Supporting Information

Murphy et al. 10.1073/pnas.1106212108

SI Methods

Mapping Populations. To identify polymorphisms between parental genotypes, simple sequence repeat (SSR) markers were developed using *SSRIT* (1). De novo sequencing and analysis of promoter and intron regions flanking predicted ORFs identified additional PCR-based SNP and insertion–deletion (INDEL) markers used for high-resolution linkage analysis. PCR primers were designed with Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) or PrimerQuest (Integrated DNA Technologies, Inc.) to amplify selected regions of the sorghum genome (http://www.phytozome.net/sorghum). Purified PCR products were used in sequencing reactions with Big Dye Terminator v3.1 (Applied Biosystems) and run on an ABI 3130xl Genetic Analyzer. Sequence assembly and analysis was carried out using Sequencher v4.8 (Gene Codes) or Phred/Phrap and Consed (http://www.phrap.org/phredphrapconsed.html). Physical locations of genetic markers and associated high-resolution map units were based on the whole-genome sequence of *Sorghum bicolor* (http://www.phytozome.net/sorghum; v5.0). All markers used in this study are listed with their physical coordinates in Table S4.

Sequencing of *SbPRR37* Alleles. For cDNA sequence determination of *SbPRR37* alleles from genotypes in Table S3, RNA was extracted from leaf tissue of 2-wk-old plants using the miRNeasy Mini Kit (QIAGEN). After extraction, RNA concentration was determined spectrophotometrically, and RNA integrity was visually assessed by denaturing agarose gel electrophoresis. Ten micrograms of total RNA was DNase-treated with a Turbo DNA-Free Kit (Applied Biosystems/Ambion), and this RNA (1 μg) was used for first-strand cDNA synthesis (SuperScript III First-Strand Synthesis System, Invitrogen) primed with either oligo$(dT)_{20}$ or random hexamers. After inactivation of reverse transcriptase, the cDNA was diluted to a final concentration of 5 ng/μL with dH_2O. On the basis of the full-length sequence of *SbPRR37* ESTs from 100M, SM100, and Blackhull Kafir, PCR primers (5′-ACAAGATCATCGTCCCATCTC-3′ and 5′-TGAAACTGACACACGGCACTA-3′) were designed adjacent to the transcription start and stop sites of *SbPRR37*, and cDNA from the remaining sorghum genotypes listed in Table S3 were PCR-amplified from first-strand cDNA template. PCR products were cloned into the pCRII-TOPO vector and One Shot TOP10 Competent cells (Invitrogen). Forty-eight independent clones were selected from each genotype for sequence determination of the *SbPRR37* alleles. Phred and Phrap were used for sequence base calling and sequence assembly, respectively, and Consed Graphical Tool was used for sequence finishing (http://www.phrap.org/phredphrapconsed.html). Intron–exon borders were annotated by aligning the genomic and cDNA sequences.

Sequence Comparison and Homology Modeling. *SbPRR37* alleles were analyzed for conserved protein domains using National Center for Biotechnology Information BLAST against the Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). The entire coding sequence of different *SbPRR37* alleles was compared with the GenBank nonredundant (nr) database using BLASTP to identify homologs (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Nucleotide and protein sequence alignments were assessed with ClustalW2 (http://workbench.sdsc.edu/).

Gene Expression Studies. Leaf tissue was homogenized in liquid nitrogen using a mortar and pestle, and total RNA was extracted from each sample using the TRI REAGENT Protocol for samples high in polysaccharides (Molecular Research Center, Inc.). The resulting RNA was further purified using the RNeasy Mini Kit with on-column DNase digestion (QIAGEN). RNA samples were quantified twice each using a NanoDrop 1000 Instrument (Thermo Fisher Scientific, Inc.), and the average value for each sample was used. Five micrograms of each RNA sample was vacuum-dried and resuspended in denaturing buffer, and RNA integrity was visualized on a 1% MOPS buffer gel using the Molecular Imager Gel Doc XR running Quantity One v4.6.8 software (Bio-Rad Laboratories, Inc.). First-strand cDNA synthesis was performed using the SuperScript III First-Strand Synthesis System (Invitrogen), primed with a 9:1 ratio of random hexamer:oligo(dT) mix using 4 μg of total RNA. The reactions were diluted to 10 ng/μL cDNA in 1× TE buffer for subsequent use in quantitative RT-PCR (qRT-PCR) expression analysis. This process was repeated twice more for a total of three biological replicates.

Primers for qRT-PCR–based expression analysis of select clock and flowering-time–related genes were designed using PrimerQuest (Integrated DNA Technologies, Inc.). Orthologs of rice and maize genes were identified on the basis of protein and nucleotide sequence similarities, as well as position within colinear regions (Gramene Genome Browser release #33: http://www.gramene.org/) (Table S5). Sorghum gene and cDNA sequences were obtained through Phytozome v5.0, and primers were engineered to span an intron or UTR sequences of each of the following genes: *PRR37*, *TOC1*, *LHY1*, *FT*, *GI*, *CO*, and *Ehd1*. To determine primer efficiencies, serial dilutions were constructed from purified PCR products amplified from cDNA using gene-specific primer sets over a range of 0.05 $ng \cdot \mu L^{-1}$ to $5.0E^{-7}$ $ng \cdot \mu L^{-1}$, which were subsequently used in qRT-PCR reactions. The resulting cycling threshold (Ct) values were used to obtain a standard curve by which the efficiencies for each primer pair for all genotypes was calculated (Applied Biosystems). Primers whose efficiencies were within 10% between genotypes were used for downstream analysis (Table S5). No-template control qRT-PCR reactions were also run using 18S ribosomal RNA (Applied Biosystems) for 10 $ng \cdot \mu L^{-1}$ RNA from each sample to verify that there was no genomic DNA contamination. All of these and subsequent reactions were performed on the ABI 7900HT Fast Real-Time PCR System running SDS v2.3 software.

Gene-specific reactions were performed using Power SYBR Green PCR Master Mix (Applied Biosystems). All control 18S ribosomal RNA reactions were performed using the TaqMan Universal PCR Master Mix Protocol with the rRNA Probe (VIC Probe), rRNA forward primer, and rRNA reverse primer (Applied Biosystems).

For the time-course studies, raw Ct values were collected for each gene and normalized to 18S ribosomal RNA to obtain ΔCt values. Relative expression was calculated using the comparative Ct (ΔΔCt) method (2) with the most highly expressed sample used as the calibration sample. For genes in which LD and SD values were to be directly compared (*FT*, *Ehd1*, *ZCN12*, *ZCN14*), one calibration sample was used for each replicate in both LD and SD samples. Mean values are based on three technical replicates and three biological replicates for both reference and target genes (±SEM).

Absolute quantification at 15 h after lights on was carried out using the standard curve method (2). PCR-amplified products from each primer set (*Ehd1*, *FT*, and ribosomal 18S mRNA) were

purified by gel electrophoresis and sequenced to ensure that no polymorphisms were present between genotypes. The resulting products were subsequently used to construct a dilution series spanning 0.05 $ng \cdot \mu l^{-1}$ to $5.0E^{-7} ng \cdot \mu l^{-1}$ for each gene. The number of molecules present in each dilution was calculated from the molecular weight of each product. The dilution series for each gene was then used on a qRT-PCR run in parallel with experimental samples, and absolute transcript abundance was calculated by the ABI 7900HT instrument running SDS v2.3 software. Transcript abundance was then expressed as a ratio of *Ehd1* or *FT* copy number to 18S ribosomal RNA copy number (TaqMan Universal PCR Master Mix, Applied Biosystems) to obtain the relative transcript number. Mean values are based on three technical replicates and three biological replicates for both reference and target genes (±SD).

Relative quantification of the splice variant found in intron 7 (Fig. S3*F*, *Lower*) was obtained using fragment analysis (3). Fluorescent primers (5-FAM) flanking the 6-bp pair insertion were designed (Table S5). The number of cycles used to amplify each product was based on previous non–splice-specific *SbPRR37* qRT-PCR data to target amplification to the exponential phase. The resulting PCR-amplified fragments were resuspended in a solution of Hi-Di Formamide and GeneScan 400HD ROX Size Standard and subjected to analysis on the Applied Biosystems 3130xl Genetic Analyzer. These fragments were also sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) to ensure pure and correct product. The resulting data were then analyzed using Peak Scanner Software v1.0 (Applied Biosystems) to obtain the area under the curve for the correctly sized fragment. The area was averaged over three biological replicates ± SEM. The expression of the splice variant at intron 8 (Fig. S3*F*, *Upper*) was analyzed by qRT-PCR as described above.

SI Discussion

Additional Mapping Population Details. A previous study localized Ma_1 to the long arm of chromosome SBI-06 (4) (Fig. 1*D*). Two mapping populations were created to refine the Ma_1 locus, and a third population was subsequently created for high-resolution mapping of Ma_1. The sorghum inbred ATx623, which is photoperiod-insensitive line (homozygous ma_1), was crossed to R.07007, which is a photoperiod-insensitive line (homozygous Ma_1). F_1 plants were subsequently used as the pollinator in a backcross to ATx623 to eliminate the effect of a recessive ma_5 allele in the R.07007 background. The resulting population of 1,821 plants was screened with a series of PCR-based markers, and variation in flowering time due to the Ma_1 locus was linked to marker *Xtxp434* on SBI06 (Fig. 1*D*). Additional markers in this region were used to further define the location of the Ma_1 locus. Twenty-two recombination events were detected between *Xtxp598* and *Xtxp434*. An additional three recombinants were found between *Xtxi20* and *Xtxp58*, delimiting the downstream border of the locus. In the upstream region, one recombination event was detected at *Xtxi48*, the breakpoint of which was found in the promoter of putative gene Sb06g014508, defining the Ma_1 interval to ∼700-kb between *Xtxsn1* and *Xtxi20* (Fig. 1*D*). This genomic region is one of low gene density, encoding 34 putative genes (Phytozome v5.0) (Table S1). One gene, *pseudoresponse regulator protein 37* (*PRR37*) (Sb06g014570), was identified as a likely gene candidate for Ma_1 on the basis of the known roles of *PRR* genes in flowering in *Arabidopsis*. Despite the relatively large number of offspring screened, further refinement of the locus was not possible in this population due to the lack of recombination within the genetic region. Therefore, two additional populations were created to map Ma_1. Milo maturity genotype 100M, which is photoperiod sensitive (homozygous Ma_1), was crossed to elite inbred BTx406 or to founder genotype Blackhull Kafir. Both BTx406 and Blackhull Kafir are photoperiod insensitive, but the ma_1 loci in these two genotypes are derived from unrelated ancestral genotypes from different regions of Africa (4) (Table S3). Initially, a small F_2 population (100M × BTx406) of 122 plants was screened with PCR-based markers, and the locus was delimited by markers *Xtxp696* and *Xtxp711* (Fig. 1*E*). An additional 255 F_2 plants from this population were genotyped and phenotyped, and a total of 16 plants were identified with crossovers in the interval from which F_3 progeny could be derived. This effort resulted in a set of recombinant 100M × BTx406 F_3 progeny that were used to further refine the Ma_1 locus. These 100M × BTx406 F_3 progeny were used in conjunction with a 100M × Blackhull Kafir F_2 population (n = 1,925) to define the position of the Ma_1 locus on the high-resolution map (Fig. 1*F*). Using a series of SSRs and INDELs, flanking markers *Xtxi58* and *Xtxi62* narrowed the genetic location to an approximate 86-kb interval. Within this 86-kb genomic region, a single annotated gene, *PRR37* (Sb06g014570), is present along with extensive stretches of repetitive elements. No additional annotated genes are present within this interval.

1. Temnykh S, et al. (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Res* 11:1441–1452.
2. Bookout AL, Mangelsdorf DJ (2003) Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. *Nucl Recept Signal* 1:e012–e018.
3. Whitney IE, et al. (2011) Genetic modulation of horizontal cell number in the mouse retina. *Proc Natl Acad Sci USA* 108:9697–9702.
4. Klein RR, et al. (2008) The effect of tropical sorghum conversion and inbred development on genome diversity as revealed by high-resolution genotyping. *Plant Genome* 48:S12–S26.

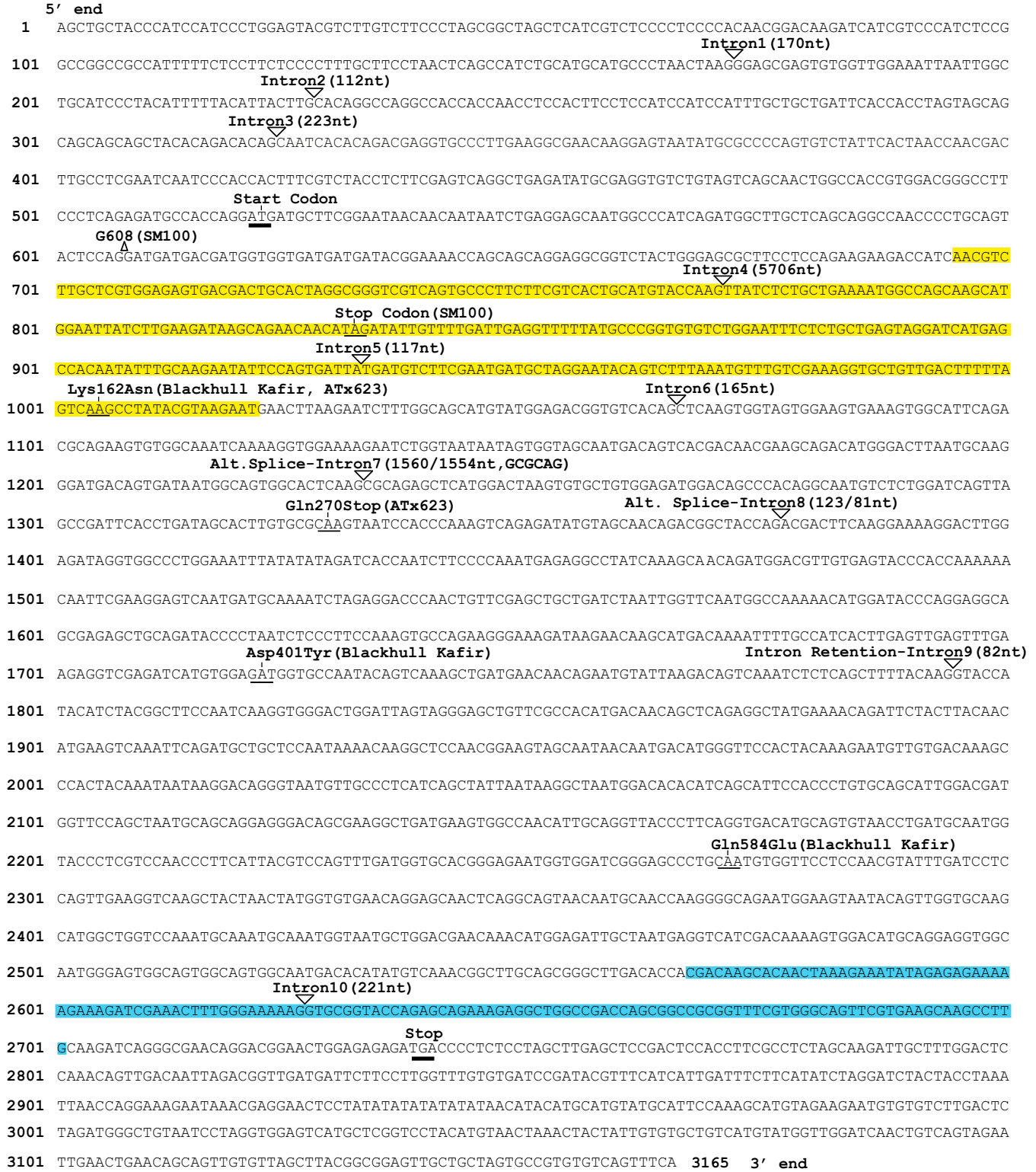


Fig. S1. Annotation of the full-length *SbPRR37* cDNA showing exon–intron splice junctions and key amino acid residues. The position of the amino acid substitution is shown above the respective codon. Downward-facing triangles indicate insertion positions of nucleotides or introns, and upward-facing triangles indicate nucleotide deletions. Nucleotides encoding pseudoreceiver domain and CCT motif residues are highlighted in yellow and blue, respectively. Alt. Splice, alternatively spliced. This annotation is based on the sequence of 100M transcript, splice variant a (GenBank accession no. JF801188).

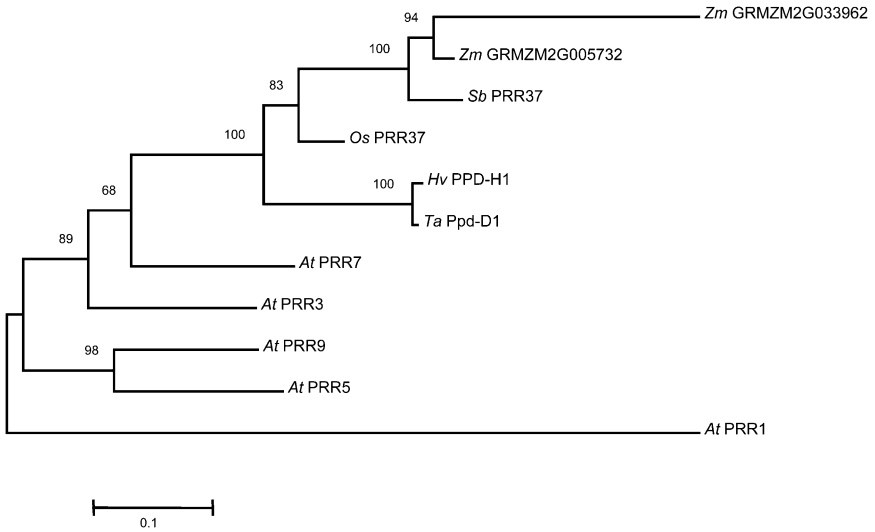


Fig. S2. Similarity of sorghum PRR37 protein to PRR proteins from maize, rice, barley, wheat, and *Arabidopsis*. The neighbor joining tree (1) is based on concatenated pseudoreceiver domain/CCT domain peptide sequences (2). Maize PRR protein sequences GRMZM2G033962 and GRMZM2G005732 were downloaded from Gramene (http://www.maizesequence.org/Zea_mays/Transcript/). Barley Ppd-H1 (AAY42109.1), wheat Ppd-D1 (ABL09464.1), rice PRR37 (NP_001060743), sorghum PRR37 (Sb06g014570.1) and *Arabidopsis* PRR1 (BAA94547.1), PRR3 (BAB13744.1), PRR5 (BAB13743.1), PRR7 (BAB13742.1), and PRR9 (BAB13741.1) protein sequences were downloaded from the National Center for Biotechnology Information. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branches (3). Evolutionary analyses were conducted in MEGA5 (4).

1. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425.
2. Turner A, Beales J, Faure S, Dunford RP, Laurie DA (2005) The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310:1031–1034.
3. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
4. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599.

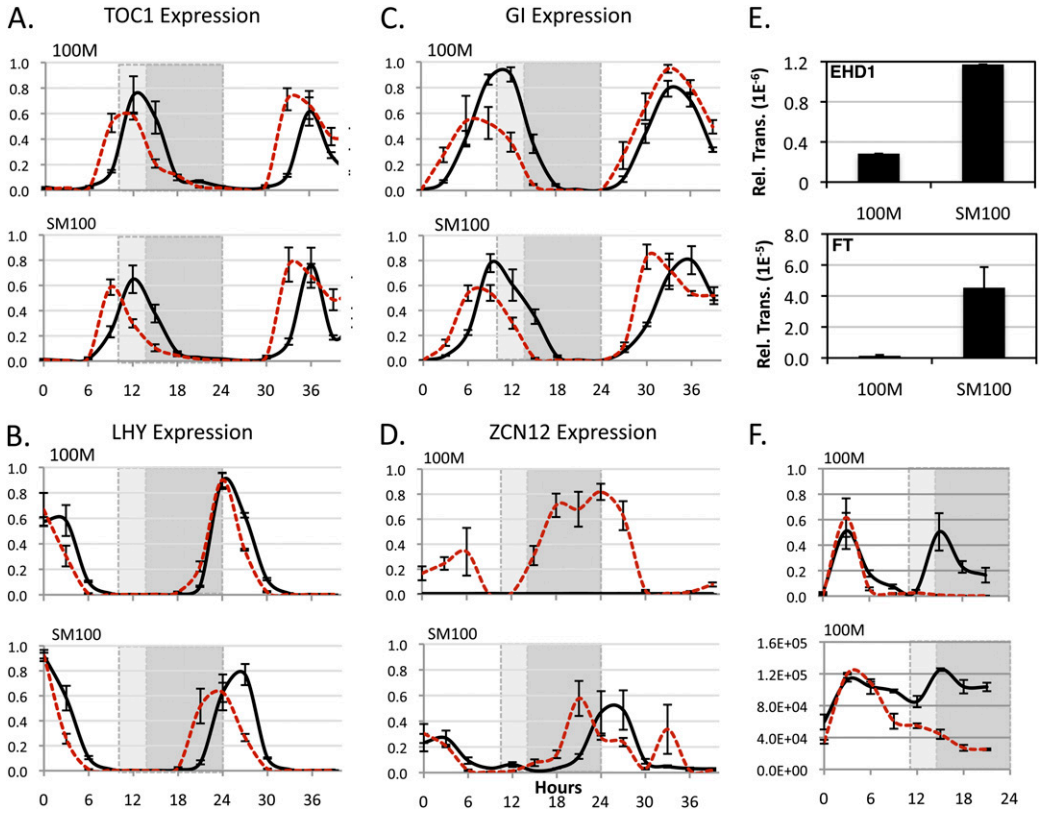
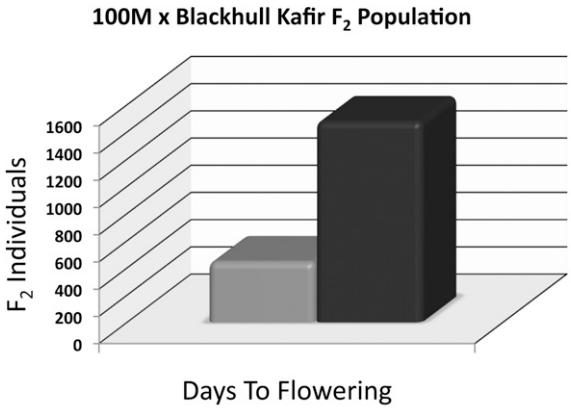


Fig. S3. Expression of flowering time genes in 100M (*SbPRR37*) and SM100 (*Sbprr37-1*). No expression differences were observed in (*A*) the core clock gene *TOC1* between 100M and *Sbprr37-1* mutant SM100 in either LD (solid black line) or SD (red dashed line). (*B*) *LHY* also is expressed in a similar manner in 100M and SM100. (*C*) Flowering-gene *GI* follows the expression pattern of *LHY* in both 100M and SM100, indicating no differential expression of *GI* between these genotypes. (*D*) *ZCN12*, a second candidate for florigen in maize, is activated in response to the SD photoperiod while remaining repressed to near-undetectable levels in LD-treated 100M plants. SD- and LD-treated SM100 shows no real differences in expression levels. The ordinate represents normalized expression relative to a calibrator sample (1) and is based on three biological replicates ± SEM. The light-gray shading within the plot area indicates darkness for SD-treated plants only, and dark-gray shading indicates darkness for both LD- and SD-treated plants. To directly compare levels of *Ehd1* and *FT* (*Hd3a*) between 100M and SM100, relative transcript abundance was calculated at 15 h into the LD, 1 h after the beginning of the dark period (*E*). This time point was selected because it corresponds to the evening peak of *SbPRR37* expression in LD (Fig. 2 *A* and *B*, arrowhead). The analysis showed that *Ehd1* (*Upper*) mRNA levels were significantly lower in the leaves of LD-treated 100M compared with SM100 (P value < 0.001). *FT* (*Lower*) expression levels were similarly decreased in LD-treated 100M plants compared with SM100 (P value < 0.01). When 100M plants were transferred to SDs, expression of *Ehd1* and *FT* increased by 17- and 7.06-fold, respectively. By contrast, upon SD transfer, the increases in *Ehd1* and *FT* levels in SM100 were only 2.35 and 0.34, respectively, consistent with the de-repression of these floral activators in the *prr37-1* background. The ordinate represents the absolute transcript abundance normalized relative to 18S ± SD (1). Statistical significance was calculated using a two-tailed Student's *t* test; actual P values were 0.005 and 0.0002 for *Ehd1* and *FT*, respectively. (*F*) Expression analysis of the alternatively spliced cDNA in intron 8 (*Upper*) and intron 7 (*Lower*) (Fig. S1 and Table S2). The expression patterns of these cDNA variants parallel that seen in overall transcript abundance in long and short days. Additionally, these splice sites are found in equal proportion in *PRR37* and *prr37* genotypes, suggesting that these differences do not necessarily contribute to overall phenotypic differences. Expression of the variant observed in intron 7 was extrapolated from the area under curve as obtained from the ABI 3130xl instrument and Peak Scanner software (2). The ordinate represents the average area under the curve ± SEM and is based on three biological replicates. For the variant observed in intron 8, expression was detected using qRT-PCR as described in *SI Methods*. The ordinate represents normalized expression relative to a calibrator sample (1) and is based on three biological replicates ± SEM.

1. Bookout AL, Mangelsdorf DJ (2003) Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. *Nucl Recept Signal* 1:e012–e018.
2. Whitney IE, et al. (2011) Genetic modulation of horizontal cell number in the mouse retina. *Proc Natl Acad Sci USA* 108:9697–9702.



Flowering Data			Population Statistics		
Classification	Range	Average	Observed	X^2	P-value
Early	72-78	75	477	1.11	0.28
Late	88-95	90	1478		

Fig. S4. Flowering-date observations and χ^2 values for the 100M × Blackhull Kafir F_2 sorghum population. Expected 3:1 late:early ratio of flowering phenotypes leads to the conclusion that Ma_1 is the only major flowering-time gene segregating in this fine-mapping population.

Table S1. Genes present in the ~700-kb interval mapped in the ATx623 by R.07007 BC_1F_1 population

Gene no.	Start (bp)	End (bp)	Annotation (Phytozome v5.0)
Sb06g014410	39,903,279	39,906,744	Similar to HCF106 family protein
Sb06g014420	39,918,495	39,921,352	Leucine-rich repeat protein/F-box
Sb06g014430	39,948,142	39,953,971	No functional annotations
Sb06g014440	39,962,396	39,969,146	Transport protein particle component
Sb06g014450	39,970,615	39,972,162	Berberine and berberine like, D-lactate dehydrogenase, FAD binding, oxidoreductase
Sb06g014460	39,994,622	39,997,657	Divergent CCT motif
Sb06g014470	40,037,697	40,043,627	Ribosomal protein L4/L1 family
Sb06g014480	40,047,428	40,054,072	Poly(A) polymerase head domain, tRNA nucleotidyltransferase activity
Sb06g014490	40,079,857	40,082,075	No functional annotations
Sb06g014500	40,082,952	40,085,195	Putative helicase related
Sb06g014504	40,102,573	40,105,282	Ulp1 protease family, C-terminal catalytic domain, Sentrin/SUMO-specific protease
Sb06g014508 (***Xtxsn1***)	40,133,864	40,693,944	General transcription factor 2-related zinc finger
Sb06g014510	40,175,447	40,175,656	No functional annotations
Sb06g014520	40,193,736	40,195,230	No functional annotations
Sb06g014530	40,210,136	40,211,600	Ribosomal protein L11; RNA-binding domain
Sb06g014550	40,216,040	40,217,587	Iron/ascorbate family oxidoreductase
Sb06g014560	40,243,699	40,244,912	Sulfotransferase
Xtxi63	40,279,273		
Sb06g014570 (***txi49***)	40,280,414	40,290,602	PRR37-like
Xtxi58	40,364,576		
Sb06g014580	40,400,791	40,403,993	Serine protease family; S10 serine carboxypeptidase
Sb06g014590	40,406,258	40,401,377	ATP-dependent CLP protease
Sb06g014630	40,434,828	40,436,984	No functional annotations
Sb06g014640	40,437,379	40,440,378	No functional annotations
Sb06g014650	40,460,341	40,462,677	No functional annotations
Sb06g014670	40,493,258	40,498,431	RNA binding; similar to Mei2
Sb06g014676	40,594,546	40,608,430	No functional annotations
Xtxi20	40,595,306		
Sb06g014680	40,688,839	40,689,723	No functional annotations
Sb06g014710	40,761,789	40,771,377	Glycyl-tRNA synthetase and related class II tRNA synthetase
Sb06g014720	40,771,644	40,776,006	Ubiquinol-cytochrome *c* reductase iron-sulfur subunit-related
Sb06g014730	40,781,189	40,783,917	Extosin family
Sb06g014740	40,784,078	40,784,927	Pollen proteins Ole e I family
Sb06g014743	40,789,072	40,795,991	No annotated domains
Sb06g014746	40,799,353	40,800,508	No functional annotations
Sb06g014750	40,827,904	40,836,238	No functional annotations
Sb06g014760	40,827,904	40,836,238	No functional annotations
Xtxp598	40,828,046		

Markers and genes used as markers are highlighted in bold type. The physical coordinates and functional annotation of each gene are given as based on Phytozome v5.0.

Table S2. Sequence variation within the *SbPRR37* coding region for the genotypes listed in Table S3

Allele*	Coding region haplotype	Nucleotide position†	Protein modification‡	PRR domain§	CCT motif§
SbPRR37¶	Alternative 3′ splice junction, intron 7	1,231	QA insert	No	No
SbPRR37¶	Alternative 5′ splice junction, intron 8	1,376	GTSNRNCMKQKYTN insert	No	No
Sbprr37-1	G deletion	608	***Frameshift, premature termination***	**Yes**	**Yes**
Sbprr37-2	G > T	1,006	***K162N***	**Yes**	No
	G > T	1,721	D401Y	No	No
	C > A	2,270	Q584E	No	No
Sbprr37-3	G > T	1,006	***K162N***	**Yes**	No
	C > T	1,329	***Q270Stop***	No	**Yes**
	G > T	1,721	NA‖	—	—
	C > G	2,270	NA‖	—	—

*Total of six founder sorghum genotypes were examined for Ma_1 coding region haplotypes. Genotypes examined are listed in Table S3.
†Position in Ma_1 mRNA based on full-length sequence of sorghum genotype 100M as shown in Fig. S1.
‡Protein modifications in boldface italic type denote mutations that alter the pseudoreceiver domain and/or CCT motif.
§Protein modification that truncates, eliminates, or alters key amino acid residues in the pseudoreceiver domain or CCT motif are denoted as **Yes** (in boldface type).
¶Alternative splicing variants were observed in wild-type and mutant alleles.
‖NA, not applicable. Mutation occurs after premature stop codon in ma_1 allele.

Table S3. Pedigree and *Maturity Locus 1* classification for sorghum genotypes used in linkage analyses and gene expression studies

Sorghum genotype	Classification at *Maturity Locus 1*	Pedigree, year/decade of germplasm development
100M Milo	*Ma_1, SbPRR37*	Derivative of Early White Milo (*ca*. 1911) × Double Dwarf Yellow Milo (*ca*. 1936). Developed by J. R. Quinby as maturity loci genetic stock for flowering-time studies
SM100 Milo	*ma_1, Sbprr37-1*	Derivative of Early White Milo (*ca*. 1911) × Double Dwarf Yellow Milo (*ca*. 1936). Developed by Quinby as maturity loci genetic stock for flowering-time studies
Blackhull Kafir	*ma_1, Sbprr37-2*	Pure line selection from Standard Kafir (*ca*. 1919)
ATx623	*ma_1, Sbprr37-3*	Selected from the cross of elite lines BTx3197 × SC170-6–4 (Kafir × Zera-zera, *ca*. 1964)
R.07007	*Ma_1, SbPRR37*	Derivative of EBA-3 (dual purpose sorghum from Argentina) (Texas AgriLife release 2007)
BTx406	*ma_1, Sbprr37-1*	Martin derivative. Martin originated from undetermined outcross of Wheatland (Kafir × Milo)

Pedigree and classification at *Maturity Locus 1* are based on germplasm release notices in historical records maintained at Texas A&M University.

Table S4. Markers used for fine mapping the *Ma_1* locus

Marker	Forward primer	Reverse primer	Marker type	SBI-06 marker location (bp)
Xtxp694	GGGCCCTGTTACATCCTTAAT	TGCCTGATCTTAGAGAAACACC	SSR	38,368,626
Xtxi48	TCAAGGCAAGATTGACGAAGCCAC	GGCTTGTAGCAGTAGCACTTGTGT	INDEL	39,098,709
Xtxsn1	ACCTTTCAGTGTGGTGCAAATGGG	TTGGCCCGATAGCAGTCCGATAAA	SNP	39,901,458
Xtxp696	TCGATCGATTCTCCTGCTTT	GTAGGTGCACCCAGTGCTTC	SSR	40,200,439
Xtxi61	GCCTTTGCAAGCAAAAATCT	TCTCGAGCCTAATCCCAATC	INDEL	40,278,618
Xtxi62	CGGTGCGTAGCAAATGTAAA	GGTCCAATGCAGAAGACGAT	INDEL	40,278,905
Xtxi63	CTCCTTTTGCTCCACGTCAT	GCATGCAGATGGCTGAGTTA	INDEL	40,279,273
Xtxi49	CGAGCCAATTTCACCTCCTA	GCCAATGCATGTTTCATAGC	INDEL	40,279,611
Xtxi55	TTTATGCCCGGTGTGTCTG	CATCACTGCACATGAACCAC	INDEL	40,286,452
Xtxi73	AATTTTCTATGCAATTAAGAAGAG	GCCTTAAGAGCCGGGAATA	INDEL	40,313,421
Xtxi58	GGAGCTATTGCTATGCTGCTT	CTCAGAGTACAGCAGCTCCAAC	INDEL	40,364,576
Xtxi68	TAGAGCCTTTGTGCAGCATTC	CAACCAATTGCCCTTGTTTAC	INDEL	40,408,262
Xtxp711	CACCTAGCAGAGGGCAAGAG	CACACTCATTGCTTGCCTGT	SSR	40,440,214
Xtxi66	GGGAGCGTTGAAACTTGATG	GCAAGCACAGACGAACTCAC	INDEL	40,491,401
Xtxi20	GCCTCCACTTGCGAATGAT	ATACATAACTTGTGGGTCGAAAG	INDEL	40,595,306
Xtxp599	TGAAAACGAACCAAACACACTC	TTTAAATACTTCCTCCATTCCAAA	SSR	40,801,774
Xtxp598	GTGGCGCACAGCTAAAAGT	TTTGGTCCGATCTTTTGGAG	SSR	40,828,046
Xtxp434	CGAGGTCCAGGAGTACACG	CGGCCTCCAGGAGGAGTAAT	SSR	42,610,344

Sequences of forward and reverse primers are given, as well as polymorphism classification. Physical coordinates are given as obtained from Phytozome v7.0.

Table S5. Primer sequences used for qRT-PCR analysis

Gene	Locus ID	Orthologous locus in rice or maize	Forward sequence	Reverse sequence	ATx623 (%)	R.07007 (%)	100M (%)	SM100 (%)
PRR37	Sb06g014570	LOC_Os07g49460	AACAGGACGGAACTGGAGAGAGAT	CCAAAGCAATCTTGCTAGAGGCGA	96	98	91	85
FT	Sb10g003940	LOC_Os06g06320	AGCATTTGGGCAAGAGGTGATCTG	AAGTCCCTGGTGTTGAAGTTCTGG	90	96	93	86
CO	Sb10g010050	LOC_Os06g16370	TAGTCCCAGACAACATGGCAACGA	AGGTCAAGTGGAGTGGCATCTGAA	95	87	92	89
Ehd1	Sb01g019980	LOC_Os10g32600	CGTCAGGGAAGCAATGTCCTTCAT	CTTCAGTTGGAAAGCACACATCGC	93	94	92	92
TOC1	Sb04g026190	LOC_Os02g40510	GAGTGCAGATGATTACTGCTCACTTTG	TGCTGCCTTGTTGCCAGTAGAAGA	91	95	88	84
LHY	Sb07g003870	LOC_Os08g06110	GGCCTGCCTCTACCATGAAGTTTA	GCACTGCATTGCAAGGTTTGAAGTCC	91	97	89	86
GI	Sb03g003650	LOC_Os01g08700	ATGCACCCGCTTCCTAGTCATCTT	TTCAGGGCTGTCATGGTTCCTCAT	92	98	92	94
ZCN8	Sb09g025760	GRMZM2G179264_T01	AACTGTCAAAGGGAAGGTGGATCG	GACTAAGCTCTCAACCCTTCAAGTC	88	90	89	90
ZCN12	Sb03g034580	GRMZM2G103666_T01	TGCATGCATGAATATCGTCGTCT	CCCGGGTAGTACATATAAGGTGGT	104	106	107	95
PRR37 splice intron 7	Sb06g014570	LOC_Os07g49460	TGACAGTCACGACAACGAAGCAGA	TCGGCTAACTGATCCAGAGACATTGC	—	—	—	—
PRR37 splice intron 8	Sb06g014570	LOC_Os07g49460	ACCAGGTACGAGCAACAGAAACTG	TTGGTGGGTACTCACAACGTCCAT	—	—	89	92

Gene locus IDs are given based on Phytozome v7.0. Forward and reverse primer sequences and amplification efficiencies for each respective genotype are given. Primer efficiencies were calculated using the standard curve method as described in *SI Methods*.