PNAS
www.pnas.org

Supplementary Information for

Cuscuta australis (dodder) parasite eavesdrops on the host plants' FT signals to flower

Guojing Shen[#1], Nian Liu[#1,2] Jingxiong Zhang[1,2], Yuxing Xu[1], Ian T. Baldwin[3], and Jianqiang Wu[1,2]*

[1]Department of Economic Plants and Biotechnology, Yunnan Key Laboratory for Wild Plant Resources, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

[2]CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing 100049, China

[3]Department of Molecular Ecology, Max Planck Institute for Chemical Ecology, Jena 07745, Germany

[#]These authors contributed equally.

[*]Correspondence to: wujianqiang@mail.kib.ac.cn.

This PDF file includes:

Supplementary text

SI Appendix, Figures S1 to S10

SI Appendix, Tables S1 to S3

Legends for *SI Appendix*, Datasets S1 to S5

SI References

www.pnas.org/cgi/doi/10.1073/pnas.2009445117

Materials and Methods

Plasmid construction and genetic transformation. For cloning the full-length genomic sequence of *C. australis FT* (named *CaFT-Genomic*) (11402 bp), the full-length genomic sequence was split into two segments (at the *StuI* enzyme digestion site located at the 5044-bp position of *CaFT-Genomic* sequence). The first 5044-bp segment was PCR amplified from *C. australis* genomic DNA using the forward primer CaFT1F with a KpnI site and the reverse primer CaFT1R with a StuI site. Another 6358-bp segment was amplified by PCR using the forward primer CaFT2F with a StuI site and the reverse primer CaFT2R with a BstEII site. Both PCR products were individually inserted into pJET1.2 (ThermoScientific) for sequencing. The plasmids were digested with KpnI and StuI and with StuI and BstEII, respectively. A three-way ligation was performed to insert the two segments into the binary vector pCAMBIA1301 vector, which was predigested with KpnI and BstEII, forming pCAMBIA1301-CaFT-Genomic. The deduced cDNA sequence of *C. australis CaFT* (named *CaFT-CDS*), which carried a NcoI site at the 5' and a FLAG sequence plus the BstEII site at the 3', was synthesized (the Beijing Genomics Institute, www.genomics.org.cn) and cloned into pJET1.2, forming pJET1.2-CaFT-CDS; pJET1.2-CaFT-CDS was digested with NcoI and BstEII, and the purified DNA fragment was cloned into the binary vector pCAMBIA3301, forming pCAMBIA-CaFT-CDS.

To control the expression of *AtFT* by an inducible promoter, the coding sequence of *AtFT* was amplified from Arabidopsis cDNA using the primers AtFT-XhoI-F and AtFT-linker-HindIII-R introducing a Gly-Gly-Gly linker into the 3' region of *AtFT*. *GFP* (*green fluorescence protein*) was amplified using the primers GFP-HindIII-F and GFP-SpeI-R. Using a three-way ligation strategy, the AtFT and GFP fragments were inserted into the pER8 vector (1) containing an estradiol inducible promoter, to form pER8-AtFT-GFP. This construct was used for stable transformation of tobacco cv Maryland Mammoth.

To knock out *NtFT5* using the CRISPR/Cas9 technology, we designed *NtFT5*-targeting sgRNA using the CRISPR-P (http://cbi.hzau.edu.cn/crispr/) (2), which displayed all optional sgRNA sequences (20 bp) immediately followed by 50-NGG (PAM, protospacer adjacent motif) in the forward or reverse strand. In this study, we selected one sgRNA

targeting *NtFT5*, named as *NtFT5*-cas9. The DNA oligonucleotides of sgRNA (Sangon Biotech; www.sangon.com) were annealed to generate double-stranded DNA before being inserted into the CRISPR/Cas9 expression vector pONRE-Cas9 (3), which was predigested with BsaI, forming pONRE-Cas9-NtFT5. This construct was introduced into *Agrobacterium tumefaciens* LBA4404 for stable transformation of cultivated tobacco cv K326.

For yeast two-hybrid (Y2H) assays, the PCR-amplified open reading frame (ORF) of *CaFD* and *AtFD* were cloned into the pGADT7 (Clontech) vector, forming pGADT7-CaFD, and the ORF of *AtFT* was cloned into the vector pGBKT7 (Clontech) to form pGBKT7-AtFT. The ORF of *CaFT* was PCR amplified (pJET1.2-CaFT-CDS served as the PCR template) and cloned in the pGBKT7 (Clontech) to form pGBKT7-CaFT. The vectors of pGBKT7-GmFT2a and pGBKT7-GmFT5a were generous gifts from Dr. Fanjiang Kong (Guangzhou University).

To construct the plasmids for bimolecular fluorescence complementarity (BIFC) assay, the ORF of *CaFD* and *AtFD* were PCR amplified and inserted into the pUC-SPYCE (for split YFP C-terminal fragment expression) (4) to form pUC-CaFD-CYFP, and the PCR amplified ORF of *AtFT* and *CaFT* ((pJET1.2-CaFT-CDS serves as a template for PCR) was inserted into the pUC-SPYNE (for split YFP N-terminal fragment expression) (4) to form pUC-AtFT-NYFP. The vectors pUC-GmFT2a-NYFP and pUC-GmFT5a-NYFP were gifts from Dr. Fanjiang Kong (Guangzhou University).

All the primer sequences used for cloning are listed in *SI Appendix*, Table S2.

Agrobacterium tumefaciens GV3101 and LBA4404 were used for Arabidopsis and tobacco transformation (5, 6), respectively, and homozygous T2 plants were used for all experiments. The seeds of transgenic soybean (background Williams 82) overexpressing *GmFT2a* were kindly provided by Dr. Fanjiang Kong (Guangzhou University).

Flowering time analysis. The flowering times were measured as the days from sowing to the emergence of the first bud in the plants, except that the flowering times of Arabidopsis were quantified with the numbers of rosette and cauline leaves on the main stem when the inflorescence reached 1 – 5 cm as described in (7).

RNA extraction and cDNA synthesis. Samples were ground in liquid nitrogen. Each sample was initially treated with Fruit-Mate for RNA purification (Takara) and the total RNA was isolated using TRIzol (ThermoScientific) following the manufacturers' instructions. Quantification of total RNA was done on a Nanodrop (ThermoScientific) and RNA integrity was examined by separating RNA samples on agarose gels. After quality check, 1.0 μg of total RNA was reverse transcribed using oligo(dT), dNTP, and Superscript II reverse transcriptase (ThermoScientific).

Quantitative real time-PCR and reverse transcription-PCR. Quantitative real time-PCR (qPCR) was performed on a CFX Connect Real-Time System (Bio-Rad) using the iTaqTM Universal SYBR Green Supermix kits (Bio-Rad). For each analysis, a linear standard curve, threshold cycle numbers versus $\log_{10}$ (designated transcript level), was constructed using a dilution series of a specific cDNA standard, and the levels of the transcripts in all unknown samples were calculated according to the standard curves. The 2 × TSINGKE Master Mix Taq polymerase (TSINGKE Biological Technology, www.tsingke.net) was used for reverse transcription PCR (RT-PCR) analysis, and after 35 cycles of PCR (15-μl reactions), 10 μl of each product was run on an agarose gel. All primers, including those for genes as internal controls, are listed in *SI Appendix*, Table S3.

RNA-seq and transcriptome analysis. For RNA-seq analysis, complete *C. australis* stems, including the stem apices, were excised 2 cm away from the haustorial junctions, and three biological replicates were used for each group of samples. Illumina TruSeq RNA Sample Prep Kit (Illumina) was used to construct cDNA libraries. The generated cDNA libraries were sequenced on a HiSeq2500-PE125 platform (Illumina) to acquire sequence reads (5 G depth). Trimmomatic (v0.39) (8) with default parameters was used to remove the adaptor sequences from the raw reads and low quality raw reads which contained more than 10% unknown bases (N). Clean reads were aligned to the reference *C. australis* genome using Hisat2 (v2.1.0) (9), and featureCounts from Subread package (v1.6) (10) was used to count reads for all genes in the reference genome. The Bioconductor DEseq2 (v3.9) package (10) was employed to infer differential gene expression according the gene count matrices. Transcripts with adjusted p value < 0.05 and absolute value of log2 (fold change) > 1 were selected as differentially expressed

genes (DEGs). The *C. australis* genome sequences (11) was used to conduct Gene Ontology and over-representation analysis and gene annotation to reveal the functional classification, significantly changed pathways, and gene functions in the DEGs between samples. For analyzing the inter-plant mobile mRNAs, only the mRNAs that matched the soybean genome but not the *C. australis* genome were defined as the mobile soybean mRNAs. All the transcriptome data can be accessed at the National Genomics Data Center (NGDC) (https://bigd.big.ac.cn/) under the BioProject ID PRJCA001885.

Yeast 2-hybrid assays. Yeast 2-hybrid (Y2H) assay was performed following the manufacturer's instructions (Clontech). In brief, the bait and prey constructs were co-transformed into the yeast strain AH109 following a PEG (poly (ethylene glycol))/lithium acetate method and the transformed yeast was selected on the SD/-Trp/-Leu medium. The positive clones were further selected on the SD/-Trp/-Leu/-His and subsequently on SD/-Trp/-Leu/-His/-Ade medium containing 200 ng/ml Aureobasidin A and 20 mg/ml X-α-Gal.

Bimolecular fluorescence complementation assay. Twenty-eight-day-old Arabidopsis leaves were harvested for protoplast isolation. Protoplast isolation and plasmid transfection were performed following a protocol established for Arabidopsis (12). The pUC-CaFD-YCE plasmids were co-transfected with pUC-AtFT-YNE, pUC-GmFT2a-YNE, or pUC-GmFT5a-YNE into *Arabidopsis* protoplasts. And the pUC-CaFT-YNE plasmids were co-transfected with pUC-CaFD-YCE, or pUC-AtFD-YCE into *Arabidopsis* protoplasts. About 18 – 20 h after transfection, protoplasts were imaged under a confocal laser scanning microscope (Olympus FV1000).

Estradiol treatment for induction of *AtFT-GFP* expression in Maryland Mammoth tobacco. Estradiol (Sigma) was dissolved in DMSO (100 mM) as a stock solution, and the working solution was prepared by 1000-fold dilution of the stock solution in water (100 μM). To verify the induction of *AtFT-GFP* mRNA and protein expression, an experiment was conducted with 30-day-old transgenic tobacco (cv Maryland Mammoth; harboring the T-DNA from pER8-AtFT-GFP) plants, which had been parasitized by *C. australis* and grown under LD conditions. For each plant, using 1-ml syringes, the same region of the oldest leaf was pressure infiltrated with the estradiol working solution or the

solvent on two consecutive days (once a day; about 150 μl for each infiltration), and these regions and the parasitizing *C. australis* stems were harvested 24 h after the second infiltration. These samples were used for total RNA and protein extraction and further analyses.

To record the induction of flowering in the transgenic tobacco and the parasitizing *C. australis*, 30-day-old transgenic tobacco (cv Maryland Mammoth; harboring the T-DNA from pER8-AtFT-GFP) plants, which had been parasitized by *C. australis* and grown under LD conditions, were used. Each plant was treated with a series of syringe infiltrations with the estradiol working solution (treatment) or solvent (control) as follows: the oldest rosette leaf was infiltrated with approximately 150 μl of working solution on day 1, and again every 2 days for a total of 8 infiltrations. The same leaf was repeatedly infiltrated but the laminal location moved to the opposite side of the midrib; after both sides of the leaf had been infiltrated (after two infiltrations), the next younger leaf was similarly treated. The flowering times of all plants were recorded.

Western blotting. Plant tissues were ground in liquid nitrogen and approximately 100 mg aliquots were extracted in 200 μl of ice-cold protein extraction buffer (100 mM Hepes, pH 7.5, 5 mM EGTA, 5 mM EDTA, 10 mM DTT, 1 mM PMSF, 10% glycerol, proteinase inhibitor cocktail (Sigma)) to extract the total protein. Protein concentrations were determined using a Quick Start Bradford Protein Assay kit (Bio-Rad). Fifteen micrograms of total proteins were loaded onto 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels. Proteins were transferred to a polyvinylidene fluoride (PVDF) membrane using the Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad) at 4 ^{o}C for 1 h under a 100-V constant voltage. The membrane was blocked with 5% not-fat milk for 2 h and was incubated with the rabbit polyclonal anti-GFP antibody (1:1000, Abcam) at 4°C overnight, and then followed by incubation with secondary HRP antibody which was raised against rabbit antibodies (ThermoFisher Scientific). The same procedure was used for the immuno-detection of actin, except that an anti-actin (1: 1000, Abmart) antibody was used as the primary antibody. The chemiluminescence gel imager system (MicroChemi, DNR) was used for imaging.

Proteomic analysis. Exactly the same *C. australis* samples that were used for RNA-seq analysis were used for proteomic analysis. Plant tissues (100 mg) were ground in liquid nitrogen and the total proteins were extracted with a lysis buffer (7 M urea, 2 M thiourea, 0.1% CHAPS, protease inhibitors). Twenty micrograms of each protein sample were loaded onto a 10% SDS-PAGE gel to separate proteins with different molecular weights (MWs) and each lane was divided into four fractions according to their MWs. Each fraction was cut into 1-mm^2 slices. The slices of each fraction were collected into one 1.5 ml Eppendorf tube and were digested with 0.01 μg/μl trypsin at 37 ^oC overnight. The digested peptides were centrifuged, and the supernatants were collected and subsequently dehydrated under vacuum. The peptides of each sample were dissolved in 20 μl of 2% methanol containing 0.1% formic acid, and 10 μl of the digested peptides were loaded onto a ThermoFisher Scientific EASY-nLC 1000 System (Nano HPLC) coupled with an Orbitrap Fusion Mass Spectrometer (ThermoFisher Scientific). The full-scan mass spectra were scanned at the range of 300-1400 m/z with the resolution of 120000 FWHM. The automatic gain control (AGC) and maximum injection time of full-scan were set to 5.0 e^5 and 50 ms, respectively. The dd-MS2 was obtained using a high energy collision induced dissociation (HCD) method, and the normalization collision energy (NCE) was set to 30%. The AGC and maximum injection time of HCD were set to 5.0 e^3 and 35 ms, respectively.

The mass spectrometry raw data were processed using the Proteome Discoverer (PD) 2.1 software (ThermoFisher Scientific). The parameters of PD 2.1 were set as following: the precursor ion mass tolerance was set to 15 ppm and the fragment ion mass tolerance was set to 0.5 Da. The cleaned MS data were searched against the *Cuscuta australis* database (http://www.dodderbase.org/) first and the remaining unmatched ones were searched against the UniProt soybean database (https://www.uniprot.org/). Only the proteins that matched the soybean genome but not the *C. australis* genome were defined as “mobile soybean” proteins.

Identification of flowering-related genes in *C. australis* and *C. campestris* genome.

In order to identify the flowering genes in *C. australis* and *C. campestris* genome, we combined the amino acid sequences of the protein-coding genes in *C. australis* and *C.*

campestris genome and those from the seven reference autotrophic plant genomes, which were used in Sun et al.(11) into a gene set. After clustering using OrthoFinder2 (v 2.3.3) (13), we obtained the orthogroups containing the Arabidopsis genes associated with the flowering, which were retrieved from the FLOR-ID database (14). For sequences in these orthogroups, we used Mafft (v 7.407) (15) for multiple sequence alignments and constructed phylogenetic gene trees using FastTree2 (v 2.1.10) (16). OrthoFinder2 (v 2.3.3) (17) was used to identify orthologous relationships among genes. Arabidopsis flowering genes that lacked an orthologous dodder gene were considered to have had been lost from the dodder genomes.

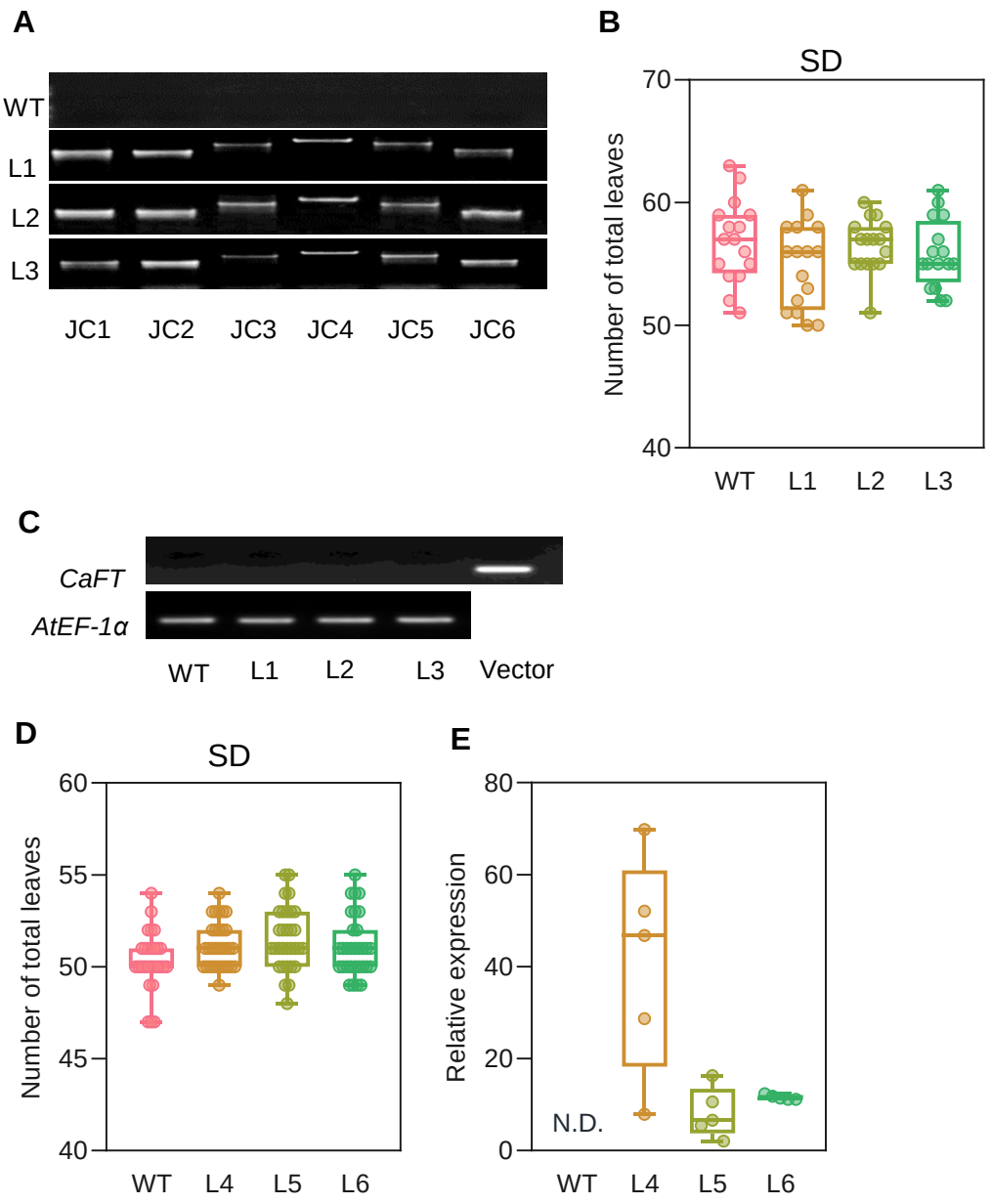

Fig. S1. Flowering times of Arabidopsis WT plants and transgenic plants expressing *C. australis* FT genomic and coding sequence. The sequences of *CaFT-Genomic* and *CaFT-CDS* driven by a 35S promoter were transformed into Arabidopsis, and three independently transformed lines for each construct, L1 to L3 for *CaFT-Genomic* and L4 to L6 for *CaFT-CDS*, were used for the experiments. **(A)** The positive identification of the *CaFT-Genomic* transgene in the transgenic plants. Six pairs of *CaFT-Genomic*-specific primers (from JC1 to JC6) were used to amplify partial but overlapping fragments of *CaFT-Genomic* DNA in WT and transgenic lines. **(B)** The flowering times

of the WT (Col-0) and *CaFT-Genomic* transgenic lines under SD, indicated by the numbers of rosette and cauline leaves on the main stem when inflorescences reached 1 to 5 cm. **(C)** RT-PCR to detect the expression of *CaFT* transcripts in the WT and *CaFT-Genomic* three transgenic lines. The "Vector" lane (positive control for *CaFT* amplification) indicates the PCR reaction using the plasmid pJET1.2-CaFT-CDS as the template and *CaFT*-specific primers. PCR products were separated on an agarose gel. **(D)** The flowering times of the WT (Col-0) and *CaFT-CDS* transgenic lines under LD, indicated by the numbers of rosette and cauline leaves on the main stem when inflorescences reached 1 to 5 cm. **(E)** qPCR to detect the expression of *CaFT* transcripts in the WT and *CaFT-CDS* three transgenic lines. N.D. = not detected. For **(B**), **(D)**, and **(E)**, the horizontal bars within boxes indicate medians. The tops and bottoms of the boxes indicate upper and lower quartiles, respectively. The upper and lower whiskers represent maximum and minimum, respectively. **(B)** and **(D):** No statistically significant differences were found between WT and any transgenic line (Student's t-test).

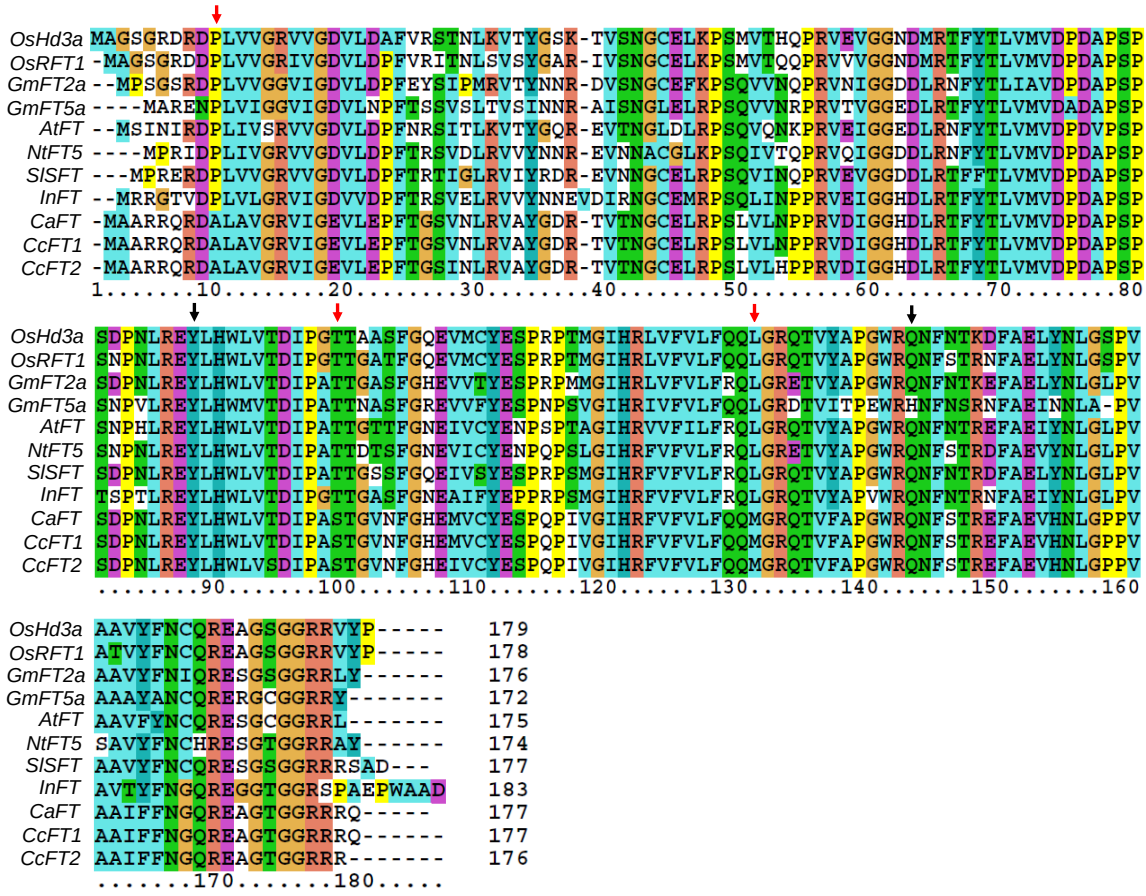

Fig. S2. The alignment of FT proteins from *C. australis* and *C. campestris* with FT/FT-like sequences from a few autotrophic plants. The FT/FT-like proteins are from rice *Oryza sativa* (*OsHd3a* (LOC_Os06g06320) and *OsRFT1* (LOC_Os06g06300)), soybean *Glycine max* (*GmFT2a* (Glyma.16G150700) and *GmFT5a* (Glyma.16G044100)), thale cress *Arabidopsis thaliana* (*AtFT* (AT1G65480)), tobacco *Nicotiana tabacum* (*NtFT5* (KY306471)), tomato *Solanum lycopersicum* (*SlSFT* (Solyc03g063100.2)), morning glory *Ipomoea nil* (*InFT* (INIL09g31483)), dodder *C. australis* (*CaFT* (RAL41934)), and dodder *C. campestris* **(***CcFT1* (Cc042091) and *CcFT2* (Cc008736)). The red arrows indicate the positions where the amino acids of CaFT, CcFT1, and CcFT2 are different from the FT proteins from all the autotrophic species, and the black arrows indicate the amino acids that had been shown to be essential for AtFT activity(18).

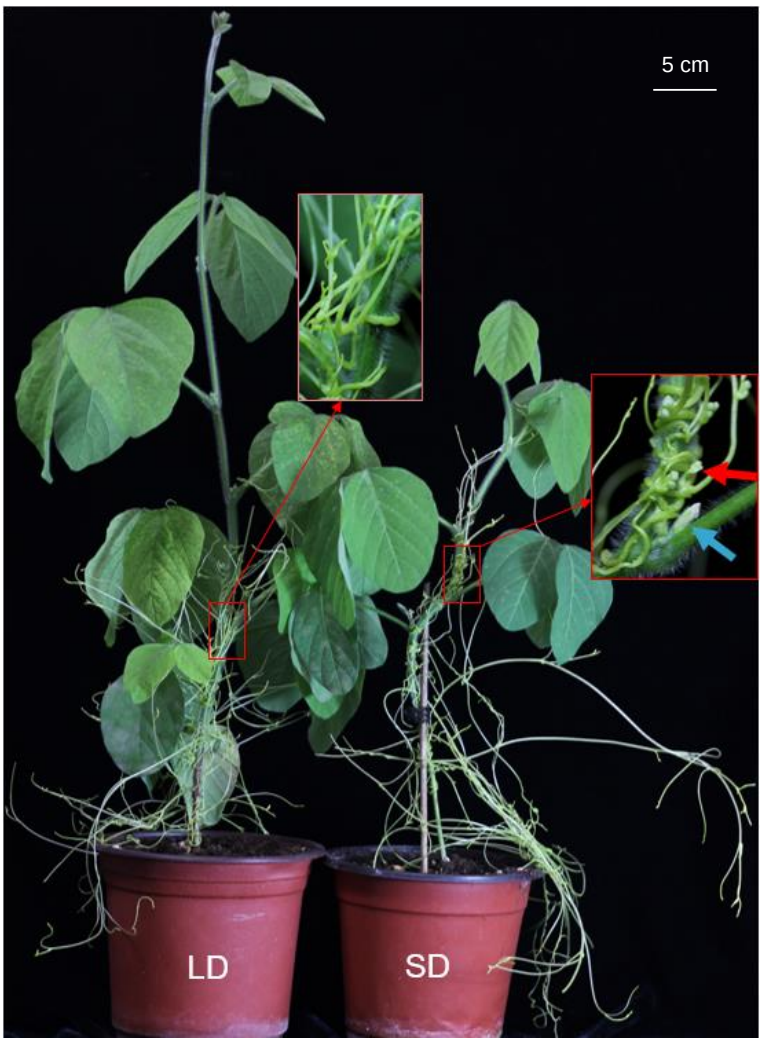

Fig. S3. The flowering phenotypes of *C. australis* and the host soybean. Two groups of soybean (Williams 82) and the parasitizing *C. australis* plants were cultivated under LD (left) and SD (right) conditions, respectively. The photograph was taken on day 35. At this time, under SD conditions (right), soybean and *C. australis* both flowered; inset: magnified view, the red and blue arrow indicate floral buds of *C. australis* and soybean, respectively. Under LD conditions (left), soybean and *C. australis* did not flower; inset: magnified view. Bar = 5 cm.

Fig. S4. The phenotypes of *C. australis* flowering on WT and *Ntft5* mutant tobacco plants. *C. australis* was infested on WT (right) and *Ntft5* mutant (left) tobacco plants (cv K326), and the photograph was taken when the tobacco plants were 150 days old. The *Ntft5* mutants were smaller than were the WT tobacco and neither *Ntft5* nor the parasitizing *C. australis* flowered (left), while WT tobacco and the parasitizing *C. australis* had mature seeds (blue and red arrows respectively indicate *C. australis* and tobacco seed capsules). Bar = 10 cm.

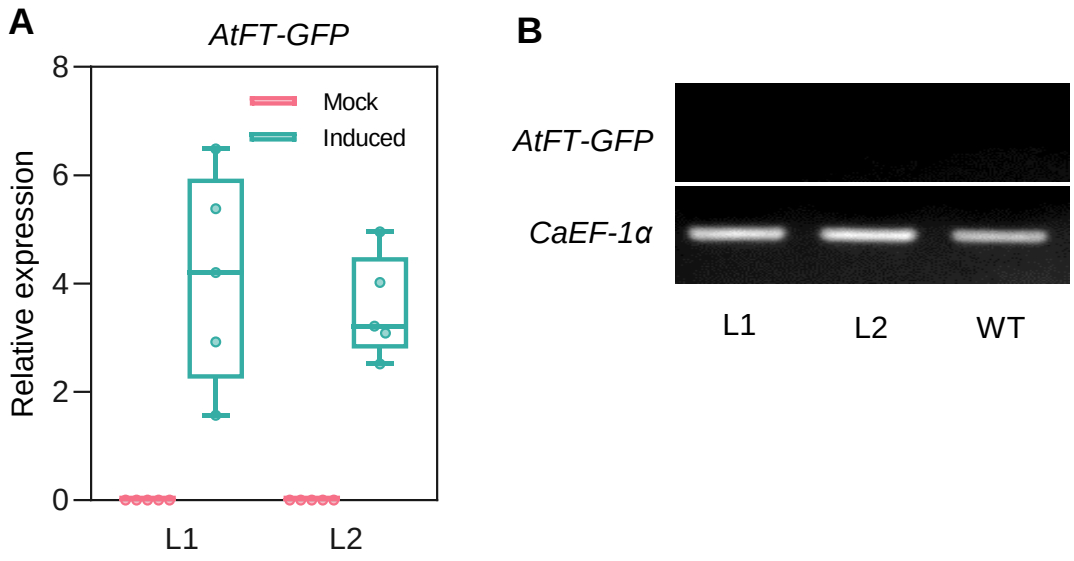

Fig. S5. Detection of transgenic *AtFT-GFP* in tobacco and *C. australis* after estradiol induction. Transgenic tobacco plants (cv Maryland Mammoth; two lines L1 and L2) harboring the T-DNA from pER8-AtFT-GFP were treated with estradiol or mock treated. The treated tobacco leaf samples and *C. australis* stems were harvested 48 h after treatments. Plants were grown under LD conditions for 30 days before being treated. **(A)** qPCR analysis of the relative transcript levels of *AtFT-GFP* in tobacco leaves. Data were normalized to the transcript levels of *NtEF-1α* (n = 5). **(B)** RT-PCR to detect the transcripts of *AtFT-GFP* in dodders grown on WT and two transgenic lines of tobacco. *CaEF-1α* was used as the internal control. PCR products were separated on an agarose gel. The same pair of primers were used for **(A)** and **(B)**.

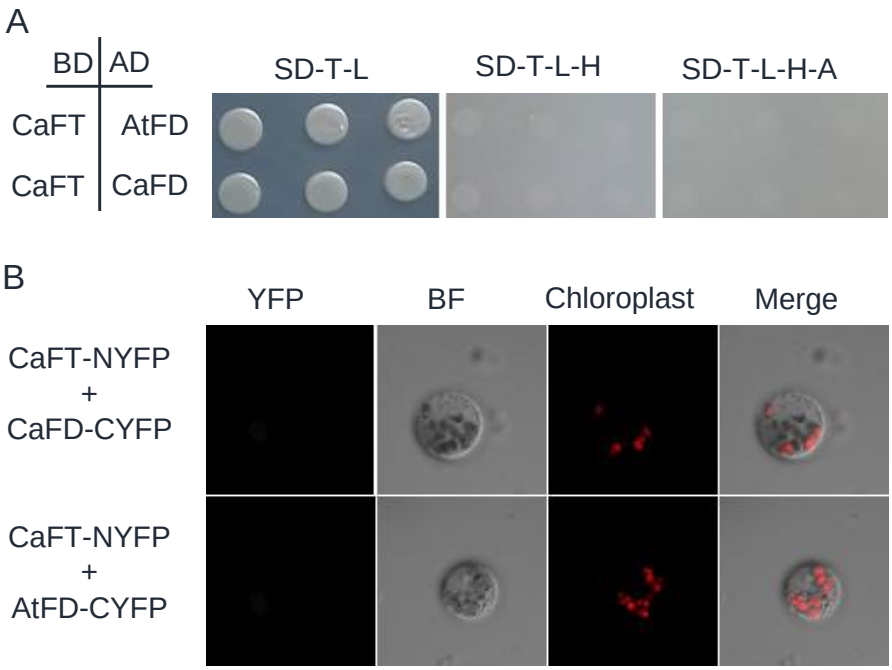

Fig. S6. Protein interaction assays between *C. australis* CaFT and FD proteins. (A) Y2H assay for interactions between CaFT and CaFD and between CaFT and AtFD. pGBKT7-CaFT and pGADT7-CaFD, pGADT7-AtFD were co-transformation into yeast AH109 respectively. The transformants were sequentially screened on plates without Trp (T) and Leu (L) first, and then on the plates without T, L, His (H), and adenine (A). To all plates 200 ng/ml Aureobasidin A and 20 mg/ml X-α-Gal were added. AD, GAL4 activation domain; BD, GAL4 DNA binding domain. **(B)** BiFC assays of the interactions between CaFT and CaFD and between CaFT and AtFD in Arabidopsis protoplasts. BF = bright field. Chloroplast = chloroplast autofluorescence.

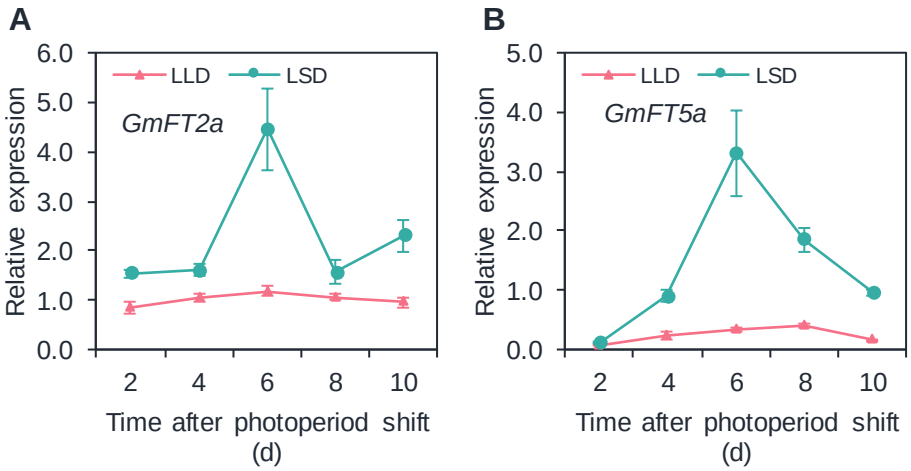

Fig. S7. Expression profiles of soybean *GmFT2a* and *GmFT5a* under different photoperiod regimes. *GmFT2a* **(A)** and *GmFT5a* **(B)** relative expression levels in soybean (Williams 82) under LLD and LSD. Fifteen-day-old Williams 82 plants parasitized by dodder were divided into two groups, one of which remained under LD condition (LLD), while the other was transferred to SD condition (LSD). Relative transcript levels were analyzed with qPCR and normalized to the levels of *Gmß-tubulin* (*GmTUB*) (n = 5).

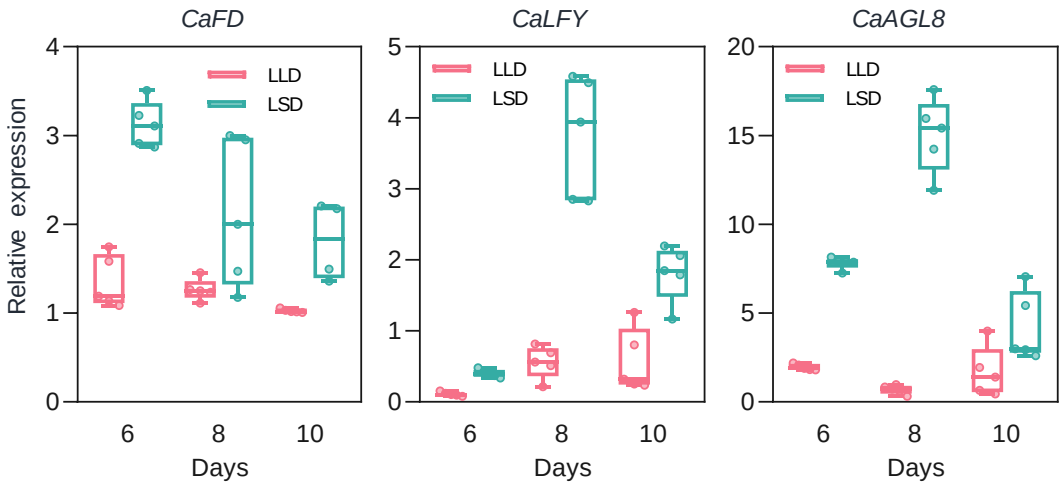

Fig. S8. Expression profiles of *C. australis CaFD*, *CaLFY*, and *CaAGL8* under different photoperiod regimes. *CaFD*, *CaLFY*, and *CaAGL8* relative expression levels in dodder (*C. australis*) under LLD and LSD. Fifteen-day-old Williams 82 plants parasitized by dodder were divided into two groups, one of which remained under LD condition (LLD), while the other was transferred to SD condition (LSD). Relative transcript levels were analyzed with qPCR and normalized to the levels of *CaEF-1α* (n = 5).

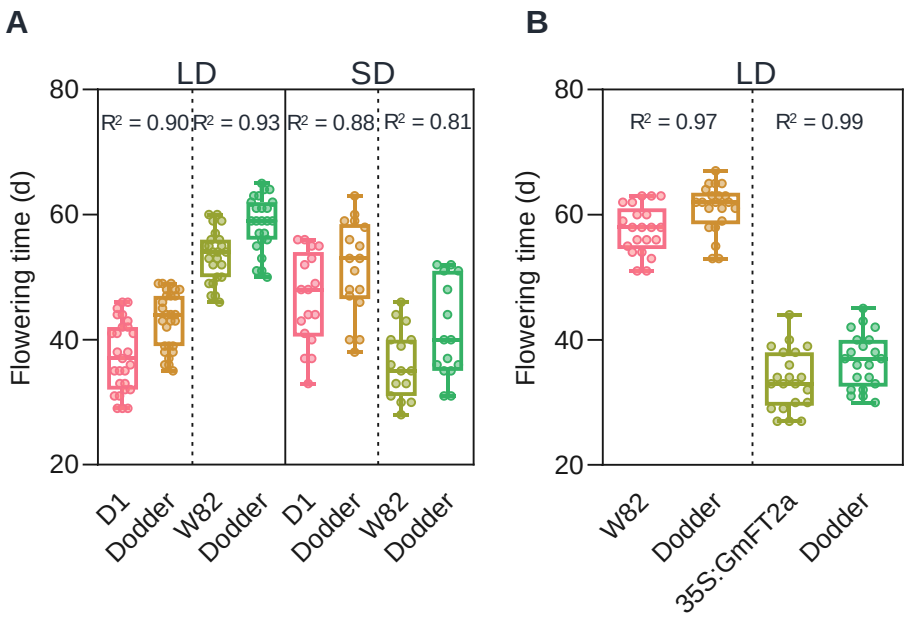

Fig. S9. The flowering times of dodder *C. campestris* on different hosts under different photoperiod conditions. (A) The flowering times of soybean Dongsheng 1 (D1) (left panels) and Williams 82 (W82) (right panels) and their parasitizing dodder *C. campestris* under long day (LD) and short day (SD) (n = 27, 27, 24, 24 under LD; n = 17, 17, 15, 15 under SD; from left to right). **(B)** The flowering times of WT soybean (Williams 82; W82) and transgenic soybean overexpressing *GmFT2a* (35S:GmFT2a) and the parasitizing dodder, which were grown under LD (n = 21, 21, 21, 21 from left to right). The horizontal bars within boxes indicate medians. The tops and bottoms of boxes indicate upper and lower quartiles, respectively. The upper and lower whiskers represent maximum and minimum, respectively. R^2 indicates correlation coefficients, which were obtained from correlation analysis using the flowering times of each host-dodder pair.

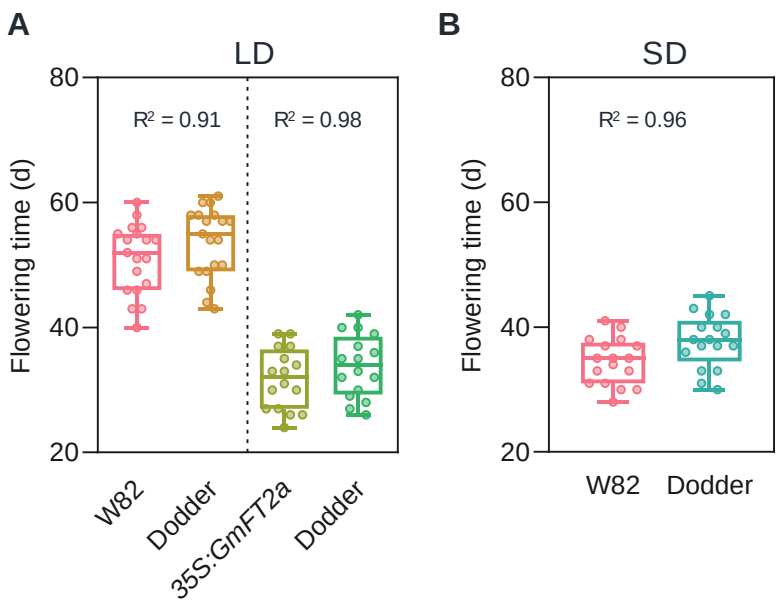

Fig. S10. The flowering times of dodder *C. europaea* on different hosts and under different photoperiod regimes. (A) The flowering times of WT soybean (Williams 82; W82) and transgenic soybean *GmFT2a* (35S:GmFT2a) and the parasitizing dodder, which were grown under LD (n = 19, 19, 16, 16; from left to right). **(B)** The flowering times of WT soybean (Williams 82) and the parasitizing dodder, which were grown under SD (n = 17, 17; from left to right). The horizontal bars within boxes indicate medians. The tops and bottoms of boxes indicate upper and lower quartiles, respectively. The upper and lower whiskers represent maximum and minimum, respectively. R^2 indicates correlation coefficients, which were obtained from correlation analysis using the flowering times of each host-dodder pair.

Table S1. The top 30 enriched GO (gene ontology) terms

GO.ID	Term	Annotated	Significant	Expected	p-values
GO:0009873	ethylene-activated signaling pathway	89	5	0.93	0.00020
GO:0055114	oxidation-reduction process	860	18	9.03	0.00029
GO:0007623	circadian rhythm	70	6	0.74	0.00071
GO:0006357	regulation of transcription from RNA polymerase II	118	5	1.24	0.00074
GO:2000652	regulation of secondary cell wall biogenesis	7	2	0.07	0.00074
GO:0080110	sporopollenin biosynthetic process	7	2	0.07	0.00074
GO:0009737	response to abscisic acid	329	8	3.46	0.00085
GO:0016115	terpenoid catabolic process	8	2	0.08	0.00099
GO:0042773	ATP synthesis coupled electron transport	35	3	0.37	0.00121
GO:0045893	positive regulation of transcription, DNA-templated	136	5	1.43	0.00140
GO:0006355	regulation of transcription, DNA-templated	1114	21	11.7	0.00182
GO:0055062	phosphate ion homeostasis	12	2	0.13	0.00229
GO:0009813	flavonoid biosynthetic process	47	3	0.49	0.00286
GO:0009908	flower development	325	7	3.41	0.00355
GO:0009690	cytokinin metabolic process	16	2	0.17	0.00410
GO:0019748	secondary metabolic process	150	9	1.58	0.00439
GO:0048731	system development	1245	20	13.08	0.00523
GO:1900033	negative regulation of trichome patterning	1	1	0.01	0.00604

GO:0010378	temperature compensation of the circadian clock	1	1	0.01	0.00604
GO:0070544	histone H3-K36 demethylation	1	1	0.01	0.00604
GO:0019379	sulfate assimilation, phosphoadenylyl sulfate	1	1	0.01	0.00604
GO:0090143	nucleoid organization	1	1	0.01	0.00604
GO:0042754	negative regulation of circadian rhythm	1	1	0.01	0.00604
GO:0051792	medium-chain fatty acid biosynthetic process	1	1	0.01	0.00604
GO:0048574	long-day photoperiodism, flowering	20	2	0.21	0.00639
GO:0010200	response to chitin	65	3	0.68	0.00711
GO:0000165	MAPK cascade	22	2	0.23	0.00770
GO:0050832	defense response to fungus	131	4	1.38	0.00870
GO:0010074	maintenance of meristem identity	23	2	0.24	0.00840
GO:0042744	hydrogen peroxide catabolic process	23	2	0.24	0.00840

Table S2. The primer sequences used for cloning or PCR verification

Gene name	Acession number	Primer name	Primers sequence (5'-3')
CaFD	RAL43730	puc_SPYCE-CaFD	F: TCTAGAATGCTCGAATGGAAGCAGTTAG R: GGTACC CAACTGGAAATCAAAGTTCAC
AtFT	NP_001320342	puc_SPYNE-AtFT	F: TCTAGA ATGTCTATAAATATAAGAGA R: CCCGGG CTTAAGTCTTCTTCCTCCGC
AtFT	NP_001320342	pGBKT7-AtFT	F: CCCGGG ATGTCTATAAATATAAGAGA R: GTCGAC CTAAAGTCTTCTTCCTCCGC
CaFD	RAL43730	pGADT7-CaFD	F: CCCGGGTGCTCGAATGGAAGCAGTTAG R: GGTACC CTACTGGAAATCAAAGTTCAC
CaFT	RAL41934	pGBKT7-CaFT	F: ACTG CCATGG ATGGCGGCGAGGCGGCAGAG R: ATGC CTGCAG CTATTGACGGCGCCTGCCGC
AtFD	NP_195315	pGADT7-AtFD	F: GTCA CCCGGG ATGTTGTCATCAGCTAAGCATCAG R: GCTA GGATCC TCAAAATGGAGCTGTGGAAGACCG
CaFT	RAL41934	puc_SPYNE-CaFT	F: ATCG TCTAGA ATGGCGGCGAGGCGGCAGAG R: ATGG GGTACC CTATTGACGGCGCCTGCCGC
AtFD	NP_195315	puc_SPYCE-AtFD	F: ATCG TCTAGA ATGTTGTCATCAGCTAAGCATCAG R: ATGG GGTACC TCAAAATGGAGCTGTGGAAGACCG
CaFT	RAL41934	pCAMBIA1301-CaFT-genomic	1F: ATGGGGTACCATGGCGGCGAGGCGGCAGAG 1R: CCCATCACGTATATGTGTAGGCCTTTTGGC 2F: GCCAAAAGGCCTACACATATACGTGATGGG 2R: AGCTGGTACCCTATTGACGGCGCCTGCCGCCGGT
NtFT5	KY306471	NtFT5- Cas9-sgRNA	F: GATTgCACCAACTTCAACCCTGGGC R: AAACGCCCAGGGTTGAAGTTGGTGc

CaFT	RAL41934	JC1 (for CaFT-genomic)	F: ATGGCGGCGAGGCGGCAGAGA R: GTGTCATCCAACATCGGAAG
CaFT	RAL41934	JC2 (for CaFT-genomic)	F: CTTCCGATGTTGGATGACAC R: GAATGTCATACTCAAGGGCC
CaFT	RAL41934	JC3 (for CaFT-genomic)	F: GGCCCTTGAGTATGACATTC R: GGTCACCGTGTAGGCCTTTTGGCTATA
CaFT	RAL41934	JC4 (for CaFT-genomic)	F: TATAGCCAAAAGGCCTACACATATACGTGAT R: CTACTTCATGTCAAGCACGC
CaFT	RAL41934	JC5 (for CaFT-genomic)	F: GCGTGCTTGACATGAAGTAG R: AGTTGAGTGCTTGGTCAGAA
CaFT	RAL41934	JC6 (for CaFT-genomic)	F: TTCTGACCAAGCACTCAACT R: CTATTGACGGCGCCTGCCGCCGGT
AtFT	NP_001320342	pER8-AtFT-GFP	AtFT-XhoI-F: AGTCCTCGAGATGTCTATAAATATAAGAGA AtFT-linker-HindIII-R: AGTCAAGCTTACCTCCTCCAAGTCTTCTTCCTCCGCAGC GFP-HindIII-F: AGTCAAGCTTGTGAGCAAGGGCGAGGAGCT GFP-SpeI-R: ATAGACTAGTTTACTTGTACAGCTCGTCCA

Table S3. The primer sequences for real-time PCR or RT-PCR analysis

Gene name	Acession number	Primer name	Primers sequence (5'-3')
GmFT2a	AB550122	qPCR GmFT2a-F	TAATTCATAACAAAGCAAACGAGTA
		qPCR GmFT2a-R	GCTGACATCTCTGTTATTGTAGGTA
GmFT5a	AB550126	qPCR GmFT5a-F	GCCTTACTCCAGCTTATACT
		qPCR GmFT5a-R	GGCATGCTCTAGCATTGCAA
GmTUB	XM_028387761	qPCR GmTUB-F	TGCCACCATCAAGACTAAGAGG
		qPCR GmTUB-R	ACCACCAGGAACAACAGAAGG
NtEF-1α	XM_009595030	qPCR NtEF-1α-F	TGAGATGCACCACGAAGCTC
		qPCR NtEF-1α-R	CCAACATTGTCACCAGGAAGTG
AtFT	NP_001320342	qPCR AtFT-F	AGGCCTTCTCAGGTTCAAAACAAGC
		qPCR AtFT-R	TGCCAAAGGTTGTTCCAGTTGTAGC
CaEF-1α	RAL46102	qPCR CaEF-1α-F	CCAGGTCAACGAGCCAAAGA
		qPCR CaEF-1α-R	AGTCTCCACACGTCCAACAG
CaFT	RAL41934	qPCR CaFT-F	CGCAACCAATCGTCGGAATC
		qPCR CaFT-R	GGCGAATTCTCGAGTGCTGA
CaFT	RAL41934	RT-PCR CaFT-F	AAACCTCAGGGTGGCTTACG
		RT-PCR CaFT-R	GGGTTGAGCACGAGAGATGG
AtEF-1α	NC_003070	RT-PCR AtEF-1α-F	GGTCGCGTGTTATAGCTTCG
		RT-PCR AtEF-1α-R	TCAAAAGCAAAATAGGTGCAAAGT
CaFD	RAL43730	qPCR CaFD-F	ACTTCAACGTTCACCCCTTCA
		qPCR CaFD-R	GCCACTTCTATCGTCCCCAA
CaLFY	RAL49932	qPCR CaLFY-F	ATAGCGGAGCTGGGCTTTAC
		qPCR CaLFY-R	AGTAGCTCCCACCGGAAGAT

CaAGL8	RAL41220	qPCR CaAGL8-F	GATTTCGGTTCTCTGCGACG
		qPCR CaAGL8-R	GCAGGATTGCTTTGCCTCTC

***SI Appendix*, Dataset S1. The flowering genes in *C. australis* and *C. campestris* genomes.**

***SI Appendix*, Dataset S2. The qPCR result of *C. australis CaFT* expression.**

***SI Appendix*, Dataset S3. Mobile soybean proteins in *C. australis* stems.**

***SI Appendix*, Dataset S4. Differentially regulated genes between the dodders parasitizing soybean under LSD and LLD conditions.**

***SI Appendix*, Dataset S5. Mobile soybean mRNAs in *C. australis* stems.**

SI References

1. J. Zuo, Q. W. Niu, N. H. Chua, Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J* **24**, 265-273 (2000).
2. Y. Lei *et al.*, CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol Plant* **7**, 1494-1496 (2014).
3. J. Gao *et al.*, CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. *Plant Mol Biol* **87**, 99-110 (2015).
4. M. Walter *et al.*, Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *Plant J* **40**, 428-438 (2004).
5. S. J. Clough, A. F. Bent, Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**, 735-743 (1998).
6. R. B. Horsch *et al.*, Analysis of Agrobacterium tumefaciens virulence mutants in leaf discs. *Proc Natl Acad Sci U S A* **83**, 2571-2575 (1986).
7. Y. H. Song, R. W. Smith, B. J. To, A. J. Millar, T. Imaizumi, FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* **336**, 1045-1049 (2012).
8. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
9. D. Kim, B. Langmead, S. L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-360 (2015).
10. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
11. G. Sun *et al.*, Large-scale gene losses underlie the genome evolution of parasitic plant Cuscuta australis. *Nat Commun* **9**, 2683 (2018).
12. S. D. Yoo, Y. H. Cho, J. Sheen, Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc* **2**, 1565-1572 (2007).
13. D. M. Emms, S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**, 157 (2015).
14. F. Bouche, G. Lobet, P. Tocquin, C. Perilleux, FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Res* **44**, D1167-1171 (2016).
15. K. Katoh, D. M. Standley, A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* **32**, 1933-1942 (2016).
16. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).
17. D. M. Emms, S. Kelly, OrthoFinder2: fast and accurate phylogenomic orthology analysis from gene sequences. 10.1101/466201 (2019).
18. J. H. Ahn *et al.*, A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. *EMBO J* **25**, 605-614 (2006).