PNAS
www.pnas.org

Supplementary Information for

Mutations in *EID1* and *LNK2* caused light-conditional clock deceleration during tomato domestication

Niels A. Müller, Lei Zhang, Maarten Koornneef, José M. Jiménez-Gómez

Corresponding authors: José M. Jiménez-Gómez and Maarten Koornneef
Email: jose.jimenez-gomez@inra.fr or koornneef@mpipz.mpg.de

This PDF file includes:

SI Methods
Figs. S1 to S13
Tables S1 to S5
References for SI citations

Other supplementary materials for this manuscript include the following:

Datasets S1 to S4

www.pnas.org/cgi/doi/10.1073/pnas.1801862115

SI Methods

Growth conditions. All tomato seeds were treated with saturated tri-sodium phosphate (Na_3PO_4) for 15 minutes. After three washes, seeds were kept in water for three days in the dark. On the third day they were sown on standard soil in 7x7 cm square pots. We positioned genotypes based on a completely randomized design.
For leaf movements analyses, seedlings were entrained in a controlled environment chamber (Elbanton, Kerkdriel, Netherlands) for two to four days under cool white fluorescent tubes (~100 µmol m^{-2} s^{-1}) in 12:12 light/dark and 20:18 °C temperature cycles. On the last day of entrainment a polystyrene ball was attached to the tip of one cotyledon of each seedling using petroleum jelly (1). At the dark:light transition, we transferred the seedlings to an identical chamber set to constant light and temperature (25 °C and ~100 µmol m^{-} s^{-1}) and started the image capture (2).
For the RNA-seq and qRT-PCR time-courses, seedlings were entrained in an controlled environment chamber (Elbanton, Kerkdriel, Netherlands) for ten days under cool white fluorescent tubes (~100µmol m^{-2} s^{-1}) in 12:12 light/dark and 25:18 °C temperature cycles. For the constant light time-course (RNA-seq) lights were kept on constantly from the 11th day onwards and leaf samples were collected every 12 hours for 2.5 days starting at ZT12. For the two constant dark time-courses (RNA-seq and qRT-PCR) lights were kept off after the 10th night. For the RNA-seq time-course leaf samples were collected every 12 hours for 2.5 days starting at ZT12, for the qRT-PCR time-course leaf samples were collected every four hours for two days starting at ZT0.
For the phyB mutant RNA-seq experiment, including the *phyB1*, *phyB2*, *phyB1B2* and Moneymaker lines (3), seedlings were grown in an environmental chamber (Elbanton, Kerkdriel, Netherlands) for ten days under cool white fluorescent tubes (~100µmol m^{-2} s^{-1}) in 12:12 light/dark and 20:18 °C temperature cycles, and single leaves collected from three independent plants per genotype.

Genome comparison. To identify expressed genes that have been completely or partially deleted during tomato domestication, we took advantage of the high quality reference genome of the wild tomato species *S. pennellii* (4). We mapped RNA-seq reads from the wild species *S. pennellii* LA0716 and *S. pimpinellifolium* LA1589 and the *S. lycopersicum* M82 cultivar (5) against the genomes of the wild species *S. pennellii* and the Heinz 1706 cultivar (4, 6). First, we scanned the *S. pennellii* genome for 1000 bp windows in which all uniquely mapping RNA-seq reads are unmapped when using the cultivated tomato genome as reference. We did this for all three species, considering only windows with an RPKM value > 1. Then, genes annotated in the *S. pennellii* genome that overlapped any of these windows in both of the wild species (*S. pennellii* and *S. pimpinellifolium*) but not in the cultivar M82 were defined as deleted in cultivated tomato. The final list includes only 11 genes, half of them annotated as hypothetical proteins (Table S1).

Cloning of *LNK2*. We cloned the *LNK2* cDNA of *S. pimpinellifolium* LA1589 and *S. lycopersicum* cv. MM, including 290 bp of its 3' UTR, under control of the MM native promoter (~2.4 kb upstream sequence from the *LNK2* start codon) using the MultiSite Gateway® Pro 2.0 Kit (Life Technologies), the destination vector pGWB1 (7) and Phusion High-Fidelity DNA polymerase (New England Biolabs). The primers used for cloning are listed in Table S4. The two constructs were transformed into the cultivar Moneymaker using the *Agrobacterium thumefaciens* mediated leaf disc transformation method (8). T2 plants were analyzed for leaf movements and the presence of the transgene was validated via PCR. Only T2 populations exhibiting segregation ratios expected for a single transgene insertion event (1:3) were used for further analysis. Additionally, expression of the wild species allele was confirmed in the transgenics via qRT-PCR. One T2 population exhibited very low expression of the transgene and was therefore excluded (*SpiLNK2*_6 in Fig. S3).

Genotyping the deletions in *LNK2* and *EID1*. We genotyped the large deletion in *LNK2* and the 3 bp deletion in *EID1* in 426 tomato accessions for which short read re-sequencing data are available (9, 10).
To score the large deletion in *LNK2*, short reads were downloaded from NCBI's SRA and aligned to the *S. pennellii* genome reference sequence v2 (4) using Bowtie2 version 2-2.0.0-b5 with default parameters (11). Coverage for the region of *LNK2* (Sopen01g030520.2, Spenn-ch01:87,686,901-

87,790,400) was plotted using custom R scripts and the existence of the large deletion in *LNK2* was scored visually (Dataset S1 and S3). Accessions TS-226, TS-235, TS-107, TS-257 and TS-260 were not scored due to low coverage.
To score the 3 bp deletion in *EID1*, short reads were downloaded from NCBI's SRA and aligned to the *S. lycopersicum* reference genome sequence v2.50 using Bowtie2 version 2-2.0.0-b5 with default parameters (11). Duplicated reads were removed using Picard version 1.65 (http://picard.sourceforge.net), and indels realigned using GATK v2.2-8 (12). All alignments were scored for the presence or absence of the causative indel at position 66,893,249 in chromosome 9 using GATK v2.2-8's UnifiedGenotyper with default parameters (12) (Dataset S3). Accession with no reads overlapping the indel position (TS-195, TS-206, TS-232, TS2-94 and TS-96) as well as those for which the indel was heterozygous (LA0113, LA1324, TS-137, TS-319, TS-216, TS-238, TS-295, TS-300, TS-430, TS-431, TS-436, TS-71, TS-124) were removed from the analysis.

Classification of re-sequenced tomato accessions into phylogenetic groups. More than 1000 tomato accessions have been previously classified into phylogenetic groups using 8700 genome-wide SNPs genotyped with the SolCAP Infinium Chip (13). We integrated available re-sequencing data (9, 10) (aligned to the tomato genome reference version 2.50) with the previously classified accessions (13). We first translated the genomic positions of the SNPs in the SolCAP Infinium array into their positions in the *S. lycopersicum* reference genome version 2.50. For this, we used BLAST on each probe in the array and took the best hit as the most likely position of the probe in each reference genome. We removed probes that did not present a single best hit within 7 Mb of the expected position in the array annotation and those for which the reference genome sequence did not match the alleles present in the more than 1000 accessions reported (13). This method resulted in 8522 SNPs from the SNP array matching the reference genome sequence version 2.50. Then, short read alignments from the published re-sequencing data (9, 10) were genotyped for the SolCAP SNP set using the UnifiedGenotyper tool from GATK v2.2-8 (12). The resulting genotypes from all accessions in the three datasets were merged and processed using the following filters. First, we removed markers that exhibited non-reference allele frequencies in the cultivated tomatoes differing by more than 30% among the three datasets. Then, we removed markers and accessions that presented more than 10% missing data. Finally, we removed markers based on their positions, by choosing one marker every 0.1 centiMorgan based on a high-density genetic map (14). This filtering resulted in a final set 1412 tomato accessions genotyped for 1956 variants. These variants were used to estimate a phylogenetic tree with the neighbor-joining function in the Bioconductor's package ape (15) (Dataset S2). Based on their neighbors in the tree and the classification of the previously genotyped accessions (13), we classified 307 tomato accessions from the two re-sequencing datasets (9, 10) (Dataset S2 and S3).

Genetic interaction analyses between the circadian rhythm QTL and *PHYB1*. To test for genetic interactions between the circadian rhythm QTL *EID1* and *LNK2* and the *PHYB1* gene, we crossed the near isogenic lines (NILs) harboring the wildtype allele of either of the two QTL (*EID1* NIL = rec47 and *LNK2* NIL = BIL497), both described in detail before (16), with a *phyB1* mutant line (3). The resulting F1 lines were self-pollinated to generate segregating F2 populations. For the *EID1* x *phyB1* cross, F2 seedlings were genotyped for *EID1* and *PHYB1* using SNP markers (3, 16) (Table S5). For each of the four allelic combinations, three plants were selected and self-pollinated to obtain fixed F3 seeds, which were used for leaf movement analysis. *LNK2* and *PHYB1* are genetically linked. It was therefore not possible to generate F2 individuals fixed for the four allelic combinations. Instead F2 individuals fixed for one but segregating for the other locus were selected and self-pollinated. The resulting F3 populations were phenotyped for circadian leaf movements and subsequently genotyped for the segregating locus. The markers used for the genotyping are listed in Table S5.

RNA sequencing of *phyB* mutants. Total RNA from leaves from three independent plants per genotype (Moneymaker, *phyB1*, *phyB2* and *phyB1B2*) was extracted separately with the RNeasy Plant Mini Kit (QIAGEN). Libraries were prepared according to the Illumina TruSeq RNA protocol and sequenced on the Illumina HiSeq platform (Illumina, Inc., San Diego, CA) at the Genome Center of the Max Planck Institute for Plant Breeding Research Cologne. We obtained a total of 266 million 100bp single end reads (average 22,1 million, minimum of 19,8 million reads) that were aligned to the

tomato reference sequence v2.5 using HISAT2 v 2.1.0 (17) with a maximum intron length of 115400. An average of 96,9% of the reads were aligned to the reference. The number of reads per transcript was counted with custom R scripts based on the Bioconductor's packages Rsamtools and ShortRead (18). We surveyed the homogeneity of the samples with the PoissonDistance function in the Bioconductor PoiClaClu package (19). This analysis revealed one of the *phyB1B2* samples as an outlier, which was consequently removed from the following analysis (Fig. S13). Differential expression between each of the three phytochrome mutants and Moneymaker was calculated with the DEseq2 package in R (20). Genes with a q-value lower than 0.05 were considered as differentially expressed.

RNA sequencing time-course. Total RNA from leaf samples was extracted with the RNeasy Plant Mini Kit (QIAGEN). Libraries were prepared according to the Illumina TruSeq RNA protocol and sequenced on the Illumina HiSeq platform (Illumina, Inc., San Diego, CA) at the Genome Center of the Max Planck Institute for Plant Breeding Research Cologne. RNA-seq single end reads (48 libraries (2 species x 6 time points x 2 conditions x 2 biological replicates) adding up to 89 Gb of sequence) were aligned to the tomato reference sequence v2.5 using TopHat2 v2.0.6 (21) including the information from the annotation v2.4 and the following parameters --read-realign-edit-dist 0 -g 1 --no-coverage-search. The number of reads per transcript was counted with custom R scripts based on the Bioconductor's packages Rsamtools and ShortRead (18). Read counts were normalized using the rlog function of the DESeq2 package (20) and scaled by subtracting the mean and dividing by the standard deviation for each condition and genotype.
In order to select marker genes whose expression reflects the internal circadian clock, we made use of the high-resolution time-course RNA-seq data available for *S. lycopersicum* and *S. pennellii* (16). For this, reads were downloaded from NCBI's SRA BioProject PRJNA295848, aligned to the *S. lycopersicum* reference sequence v2.50 using Tophat2 v2.0.6 (21) with the following parameters: --max-insertion-length 12 --max-deletion-length 12 -g 1 --read-gap-length 12 --read-edit-dist 20 --read-mismatches 12 --no-coverage-search --read-realign-edit-dist 0 --segment-mismatches 3 --splice-mismatches 1.
The numbers of reads per transcript were quantified based on the *S. lycopersicum* ITAG annotation v2.4 using the same R script mentioned above. This experiment is divided in two parts: One day in diurnal conditions and two days in circadian conditions. For each part and species we normalized read counts using the rlog function in Bioconductor's DESeq2 package (20) and identified cycling transcripts using ARSER (22) as described previously (16). Shortly, we ran a modified version of ARSER that outputs period and phase estimates using exclusively the "mle"-method with a period window from 25 to 42 (default of 34) for *S. lycopersicum* under circadian conditions, from 13 to 26 (default of 20) for *S. pennellii* under circadian conditions, and from 20 to 28 (24 default) for both *S. lycopersicum* and *S. pennellii* under diurnal conditions. To run ARSER with the diurnal data, we generated an extra day of data by randomly choosing one replicate from every time-point and adding 24 hours to their collection time. For further analysis, we considered genes with q-values below 0.05 (fdr_BH < 0.05) in all four parts (*S. lycopersicum* diurnal, *S. lycopersicum* circadian, *S. pennellii* diurnal and *S. pennellii* circadian). This resulted in 1218 cycling genes that were grouped into bins based on the time of the day where their expression peaks as reported by ARSER for *S. pennellii* in circadian conditions (Dataset S4).
For each group of marker genes the average normalized and scaled expression ± s.e.m. was calculated and plotted to estimate the functioning of the circadian clock in each species and condition (Fig. S11).

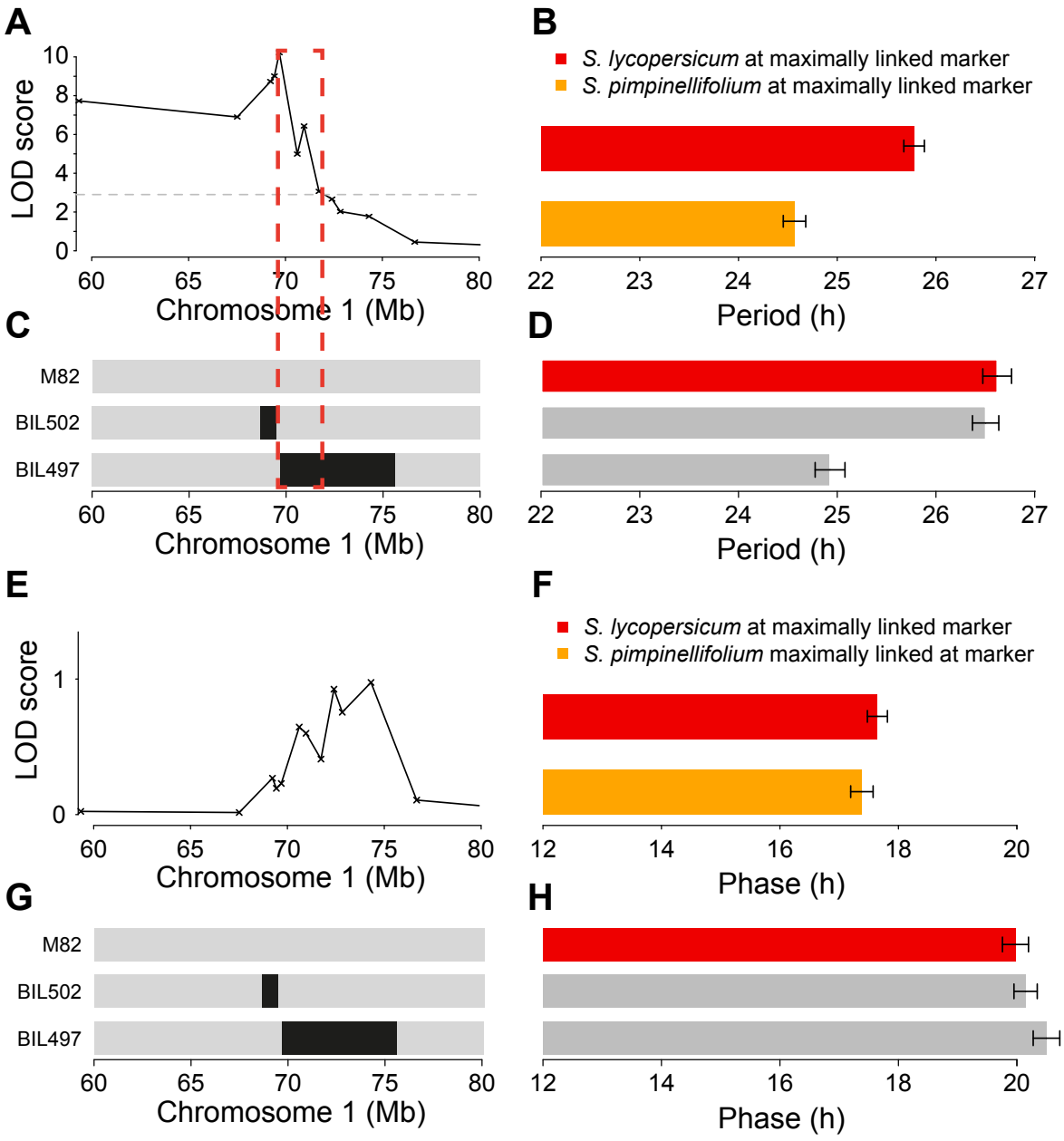


Fig. S1. Localization and effect of the circadian period QTL on chromosome 1. (*A*) and (*E*) Logarithm of the odds (LOD) scores for circadian period (*A*) and phase (*E*) in the *S. pimpinellifolium x S. lycopersicum* RIL population. The dashed horizontal line indicates the 5% significance threshold. (*B*) and (*F*) Mean period (*B*) and phase (*F*) ± s.e.m. of all RILs grouped by the maximally linked marker on chromosome 1. (*C*) and (*G*) Genotypic representation of *S. lycopersicum* cv. M82 and two BILs from the *S. pennellii* x *S. lycopersicum* BIL population harboring introgressions of the region of the period QTL. *S. lycopersicum* = gray, *S. pennellii* = black. (*D*) and (*H*) Mean period (*D*) and phase (*H*) estimates ± s.e.m. of the lines shown in (*C*) (n = 10-22). The red rectangle in (*A*) and (*C*) highlights the candidate region for the period QTL, delimited on the right by the 5% significance threshold on the *S. pimpinellifolium* x *S. lycopersicum* QTL and on the left by the border of the introgression of BIL497.

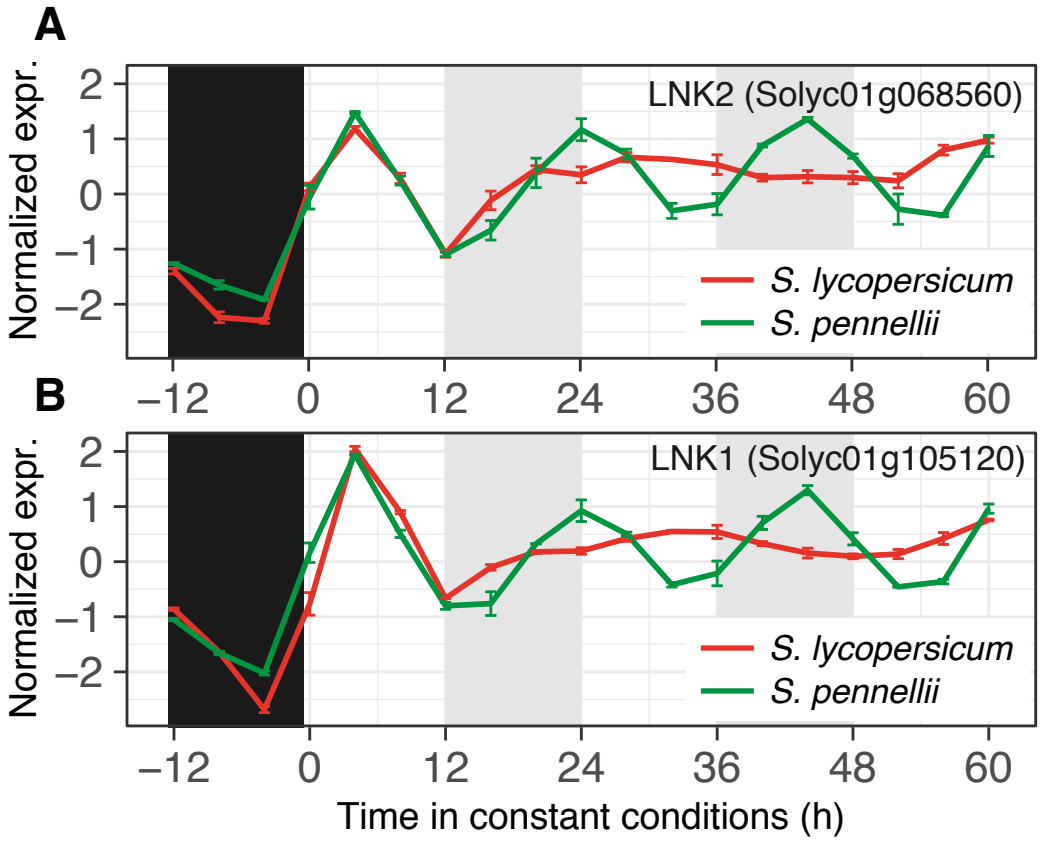


Fig. S2. Temporal expression patterns of *LNK2* and *LNK1* in cultivated tomato and its wild relative *S. pennellii*. Expression profiles of *LNK2* (Solyc01g068560) (*A*) and *LNK1* (Solyc01g105120) (*B*) were extracted from available transcriptomic data (SRA Bioproject PRJNA295848) (16) from *S. lycopersicum* cv. M82 (red) and *S. pennelii* LA0716 (green). Seedlings were grown for 6 days in diurnal conditions and then released into constant light and temperature. Samples were taken every four hours on the last day under diurnal conditions and the first two days under circadian conditions. Black, white and gray backgrounds represent night, day and subjective night conditions, respectively. Normalized read counts for each gene and species were scaled by subtracting the mean and dividing by the standard deviation. Each point represents the average of two biological replicates ± s.e.m.

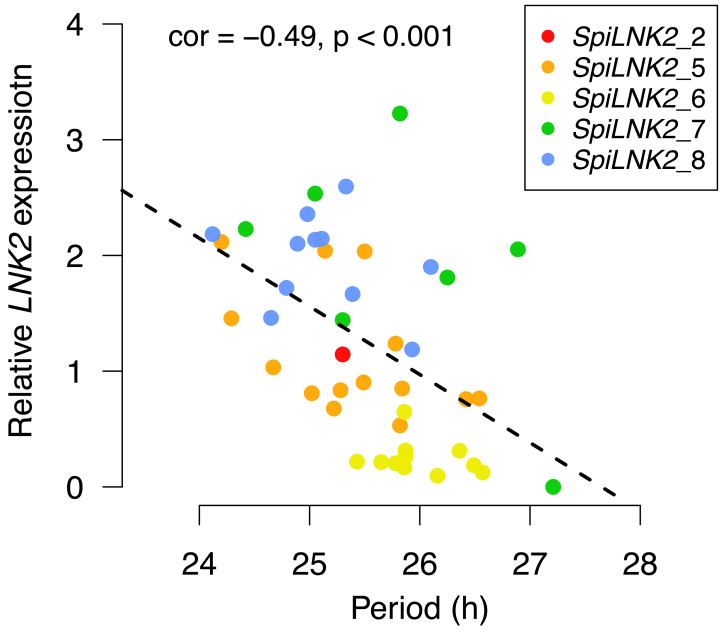


Fig. S3. Relationship between expression of the *LNK2* transgene and circadian period. Seedlings from one of the three independent experiments shown in Fig. 2 were tested for the expression of the *LNK2* transgene by qRT-PCR. *LNK2* expression is relative to the AP-2 complex subunit mu (Solyc08g006960, *CAC*). Relative *LNK2* expression of each seedling is plotted against its estimated period. The correlation coefficient (Pearson's *r*) and the p-value of the correlation test are shown in the top left corner. Different dot colors reflect different T2 families.

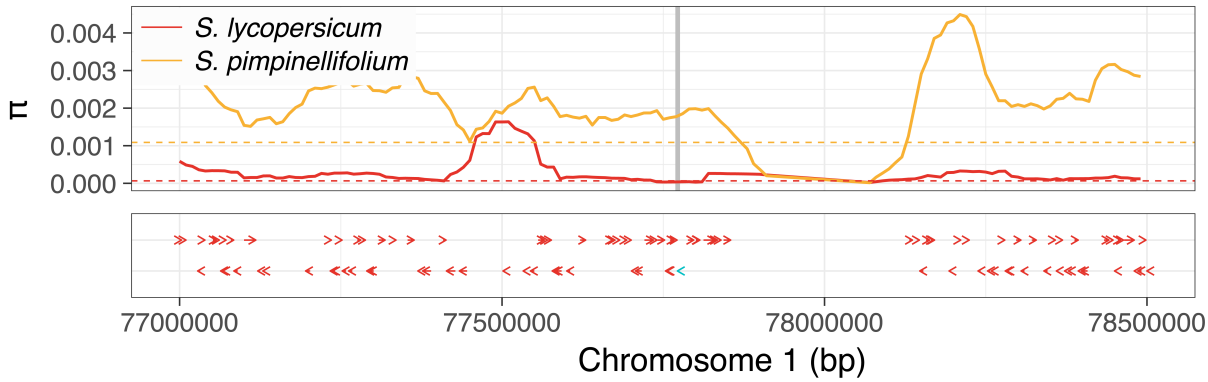


Fig. S4. Nucleotide diversity in *S. lycopersicum* and *S. pimpinellifolium* along the *LNK2* genomic region. Top: Nucleotide diversity (pi) along the *LNK2* genomic region for 144 re-sequenced *S. lycopersicum* (classified as SLL fresh, SLL_processing or SLL_vintage in Dataset S3) and for 32 *S. pimpinellifolium* accessions (classified as Spim Ecuador or Spim Peru in Dataset S3) (19). Pi was calculated in windows of 100 kb with 10 kb steps. Horizontal dashed lines mark the chromosome-wide threshold for the bottom 5% windows in each group. The vertical shadowed region represents the position of *LNK2* in the tomato reference genome. Bottom: Annotated genes in the region of *LNK2* (genome version v2.5). Each gene is represented by a red arrow. *LNK2* is colored in blue.

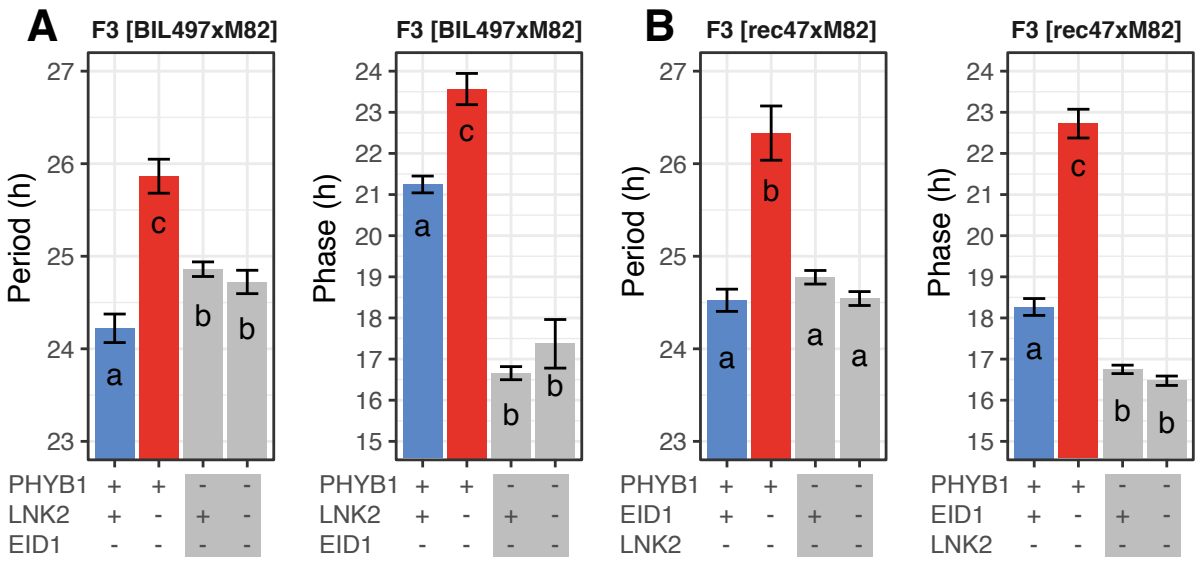


Fig. S5. The *phyB1* mutant is epistatic to the circadian period lengthening and phase delay caused by the mutated *lnk2* and *eid1* alleles. (*A*) Mean period (left) and phase (right) estimates (± s.e.m.) from leaf movements in constant light and temperature from F3 lines derived from a cross of the *LNK2* introgression line BIL497 (*S. pennellii* introgression in cv. M82 background) and a *phyB1* mutant (in the cv. MM background) (n = 13-23) (*B*) Mean period (left) and phase (right) estimates (± s.e.m.) from leaf movements in constant light and temperature from F3 lines derived from a cross of the *EID1* introgression line rec47 (*S. pennellii* introgression in cv. M82 background) and the *phyB1* mutant (n = 4-17). Different letters in each bar indicate significant differences ($P < 0.05$, two-way ANOVA and Tukey's *post hoc* HSD test).

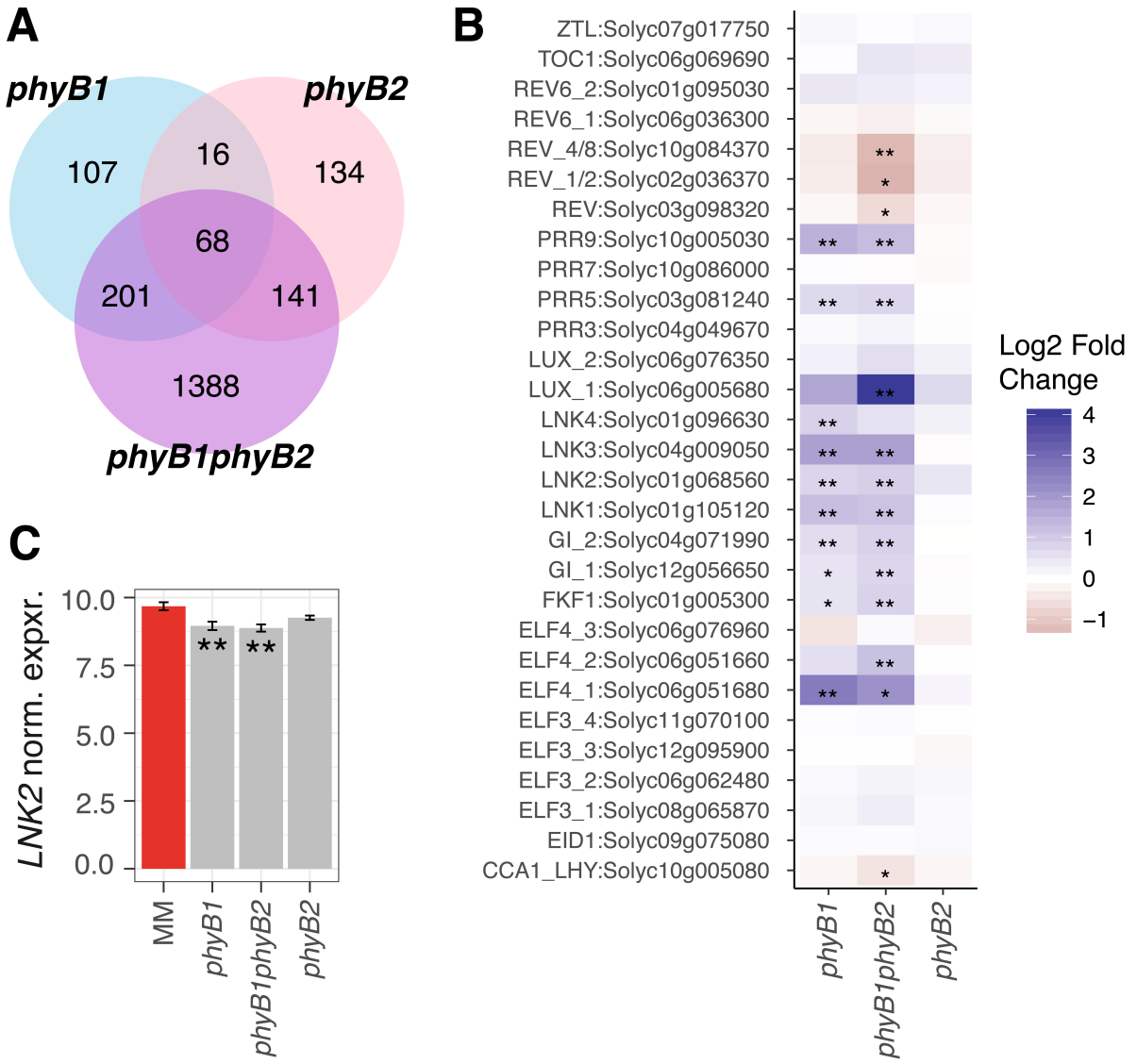


Fig. S6. RNA-seq of *phyB* mutants reveals effects of *PHYB1* on clock gene expression. (*A*) Venn diagram of differentially expressed genes between Moneymaker (MM) and three *phyB* mutant lines. (*B*) Heatmap representing the log2 fold change in expression between MM and the three *phyB* mutants for tomato homologs of Arabidopsis circadian clock genes. (*C*) rlog normalized expression of *LNK2*. Asterisks in (*B*) and (*C*) represent significant differences in expression between MM and the mutant lines (DESeq2, adjusted $P < 0.05$*, adjusted $P < 0.01$**).

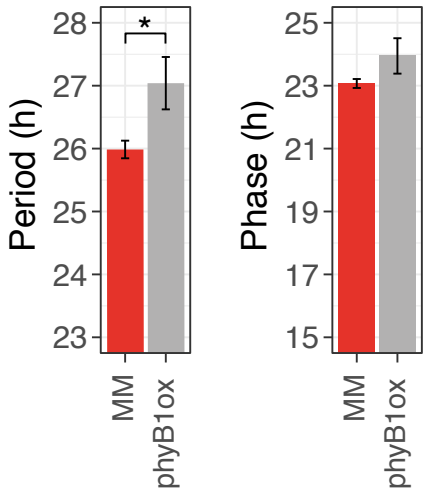
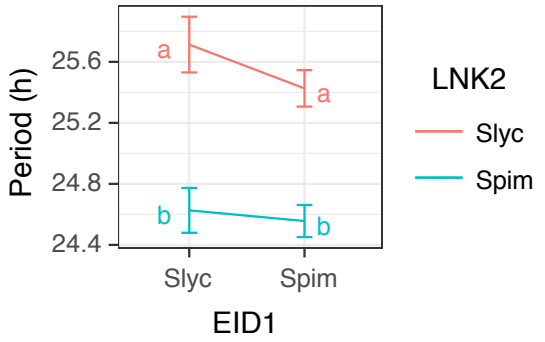


Fig. S7. Overexpression of *PHYB1* in cultivated tomato leads to a further circadian period lengthening. Mean period and phase estimates (± s.e.m.) from leaf movements in constant light and temperature of *S. lycopersicum* cv. Moneymaker (MM) and PHYB1ox lines (n > 5). Asterisks indicate significant differences ($P < 0.05$*, one-way ANOVA).

A

	Df	Sum.Sq	F.value	Pval
EID1	1	2.48	3.808	0.054
LNK2	1	20.883	32.073	1.8e-07 ***
EID1:LNK2	1	0.252	0.387	0.536
Residuals	89	57.949		



B

	Df	Sum.Sq	F.value	Pval
EID1	1	43.516	41.405	6.06e-09 ***
LNK2	1	3.762	3.58	0.062
EID1:LNK2	1	0.814	0.775	0.381
Residuals	89	93.537		

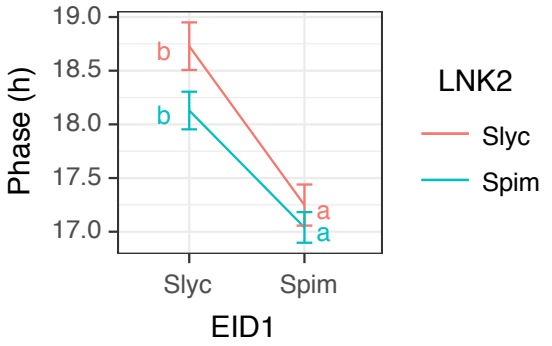


Fig. S8. *EID1* and *LNK2* do not genetically interact with each other. Ninety-three lines from the *S. lycopersicum* x *S. pimpinellifolium* RIL population were previously analyzed for leaf movements under constant light conditions (16). The genetic interaction between *EID1* and *LNK2* for period (*A*) and phase (*B*) was investigated by grouping the RILs by their genotype at the most closely linked marker. Tables show the results of a two-way ANOVA with the alleles of *LNK2* and *EID1* and their interaction as factors. Line plots show the average ± s.e.m. for each group (n = 17-36). Different letters indicate significant differences between genotypes ($P < 0.05$, two-way ANOVA and Tukey's *post hoc* HSD test).

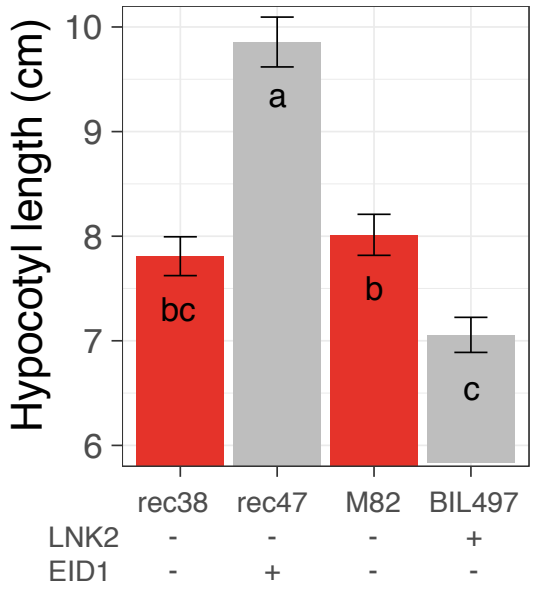


Fig. S9. The mutated *eid1* and *lnk2* alleles have opposite effects on hypocotyl growth. Tomato seedlings were germinated and grown in constant red light. Hypocotyls were measured 10 days after germination. Each bar represents the average ± s.e.m. (n = 20-23). Different letters in each bar indicate significant differences between genotypes ($P < 0.05$, two-way ANOVA and Tukey's *post hoc* HSD test). The left two bars represent NILs only differing for their allelic state of *EID1* and 12 neighboring genes and were described previously as 'rec38' and 'rec47' (16). The right two bars represent the cultivar M82 and the backcrossed inbred line 'BIL497'. 'BIL497' contains an introgression from *S. pennellii*, overlapping the chromosomal region of *LNK2*, in an otherwise M82 background. The exact borders of the introgression were reported previously (16).

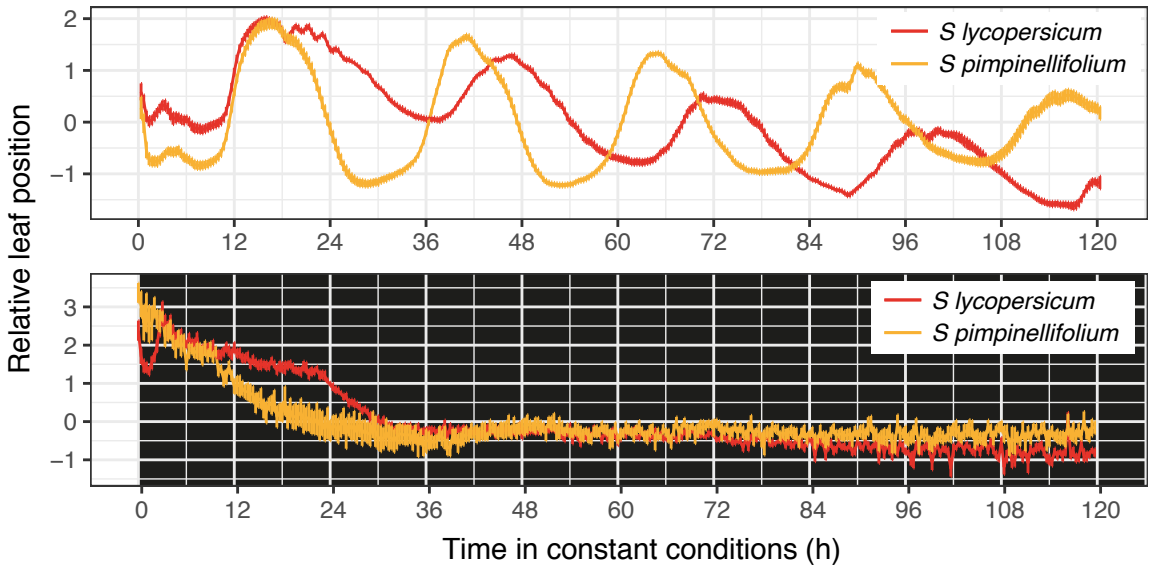


Fig. S10. Circadian leaf movements in constant light or dark. *S. lycopersicum* cv. MM and *S. pimpinellifolium* (LA1589) seedlings were entrained in 12 h light/12 h dark photoperiods for four days and released into constant light (top) or dark (bottom) conditions. Leaf movements were monitored every 20 minutes with point-and-shoot cameras and flashes of green light given at the time of image acquisition. Average relative leaf positions ± s.e.m. scaled to the mean and standard deviation of each experiment and genotype are represented on the y-axis (n = 10-31).

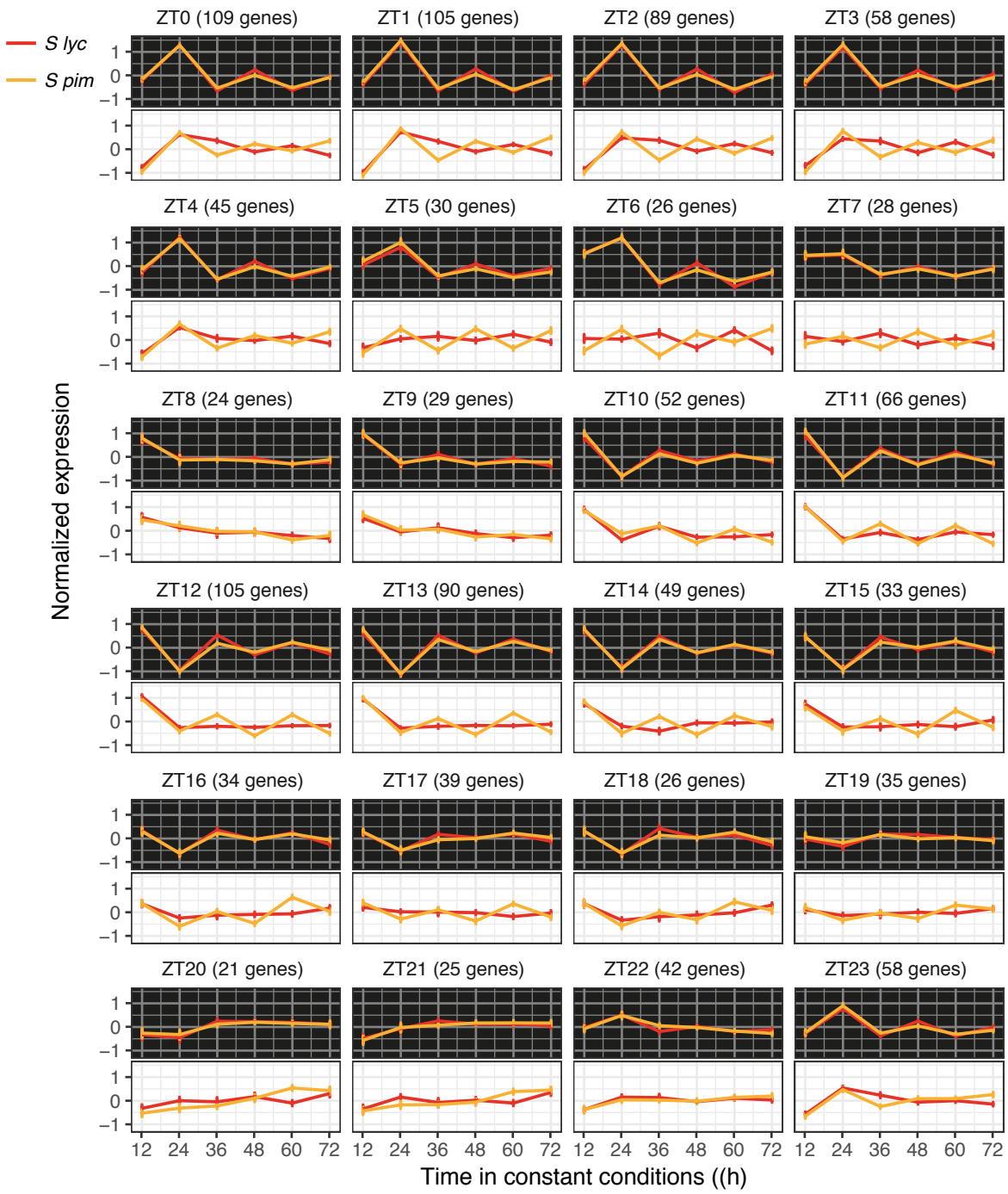


Fig. S11. RNA-seq time-course in constant light and darkness for *S. lycopersicum* and *S. pimpinellifolium* demonstrates light-conditionality. Genes were selected based on their rhythmic gene expression under diurnal and circadian conditions in tomato and grouped into 24 bins according to their time of maximal expression during the 24-hour daily cycle, relative to dawn (ZT0-ZT23). Samples for RNA-seq were collected every 12 hours for 3 days from *S. lycopersicum* cv. MM and *S. pimpinellifolium* (LA1589) seedlings in constant light or dark conditions. Normalized read counts for the genes in each bin were scaled to the mean and standard deviation and their average ± s.e.m. represented in the y-axis. Black or white background represents constant dark or light conditions, respectively.

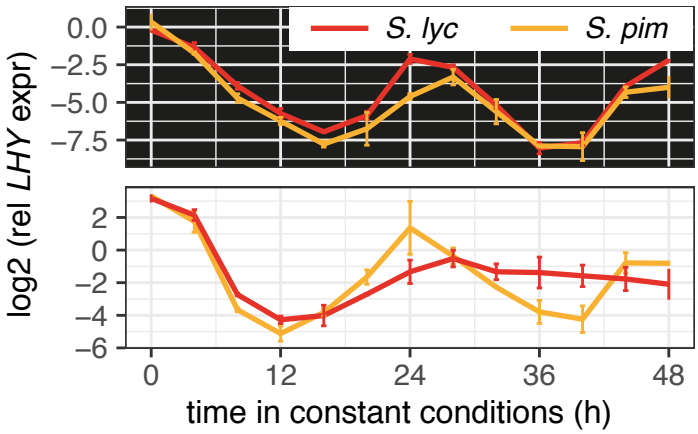


Fig. S12. Temporal gene expression of *LHY* in constant light and dark shows light-conditionality of the clock modulation in cultivated tomato. *LHY* (Solyc10g005080) expression in *S. lycopersicum* cv. MM and *S. pimpinellifolium* (LA1589) relative to the AP-2 complex subunit mu (Solyc08g006960, *CAC*) during two days in constant dark (top) or light (bottom). Plants were entrained in 12 h light/12 h dark photoperiods and released into constant conditions after 7 days. Points represent scaled averages ± s.e.m. for two biological replicates.

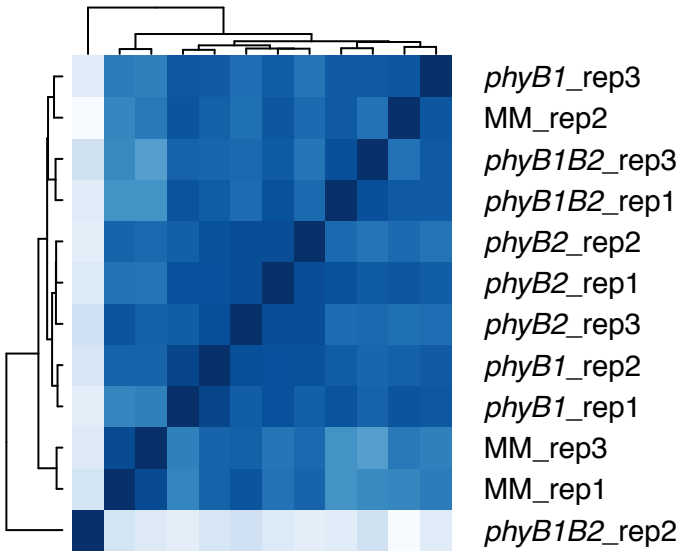


Fig. S13. Poisson distances between RNA-seq samples. The heatmap shows the Poisson dissimilarity matrix calculated from the RNA-seq experiment including *S. lycopersicum* cv. Moneymaker (MM) and the *phyB1*, *phyB2* and *phyB1B2* mutants. The heatmap was calculated using the PoissonDistance function in the Bioconductor's PoiClaClu package on the raw read gene counts for each sample.

Table S1. Genes deleted from the tomato genome but present in *S. pimpinellifolium* and *S. pennellii*.

Chr	Start	End	Gene ID	Functional description
ch01	69450900	69457108	Sopen01g022790.1	A member of ARF GTPase family. | BEST Arabidopsis thaliana protein match is: Ras-related small GTP-binding family protein
ch01	87729638	87746008	Sopen01g030520.1	molecular_function unknown | BEST Arabidopsis thaliana protein match is: dentin sialophosphoprotein-related .
ch02	59632183	59635916	Sopen02g039260.1	Involved in a SNM-dependent recombinational repair process of oxidatively induced DNA damage. SENSITIVE TO NITROGEN MUSTARD 1 (SNM1) | BEST Arabidopsis thaliana protein match is: sterile alpha motif (SAM) domain-containing protein
ch04	73372112	73375059	Sopen04g031670.1	hypothetical protein
ch06	552235	556391	Sopen06g001450.1	Encodes AtOEP16, involved in plastid import of protochlorophyllide oxidoreductase A | BEST Arabidopsis thaliana protein match is: Mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein
ch08	127637	129527	Sopen08g001170.1	hypothetical protein
ch09	311484	318754	Sopen09g001390.1	encodes 3-phosphoshikimate 1-carboxyvinyltransferase / 5-enolpyruvylshikimate-3-phosphate / EPSP synthase. BEST Arabidopsis thaliana protein match is: RNA 3-terminal phosphate cyclase/enolpyruvate transferase, alpha/beta
ch09	2779112	2783016	Sopen09g004060.1	hypothetical protein
ch09	2779125	2783048	Sopen09g004070.1	hypothetical protein
ch10	58576863	58579393	Sopen10g021230.1	hypothetical protein
ch10	61861448	61871038	Sopen10g022170.1	GIGANTEA (GI)

Table S2. BLAST results for tomato LNK2 against Arabidopsis. Results from blasting the *S. pennellii* LNK2 protein (Sopen01g030520) against all Arabidopsis proteins (TAIR10) using BLASTP v2.2.24.

Gene id	Gene symbol	Score	E value
AT3G54500.1	LNK2	295	9,00E-80
AT3G54500.3	LNK2	260	3,00E-69
AT3G54500.2	LNK2	244	2,00E-64
AT3G54500.4	LNK2	243	2,00E-64
AT5G64170.1	.	82	1,00E-15
AT5G64170.2	.	82	1,00E-15
AT3G12320.1	.	47	3,00E-05
AT5G06980.3	.	35	0.14
AT5G06980.4	.	35	0.15
AT5G06980.2	.	34	0.42
AT3G48610.1	NPC6	33	0.77
AT5G06980.1	.	32	1.1
AT2G03070.1	MED8	29	8.2
AT3G09930.1	.	29	9.4

Table S3. Primers used for the qRT-PCR experiments. Primers span exon-exon junctions in order to only amplify cDNA synthesized from mRNA. Primers were tested to have the same efficiency for *S. lycopersicum* and *S. pimpinellifolium*.

Gene	**Left primer**	**Right primer**
LNK2 (Solyc01g068560)	GCACACATTCTGGTCCTTGA	TTGGAGACACAAGCCTTCCT
LHY (Solyc10g005080)	TTTACAAAGTTAGAAAAGGAGGCTCT	TAAGTTCCTTTCCTCACAGATGG
CAC (Solyc08g006960)	CCTCCGTTGTGATGTAACTGG	ATTGGTGGAAAGTAACATCATCG

Table S4. Primers used for cloning of *LNK2*. The colored parts of the primers represent the specific *att*-sites needed for the recombination reactions when using the MultiSite Gateway® Pro 2.0 Kit.

	Species	**Primer**	**Sequence**
***LNK2* cDNA**	*S. lycopersicum* + *S. pimpinellifolium*	fwd	GGGGACAACTTTGTATACAAAAGTTGCGGCGGAAGTTTCTCTGTG
		rev	GGGGACCACTTTGTACAAGAAAGCTGGGTACATAGGGAACACATGGAGCAT
***LNK2* promoter**	*S. lycopersicum*	fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTAGTTGCACAAGTCCTCAGTTCAG
		rev	GGGGACAACTTTTGTATACAAAGTTGCGCCGGAGATTCACATAGTAA

Table S5. Primers used for genotyping. For CAPS markers the enzyme needed to digest the amplified sequence is indicated. The expected fragment size for mutant (Mut.) and wildtype (WT) alleles are given in base pairs.

Line	**Left primer**	**Right primer**	**Enzyme**	**Mut. / WT (bp)**
EID1 NIL	AACCCACAGTTATTACCAAAGCTC	ATGGCAACCTACAATGATACACC	EcoRI	197+126 / 323
LNK2 NIL	GCACAGATTGTACACAAACCAAA	GTTTGGGGAAAATACGTACCAGT	n/a	369 / 312
phyB1 mut.	CTAAAATTCAAAGAGGAGGTCAgATT	GAAGGGGTAAAAAGGGTCCTAA	HinfI	192 / 166+26

References

1. Salathia N, Lynn JR, Millar AJ, & King GJ (2007) Detection and resolution of genetic loci affecting circadian period in *Brassica oleracea*. *Theor Appl Genet* 114(4):683-692.
2. Müller NA & Jimenez-Gomez JM (2016) Analysis of circadian leaf movements. *Methods Mol Biol* 1398:71-79.
3. Weller JL, Schreuder ME, Smith H, Koornneef M, & Kendrick RE (2000) Physiological interactions of phytochromes A, B1 and B2 in the control of development in tomato. *Plant J* 24(3):345-356.
4. Bolger A, et al. (2014) The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nat Genet* 46(9):1034-1038.
5. Koenig D, et al. (2013) Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc Natl Acad Sci USA* 110(28):E2655-2662.
6. Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485(7400):635-641.
7. Nakagawa T, et al. (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J Biosci Bioeng* 104(1):34-41.
8. McCormick S, et al. (1986) Leaf disc transformation of cultivated tomato (*L. esculentum*) using *Agrobacterium tumefaciens*. *Plant Cell Rep* 5(2):81-84.
9. Aflitos S, et al. (2014) Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *Plant J* 80(1):136-148.
10. Lin T, et al. (2014) Genomic analyses provide insights into the history of tomato breeding. *Nat Genet* 46(11):1220-1226.
11. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357-359.
12. DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491-498.
13. Blanca J, et al. (2015) Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics* 16:257.
14. Sim SC, et al. (2012) Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS One* 7(7):e40563.
15. Paradis E, Claude J, & Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20(2):289-290.
16. Müller NA, et al. (2016) Domestication selected for deceleration of the circadian clock in cultivated tomato. *Nat Genet* 48(1):89-93.
17. Kim D, Langmead B, & Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12(4):357-360.
18. Morgan M, et al. (2009) ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* 25(19):2607-2608.
19. Witten DM (2011) Classification and clustering of sequencing data using a Poisson model. *Ann Appl Stat* 5(4):2493-2518.
20. Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.
21. Kim D, et al. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14(4):R36.
22. Yang R & Su Z (2010) Analyzing circadian expression data by harmonic regression based on autoregressive spectral estimation. *Bioinformatics* 26(12):i168-174.