

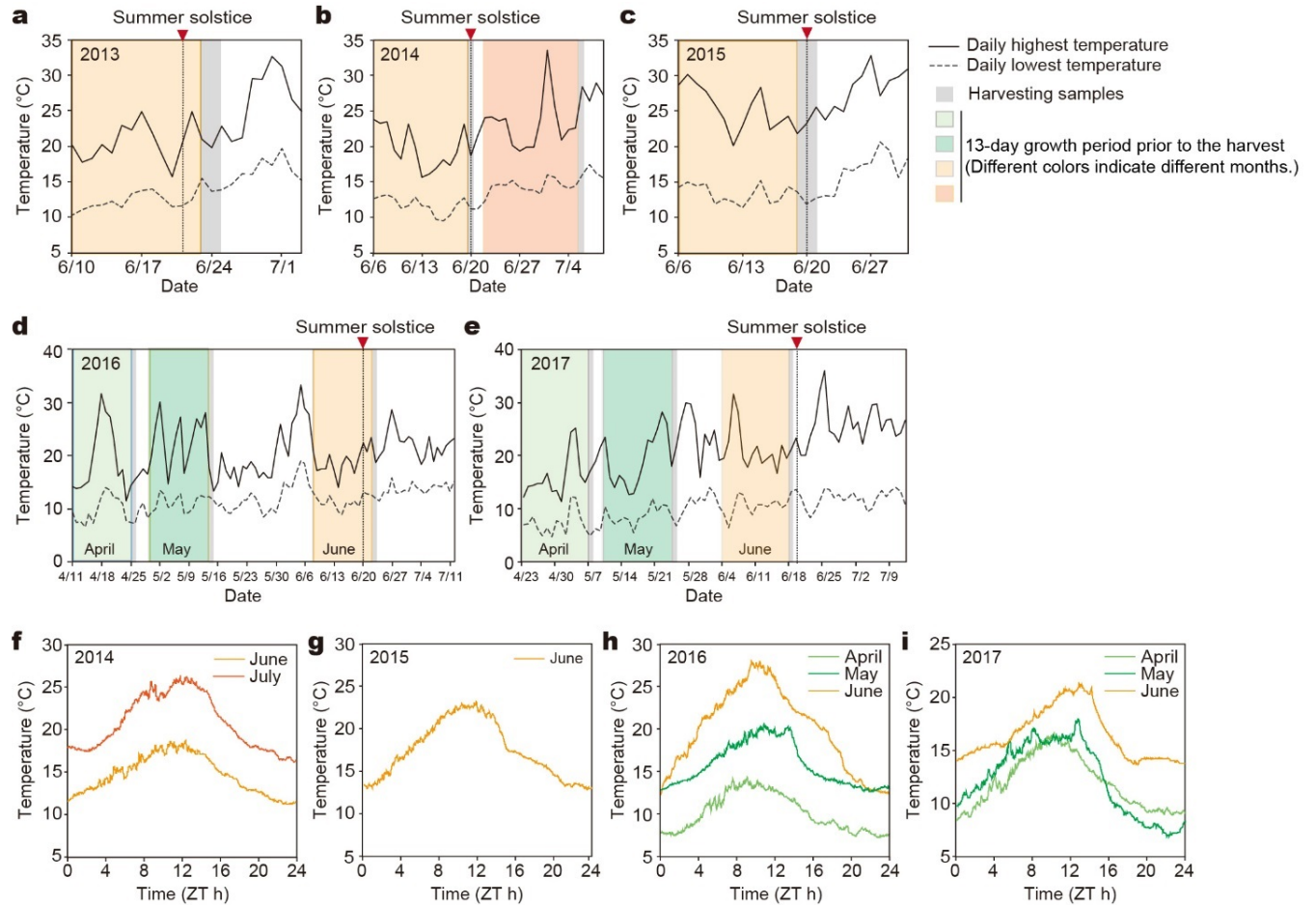
In the format provided by the authors and unedited.

Molecular basis of flowering under natural long-day conditions in *Arabidopsis*

Young Hun Song ^{1,2,8*}, Akane Kubota^{1,8}, Michael S. Kwon¹, Michael F. Covington³, Nayoung Lee¹, Ella R. Taagen¹, Dianne Laboy Cintrón¹, Dae Yeon Hwang ², Reiko Akiyama⁴, Sarah K. Hodge⁵, He Huang⁶, Nhu H. Nguyen¹, Dmitri A. Nusinow ⁶, Andrew J. Millar ⁵, Kentaro K. Shimizu ^{4,7} and Takato Imaizumi ^{1*}

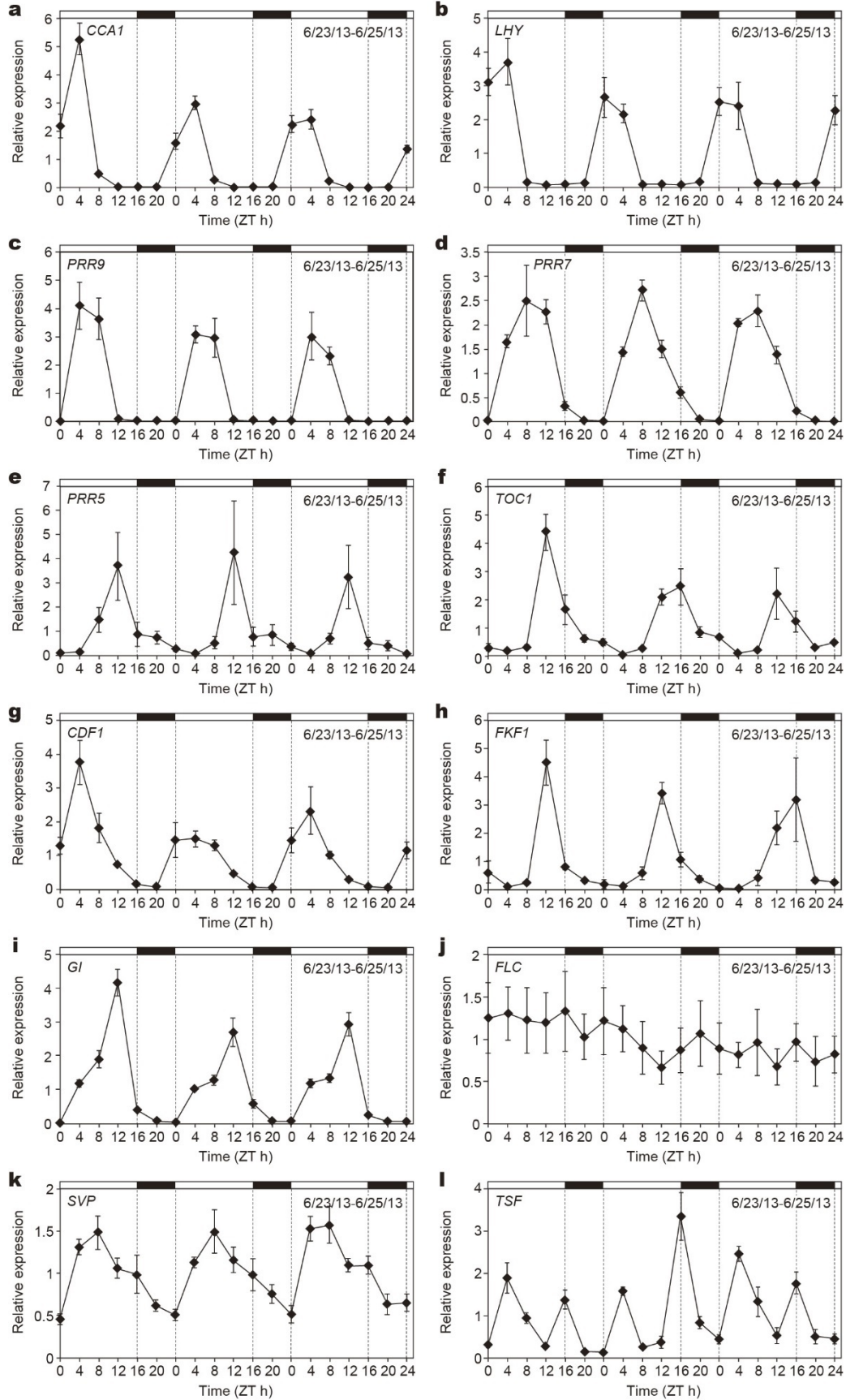
¹Department of Biology, University of Washington, Seattle, WA, USA. ²Department of Life Sciences, Ajou University, Suwon, Korea. ³Amaryllis Nucleics, Oakland, CA, USA. ⁴Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland. ⁵School of Biological Sciences and SynthSys, University of Edinburgh, Edinburgh, UK. ⁶Donald Danforth Plant Science Center, St. Louis, MO, USA. ⁷Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan. ⁸These authors contributed equally: Young Hun Song, Akane Kubota.

*e-mail: youngsong@ajou.ac.kr; takato@uw.edu



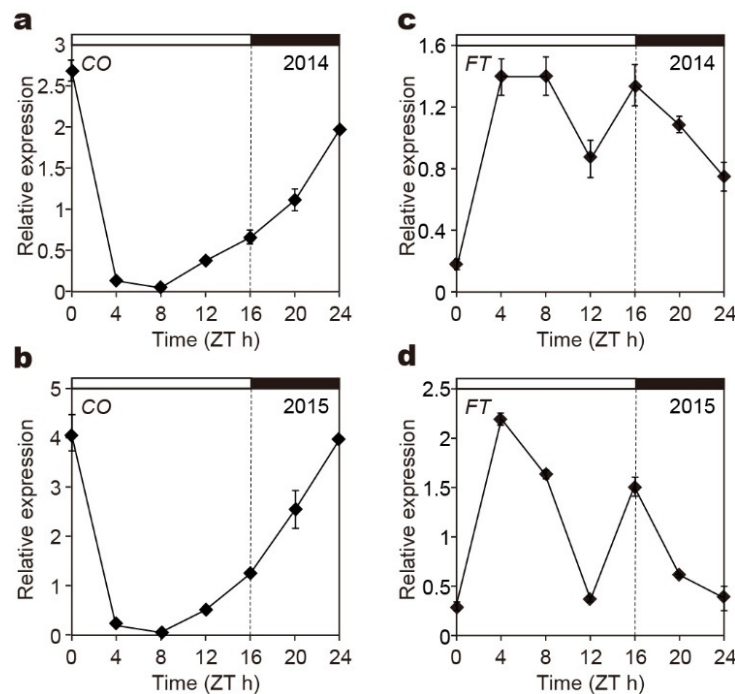
Supplementary Figure 1. Daily changes in high and low temperature from spring to early summer from 2013 to 2017 in Seattle when *Arabidopsis* plants were grown.

a-e, Air temperature records (daily high and low temperature) from 2013 to 2017 during the period when plants were grown outside are shown. Each panel shows daily high and low temperature changes in 2013 (**a**), 2014 (**b**), 2015 (**c**), 2016 (**d**), and 2017 (**e**). Solid and broken lines indicate the highest and lowest daily temperature, respectively. Colored boxes indicate the period when plants were grown outside before harvesting samples for gene expression analysis, and gray boxes indicate the day of harvesting. Some plants were allowed to continue growing to analyze flowering time after harvesting the samples. Red triangles indicate the summer solstice in each year. **f-i**, Twenty four-hour records of temperature changes on the days of harvesting indicated in (**a**) to (**e**): 2014 (**f**), 2015 (**g**), 2016 (**h**), and 2017 (**i**). Time indicates hours (h) after actual sunrise, set as Zeitgeber time 0 (ZT0).



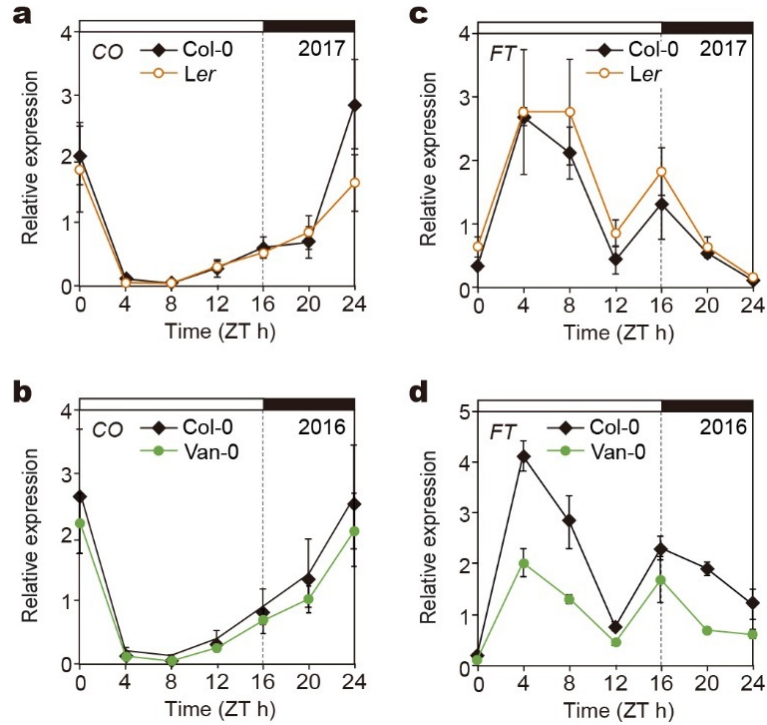
Supplementary Figure 2. Diurnal expression profiles of genes involved in the circadian clock and flowering time regulation in wild-type plants under natural long-day (LD) conditions.

a-l, Results of Q-PCR analysis show the expression profiles of genes involved in the photoperiodic flowering. The Col-0 plants were grown on soil outside in Seattle for 13 days, and harvested for three consecutive days (June 23rd to June 25th in 2013), every four hours, from dawn on day 14. For the circadian clock genes, the expression patterns of *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* (**a**), *LATE ELONGATED HYPOCOTYL (LHY)* (**b**), *PSUEDO RESPONSE REGULATOR 9 (PRR9)* (**c**), *PRR7* (**d**), *PRR5* (**e**), and *TIMING OF CAB EXPRESSION 1 (TOC1)* (**f**) genes were analyzed. We also analyzed the expression patterns of the genes important for photoperiodic flowering, such as *CYCLING DOF FACTOR 1 (CDF1)* (**g**), *FKF1* (**h**), and *GI* (**i**), the floral repressor genes, *FLOWERING LOCUS C (FLC)* (**j**) and *SHORT VEGITATIVE PHASE (SVP)* (**k**), and another florigen gene *TSF* (**l**). Bars above the traces represent light conditions: white and black bars represent day and night, respectively. All expression data were normalized against *IPP2* and *PP2A* and shown relative to the average expression of each gene. The results represent means $\pm$ SEM (n=3 biologically independent samples).



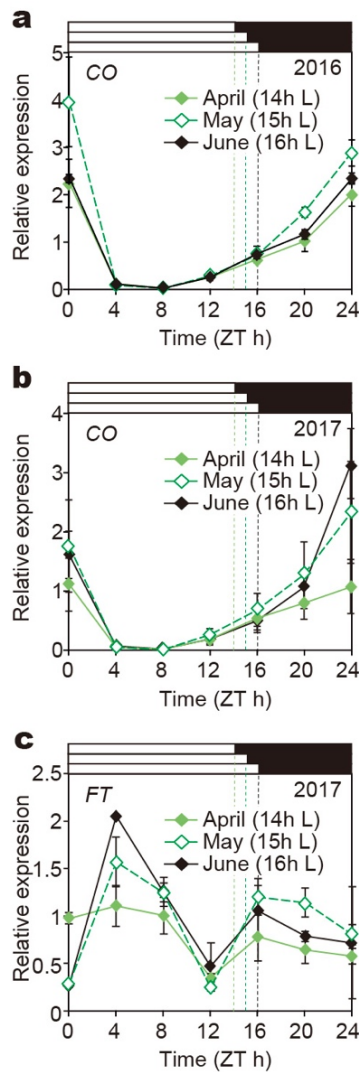
Supplementary Figure 3. Gene expression profiles of *CO* and *FT* in wild-type plants grown under natural LD conditions in 2014 and 2015.

a-d, Results of Q-PCR analysis show the gene expression profiles of *CO* (**a**, **b**) and *FT* (**c**, **d**) in the Col-0 plants grown outside in Seattle around the summer solstice of 2014 (**a**, **c**) and 2015 (**b**, **d**). The results represent means $\pm$ SEM (n=3 biologically independent samples).



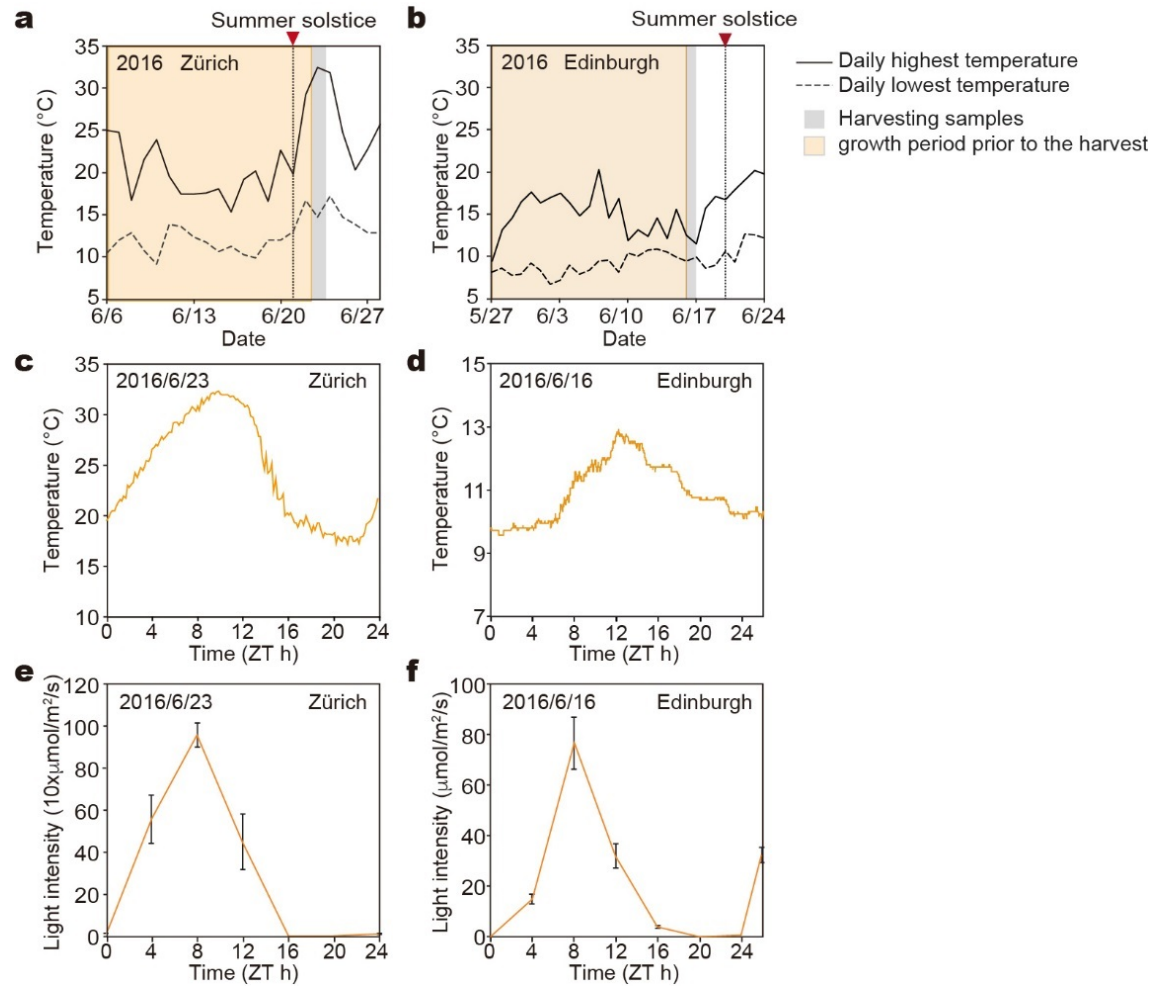
Supplementary Figure 4. Gene expression profiles of *CO* and *FT* in the common lab and the wild-type accessions under natural LD conditions.

a-d, Results of Q-PCR analysis show the gene expression patterns of *CO* (**a**, **b**) and *FT* (**c**, **d**) in the common lab accession, *Ler* plants (**a**, **c**) and the wild-type accession, *Van-0* plants (**b**, **d**) in comparison to the *Col-0* plants grown outside in Seattle around the summer solstice of 2016 (**b**, **d**) and 2017 (**a**, **c**). The results represent means $\pm$ SEM ($n=3$ biologically independent samples).



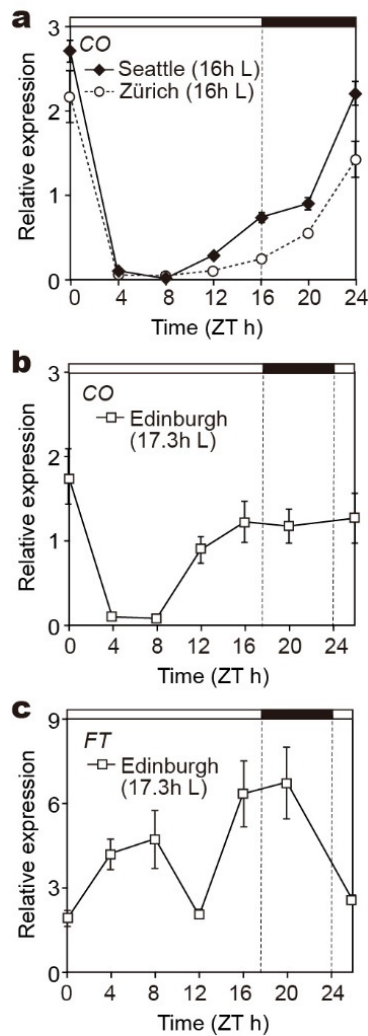
Supplementary Figure 5. Gene expression profiles of *CO* and *FT* in wild-type plants grown in each month from April to June.

a-c, Results of Q-PCR analysis show the gene expression profiles of *CO* (**a**, **b**) and *FT* (**c**) in the Col-0 plants grown outside in April, May, or June in 2016 (**a**) and 2017 (**b**, **c**). The results represent means $\pm$ SEM [n=3 biologically independent samples, except for the ZT4 samples harvested in June, which are n=2 biologically independent samples. The results of this time point (ZT4, June) only show the mean values]. The day length (hours (h) in light (L)) on the day of harvesting is indicated for each month.



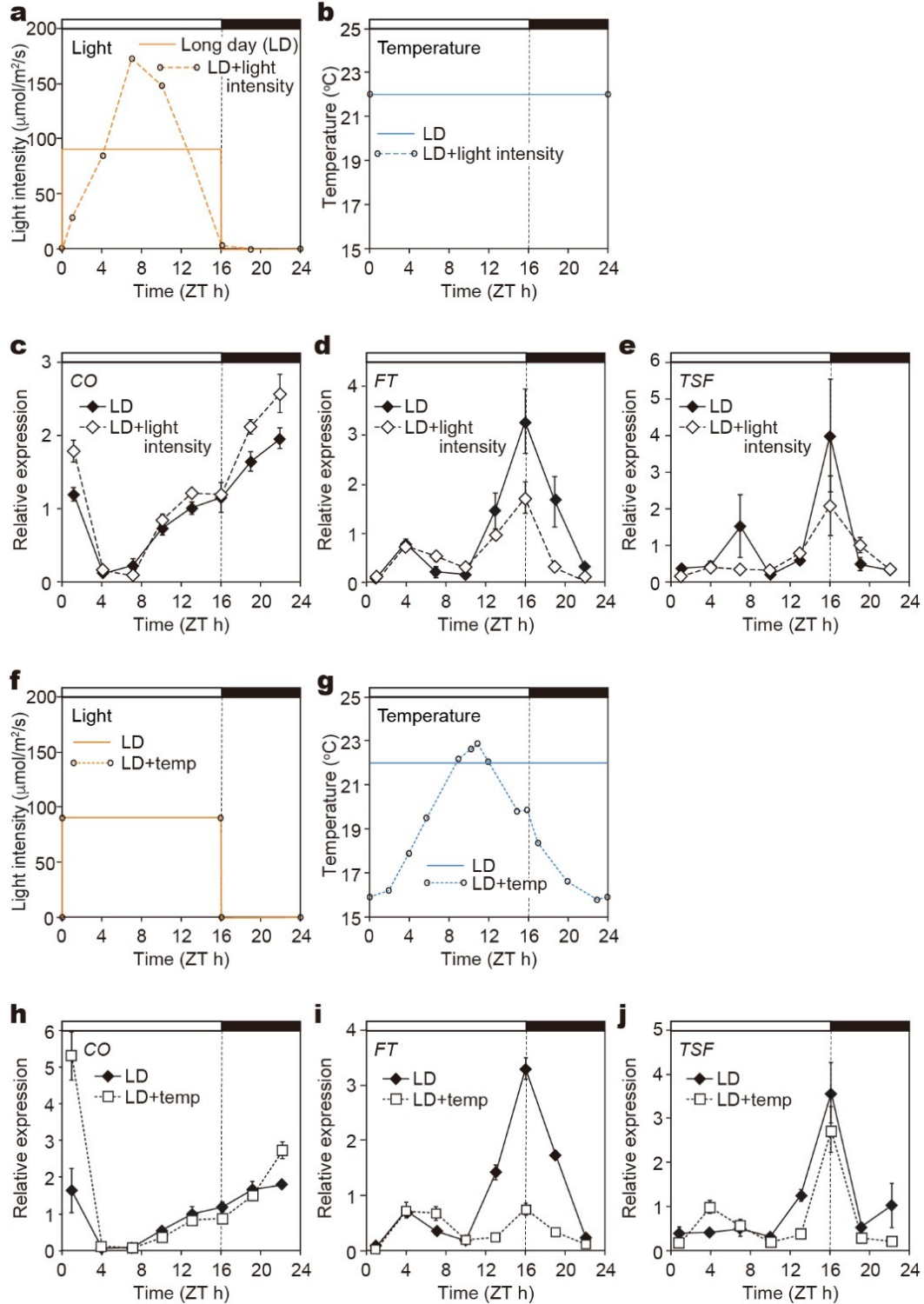
Supplementary Figure 6. Daily high and low temperature changes and light conditions around the summer solstice in Zürich and Edinburgh.

a, b, Records of daily high and low temperature in Zürich (**a**) and Edinburgh (**b**) during the period when plants were grown outside in 2016 are shown. Solid and broken lines indicate the highest and lowest daily temperature, respectively. Colored boxes indicate the period when plants were grown outside before harvesting samples for gene expression analysis, and the gray boxes indicate the day of harvesting. Red triangles indicate the day of the summer solstice. **c, d**, Twenty four-hour records of temperature changes on the days of harvesting in Zürich (**c**) and Edinburgh (**d**). **e, f**, Observed light intensity throughout the days of harvesting in Zürich (**e**) and Edinburgh (**f**). Data represent means $\pm$ SEM measured in three different spots.



Supplementary Figure 7. Gene expression profiles of *CO* and *FT* in wild-type plants under natural LD conditions in different geographic locations.

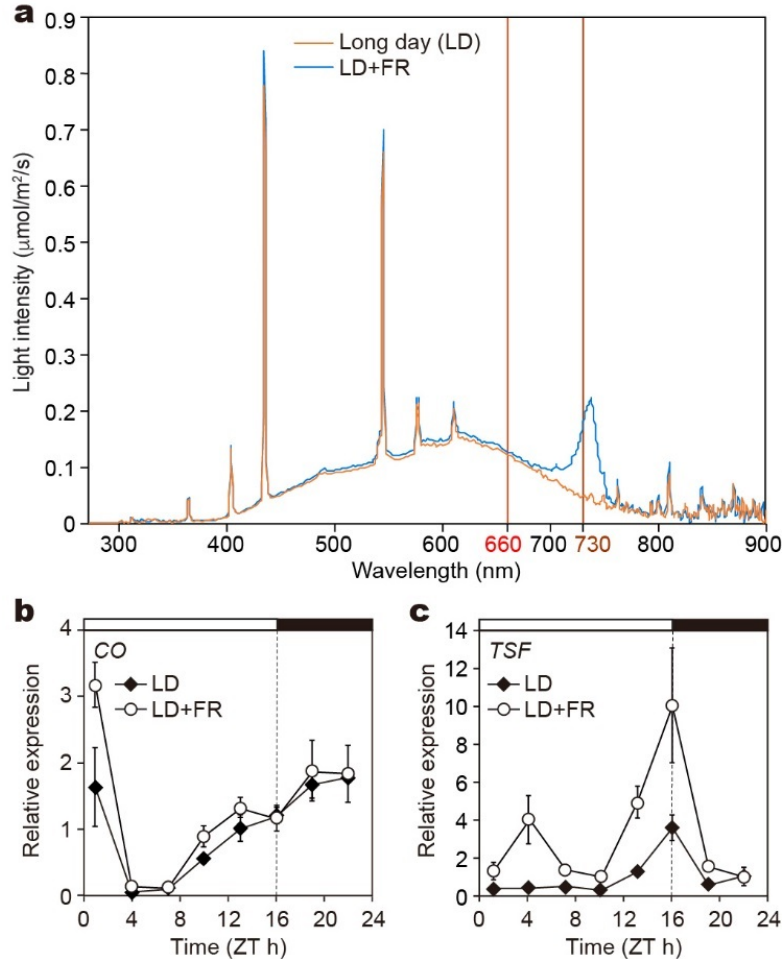
a-c, Results of Q-PCR analysis show the gene expression profiles of *CO* (**a**, **b**) and *FT* (**c**) in the Col-0 plants grown outside in Zürich (**a**) and Edinburgh (**b**, **c**) around the summer solstice in 2016. *CO* expression profile observed in Seattle (Supplementary Fig. 4b) is shown for comparison with that observed in Zürich. Plants were harvested on day 17 (**a**) and day 20 (**b**, **c**). The results represent means $\pm$ SEM ($n=3$ biologically independent samples). The day length (hours (h) in light (L)) on the day of harvesting is indicated.



Supplementary Figure 8. Gene expression profiles of *CO*, *FT*, and *TSF* under LD conditions with light or temperature changes.

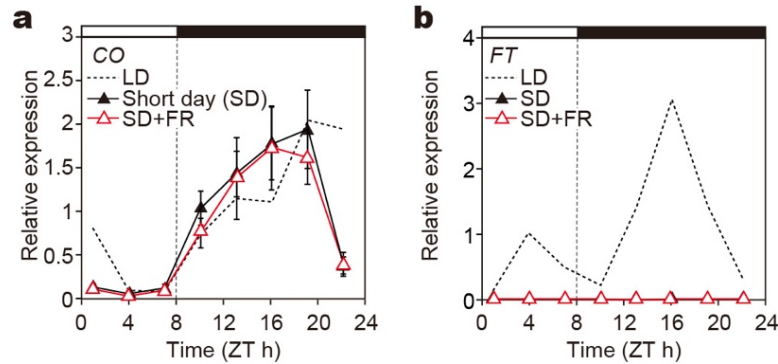
a, b, f, g, Experimental lab LD conditions with fluctuations in light intensity during the daytime (**a, f**) or temperature throughout the day (**b, g**). Solid lines indicate light or temperature changes

in LD, and broken lines indicates those in LD with oscillations of light intensity (LD+light intensity) or temperature (LD+temp). **c-e, h-j**, Results of Q-PCR analysis show the gene expression profiles of *CO* (**c, h**), *FT* (**d, i**), and *TSF* (**e, j**) in the Col-0 plants grown under the conditions shown in (**a**) and (**b**), or under (**f**) and (**g**). The results represent means $\pm$ SEM (n=3 biologically independent samples).



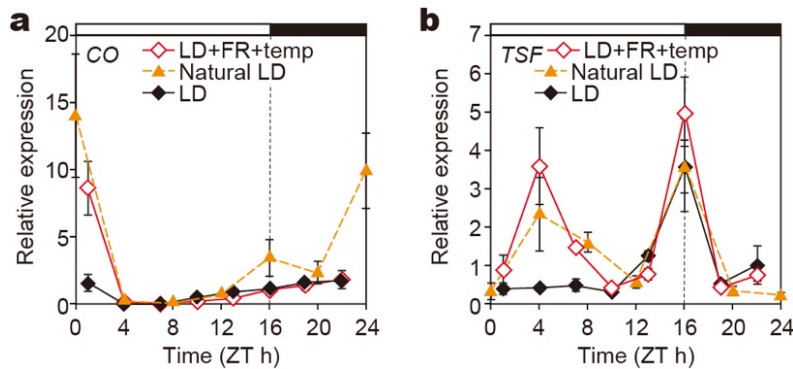
Supplementary Figure 9. Gene expression profiles of *CO* and *TSF* under lab LD conditions supplemented with far-red (FR) LED.

a, Light spectrum of fluorescent lamps used in LD (orange line) conditions and that of fluorescent lamps supplemented with dim FR LEDs used in LD+FR conditions (blue line). A vertical line indicates the absorption peaks for the Pr (660 nm) and Pfr forms (730 nm) of phytochrome. The light intensities of these two wavelengths (± 15 nm) were used to calculate the R/FR ratios. **b, c**, Results of Q-PCR analysis show the gene expression profiles of *CO* (**b**) and *TSF* (**c**) in the Col-0 plants grown in LD and LD+FR conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).



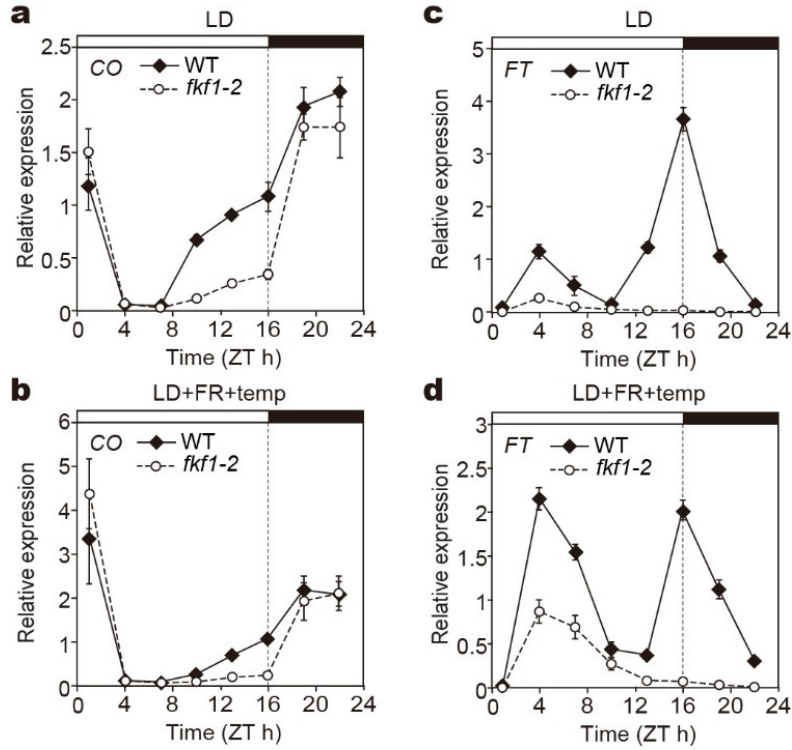
Supplementary Figure 10. Gene expression profiles of *CO* and *FT* under lab short-day (SD) conditions supplemented with FR light.

a, b, Results of Q-PCR analysis show the gene expression profiles of *CO* (**a**) and *FT* (**b**) in the Col-0 plants grown under SD (black triangles with solid lines) and SD+FR (red open triangles with solid lines) conditions in comparison to expression profiles under LD conditions (dotted line). Expression values in SD conditions are shown relative to the average expression value of each gene from LD conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).



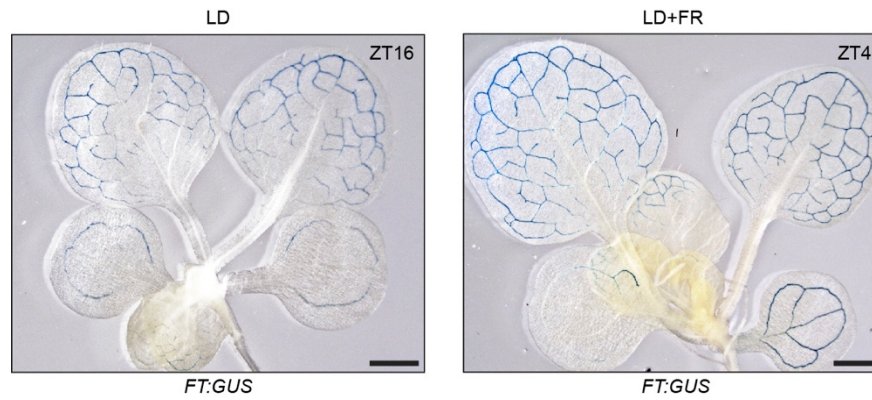
Supplementary Figure 11. Gene expression profiles of *CO* and *TSF* under recreated natural LD (LD+FR+temp) conditions.

a, b, Results of Q-PCR analysis show the gene expression profiles of *CO* (**a**) and *TSF* (**b**) in the Col-0 plants grown under lab LD conditions, outside samples from July in 2014, and LD+FR+temp conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).



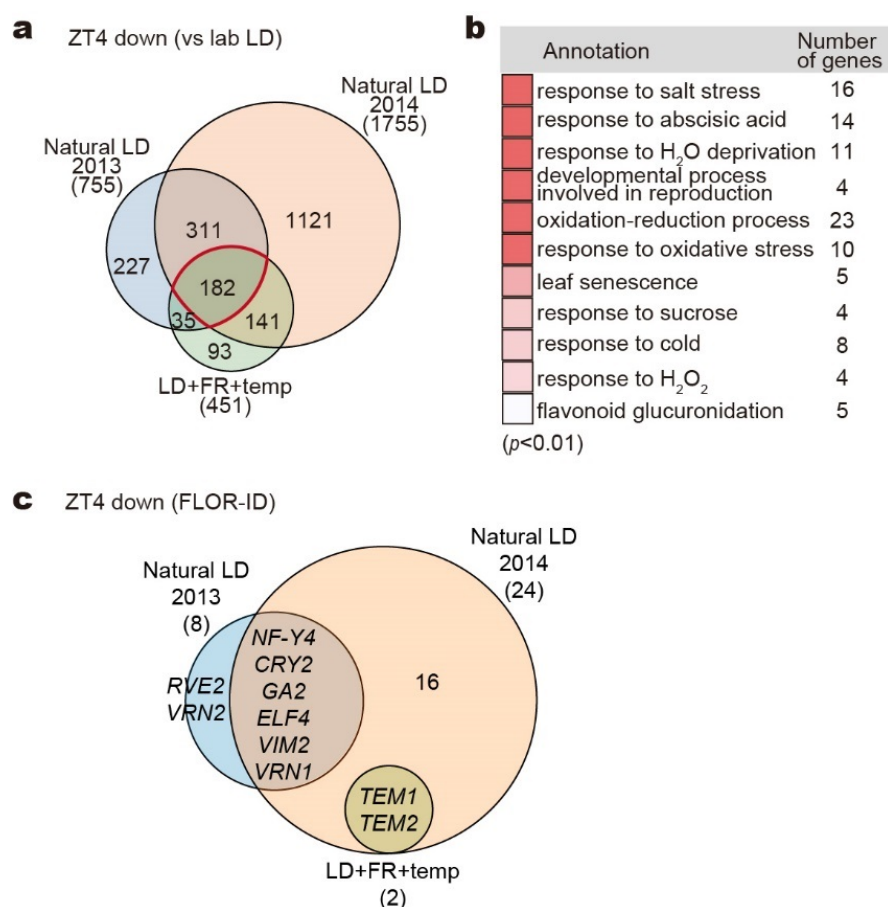
Supplementary Figure 12. Gene expression profiles of *CO* and *FT* under LD and LD+FR+temp conditions.

a-d, Results of Q-PCR analysis show the gene expression profiles of *CO* (**a**, **b**) and *FT* (**c**, **d**) in the Col-0 plants and the *fkl1-2* mutant grown under LD (**a**, **c**) and LD+FR+temp (**b**, **d**) conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).



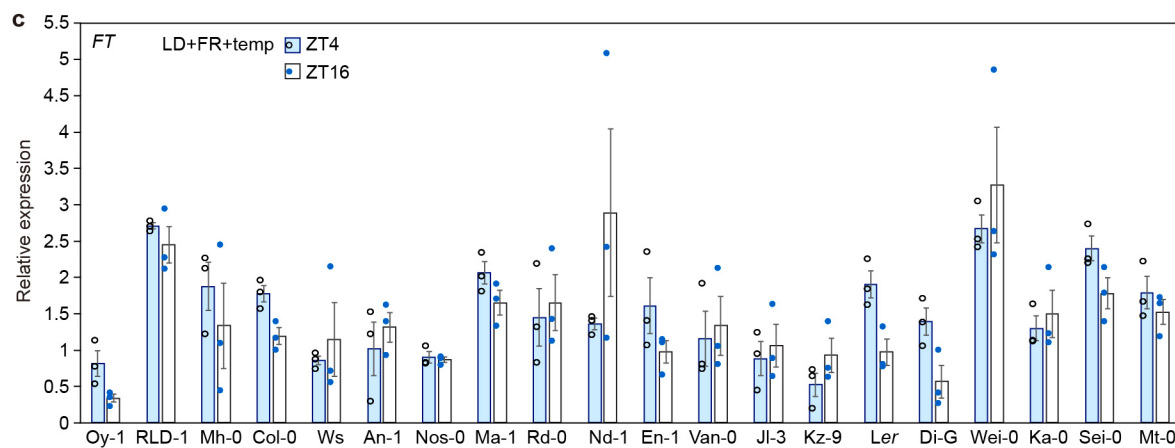
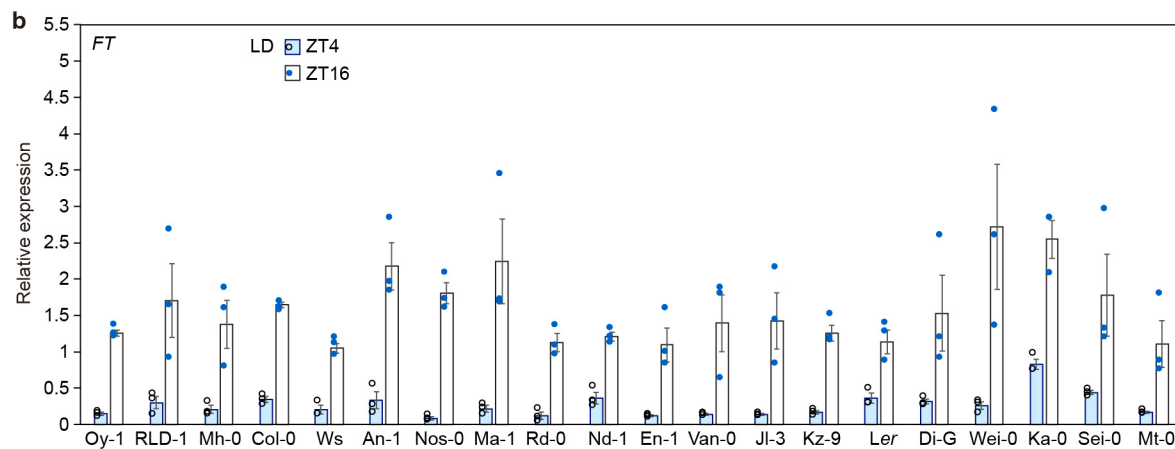
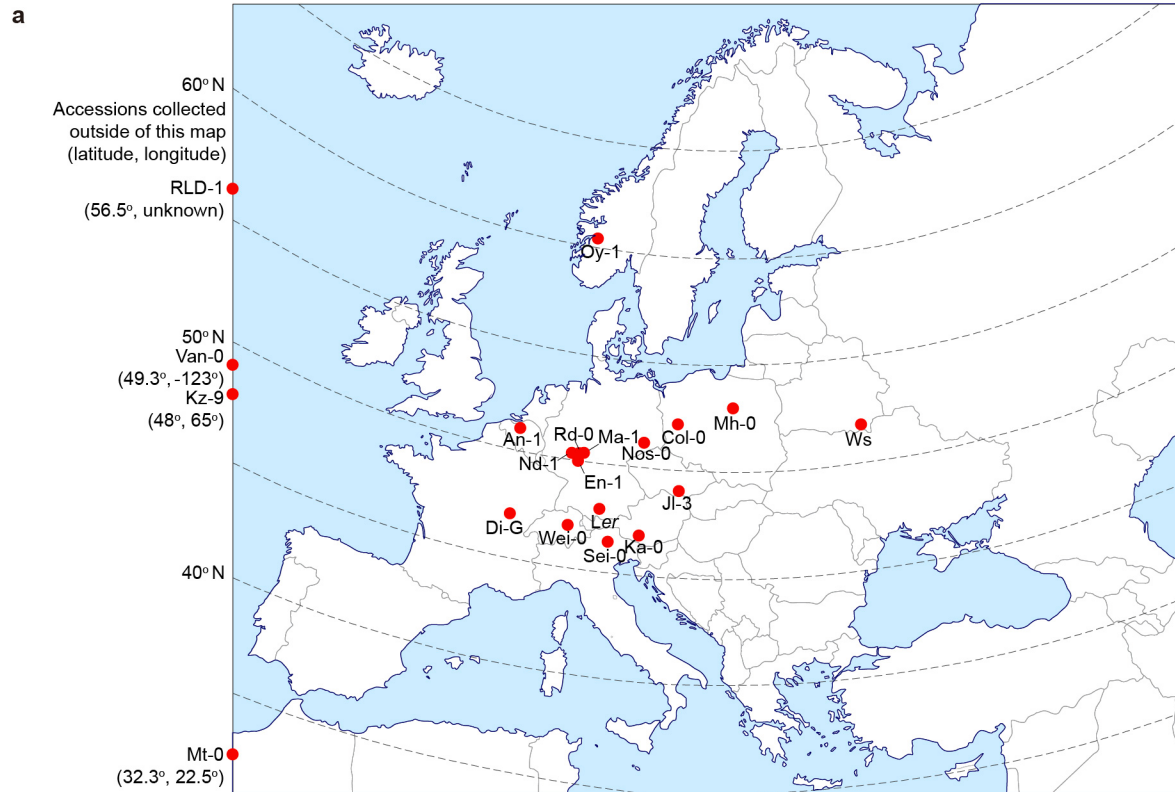
Supplementary Figure 13. GUS staining image from the plant harboring a *GUS* reporter controlled by the *FT* promoter.

Representative images of the *FT:GUS* plant grown under LD and LD+FR conditions for two weeks are shown. The *FT:GUS* plant grown under LD and LD+FR conditions were fixed and stained at ZT16 and ZT4, respectively. We analyzed 4-5 plants for each condition, and the experiments were repeated twice independently with similar results. Scale bar, 1 mm.



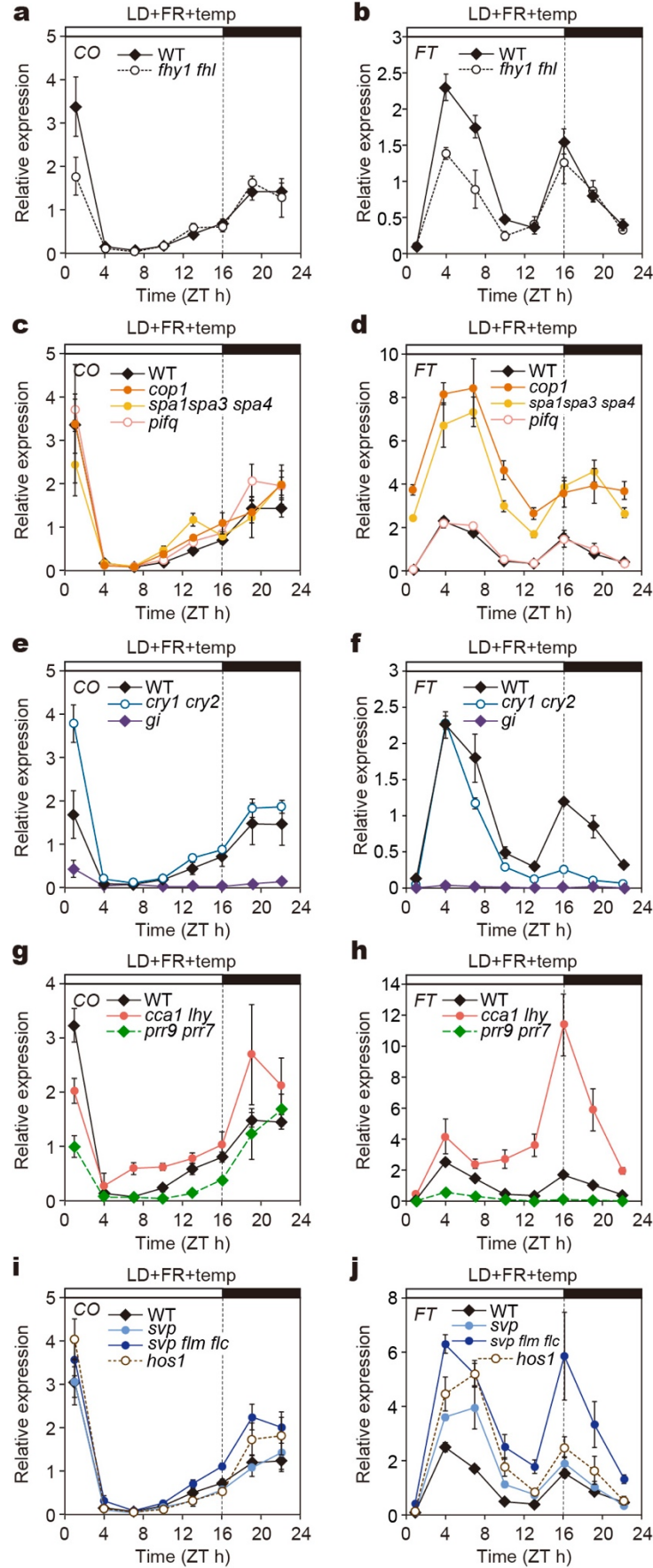
Supplementary Figure 14. RNA-seq analysis using the samples grown in natural LD and LD+FR+temp conditions.

a, A Venn diagram that shows the number of genes downregulated under natural LD condition from 2013, 2014, and LD+FR+temp conditions in comparison to LD conditions at ZT4 ($n=3$ biologically independent samples). Genes that showed more than a two-fold difference with less than 0.05 of false discovery rate (FDR) were included. **b**, A list of significantly enriched GO terms derived from 182 genes that are commonly downregulated under natural LD and LD+FR+temp conditions, compared to LD conditions at ZT4, is shown. The list is ranked by modified p -value (one-tail Fisher Exact Probability Values). **c**, A modified Venn diagram of (**a**) generated by the subset of genes annotated as flowering regulators in FLOR-ID.



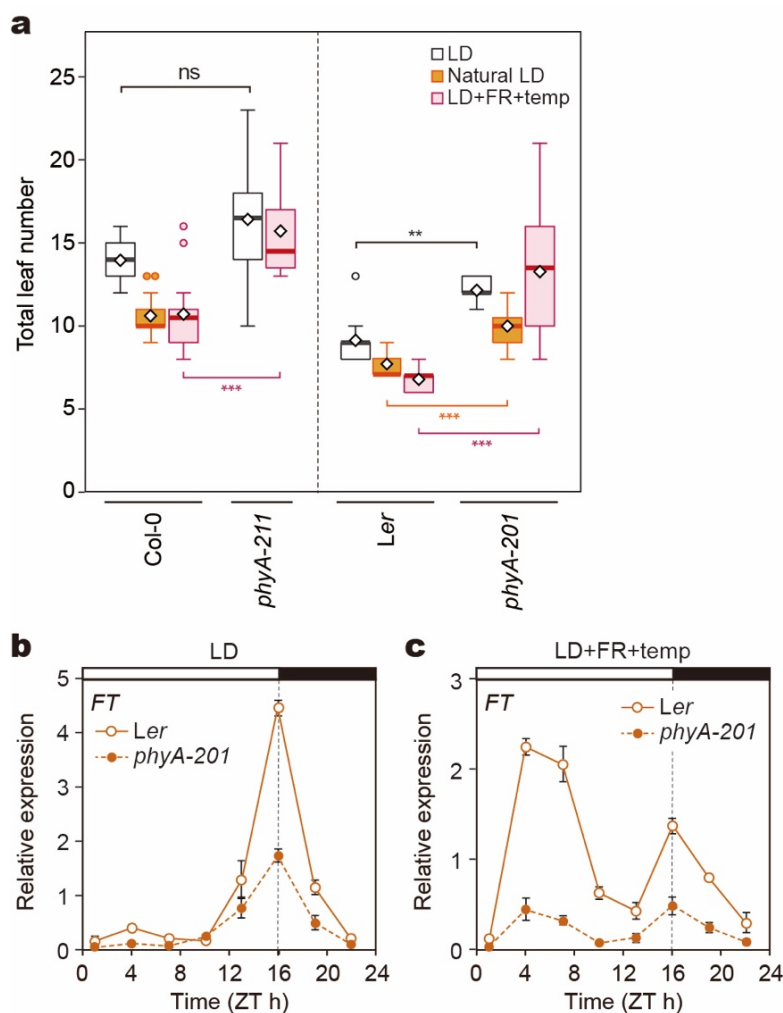
Supplementary Figure 15. The expression levels of *FT* from individual accessions under LD and LD+FR+temp conditions.

a, Geographic locations of the origins of the *Arabidopsis* accessions used for the analysis are shown. Latitude information for the accessions derived from outside this map is indicated on the left side of the map outline. **b, c**, The expression levels of *FT* at ZT4 and ZT16 under LD (**b**) and LD+FR+temp (**c**) conditions from each accession. All expression data are shown relative to the average expression value in the Col-0 plants grown under LD conditions. The accessions are aligned based on the latitudes (from high to low) of their isolated locations. We did not observe any clear correlations between original latitudes and *FT* patterns. Mean values without error bars are shown in Fig. 3c. The results shown by bar graphs represent means $\pm$ SEM (n=3 biologically independent samples), and each individual data point is overlaid on top of the bar graphs (open circles; ZT4 data points, and closed circles; ZT16 data points).



Supplementary Figure 16. Gene expression profiles of *CO* and *FT* from various mutants under LD+FR+temp conditions.

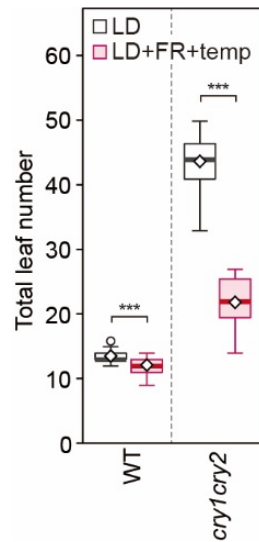
a-j, Results of Q-PCR analysis show the gene expression profiles of *CO* (**a**, **c**, **e**, **g**, and **i**) and *FT* (**b**, **d**, **f**, **h**, and **j**) in the mutants of known regulators involved in light-signaling: *phyA* *phyB*, *cop1*, *spa1* *spa3* *spa4*, and *pif4* mutants (**a** to **d**), photoperiod-dependent flowering: *cry1* *cry2* and *gi* mutants (**e** and **f**), circadian clock: *cca1* *lhy* and *prr9* *prr7* mutants (**g** and **h**), and ambient temperature-dependent flowering: *svp*, *svp* *flm* *flc*, and *hos1* mutants (**i** and **j**), in comparison to the WT (Col-0) plants grown under LD+FR+temp conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).



Supplementary Figure 17. Flowering phenotypes of *phyA* mutants grown under natural LD, lab LD and LD+FR+temp conditions.

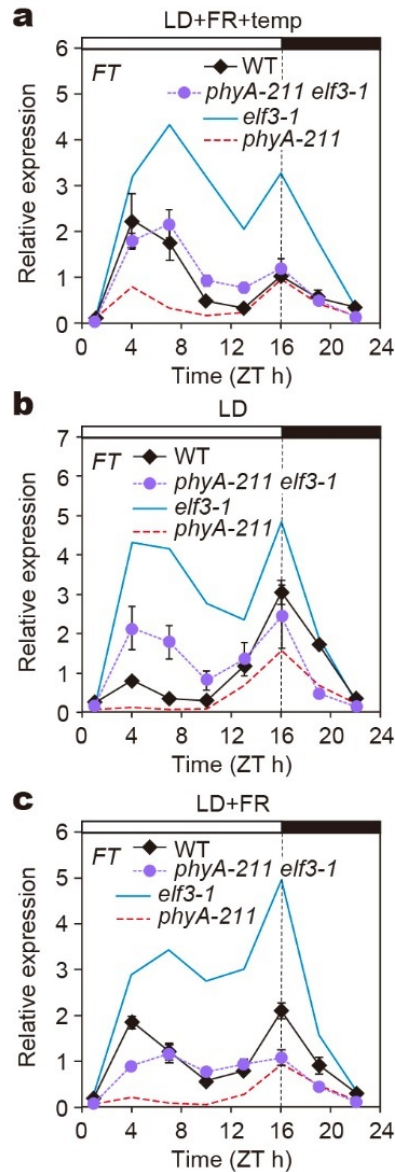
a, Flowering phenotypes of *phyA* mutants in different backgrounds (Col-0 and *Ler*) under natural LD, lab LD and LD+FR+temp conditions. The flowering time experiments for *Ler* and *phyA-201*

under natural LD conditions were performed in June 2017 (Supplementary Fig. 1e). Total leaf numbers were counted when plants bolted. Data are from at least 12 individual plants. Each box is located between the upper and the lower quartiles, and the whiskers indicate 1.5-times interquartile ranges. The thick horizontal lines in the boxes represent the median, and open diamonds represent the mean. Outliers are indicated by circles. The same data from Fig. 2c was used for Col-0 and *Ler* under natural LD. (** $p < 0.01$, *** $p < 0.001$, ns: non-significant, statistical information in Supplementary Table 3) **b, c**, Results of Q-PCR analysis show *FT* expression profiles under LD (**b**) and LD+FR+temp (**c**) conditions. The results represent means $\pm$ SEM from three biologically independent samples.



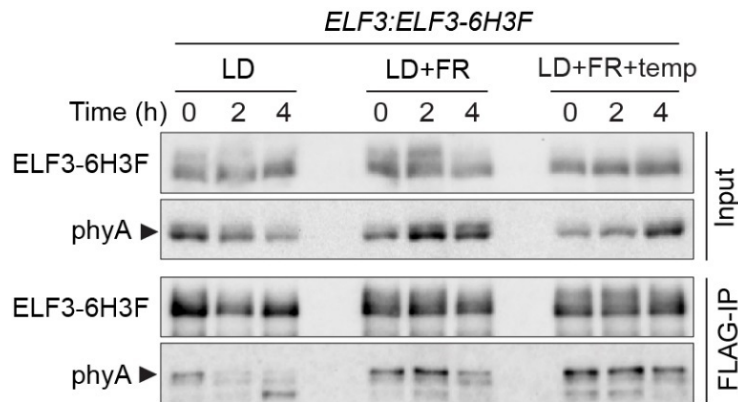
Supplementary Figure 18. Flowering phenotype of *cry1cry2* under LD and LD+FR+temp conditions.

Flowering phenotypes of WT (Col-0) and the *cry1cry2* mutant under LD and LD+FR+temp conditions. Total leaf numbers were counted when plants bolted. Data are from at least 16 individual plants. Each box is located between the upper and the lower quartiles, and the whiskers indicate 1.5-times interquartile ranges. The thick horizontal lines in the boxes represent the median, and open diamonds represent the means. (***) $p < 0.001$, statistical information in Supplementary Table 3).



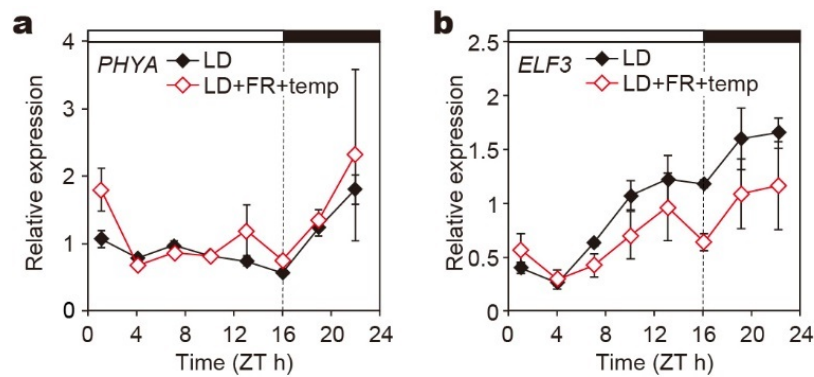
Supplementary Figure 19. Antagonistic function of *phyA* and *ELF3* on the regulation of *FT* expression profiles under different conditions.

a-c, Results of Q-PCR analysis show *FT* expression profiles from WT (Col-0), *phyA-211*, *elf3-1*, and *phyA-211 elf3-1* grown under LD+FR+temp (**a**), LD (**b**), and LD+FR (**c**) conditions. The results of WT plants and *phyA-211 elf3-1* represent means $\pm$ SEM from three biologically independent samples. The mean value traces of *FT* levels in *phyA-211* and *elf3-1* shown in Fig. 4 are included for reference.



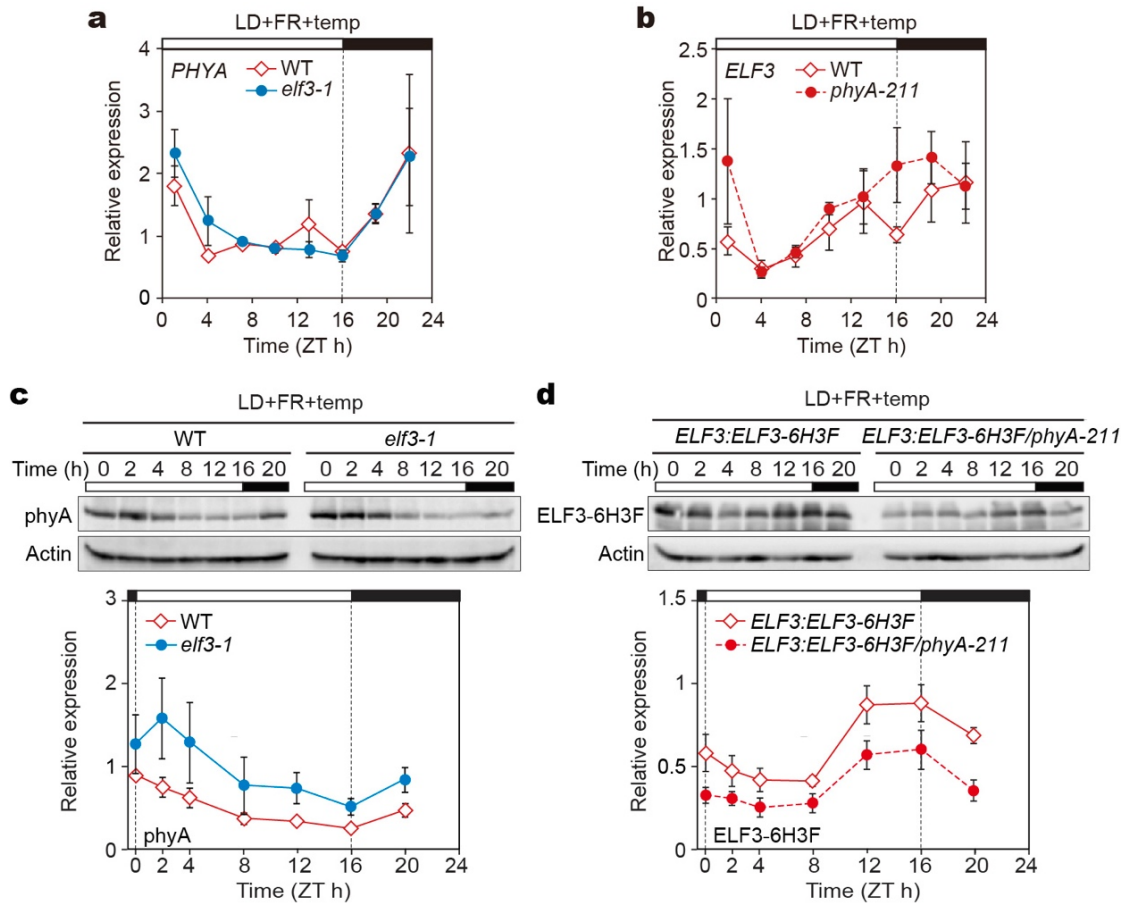
Supplementary Figure 20. Protein complex formation analysis between phyA and ELF3 *in vivo*.

Results of co-immunoprecipitation (co-IP) experiments between ELF3 and phyA proteins using the transgenic *ELF3:ELF3-6H3F* plants grown under LD, LD+FR, and LD+FR+temp conditions. The plant tissues were harvested at ZT0, ZT2, and ZT4 to analyze the potential interaction of ELF3 and phyA during the morning and used for whole protein extract. ELF3-6H3F was immunoprecipitated by anti-FLAG antibody, and the presence of phyA in the anti-FLAG antibody IPed samples was analyzed by anti-phyA antibody. Similar results were observed in two biologically independent experiments.



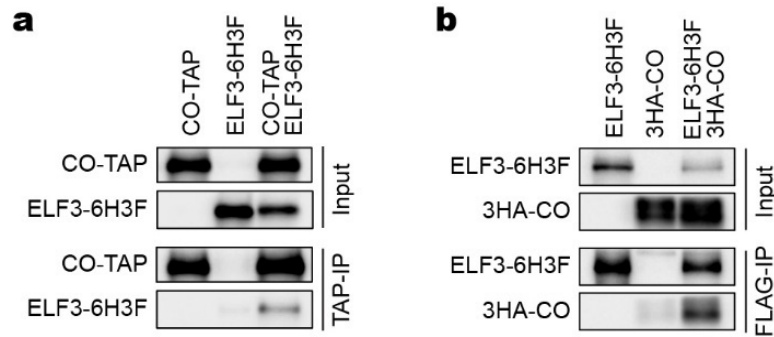
Supplementary Figure 21. Gene expression profiles of *PHYA* and *ELF3* under LD and LD+FR+temp conditions.

a, b, Results of Q-PCR analysis show *PHYA* (**a**) and *ELF3* (**b**) expression profiles in the Col-0 plants grown under LD and LD+FR+temp conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).



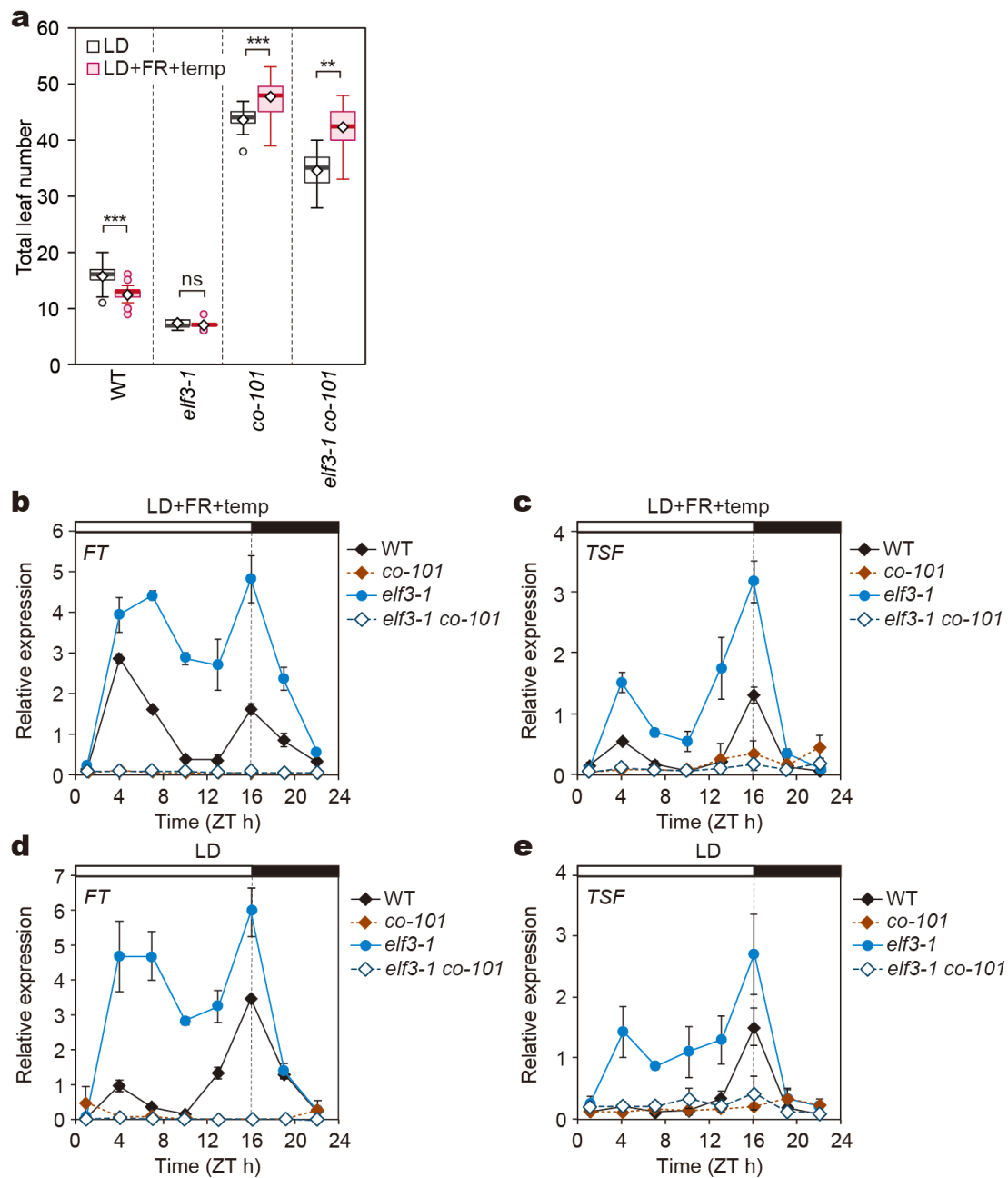
Supplementary Figure 22. mRNA and protein expression profiles of phyA and ELF3 under LD+FR+temp conditions.

a, b, Results of Q-PCR analysis show expression profiles of *PHYA* gene in *elf3-1* (**a**) and *ELF3* gene in *phyA-211* (**b**) compared to the Col-0 plants under LD+FR+temp conditions. **c, d,** Representative images of western blot analysis and quantification graphs that show phyA (**c**) or ELF3 (**d**) protein abundance in *elf3-1* (**c**) or *phyA-211* (**d**) mutants, respectively, compared to Col-0 plants under LD+FR+temp conditions. Actin protein was used for an internal control. The results represent means $\pm$ SEM [$n=3$ (**a, b**) and $n=4$ (**c, d**) biologically independent samples].



Supplementary Figure 23. Reciprocal co-immunoprecipitation assays between CO and ELF3 proteins *in planta*.

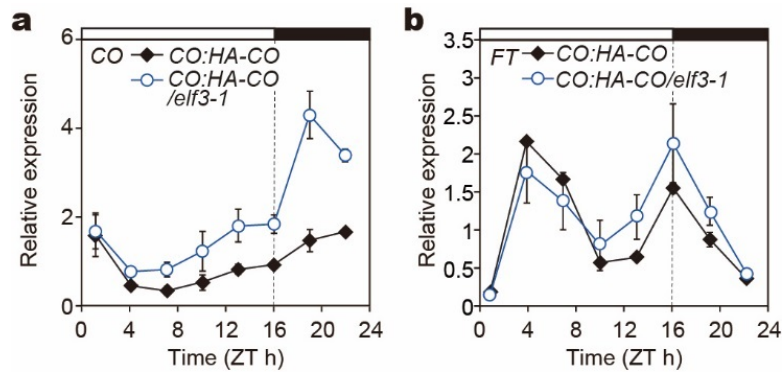
a, b, Co-IP experiments using tobacco-based transient expression. Agrobacterium harboring *CO* and *ELF3* overexpression constructs were individually or together infiltrated into *N. benthamiana* leaves. The CO-ELF3 protein complex formation was analyzed using co-IP methods when CO-TAP and ELF3-6H3F proteins (**a**) and 3HA-CO and ELF3-6H3F proteins (**b**) were co-expressed. The experiments were repeated three times independently, and we obtained similar results.



Supplementary Figure 24. Functional epistasis between *CO* and *ELF3* on flowering regulation.

a, Flowering phenotypes of *elf3-1 co-101* mutants under LD and LD+FR+temp conditions. Total leaf numbers were counted when plants bolted. Data are from 16 individual plants. Each box is located between the upper and the lower quartiles, and the whiskers indicate 1.5-times interquartile ranges. The thick horizontal lines in the boxes represent the median, and open diamonds represent the mean. Outliers are indicated by circles. The same data from Fig. 2c was

used for *co-101*. (** $p < 0.01$, *** $p < 0.001$, ns: non-significant, , statistical information in Supplementary Table 3) **b-e**, Results of Q-PCR analysis show *FT* (**b, d**) and *TSF* (**c, e**) expression profiles under LD (**d, e**) and LD+FR+temp (**b, c**) conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).



Supplementary Figure 25. Gene expression profiles of *CO* and *FT* in *CO:HA-CO* and *CO:HA-CO/elf3-1* plants under LD+FR+temp conditions.

a, b, Results of Q-PCR analysis show *CO* (**a**) and *FT* (**b**) expression profiles in *CO:HA-CO* and *CO:HA-CO/elf3-1* plants grown under LD+FR+temp conditions. The results represent means $\pm$ SEM (n=3 biologically independent samples).

Supplementary Table 2. ELF3 and its interacting proteins identified by TAP-MS analysis

AGI number	Protein name	Molecular Weight	ELF3 TAP-MS ²⁾	Exclusive unique peptide count/Percent coverage ¹⁾					
				ELF3-3F6H in <i>phyB-9</i>		ELF3-3F6H in <i>phyA-211</i>		ELF3-3F6H	
				Rep1	Rep2	Rep1	Rep2	Rep1	Rep2
AT2G25930	ELF3	77 kDa		5/11%	5/11%	14/28%	13/26%	14/29%	11/23%
AT2G18790	phyB	129 kDa	Y	-	-	31/31%	26/27%	33/33%	31/31%
AT5G35840	phyC	124 kDa	Y	-	-	22/27%	15/18%	17/22%	22/29%
AT4G18130	phyE	123 kDa	Y	-	-	10/13%	4/5%	15/20%	13/18%
AT5G43630	TZP	91 kDa	Y	-	-	12/16%	9/14%	12/16%	11/16%
AT2G16365.2	PCH1 ³⁾	51 kDa ³⁾	Y	-	-	7/21%	5/16%	10/28%	7/20%
AT2G40080	ELF4	12 kDa	Y	-	-	3/32%	2/24%	3/32%	3/42%
AT2G32950	COP1	76 kDa	Y	-	-	-	1/2% ⁴⁾	3/5%	1/2% ⁴⁾
AT1G09570	phyA	125 kDa	Y	-	-	-	-	3/4%	-
AT5G18190	MLK1	77 kDa	Y	-	-	-	-	1/6% ^d	-
AT3G13670	MLK4	79 kDa	Y			-	-	1/5% ^d	-
AT1G47128	RD21a	51 kDa	Y	-	-	-	-	2/4%	-
AT5G06850	FTIP1	91 kDa	Y	3/7%	2/4%	-	-	2/4%	1/2% ⁴⁾
AT1G09340	CRB	43 kDa	Y	-	1/4% ⁴⁾	1/4% ⁴⁾	-	3/12%	3/12%
AT3G44310	NIT1	38 kDa	N	3/13%	6/28%	5/23%	4/18%	3/14%	1/6% ⁴⁾
AT4G35630	PSAT1	47 kDa	N	1/3% ⁴⁾	1/6% ⁴⁾	-	1/3% ^d	1/3% ⁴⁾	-
AT3G14420	GOX1	40 kDa	N	-	-	2/5%	2/3%	2/3%	-

¹⁾All listed proteins match 99% protein threshold, minimum number peptides of 2 and peptide threshold as 95%. Proteins not matching the criteria were marked with "-".

²⁾ELF3 TAP-MS³⁶ in the *ELF3:ELF3-3F6H* plants harvested at dusk was used for comparison. "Y" means present and "N" means absent in the dataset.

³⁾Molecular weight and percent coverage for PCH1 were calculated using protein encoded by At2g16365.2

⁴⁾Only one exclusive unique peptide was detected.