Supporting Information

Hosokawa et al. 10.1073/pnas.1019612108

SI Materials and Methods

Bacterial Strains. *Synechococcus elongatus* PCC 7942 and its derivatives used in this study are listed in Table S1. NUC42 (WT; ref. 1), NUC43 (*kaiABC*-null; ref. 1), IILC357 (*rpoD5*-null), ILC48 (*digA::flag*), ILC56 (*hspA::flag*), *kaiC*^{S157P} (ref 2.), and ILC504 (*rpoD5*-null;*kaiC*^{S157P}) were used, each containing a bacterial luciferase reporter gene set, *luxAB*, fused to the *kaiBC* promoter (P$_{kaiBC}$::*luxAB*), with the chloramphenicol-resistance gene at Neutral Site I in the genome. ILC357 was generated by the transformation of NUC42 with plasmid pIL470 carrying the downstream and upstream sequences of the *rpoD5* gene, with a kanamycin-resistance gene (*km*^{r}) in the pGEM-T Easy Vector (Promega) background, to substitute each σ factor gene with *km*^{r}. The complete segregation of the null mutant strains was confirmed by Southern blotting and PCR analyses. The loss of each σ factor mRNA was confirmed by Northern blotting analysis. The FLAG epitope-tagged strains ILC48 and ILC56 were generated by the insertion of the FLAG-encoding sequence (5′-GACTACAAGGACGATGACGACAAG-3′) at the 3′ terminus of the ORFs of *digA* and *hspA*, respectively (for the induction of the C-terminal FLAG-tagged protein) and *km*^{r}. ILC504 was generated by the transformation of *kaiC*^{S157P} with pIL470.

DNA Microarray Analysis. Total RNA was subjected to microarray analysis using an Affymetrix GeneChip designed on the basis of the *Synechococcus* genome, and the resulting raw signals (deposited in the NCBI GEO database GSE22468) were divided according to their genomic DNA-normalized values, as described previously (3). For each series of the time course experiment, the expression level x of gene g in strain s at t h after transfer to dark conditions in the nth experiment was then further normalized as:

$$x(g,s,t,n) = x(g,s,t,n)/Z(s,n),$$

where x represents the genomic DNA-normalized expression of gene g in strain $s \in \{WT, kaiABCnull\}$ at t h after transfer to dark conditions by the procedure described above, $n \in \{1,2\}$, and a normalization factor $Z(s,n)$ was used so that the total RNA signal at hour 12 in the light ($t = 0$) was deemed to be 1,000, as follows:

$$Z(s,n) = \sum_g (\gamma(g,s,0,n))/1000.$$

The values for the RNA signals given hereafter are these normalized values. The procedure for the normalization of the time course data is shown schematically in Fig. S6.

Identifying Dark-Induced Genes. We identified the expression profiles of the WT and *kaiABC*-null strains that were significantly induced under the dark conditions from among 2,515 *Synechococcus* ORFs. "Dark-induced" genes are defined in this study as those genes that are expressed more abundantly in the WT or *kaiABC*-null strains at at least one time point under the dark conditions than at hour 12 in the light (= dark onset). We statistically tested whether there is at least one point in which the average of *x(g, s, 0, 1)* and *x(g, s, 0, 2)* is higher than the average of *x(g, s, t, 1)* and *x(g, s, t, 2)*, using the MulCom test, a modified Dunnett's multiple comparison test (4). The two parameters in MulCom test, m and t, were deterrmined by *mulOpt function* in the Bioconductor package (version 1.2.0), so that the optimal combination of m and t was chosen to obtain the maximum number of differentially induced genes satisfying a false discovery rate threshold of <0.005. The parameters and t value of the modified Dunnett's test are described in Dataset S1. In this way, 198 "dark-induced" genes were extracted. Note that the transition of the index for dark induction is continuous between the dark-induced and noninduced genes, and there is no clear cutoff point that distinguishes the two groups. Therefore, different thresholds can increase or reduce the number in each group.

Evaluating the *kai* Dependence of Dark Induction. We next identified the genes regulated by *kai* genes from these 198 dark-induced genes by comparing their expression profiles in the WT and *kaiABC*-null strains. Two values were used to evaluate *kai* dependence: the correlation between the time courses of the two strains (hereafter, *Cor*), and the statistical difference between the temporal averages of the two expression profiles (hereafter, *Avg-Pval*). The former value reflects the waveform of the expression profile, and the latter value evaluates the difference in the absolute expression levels. We defined *Cor* for gene g, as follows:

$$Cor = \frac{1}{4} \sum_{i,j \in \{1,2\}} Correlation(\mathbf{x}(g, WT, i), \mathbf{x}(g, kaiABCnull, j)),$$

where x(g,s,n) = [x(g,s,0.5,n), x(g,s,1,n), x(g,s,2,n)…x(g,s,12,n)] and "*Correlation*" represents Pearson's correlation function.

Avg-Pval was defined as the P value calculated with a t test when the control groups, including *Average*(**x**(g,*WT*,1)) and *Average*(**x**(g, *WT*,2)), were compared with the sample groups, including *Average* (**x**(g,*kaiABCnull*,1)) and *Average*(**x**(g, *kaiABCnull*,2)). If the expression profiles of a gene satisfied *Cor* of >0.754 (corresponding to the upper-limit value for the decorrelation test; $P < 0.05$) and *Avg-Pval* of greater than the threshold 0.05, we defined it as *kai*-independent; if not, we regarded it as *kai*-dependent. Each value of *Cor* or *Avg-Pval* for the 198 dark-induced genes is shown in Dataset S1.Thus, 167 genes were categorized as "*kai*-dependent" dark-induced genes with this method. Again, the transition of the index used to estimate *kai* dependence was continuous between the *kai*-dependent and *kai*-independent genes. Therefore, different threshold values can increase or reduce the number of *kai*-dependent genes.

1. Nishimura H, et al. (2002) Mutations in KaiA, a clock protein, extend the period of circadian rhythm in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Microbiology* 148:2903–2909.
2. Nakajima M, et al. (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* 308:414–415.
3. Ito H, et al. (2009) Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proc Natl Acad Sci USA* 106:14168–14173.
4. Isalla, C (2011) Mulcom. Available at http://www.bioconductor.org/packages/2.8/bioc/html/Mulcom.html. Accessed May 26, 2011.

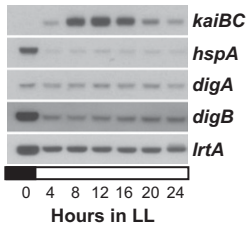

Fig. S1. Expression of dark-induced genes under continuous light (LL). Expression profiles of *kaiBC* and four dark-induced genes under LL conditions after two 12 h/12 h light/dark cycles.

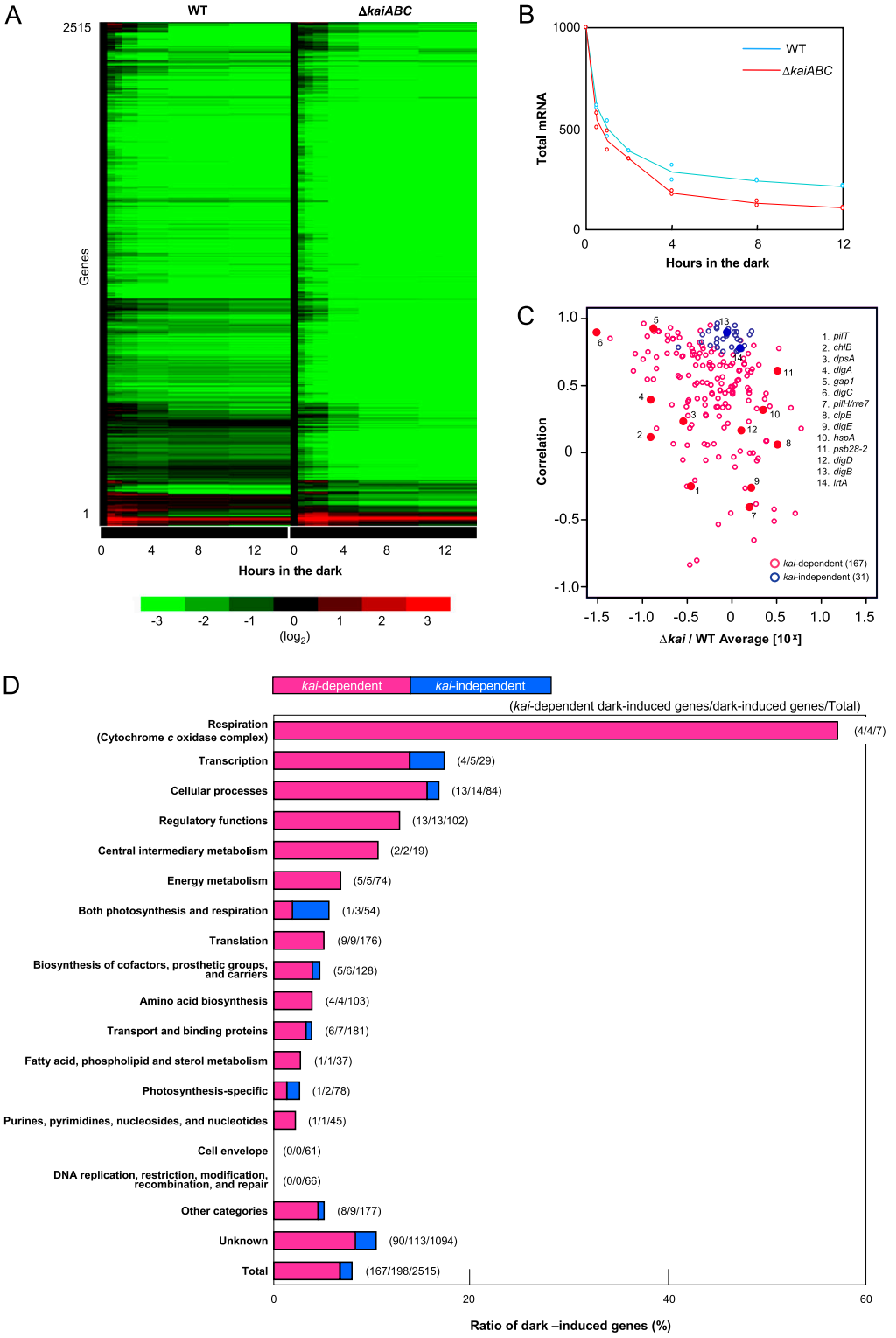

Fig. S2. Effect of *kaiABC* nullification on global expression profiles in the dark. (*A*) Genome-wide expression profiles in the dark conditions in WT and *kaiABC*-null mutant strains. Data were normalized to the value at hour 0 in the dark (= hour 12 in the light). (*B*) Total mRNA accumulation levels estimated from the sum of the mRNA hybridization signals normalized to the genomic DNA signals under the dark in the WT (blue) and *kaiABC*-null (red) strains. The results of two independent experiments (dots) and their average values (line) are shown for each strain. The maximal total signals at hour 0 in the dark (corresponding to hour 12 in the light) were deemed to be 1,000. (*C*) Identification of the *kai*-regulated genes among the dark-induced genes. The 198 dark-induced genes were classified into either "*kai*-dependent" or "*kai*-independent" group (*SI Materials and Methods*). The scatter plots indicate the distribution of the correlation value and the expression ratio (mutant/WT) for the 167 *kai*-dependent genes (red) and 31 *kai*-independent genes (blue). The filled circles indicate the 14 dark-induced genes tested in this study. (*D*) The *kai*-dependent (red) and *kai*-independent (blue) dark-induced genes were classified according to their functional categories. The subpopulation of genes relative to the population of total genes in each category is shown. The numbers of *kai*-dependent dark-induced genes, dark-induced genes, and total genes in each category are shown in parentheses.

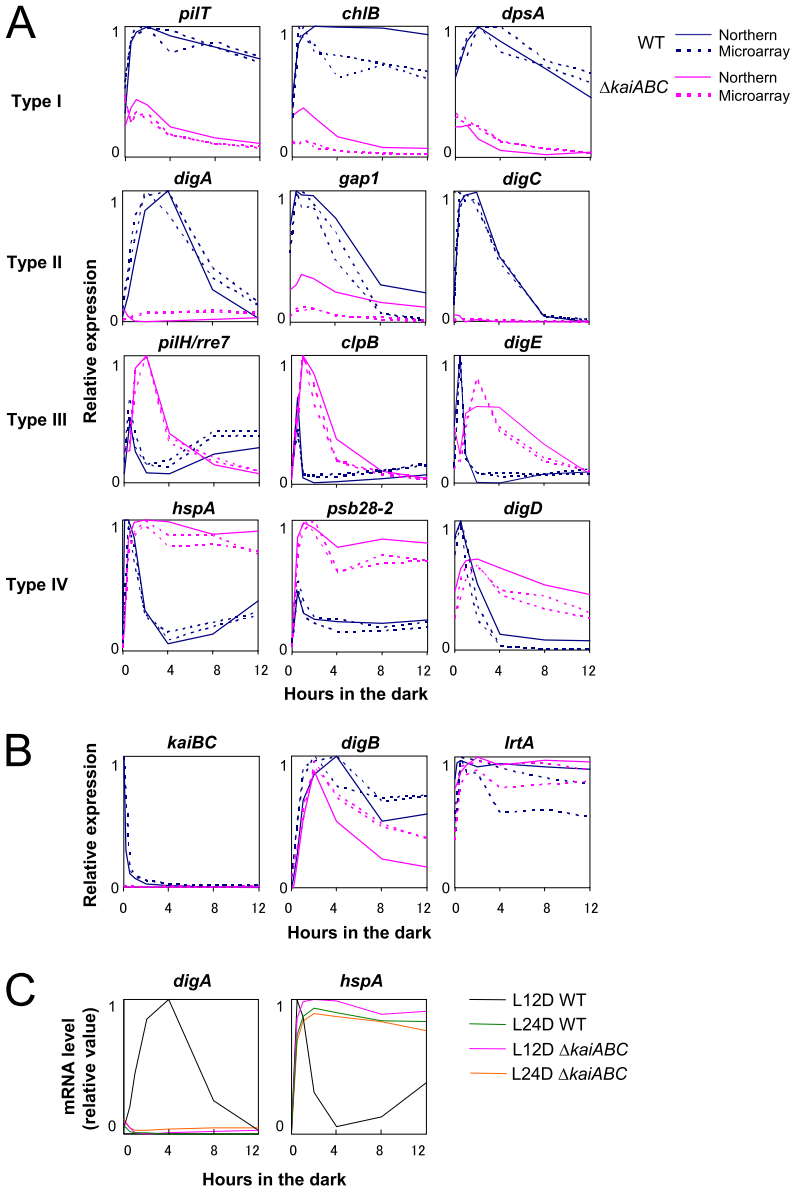

Fig. S3. Validation of the microarray analysis and the circadian gating of dark-induced gene expression. (*A* and *B*) Validation of the microarray data (dotted lines, $n = 2$) by Northern blotting analysis (solid lines) of the expression profiles of type I–IV genes (*A*) and the *kaiBC*, *digB*, and *lrtA* genes (*B*) in the dark. Blue and red lines indicate the expression profiles of the genes in the WT and *kaiABC*-null strains, respectively. Data were normalized to each maximum value in the Northern blot/DNA microarray data for either the WT or the *kaiABC*-null mutant strain. (*C*) Day-length dependence of dark-induced profiles of the *digA* and *hspA* genes. When the cells were transferred to the dark after 12 h in the light, the expression profile of each gene was quite different in the WT (L12D WT; black line) and *kaiABC*-null mutant strains (L12D Δ*kaiABC*; red line). When the WT and *kaiABC*-null mutant cells were transferred to the dark after 24 h in the light, their expression profiles became essentially similar in the WT (L24D WT; green line) and *kaiABC*-null mutant strains (L24D Δ*kaiABC*; orange line), suggesting that the Kai-based clock system gates the timing of dark-induced gene expression during subjective dusk.

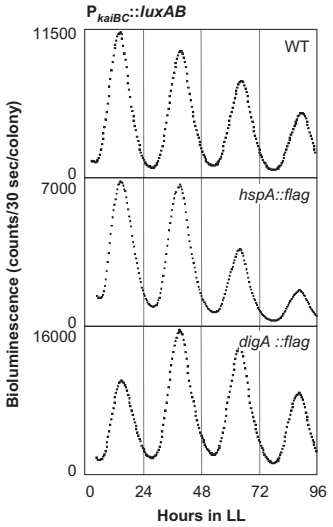

Fig. S4. Circadian *kaiBC* promoter activity rhythm monitored with a bioluminescent reporter in the FLAG-tagged strains. Bioluminescence rhythms were monitored under LL conditions after two 12 h/12 h light/dark cycles in the WT, *hspA::flag*, and *digA::flag* strains.

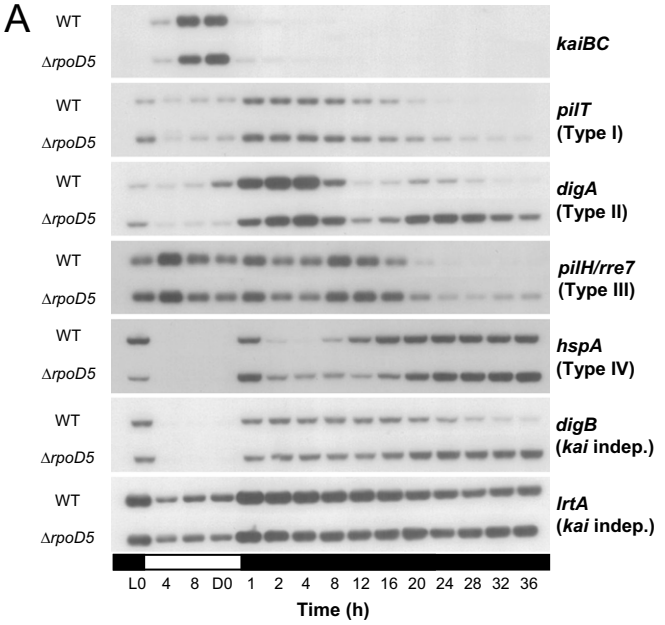

Fig. S5. RpoD5 modulates the expression profiles of dark-induced genes under continuous dark conditions. (*A*) Northern hybridization profiles for the expression of the *kaiBC* operon and representative dark-induced genes are shown in the WT and *rpoD5*-null strains. (*B*) Circadian *kaiBC* promoter activity rhythms under continuous light (LL) monitored by bioluminescence reporters in the WT, *kaiC*^{S157P}, and *rpoD5*-null (Δ*rpoD5*) strains and the *kaiC*^{S157P};Δ*rpoD5* double mutant strain.

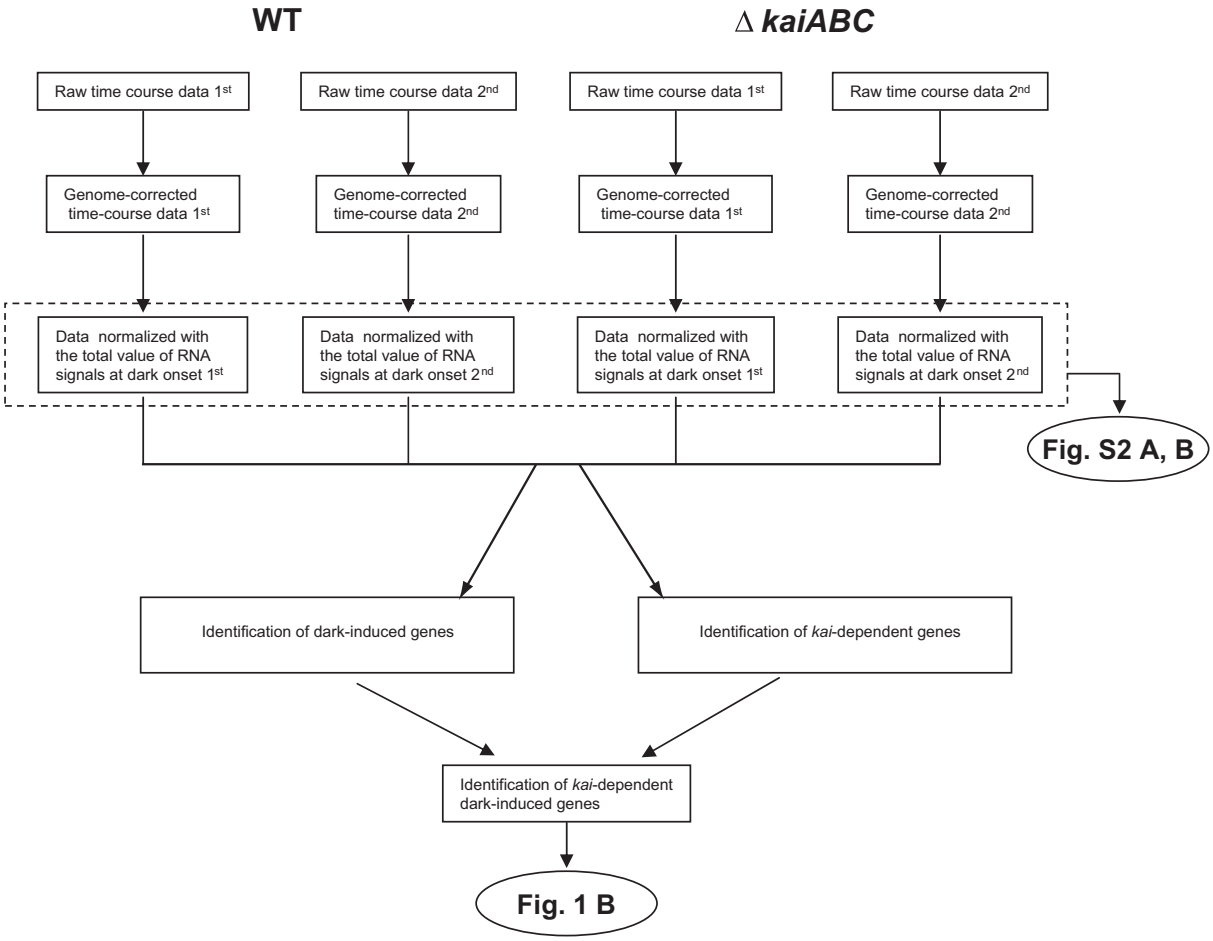

Fig. S6. Schematic diagram of data processing. Details are described in *Materials and Methods* (main text) and *SI Materials and Methods*.

Table S1. ***Synechococcus kaiBC*-reporter strains used in this study**

Strain	DNA used for transformation	Host strain	Reporter cassette	Marker*	Description	References
NUC42	—	—	P_{kaiBC}::*luxAB*::Cm^r	Cm	Wild-type	1
NUC43	—	—	P_{kaiBC}::*luxAB*::Cm^r	Cm, Km	*kaiABC*-null	1
ILC48	pIL468 (*syc2427_c*::*flag*::Km^r)	NUC42	P_{kaiBC}::*luxAB*::Cm^r	Cm, Km	*digA*::*flag*	This study
ILC56	pIL469 (*syc1704_d*::*flag*::Km^r)	NUC42	P_{kaiBC}::*luxAB*::Cm^r	Cm, Km	*hspA*::*flag*	This study
ILC357	pIL470 (Δ*syc2245_c* Km^r)	NUC42	P_{kaiBC}::*luxAB*::Cm^r	Cm, Km	*rpoD5*-null	This study
$kaiC^{S157P}$	—	—	P_{kaiBC}::*luxAB*::Cm^r	Cm, Sp	Short period (21 h)	2
ILC504	pIL470 (Δ*syc2245_c* Km^r)	$kaiC^{S157P}$	P_{kaiBC}::*luxAB*::Cm^r	Cm, Km, Sp	*rpoD5*-null; $kaiC^{S157P}$	This study

*Cm, Km, and Sp represent chloramphenicol-, kanamycin-, and spectinomycin- resistant genes, respectively.

1. Nishimura H, et al. (2002) Mutations in KaiA, a clock protein, extend the period of circadian rhythm in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Microbiology* 148: 2903–2909.
2. Nakajima M, et al. (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* 308:414–415.

Other Supporting Information Files

Dataset S1 (XLS)