruvian landrace previously reported to carry the recessive *j* allele[11,16] (USDA GRIN database). In comparison with the reference temperate cultivar Harosoy, PI 159925 flowered and matured significantly later under controlled short-day conditions (**Fig. 1a–c**). In addition, PI 159925 significantly differed from Harosoy in a number of traits associated with grain yield, including plant height, number of nodes, number of branches, internode length, number of pods per plant, and number of grains per plant (**Supplementary Fig. 1a–f**). Consistent with these differences, PI 159925 ultimately exhibited significantly higher grain yield than Harosoy (**Fig. 1d**), implying an important role for the *j* allele. To understand the molecular basis of the *J* locus, we performed quantitative trait locus (QTL) mapping using F_2 populations developed from crosses between PI 159925 and Harosoy. The results indicated a position for *J* on soybean chromosome 4 (Gm04) (**Supplementary Fig. 2a** and **Supplementary Table 1**), in agreement with a previous preliminary report[18]. In addition to PI 159925, we also characterized a second LJ accession from Brazil, BR121, which also exhibited later flowering and maturity (**Fig. 1a–c**) and higher yield performance (**Fig. 1d** and **Supplementary Fig. 1a–f**). Genetic analysis of flowering time in BR121 × Harosoy progeny also showed the presence of a major locus in the same region of chromosome 4, suggesting that the LJ trait in this line is also due to the presence of a recessive *j* allele (**Supplementary Fig. 2b** and **Supplementary Table 1**). In addition, a second QTL coinciding with *E1*, a major soybean maturity locus[19], was consistently detected in both crosses (**Supplementary Fig. 2a**,**b** and **Supplementary Table 1**) under short-day conditions. Analysis of the interaction of these loci suggested that the *E1* QTL might have an epistatic effect on the *J* locus in both F_2 populations under short-day conditions (**Supplementary Fig. 2c**,**d**).

Positional cloning of the *J* locus

We next generated a large (n = 2,816) inbred F_6 population from the BR121 × Harosoy cross for fine-mapping of *J* by recurrent selection for heterozygosity at *J* from the F_2 to the F_5 generation[20]. Analysis of this population localized *J* within a 239-kb region between markers M1 and M3 (**Fig. 2a–c**), a region that, according to the Williams 82 reference genome[21], harbors 29 genes (**Fig. 2c** and **Supplementary Table 2**). We cloned and sequenced all 29 of the predicted genes in the two parents; of these, only one (*Glyma.04G050200.1*) differed in sequence between BR121 and Harosoy (**Fig. 2d**). In addition to a SNP, BR121 harbored a 10-bp deletion predicted to cause a frameshift resulting in premature termination of translation after 195 amino acids in the 714-amino-acid Glyma.04G050200.1 protein and addition of 20 missense amino acids (**Fig. 2d**,**f** and **Supplementary Fig. 3**). Sequencing of the coding region of *Glyma.04G050200.1* in PI 159925, our other initial reference line for *j*, identified four polymorphisms relative to Harosoy, including three SNPs and the deletion of a single cytosine, which also directed a frameshift and truncation of the predicted protein (**Fig. 2d**,**e** and **Supplementary Fig. 3**). The presence of different deleterious mutations in these two lines suggested that the *Glyma.04G050200.1* gene was a strong candidate for the *J* locus.

To further test this possibility, we transformed the recessive *j* accession BR121 with the coding sequence (CDS) of *Glyma.04G050200.1* from Harosoy under the control of its native promoter and characterized T_4 progeny from two independent primary transformants (**Fig. 3a**). This analysis showed that the presence of the Harosoy gene fully rescued the late flowering and maturity (**Fig. 3b**,**c**) of BR121, strongly indicating that *Glyma.04G050200.1* is the basis for the *J* locus. We therefore refer hereafter to *Glyma.04G050200.1* as *J* and to the loss-of-function variants in PI 159925 and BR121 as *j-1* and *j-2*, respectively (**Fig. 2d**). Consistent with the comparison between BR121 and Harosoy (**Fig. 1** and **Supplementary Fig. 1**), the transformants also had significantly altered yield-related traits and reduced overall grain yield relative to untransformed BR121 (**Fig. 3d** and **Supplementary Fig. 4a–f**). As an additional approach to examine the specific effects of the *J* locus, we also compared the phenotypes of two F_6 near-isogenic lines (NILs) carrying either the functional Harosoy *J* allele (NIL-*J*) or the non-functional BR121 *j-2* allele (NIL-*j*) (**Fig. 3e** and **Supplementary Fig. 5a**,**b**). The results showed that NIL-*J* had significantly promoted flowering time and maturity (**Fig. 3f**,**g**), along with reduced plant height, node number, pod number, total grain number (**Supplementary Fig. 4g–j**), and grain yield (**Fig. 3h**), confirming that the *j* allele could greatly enhance grain yield in soybean under short-day conditions.

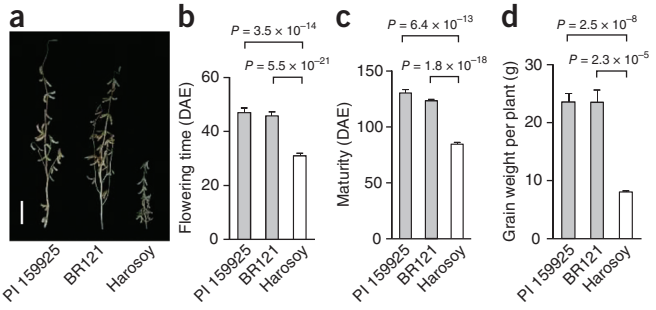

Figure 1 The long-juvenile trait delays flowering and increases yield under short-day conditions. (**a**) Phenotypes of two representative LJ lines, PI 159925 and BR121. Relative to the standard cultivar Harosoy, both LJ lines showed increased plant height, node number, and pod number. Scale bar, 10 cm. (**b**) Flowering time. Flowering time was recorded at the R1 stage (days from emergence to the appearance of the first open flower on 50% of the plants). DAE, days after emergence. (**c**) Maturity. Maturity was recorded at the R8 stage (days from emergence to the date at which 95% of pods had mature color). (**d**) Grain yield (total seed weight) per plant. The plants were grown in a standard field with artificially controlled short-day conditions (12 h light/12 h dark). All data are given as means ± s.e.m. (n = 10 plants). One-tailed, two-sample t-tests were used to generate the P values.

Phylogenetic investigation indicated that *J* is a co-ortholog of the *Arabidopsis* flowering-time gene *ELF3* (**Supplementary Fig. 6a**). ELF3 is a highly conserved plant-specific nuclear protein that has important roles in maintaining circadian rhythms and controlling flowering time in different species[22–24]. When the CDS of *J* driven by the *Arabidopsis ELF3* promoter was transformed into the *Arabidopsis elf3-8* mutant, it rescued the early-flowering mutant phenotype (**Supplementary Fig. 6b**,**c**). Overall, these results are consistent with several previous reports that *ELF3* has a highly conserved role in flowering time between *Arabidopsis* and crop species[23–26]. They are also consistent with recent reports demonstrating that ELF3 has a flower-promoting function in rice[24,27], an SDP species, in contrast to its inhibitory role in long-day plant (LDP) species[23,25,26,28]. This lends further support to an emerging picture that 'upstream' components of photoperiod response pathways (such as photoreceptor- and circadian clock–related genes) have opposite functions in SDPs[24–27] and LDPs[23,25,26,28], which implies that the molecular basis for the SDP–LDP difference likely resides in the regulation or activity of downstream integrator components.

J acts upstream of the legume-specific flowering repressor *E1*

Previous detailed genetic characterizations of early flowering in temperate soybean cultivars have identified a number of important genes

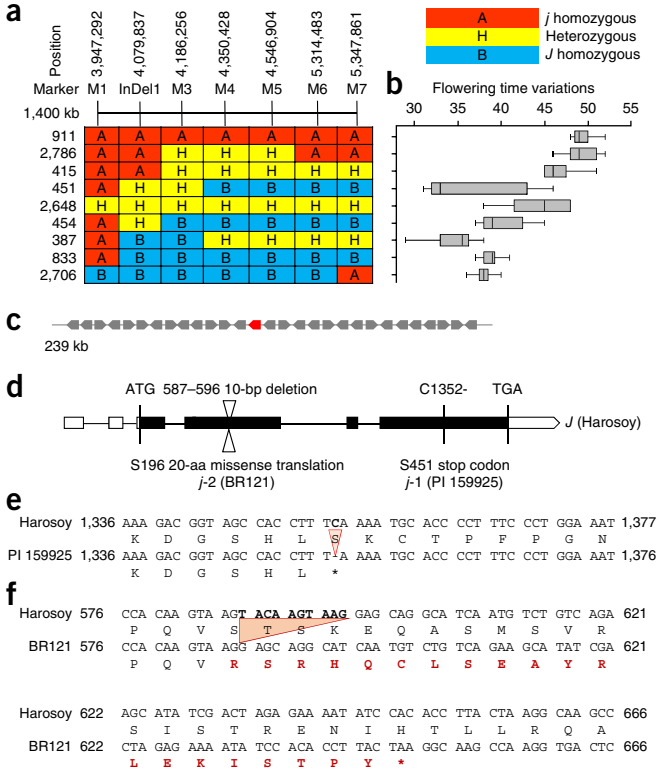

Figure 2 Identification of the *J* gene by map-based cloning. (**a**) Delimitation of the *J* locus to a 239-kb region of chromosome 4 in a segregating-heterozygous inbred family ($n = 2{,}816$) from a cross between BR121 and Harosoy. A, homozygous for the allele from BR121; B, homozygous for the allele from Harosoy; H, heterozygous. (**b**) Segregation of flowering time (d) is shown in box-plot format, where the interquartile region, median, and range are represented by the box, the bold vertical line, and the horizontal line, respectively ($n = 20$ plants). Each box plot corresponds to the segregant on the same row in **a**. (**c**) The 239-kb genomic region between markers M1 and M3 contains 29 predicted genes in the Williams 82 reference genome. (**d**) Allelic variation in the *J* candidate gene *Glyma04G050200* in PI 159925, BR121, and Harosoy. (**e**) Sequence comparison of Harosoy and PI 159925. The arrowhead indicates the position of the single deleted nucleotide in PI 159925 (shown in bold). An asterisk indicates the termination of translation. (**f**) Sequence comparison of Harosoy and BR121. The arrowhead indicates the position of the ten nucleotides deleted in BR121 (shown in bold). Missense amino acids resulting from the frameshift are shown in red. An asterisk indicates the termination of translation.

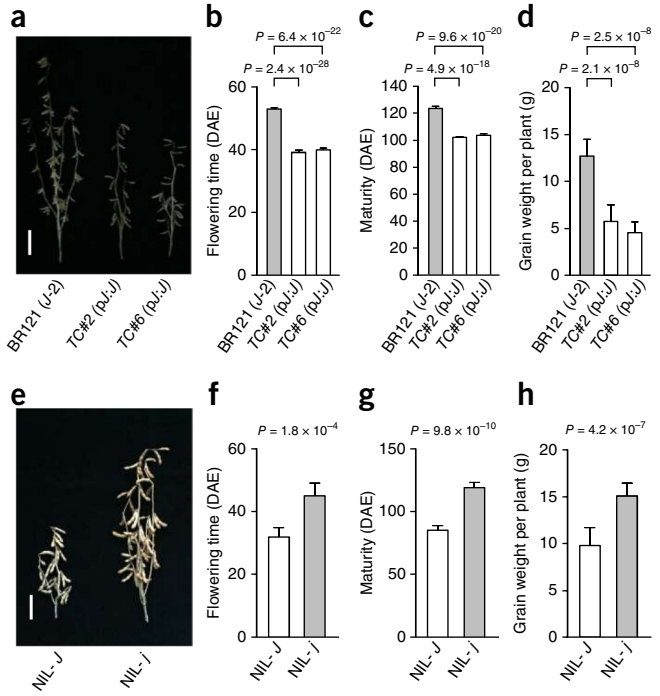

Figure 3 Confirmation of *J* identity and function by transgenic complementation and NIL analysis. (**a**–**d**) Transgenic complementation. Two independent transformants *TC#2* and *TC#6* were generated by genetic transformation of the CDS of *J* driven by its 3-kb native promoter from Harosoy into BR121, which possesses the deletion allele *j-2*. (**a**) T_4 progeny homozygous for the *J* transgene showed reduced plant height, number of nodes, and number of pods relative to untransformed BR121. Scale bar, 10 cm. (**b**) Flowering time. (**c**) Time to maturity. (**d**) Grain yield per plant. (**e**–**h**) Analysis of NILs differing at the *J* locus. (**e**) The presence of the *J* allele in NIL-*J* is associated with reduced plant height, number of nodes, and number of pods relative to NIL-*j* carrying the mutant form of *j*. Scale bar, 10 cm. (**f**) Flowering time. (**g**) Time to maturity. (**h**) Grain yield per plant. All plants were grown in a standard field with artificially controlled short-day conditions (12 h light/12 h dark). DAE, days after emergence. All data are given as means ± s.e.m. ($n = 10$ plants). One-tailed, two-sample *t*-tests were used to generate the *P* values.

needed for delay of flowering under non-inductive long-day conditions[2,29]. Among these, the legume-specific *E1* gene, which encodes a putative B3-like transcription factor, appears to have a central role in photoperiod-regulated flowering[19]. To understand the regulation of *J* and the functional relationship between *J* and *E1*, we first investigated their expression using independent NIL pairs for each locus (**Fig. 4a**,**b** and **Supplementary Fig. 5a**,**b**). In NIL-*J* plants grown under short-day conditions, expression of *J* showed a clear diurnal rhythm with a peak at dusk (**Fig. 4a**), consistent with the expression pattern of *ELF3* orthologs in other species[22–25]. In contrast, the *j* allele abolished the diurnal expression pattern, which is also similar to observations in other species[23,25,30] and is consistent with the possibility that *J* might sustain its own rhythmic expression[23,30]. In the same material, *E1* expression in NIL-*J* was very low, consistent with a previous report[19], but a striking derepression was observed in NIL-*j* (**Fig. 4b**). A similar effect of *J* was also observed in a comparison of BR121 (*j-2*) with the transgenically complemented line *TC#2* (*J*) (**Supplementary Fig. 7a**,**b**). Together, these results indicate that *J* acts to suppress *E1* expression under inductive short-day conditions. In contrast, the expression of *J* was not regulated by *E1* (**Supplementary Fig. 8a**), implying that *J* acts upstream of *E1* in the soybean flowering pathway. To further address the genetic interaction of *J* and *E1*, we developed a NIL set for the four different homozygous allelic combinations at the two loci from the cross between BR121 and Harosoy (**Supplementary Fig. 5a**,**c**) and evaluated flowering time and maturity differences under short-day conditions (**Fig. 4c**,**d**). The recessive Harosoy allele, $e1^{as}$, is a weak allele and carries a mutation directing an amino acid substitution that interferes with the nuclear localization of the E1 protein[19]. We found that, while the *j* allele delayed flowering and maturity on both an *E1* and an $e1^{as}$ genetic background, this effect was weaker in the $e1^{as}$ background, implying that *J* may depend on *E1* for its full activity. In contrast, the repression of flowering by *E1* was no less effective in the presence of a non-functional *j* allele, suggesting that *E1* does not depend functionally on *J*. These results are consistent with the allelic interaction between *J* and *E1* in two F_2 populations (**Supplementary Fig. 2c**,**d**) and the regulatory relationship indicated by expression analyses (**Fig. 4b** and **Supplementary Figs. 7b** and **8a**).

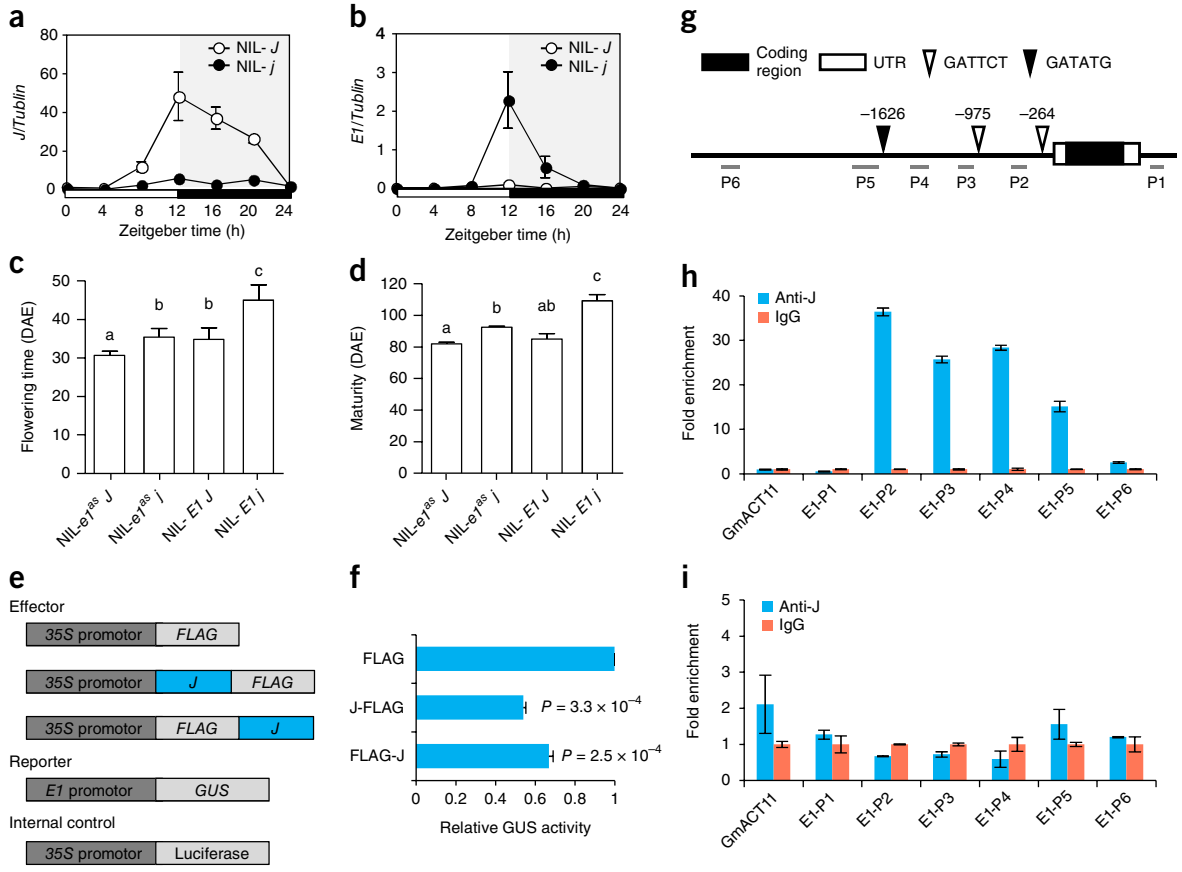

Figure 4 Genetic and regulatory interactions of *J* with *E1*. (**a**,**b**) Diurnal expression patterns of *J* (**a**) and *E1* (**b**) in leaf tissue from 20-d-old seedlings of NILs differing at *J*. Data shown are relative to the control gene *Tubulin* and represent means ± s.e.m. for three biological replicates. (**c**,**d**) Flowering time (**c**) and maturity (**d**) for NILs differing in genotype at the *E1* and *J* loci. Plants were grown in a standard field with artificially imposed short-day conditions (12 h light/12 h dark). All data are shown as means ± s.e.m. ($n = 10$ plants). The presence of the same lowercase letter above the histogram bars in **c** and **d** denotes non-significant differences across the two panels ($P > 0.05$). (**e**) Constructs of *J* and *E1* used for the transient transfection assay in *Arabidopsis* protoplasts. (**f**) J protein suppresses transcription from the *E1* promoter in *Arabidopsis* protoplasts. (**g**) Schematic of the *E1* gene and regions tested for enrichment in the ChIP assay. Filled and unfilled arrowheads represent the LUX-binding sites, with coordinates given as nucleotide position relative to the start codon (+1). Rectangles represent ChIP amplicons. (**h**) ChIP assay of *E1* amplicons in transformant *TC#2*. (**i**) ChIP assay of *E1* amplicons in BR121. Native monoclonal antibody raised against J was used for ChIP assays. One-tailed, two-sample *t*-tests were used to generate the *P* values.

J is a direct transcriptional repressor of *E1*

To investigate the molecular nature of the relationship between *J* and *E1*, we tested whether J might directly repress *E1* expression using a transient transfection approach in *Arabidopsis* protoplasts. The results showed that, when a p*35S:J* construct was cotransfected with a p*E1:GUS* construct (**Fig. 4e**), GUS activity was significantly suppressed by *J* (**Fig. 4f**), suggesting that the J protein might interact with the *E1* promoter to suppress its activity. To further determine whether J protein was bound to the *E1* promoter *in vivo*, we generated an antibody that binds the J protein (**Supplementary Fig. 9**) and used it to perform a chromatin immunoprecipitation (ChIP)–PCR assay comparing the relative enrichment of specific *E1* sequences (**Fig. 4g**) in the transgenic line *TC#2* (*J*) (**Fig. 4h**) and the BR121 (*j-2*) control line (**Fig. 4i**). In *Arabidopsis*, ELF3 physically interacts with ELF4 and LUX ARRYTHMO (LUX)[31,32] proteins to form the so-called evening complex (EC) that regulates transcription through binding to GATWCG sequence motifs (where W = A or T; referred to as LUX-binding sites) in the promoters of target genes[31]. We found that the J protein was associated with strong enrichment near three such sites (**Fig. 4g**) in the promoter of *E1* in the *TC#2* (*J*) line (**Fig. 4h**), while enrichment was extremely low in the control line, BR121 (*j-2*) (**Fig. 4i**), which lacks the sequence specifically interacting with the antibody owing to truncation of the J protein (**Supplementary Fig. 3**). Disruption of all three GATWCG motifs in the *E1* promoter in the mp*E1:GUS* construct abolished the suppressive effect of *J* on GUS activity (**Supplementary Fig. 10a**,**b**). These findings further support a model in which J influences flowering through transcriptional repression of *E1*, possibly acting through association with LUX-like proteins.

J may mediate the control of *E1* by *PHYA* genes *E3* and *E4*

Previous investigations of *E1* regulation and function have shown that *E1* expression is promoted by two *PHYTOCHROME A* (*PHYA*) homologs, *E3* (*GmPHYA3*) and *E4* (*GmPHYA2*)[33–35], and *E1* in turn acts to repress the expression of two *FT* homologs, *FT2a* and *FT5a*[19,36]. Relative to the NIL-*J* and *TC#2* lines, which both carry a functional *J* gene and strongly express *FT2a* and *FT5a* under short-day conditions, the corresponding *j* lines NIL-*j* and BR121 showed substantially reduced expression of both *FT* genes (**Supplementary Fig. 7c–f**), which is consistent with the higher expression of *E1* (**Fig. 4b** and **Supplementary Fig. 7b**). We observed that the BR121 *j-2* allele also caused a delay in flowering under long-day conditions (**Supplementary Fig. 11a**,**b**), which was associated with higher peak

expression of *E1* and lower peak expression of *FT2a* and *FT5a* under long-day conditions (**Supplementary Fig. 11c–f**). Using NILs for *E3* and *E4* in the Harosoy background[36], we found that expression of *J* was suppressed by the combined action of *E3* and *E4* (**Supplementary Fig. 8b**), As expected, *E3* and *E4* promoted the expression of *E1* (**Supplementary Fig. 8c**) and delayed flowering under short-day conditions (**Supplementary Fig. 8d**). These results suggest that *J* might function at least partially downstream of *E3* and *E4* under short-day conditions.

Taken together, these results suggest a working model for the role of the *J* gene in modulating photoperiod-regulated flowering in soybean under short-day conditions (**Supplementary Fig. 12**): *J* is suppressed by two PHYA proteins, E3 and E4, and the J protein physically binds to the promoter of *E1* near the LUX-binding motif to suppress *E1* transcription. This relieves the *E1*-dependent transcriptional repression of *FT2a* and *FT5a*, thereby promoting flowering. When the function of the *J* gene is impaired, *E1* itself is released from repression and is able to repress *FT2a* and *FT5a*, resulting in later flowering. These interactions outline a simple genetic pathway that is not conserved in model species such as *Arabidopsis* and rice. They also indicate a fundamental difference in ELF3 function between LDPs and SDPs. In the LDP *Arabidopsis*, *ELF3* delays flowering by indirectly suppressing expression of the key flowering activator *CONSTANS* and its main target, *FT*[37]. In contrast, it appears that, in the SDP species rice and soybean, *ELF3* instead promotes flowering by suppressing expression of the key *FT* repressors *Grain number, plant height and heading date 7* (*Ghd7*)[24,28] and *E1*, respectively.

Identification and functional analysis of novel *j* alleles

The earliest records documenting the broad utilization of *j* alleles in soybean breeding are from South America and date to the late 1970s[9–11]. However, soybean has been grown in tropical regions of South and Southeast Asia for over 2,000 years[5,8], and successful production also depends on delay of flowering under short-day conditions. This raises the question of whether other loss-of-function *J* alleles might exist in nature and might have been used in breeding for tropical regions. To investigate this question, we examined variation in the *J* coding sequence in our previously described collection of 302 resequenced accessions[38] and in an additional 125 non-resequenced accessions collected from low-latitude regions in Brazil, South and Southeast Asia, and southern China (**Supplementary Tables 3** and **4**). These sequence comparisons identified a total of 40 polymorphisms, including 34 SNPs and 6 indels, and defined 34 haplotypes (**Supplementary Table 5**), of which 2 (haplotypes H3 and H6, respectively) corresponded to the *j-1* and *j-2* frameshift alleles described above. The four other distinct indels (haplotypes H13, H4, H2, and H12) also generated frameshifts in the coding sequences (**Fig. 5a** and **Supplementary Fig. 3**) and were all associated with later flowering under short-day conditions (**Supplementary Fig. 13** and **Supplementary Table 6**). Therefore, these four haplotypes were also considered to be loss-of-function alleles and were named *j-3* to *j-6*, respectively (**Fig. 5a** and **Supplementary Table 6**). In addition to these six clear loss-of-function alleles, several nonsynonymous SNPs directing amino acid substitutions (p.Arg48Thr, p.Arg73Gly, p.Gln173Arg, p.Arg307Met, and p.Lys423Met) were also identified in the low-latitude accessions in haplotypes H15, H18, H19, and H20 (**Supplementary Table 4**). Of these, the p.Arg48Thr substitution in particular affects a highly conserved residue in the N-terminal conserved domain[39] of the J protein (**Supplementary Fig. 3**), and this conversion may disrupt the function of J. Transient assays showed that the nonsynonymous amino acid substitutions from H15 (p.Arg73Gly and p.Arg307Met) and H20 (p.Arg73Gly and p.Gln173Arg) had no effect on J activity relative to the functional Williams 82 haplotype (H1) (**Supplementary Fig. 10c** and **Supplementary Table 4**). However, the additional substitutions in H18 (p.Lys423Asn) and H19 (p.Arg48Thr) significantly impaired the ability of J to repress expression from the *E1* promoter, relative to H20 and H1 (**Supplementary Fig. 10c**). H18 and H19 thus appear to be weak loss-of-function alleles and were therefore named *j-7* and *j-8*, respectively (**Fig. 5a** and **Supplementary Table 6**).

An examination of the geographical origins of *J* haplotypes demonstrated that all eight mutational alleles (*j-1* to *j-8*) were restricted to accessions from low-latitude regions (**Fig. 5b** and **Supplementary Table 6**) and were not found in landraces originally domesticated from China[38] (**Supplementary Fig. 14** and **Supplementary Table 3**). This suggests that variation at *J* was not important for domestication of soybean but supports the idea that loss of *J* function confers advantage at lower latitude and has arisen independently several times as one important means of adaptation during expansion to these regions (**Fig. 5b**,**c**).

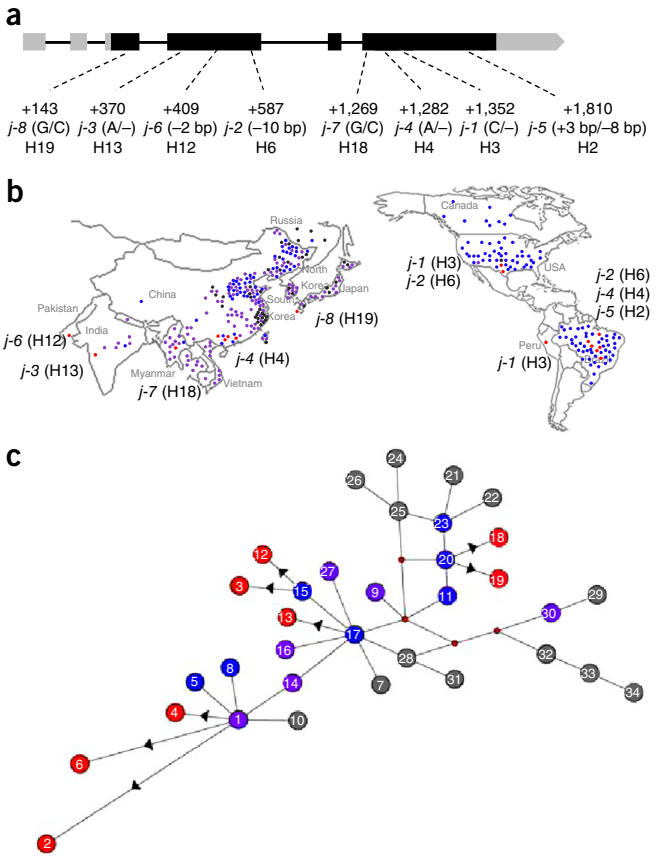

Figure 5 Sequence variation in *J* in relation to geographical distribution and origin. (**a**) Diagram of the *J* gene showing details of the eight identified mutations. (**b**) Geographical distribution of 376 accessions included in sequence analyses of *J*. Details of accessions are provided in **Supplementary Table 3**. Dark gray circles represent wild soybean, violet circles represent landraces, blue circles represent improved cultivars, and red circles represent accessions carrying the *j-1* to *j-8* mutational alleles. (**c**) Median-joining network representing the relatedness of 34 *J* haplotypes, each represented by a numbered circle. Details of the 34 haplotypes and their effects on J protein sequences are provided in in **Supplementary Table 5**. Dark gray circles represent wild soybean, violet haplotype circles represent landraces, blue haplotype circles represent improved cultivars, and red haplotype circles represent the eight mutational alleles.

Median-joining network analysis indicated that the loss-of-function haplotypes derived from different haplotypes (**Fig. 5c**), consistent with reports that cultivar Bossier (*j-2*; haplotype H6) was derived from Lee (haplotype H1) by induced mutation (USDA GRIN database) and that accession SS-1 (*j-4*; haplotype H4) was a naturally arising derivative of cultivar Parana (haplotype H1)[40] (**Fig. 5a**,**c** and **Supplementary Table 3**). We determined that the four lines from China carrying the *j-4* (haplotype H4) allele were all recently developed elite cultivars, suggesting that they were originally derived from SS-1 (**Supplementary Table 6**). These cultivars have greatly improved soybean production in southern China and include Huaxia 3, which shows greatly enhanced productivity relative to its local parent Guizao 1 under short-day conditions (**Supplementary Fig. 15a–g**) and recently achieved grain yield over 4,500 kg/ha as compared to the average yield of 3,000 kg/ha (ref. 41). Transgenic complementation of the *j-4* allele in Huaxia 3 with a functional copy of *J* further confirmed that the *j-4* allele significantly delayed flowering time (**Supplementary Fig. 15h**), maturity (**Supplementary Fig. 15i**), and enhanced grain yield in Huaxia 3 contrasting to its transgenic T_4 line, *TC#H7* (p*J*:*J*) (**Supplementary Fig. 15j–o**). Huaxia 3 has recently been successfully introduced to several countries at low latitudes in Africa, including Mozambique, Zambia, Angola, and Democratic Republic of the Congo[41]. Introgression of the loss-of-function alleles *j-2*, *j-4*, and *j-5* to temperate cultivars has also been important for development of LJ elite cultivars in Brazil (six cultivars) and southern China (four cultivars) (**Supplementary Fig. 13** and **Supplementary Table 6**), the second- and fourth-ranked countries for soybean production countries, suggesting that utilization of loss-of-function *J* alleles will be a valuable strategy to facilitate adaptation and enhance grain yield in the tropics generally. Our characterization of *J* adds to the growing body of evidence indicating that alteration to *ELF3* function is a preferred route to adaptive changes in flowering time in both SDPs (rice[24,27] and soybean) and LDPs (barley[25,26], wheat[28], pea, and lentil[23]). The main basis for this adaptation ultimately appears to be the altered expression of *FT* genes—either relief from repression to provide earlier flowering in LDP crops at high latitudes or impaired induction to delay flowering in SDP crops at low latitudes. However, it remains a possibility that other unique aspects of ELF3 function may provide additional adaptive advantage relative to other upstream regulators of *FT* genes.

Potential contribution of other EC genes to the LJ trait

Although variation in the *J* gene clearly has an important role in conferring the LJ trait and has been widely deployed in several major soybean production regions[9–11,16], the existence of many late-flowering lines from low-latitude regions that carry an apparently functional *J* allele suggests that on a global scale it is not the only locus responsible for this trait (**Supplementary Table 4**). This is consistent with previous observations that in some contexts the LJ trait is inherited in a quantitative manner and may be conditioned by at least three independent loci[11,13]. In *Arabidopsis*, mutants for the three EC components *ELF3*, *ELF4*, and *LUX* have similar early-flowering phenotypes[22,32,42], and the same is true in the long-day legume pea[23,43,44], pointing to a strong possibility that other EC genes in soybean might have similar effects as *J* and could potentially contribute to the LJ trait. Both *LUX* and *ELF4* are single-copy genes in diploid legumes[43,44] and as expected are represented by two homologs in soybean (**Supplementary Table 7**). In contrast, *ELF3* has undergone an ancient legume-specific duplication into two ancient clades, *ELF3a* and *ELF3b*[45].

Sequencing of these six additional soybean EC genes (two *ELF3* homologs[45], *ELF3b-1* and *ELF3b-2*; two *ELF4* homologs[43], *ELF4a* and *ELF4b*; and two *LUX* homologs[44], *LUX1* and *LUX2*) in 37 accessions carrying the functional *J* haplotype H1 (**Supplementary Table 4**) identified several polymorphisms with the potential to affect protein structure (**Supplementary Table 7**). Of these, a 12-bp insertion in the *LUX2* CDS (designated as the *LUX2-I12* allele) introduced four amino acids to the LUX2 C-terminal domain, which has been shown to have an important role in protein interactions[31] (**Supplementary Table 7**). Genotyping of the *LUX2-I12* allele in 125 low-latitude accessions identified 11 accessions carrying this allele (**Supplementary Table 8**). These results further indicate that the natural variation at EC genes may facilitate adaption into low latitudes in soybean. In any case, further genetic and molecular characterization of EC genes will help extend understanding of flowering-time control, adaptation mechanisms, and the history of soybean expansion to low-latitude regions.

DISCUSSION

In summary, the soybean *J* gene is a major source of adaptation to low-latitude regions, and we have used map-based cloning, transgenic complementation, and diversity analyses to show that *J* is a co-ortholog of *Arabidopsis ELF3*. A similar conclusion has recently been reported by Yue *et al.*[42]. *J* is likely the first major soybean locus to be characterized with a role in promotion of flowering under inductive conditions, and our functional analyses have integrated *J* into the emerging model for soybean flowering-time control. Our data indicate that *J* is a direct repressor of the key legume-specific flowering repressor *E1* and promotes flowering by relieving the suppression of *FT* gene expression by *E1*. Multiple loss-of-function *J* alleles are distributed across the global soybean germplasm and greatly prolong soybean maturity and enhance grain yield in the tropics. Our findings provide an important perspective on soybean flowering-time control and latitudinal adaptation that will hopefully help to drive improvements in soybean breeding for tropical regions around the world and secure sustainable global production of major soybean protein and oil products.

URLs. USDA GRIN database, http://www.ars-grin.gov/; Phytozome database, https://phytozome.jgi.doe.gov/.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank R. Nelson at the University of Illinois, J. Abe at Hokkaido University, H. Liao at the Fujian Agriculture and Forestry University, and L. Qiu at the Chinese Academy of Agricultural Sciences for sharing soybean germplasm and some phenotyping data. This work was supported by National Natural Science Foundation of China (Grant No. 31430065 to F.K.), National Key Research and Development Program (Grant No. 2016YFD0100401 to F.K.), "Strategic Priority Research Program" of the Chinese Academy of Sciences (Grant No. XDA08030108 to F.K.), National Key Research and Development Program (Grant No. 2016YFD0101900 to X.Z.), National Natural Science Foundation of China (Grant Nos. 31571686, 31371643, 31071445 to F.K. and 91531304, 31525018 to Z.T.), "Strategic Priority Research Program" of the Chinese Academy of Sciences (Grant No. XDA08020202 to Z.T.), the Open Foundation of the Key Laboratory of Soybean Molecular Design Breeding of Chinese Academy of Sciences, and "One-hundred talents" Startup Funds from Chinese Academy of Sciences to B.L. and F.K.

AUTHOR CONTRIBUTIONS

F.K. and B.L. designed the experiments and managed the projects. S. Lu, X.Z., Y.H., H.N., X.L., C.F., L.K., D.C., E.R.C., T.S., F.Z., and S. Li performed experiments.

S. Lu, S. Liu, Y.H., X.S., Z.W., X.Y., J.L.W., E.R.C., X.H., Z.T., and F.K. performed data analysis. F.K., Z.T., X.H., and J.L.W. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Graham, P.H. & Vance, C.P. Legumes: importance and constraints to greater use. *Plant Physiol.* **131**, 872–877 (2003).
2. Watanabe, S., Harada, K. & Abe, J. Genetic and molecular bases of photoperiod responses of flowering in soybean. *Breed. Sci.* **61**, 531–543 (2012).
3. Destro, D., Carpentieri-Pipolo, V., Kiihl, R.A.S. & Almeida, L.A. Photoperiodism and genetic control of the long juvenile period in soybean: a review. *Crop Breed. Appl. Biotechnol.* **1**, 72–92 (2001).
4. Parvez, A.Q. & Gardner, F.P. Daylengh and sowing date responses of soybean lines with "Juvenile" trait. *Crop Sci.* **27**, 305–310 (1987).
5. Hymowitz, T. On the domestication of the soybean. *Econ. Bot.* **24**, 408–421 (1970).
6. Carter, T.E., Nelson, R., Sneller, C.H. & in *Soybeans: Improvement, Production and Uses* 3rd edn. (eds. Boerma, H.R. & Specht, J.E.) 303–416 (American Society of Agronomy–Crop Science Society of America–Soil Science Society of America, 2004).
7. Li, Y. *et al.* Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. *Theor. Appl. Genet.* **117**, 857–871 (2008).
8. Wilson, R.F. *Soybean: Market Driven Research Needs in Genetics and Genomics of Soybean* (Springer, 2008).
9. Neumaier, N. & James, A.T. Exploiting the long-juvenile trait to improve adaptation of soybeans to the tropics. *Food Legume Newsl.* **8**, 12–14 (1993).
10. Spehar, C.R. Impact of strategic genes in soybean on agricultural development in the Brazilian tropical savannah. *Field Crops Res.* **41**, 141–146 (1995).
11. Hartwig, E.E. & Kiihl, R.A.S. Identification and utilization of a delayed flowering character in soybeans for short-day condtions. *Field Crops Res.* **2**, 145–151 (1979).
12. Sinclair, T.R. & Hinson, K. Soybean flowrering in response to the long-juvenile trait. *Crop Sci.* **32**, 1242–1248 (1992).
13. Carpentieri-Pipoplo, V., Almeida, L.A.D. & Kiihl, R.A.S. Inheritance of a long juvenile period under short-day conditions in soybean. *Genet. Mol. Biol.* **25**, 463–469 (2002).
14. Carpentieri-Pipolo, V., Almeida, L.A.D., Kiihl, R.A.S. & Pagliosa, E.S. Inheritance of late flowering in natural variants of soybean under short-day conditions. *Pesquisa Agropecu. Bras.* **49**, 796–803 (2014).
15. Xu, M. *et al.* Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. *BMC Plant Biol.* **13**, 91 (2013).
16. Ray, J.D., Hinson, K., Mankono, J. & Malo, M.F. Genetic control of a long-juvenile trait in soybean. *Crop Sci.* **35**, 1001–1006 (1995).
17. Cober, E.R. Long juvenile soybean flowering responses under very short photoperiods. *Crop Sci.* **51**, 140–145 (2011).
18. Cairo, C.A. *et al.* Tagging the juvenile locus in soybean [*Glycine max* (L.) Merr.] with molecular markers. *Euphytica* **124**, 387–395 (2002).
19. Xia, Z. *et al.* Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proc. Natl. Acad. Sci. USA* **109**, E2155–E2164 (2012).
20. Watanabe, S. *et al.* A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics* **188**, 395–407 (2011).
21. Schmutz, J. *et al.* Genome sequence of the palaeopolyploid soybean. *Nature* **463**, 178–183 (2010).
22. Hicks, K.A., Albertson, T.A. & Meeks-Wagner, D.R. *EARLY FLOWERING 3* encodes a novel protein that regulates circadian clock function and flowering in *Arabidopsis*. *Plant Cell* **13**, 1281–1292 (2001).
23. Weller, J.L. *et al.* A conserved molecular basis for photoperiod adaptation in two temperate legumes. *Proc. Natl. Acad. Sci. USA* **109**, 21158–21163 (2012).
24. Matsubara, K. *et al.* Natural variation in *Hd17*, a homolog of *Arabidopsis ELF3* that is involved in rice photoperiod flowering. *Plant Cell Physiol.* **53**, 709–716 (2012).
25. Faure, S. *et al.* Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. *Proc. Natl. Acad. Sci. USA* **109**, 8328–8333 (2012).
26. Zakhrabekova, S. *et al.* Induced mutations in circadian clock regulator *Mat-a* facilitated short-season adaptation and range extension in cultivated barley. *Proc. Natl. Acad. Sci. USA* **109**, 4326–4331 (2012).
27. Saito, H. *et al. Ef7* encodes an ELF3-like protein and promote rice flowering by negatively regulating the floral repressor gene *Ghd7* under both short- and long-day conditions. *Plant Cell Physiol.* **53**, 717–728 (2012).
28. Alvarez, M.A., Tranquilli, G., Lewis, S., Kippes, N. & Dubcovsky, J. Genetic and physical mapping of the earliness per se locus *Eps-A^{m}1* in *Triticum monococcum* identifies *EARLY FLOWERING 3* (*ELF3*) as a candidate gene. *Funct. Integr. Genomics* **16**, 365–382 (2016).
29. Cao, D. *et al.* Molecular bases of flowering under long days and stem growth habit in soybean. *J. Exp. Bot.* http://dx.doi.org/10.1093/jxb/erw394 (2016).
30. McWatters, H.G., Bastow, R.M., Hall, A. & Millar, A.J. The ELF3 zeitnehmer regulates light signalling to the circadian clock. *Nature* **408**, 716–720 (2000).
31. Nusinow, D.A. *et al.* The ELF4–ELF3–LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* **475**, 398–402 (2011).
32. Hazen, S.P. *et al. LUX ARRHYTHMO* encodes a Myb domain protein essential for circadian rhythms. *Proc. Natl. Acad. Sci. USA* **102**, 10387–10392 (2005).
33. Watanabe, S. *et al.* Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics* **182**, 1251–1262 (2009).
34. Liu, B. *et al.* Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. *Genetics* **180**, 995–1007 (2008).
35. Cao, D. *et al. GmCOL1a* and *GmCOL1b* function as flowering repressors in soybean under long-day conditions. *Plant Cell Physiol.* **56**, 2409–2422 (2015).
36. Kong, F. *et al.* Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. *Plant Physiol.* **154**, 1220–1231 (2010).
37. Yu, J.W. *et al.* COP1 and ELF3 contol circadian function and photoperiodic flowering by regulating GI stability. *Mol. Cell* **32**, 617–630 (2008).
38. Zhou, Z. *et al.* Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* **33**, 408–414 (2015).
39. Herrero, E. *et al.* EARLY FLOWERING4 recruitment of EARLY FLOWERING3 in the nucleus sustains the *Arabidopsis* circadian clock. *Plant Cell* **24**, 428–443 (2012).
40. Bonato, E.R. & Vello, N.A. *E6*, a dominant gene conditioning early flowering and maturity in soybeans. *Genet. Mol. Biol.* **22**, 229–232 (1999).
41. Yue, Y. *et al.* A single nucleotide deletion in *J* encoding GmELF3 confers long juvenility and is associated with adaption of tropic soybean. *Mol. Plant* http://dx.doi.org/10.1016/j.molp.2016.12.004 (2016).
42. Doyle, M.R. *et al.* The *ELF4* gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. *Nature* **419**, 74–77 (2002).
43. Liew, L.C. *et al. DIE NEUTRALIS* and *LATE BLOOMER 1* contribute to regulation of the pea circadian clock. *Plant Cell* **21**, 3198–3211 (2009).
44. Liew, L.C., Hecht, V., Sussmilch, F.C. & Weller, J.L. The pea photoperiod response gene *STERILE NODES* is an ortholog of *LUX ARRHYTHMO*. *Plant Physiol.* **165**, 648–657 (2014).
45. Rubenach, A.J.S. *et al. ELF3* redundancy fine-tunes photoperiod sensitivity in pea. *Plant Physiol.* http://dx.doi.org/10.1104/pp.16.01738 (2017).

ONLINE METHODS

Plant materials, growth conditions, and phenotyping. The soybean parental lines PI 159925 and Harosoy were obtained from the USDA Soybean Germplasm Collection, and BR121 was obtained from the Guangxi Academy of Agricultural Sciences, Guangxi, China. For initial QTL mapping, two F_2 populations were generated from PI 159925 × Harosoy and BR121 × Harosoy crosses, consisting of 94 and 74 progenies, respectively (**Supplementary Table 1**). For map-based cloning, a heterozygous inbred progeny of 2,816 individuals segregating at the *J* locus was developed from the BR121 × Harosoy cross. NILs for both the *E1* and *J* loci were also selected from the F_6 progeny of this same cross, using molecular markers for *E1* and *J*. Soybean plants from the F_2 population, heterozygous inbred progeny, NILs, low-latitude-adapted accessions, and transformants for phenotyping were grown under artificial short-day conditions in the field (12 h light/12 h dark) from 2010 to 2016 at the Experimental Station of the Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin (45° 75′ N, 126° 63′ E). 302 whole-genome-resequenced wild soybeans, landraces, and improved cultivars reported previously[38] and 125 non-resequenced accessions collected from low-latitude regions in this study were used for analysis. Artificial short-day conditions in the field were set up using a walk-in arched iron shelf that was 20 m long, 5 m wide, and 2.5 m high covered by black plastic. Day length was achieved by opening the plastic at 6 a.m. and closing at 6 p.m. each day. Plants were spaced 0.15 m apart in rows that were 5 m long with 0.7 m between rows. Plants were sowed at the beginning of May and harvested in September or October of each year. Plants for expression analysis and ChIP assays were grown under short-day (12 h light/12 h dark) or long-day (14 h light/10 h dark) conditions in plant growth cabinets (CONVIRON ADAPTIS, A1000) with light intensity of 500 $\mu mol/m^2/s$.

Flowering time was recorded at the R1 stage (days from emergence to the first open flower appeared on 50% of the plants). Maturity was recorded at the R8 stage (days from emergence to when 95% of pods had the mature color)[46]. Plant height, number of branches, number of nodes, average internode length, number of pods per plant, number of grains per plant, and yield per plant were all recorded at the R8 stage.

Molecular mapping, map-based cloning, and genotyping. Linkage map construction was performed according to previous reports[47,48]. QTL detection and the logarithm of odds (LOD) threshold were achieved with QTL IciMapping[49]. Primer sequences of the markers for mapping are listed in **Supplementary Table 9**. For fine-mapping, six additional indel markers between markers BAR_04_0002 and PSI0526 were identified (**Fig. 2a**). Nine recombinants were identified in the fine-mapping population using seven markers (**Fig. 2a**), and the flowering time of their progeny was evaluated to delimit the genomic interval containing *J*. The genotypes of the *J* allele were analyzed by tagging marker HRM1 or M2 (**Supplementary Table 9**). The *E1* allele was genotyped by its functional markers[15,19].

DNA and RNA isolation, PCR, RT–PCR, qRT–PCR, sequencing, and alignment. Genomic DNA isolation, PCR primer design, PCR amplification, indel marker development, PCR fragment purification, total RNA isolation, cDNA synthesis by RT–PCR, qRT–PCR, and sequencing of PCR and RT–PCR fragments were carried out as described previously[36]. The *Tubulin* gene was used as the internal control in qRT–PCR experiments. The two fragments of the *J* genomic sequence from 125 low-latitude accessions were cloned and sequenced. The EC genes (two *ELF3* homologs, *ELF3b-1* and *ELF3b-2*; two *ELF4* homologs, *ELF4a* and *ELF4b*; two *LUX* homologs, *LUX1* and *LUX2*) (**Supplementary Table 7**) were cloned from 37 accessions with *J* haplotype H1 and sequenced. The *LUX2-I12* alleles from 125 accessions were genotyped using gene-specific markers. Alignment of nucleotide and amino acid sequences was performed with Clustal X[50]. The primers used for PCR, RT–PCR, qRT–PCR, sequencing, and genotyping are listed in **Supplementary Table 10**.

Plasmid construction and plant transformation. The CDS of the *J* candidate gene and the 3-kb promoter region upstream of the start codon of the candidate gene were obtained from Harosoy by PCR amplification. The 3-kb promoter and CDS fragment were amplified by overlapping PCR to obtain one fragment and then introduced into the pTF101-Gene vector (containing the *bar* gene for glufosinate resistance), replacing the p*35S*-gene fragment. This new construct (p*J-J*) was introduced into *Agrobacterium* strain EHA101, and *Agrobacterium*-mediated transformation of the soybean accession BR121 was performed as described previously[35,51]. The ~3-kb *E1* promoter sequence was amplified from NIL-*E1 E3 E4* and was introduced into pTF101-*GUS* to generate the construct p*E1-GUS*. To generate the mp*E1-GUS* construct, a ~3-kb *E1* promoter was synthesized in which three LUX-binding site motifs were mutated as described in **Supplementary Figure 10a** and replaced the ~3-kb promoter of construct p*E1-GUS*. The CDSs of *J* from Harosoy (H5) and of *J* fused with three FLAG tags at either the N- or C-terminal end were introduced into pTF101-Gene to generate the p*35S-J*, p*35S*-FLAG-*J*, and p*35S-J*-FLAG constructs, respectively. The CDSs of H1, H15, H18, H19, and H20, the nonsynonymous SNP haplotypes of *J*, were introduced into pTF101-Gene to generate the p*35S*-H1, p*35S*-H15, p*35S*-H18, p*35S*-H19, and p*35S*-H20 constructs. To generate the p*ELF3-J* construct, a ~1.8-kb fragment of the *ELF3* promoter was amplified from Col-0 and cloned into p*J-J* to replace the *J* promoter. All the primers used for vector construction are listed in **Supplementary Table 11**.

Transient expression assays. The above-described constructs p*E1-GUS* and mp*E1-GUS* were used as reporters and the p*35S-J-FLAG*, p*35S*-H1, p*35S*-H15, p*35S*-H18, p*35S*-H19, p*35S*-H20, and p*35S-FLAG-J* constructs were used as effectors, while a construct containing firefly luciferase with expression driven by the *35S* promoter in pGreen-*35S* was used as an internal control to evaluate the protoplast transfection efficiency[52]. Mesophyll protoplasts were isolated from the leaves of 3-week-old *Arabidopsis*. Subsequently, the reporter, effector, and internal control plasmids were cotransformed into the prepared mesophyll protoplast by PEG4000 (Sigma) as previously described[53] and then cultured in darkness overnight. The activities of GUS and luciferase were measured using a microfluorometer (Cary Eclipse). Relative GUS activity was calculated by normalizing against the luciferase activity, and the data presented are the averages of at least three independent replicates.

Generation of antibody to J and immunoblotting. Antibodies were generated in rabbits (AB Mart) against a J-specific peptide, QTSPADAINDTEHH, corresponding to amino acids 284–297. Antibodies were affinity purified against this peptide using a SulfoLink Immobilization kit (Thermo Scientific). Eluted antibody-containing fractions were buffer exchanged into 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 50% glycerol, and 0.02% NaN_3 by using an equilibrated PD-10 column (GE Healthcare) and then stored at −80 °C. The specificity of the antibody to J was tested by immunoblotting and is shown in **Supplementary Figure 10**. Immunoblotting was performed as described previously[52]. The use of experimental animals was approved by the Science and Technology Committee of Shanghai, China.

ChIP assays. Leaf samples were collected from 20-d-old seedlings at zeitgeber time 12 under short-day conditions from BR121 and its transformant *TC#2*. Samples were fixed on ice for 45 min in 1% formaldehyde under vacuum. Fixed tissues were homogenized, and chromatin was isolated and sonicated as previously described[54]. The solubilized chromatin was immunoprecipitated by antibody to J or mouse IgG (15381, Sigma) with Protein G PLUS agarose (sc-2002, Santa Cruz Biotechnology). The immunoprecipitated DNA was recovered and analyzed by quantitative real-time PCR in triplicate. Relative fold enrichment was calculated by normalizing the amount of a target DNA fragment against that of a genomic fragment of a reference gene, *ELONGATION FACTOR 1 GmELF1B* (*Glyma.02G276600.1*) and then by normalizing the value for immunoprecipitation using a specific antibody against that of mouse IgG. The enrichment of an *ACTIN* gene, *GmACT11* (*Glyma.18G290800.1*), genomic fragment was used as a negative control. The primers used for amplification are listed in **Supplementary Table 12**.

Statistical analyses. For phenotypic evaluation, at least five individual plants were analyzed per accession, and exact numbers of individuals (*n*) are presented in all figure legends. For expression analyses using qRT–PCR, at least three individual plants were pooled per tissue sample and at least three qRT–PCR reactions (technical replicates) were performed. The exact number of replicates is given in figure legends. Mean values for each measured

doi:10.1038/ng.3819

parameter were compared using one-way analysis of variance from SPSS (version 20, IBM) or one-tailed, two-sample Student's *t*-tests from Microsoft Excel, whenever appropriate; the statistical tests used for each experiment are given in figure legends.

Data availability. The nucleotide sequences data for the *J* allele and loss-of-function alleles *j-1* to *j-6* have been deposited in GenBank under accessions KX943537–KX943543.

46. Wright, D. & Lenssen, A.W. *Staging Soybean Development* (Agriculture and Environment Extension Publications, 2013).
47. Lu, S. *et al.* QTL mapping for flowering time in different latitude in soybean. *Euphytica* **206**, 725–736 (2015).
48. Kong, F. *et al.* A new dominant gene *E9* conditions early flowering and maturity in soybean. *Crop Sci.* **54**, 2529–2535 (2014).
49. Meng, L., Li, H., Zhang, L. & Wang, J. QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* **3**, 269–283 (2015).
50. Thompson, J.D., Gilson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882 (1997).
51. Nan, H. *et al.* GmFT2a and GmFT5a redundantly and differentially regulate flowering through interaction with and upregulation of the bZIP transcription factor GmFDL19 in soybean. *PLoS One* **9**, e97669 (2014).
52. Liu, X. *et al.* The NF-YC–RGL2 module integrates GA and ABA signaling to regulate seed germination in *Arabidopsis*. *Nat. Commun.* **7**, 12768 (2016).
53. Yoo, S.D., Cho, Y.H. & Sheen, J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* **2**, 1565–1572 (2007).
54. Hou, X. *et al.* Nuclear factor Y–mediated H3K27me3 demethylation of the *SOC1* locus orchestrates flowering responses of *Arabidopsis*. *Nat. Commun.* **5**, 4601 (2014).