

# A new promoter element associated with daily time keeping in *Drosophila*

Brandi Sharp<sup>1</sup>, Eric Paquet<sup>2</sup>, Felix Naef<sup>2</sup>, Akanksha Bafna<sup>3</sup> and Herman Wijnen<sup>1,3,\*</sup>

<sup>1</sup>Department of Biology, University of Virginia, Charlottesville, VA 22903, USA, <sup>2</sup>Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland and <sup>3</sup>Biological Sciences and Institute for Life Sciences, University of Southampton, Southampton SO17 1BJ, UK

Received November 07, 2016; Revised March 06, 2017; Editorial Decision March 27, 2017; Accepted April 11, 2017

## ABSTRACT

**Circadian clocks are autonomous daily timekeeping mechanisms that allow organisms to adapt to environmental rhythms as well as temporally organize biological functions. Clock-controlled timekeeping involves extensive regulation of rhythmic gene expression. To date, relatively few clock-associated promoter elements have been identified and characterized. In an unbiased search of core clock gene promoters from 12 species of *Drosophila*, we discovered a 29-bp consensus sequence that we designated as the Clock-Associated Transcriptional Activation Cassette or ‘CATAC’. To experimentally address the spatiotemporal expression information associated with this element, we generated constructs with four separate native CATAC elements upstream of a basal promoter driving expression of either the yeast *Gal4* or firefly *Luciferase* reporter genes. Reporter assays showed that presence of wild-type, but not mutated CATAC elements, imparted increased expression levels as well as rhythmic regulation. Part of the CATAC consensus sequence resembles the E-box binding site for the core circadian transcription factor CLOCK/CYCLE (CLK/CYC), and CATAC-mediated expression rhythms are lost in the presence of null mutations in either *cyc* or the gene encoding the CLK/CYC inhibitor, *period* (*per*). Nevertheless, our results indicate that CATAC’s enhancer function persists in the absence of CLK/CYC. Thus, CATAC represents a novel *cis*-regulatory element encoding clock-controlled regulation.**

## INTRODUCTION

The circadian clock of higher eukaryotes is understood to be a conserved transcription/translation auto-regulatory feedback mechanism controlled via rhythmic transcriptional activation and repression. Mammals as well as in-

sects have circadian clocks that operate via interlocking transcriptional feedback loops that rely on heterodimers of basic helix–loop–helix (bHLH), Per-Arnt-Sim (PAS) domain transcription factors (TFs) for transcriptional activation (1). In the clock circuit of the fruit fly, *Drosophila melanogaster*, a heterodimer consisting of CLOCK (CLK) and CYCLE (CYC), acts as the core transcription factor (2,3). CLK/CYC binds DNA by associating with a canonical CACGTG E-box or E-box-like sequences (4–6). The detection of transcripts regulated by CLK/CYC is of particular interest to the field of circadian biology. E-boxes, however, are poor predictors of potential clock-regulated genes because the sequence motif is widespread throughout the fly genome. The nucleotides flanking the E-box and/or an arrangement of closely spaced associated motifs, on the other hand, likely contribute to clock transcriptional activity, TF specificity and increased binding affinity (7,8).

In *Drosophila*, studies of the promoter regions of known clock genes, *period* (*per*) and *timeless* (*tim*), found that E-box-dependent enhancers are necessary for circadian transcriptional modulation. The *per* promoter has, arguably, the best studied circadian enhancer motif to date. The enhancer is a 69-bp sequence upstream of the transcription start site (TSS) that activates circadian gene expression (5). This enhancer relies on an E-box motif to mediate transcriptional activation and 3′ sequences near the E-box to drive strong amplitudes and tissue specific expression (5,9). However, the core of the 69-bp enhancer is actually an E-box-bearing 18-bp region that is able, when multimerized, to recapitulate *per* spatial and temporal expression (4). The *tim* promoter possesses closely spaced E- and TER (*tim* E-box-like repeats) boxes, which are a variant of the consensus E-box sequence (6). This *tim* enhancer relies on two non-canonical E-boxes, TER1 and TER2, to strongly initiate gene expression and cycling amplitude. The canonical E-box present in the enhancer appears to be non-functional on its own and seems to require the TER boxes to elicit functionality.

Increasing knowledge of the *Drosophila* genome, annotated and cloned full-length cDNAs, transcription start sites (TSSs) and transcription factor binding sites (TFBSs) have provided more tools to determine the elements re-

\*To whom correspondence should be addressed. Tel: +44 23 80594336; Fax: +44 23 8059 5159; Email: hw3a11@soton.ac.uk

sponsible for transcriptional regulation (10,11). In previous work (12), we identified a conserved motif consisting of two closely spaced E-box-like elements in the 69-bp *per* enhancer as well as in other CLK/CYC-controlled genes (*tim*, *vriille*, *Par domain protein 1* and *clockwork orange*). As a continuation of this work, we, hereby, identified a second novel and independent 29-bp motif. The present study describes spatiotemporal expression information contributed by the latter motif, which was named Clock-Associated Transcriptional Activation Cassette or ‘CATAAC’.

## MATERIALS AND METHODS

### Identification of CATAAC consensus sequence

The bioinformatics model used to predict the ‘CATAAC’ element is identical to that previously described for the discovery of the E1–E2 element (12). Briefly, we used MEME (13) to scan the genomic sequences in windows of 2.5 kb around the TSSs of *tim*, *vriille*, *Pdp1* and *cwo* (*stich1*) in 12 *Drosophila* species (MultiZ alignments were downloaded from UCSC, dm3 assembly). This identified a conserved set of CATAAC sequences (Supplementary Figure S1). In order to obtain our final position weight matrix (Figure 1A), we trained a Hidden Markov Model as in (12), but consisting only of a single 29 bp motif, using the MEME-derived consensus as a seed. A .bed file that identifies all CATAAC sites with a bit score >5 in the *Drosophila melanogaster* genome (BDGP R5/dm3) is included in the supplemental information.

### Generating CATAAC reporter constructs

Synthetic oligonucleotides were designed to match four separate, native occurrences of the CATAAC element from either the *Pdp1* or *Slob* promoters: *Slob* 4xCATAAC enhancer, 5'pGGCCGATAACGCGGCGTATGCGCAA TGTCGAAGCATATTACGCATACGCCCCATCCGC; 5'pGTTTGTGTGGCGGATGGGGCGTATGCGTAA TATGCTTCGACATTGCGCATACGCGCGGTTATC; 5'pCACAGCAAACGCTGCGTATGCGTAATACTTT GTGCACACGTTGCGTATGAGTAATGTCCT; 5'pA ATTAGGACATTACTCATACGCAACGTGTGCAC AAAGTATTACGCATACGCAGC, *Slob* mt4xCATAAC enhancer, 5'pGGCCGATATCCCGGCTTTGCCCTA TGTCGAAGCATATAAGGCAAGGCCGAACCGC; 5'pGATTGCTGTGGCGGTTCCGGGCTTTGCCTTA TATGCTTCGACATAGGGCAAAGGCCGGGATATC; 5'pCACAGCAATCCCTGCCTTTGCCT TATACT TTGTGCACTCCTTTGCCTTTGACTTATGTCCT; 5'pAATTAGGACATAAGTCAAAGGCAAGGAGT GCACAAAGTATAAGGCAAAGGCAGG, *Pdp1* 4xCATAAC enhancer, 5'pGGCCAGCACATTACGCA TACGTCACGTGTTGCAAAAATAACGTATAC GACGCGTGTC; 5'pTGGTACAGAAGACACGC GTCGTATACGTATTATTTTGC AACACGTGAC GTATGCGTAATGTGCT; 5'pTTCTGTACCATGC GCGTATGAGCAATCTGTTAATACGTTACC CATAACGCCCCGTGGGC C; 5'pAATTGGCCAC GGGGCGTATGGGTAACGTATTAACAGATTG CTCATACGCCGCA, *Pdp1* mt4xCATAAC enhancer, 5'pGGCCAGCACATAAGGCAAAGGTCAGGAGT

TGCAAAAATAAAAAGGTAAAGGACGGGAGTC; 5'pAGGTACAGAAGACTCCCGTCCCTTTACCTTT ATTTTTGCAACTCCTGACCTTTGCCTTATGTGCT; 5'pTTCTGTACCTTC CGGCCTTTGACCTATCTG TTAATACGTA AGCCAAAGGCCCGGAGGGCC; 5'pA ATTGGCCCTCCGGGCTTTGGCTTAC GTATTAACAGATAGGTCAAAGGCCGGA, *Pdp1* 4xCATAAC with *Slob*-like E-boxes enhancer, 5'pGGCCAGCACATTACGCATACGTAACGTGT TGCAAAAATAACGTATACGAAGCGTTTC; 5'pTCGTACAGA AGAAACGCTTCGTATACGTA TTATTTTGC AACACGTTACGTATGCGTAATGT GCT; 5'pTTCTGTACGATGGGGCGTATGAGCAAT CTGTTAATACGTTACCCATACGCCGCGTTGGC C; 5'pAATTGGCCAACGCGGCGTATGGGTAACGT ATTAACAG ATTGCTCATACGCCCA. Oligos were annealed to match their order and orientation in the native *Pdp1* or *Slob* promoters. Flanking *EagI* and *EcoRI* sequences allowed for the 4xCATAAC enhancers to be inserted upstream of the *hsp70* basal promoter in the pPTGAL vector (14) (obtained from the *Drosophila* Genomics Resource Center). The luciferase constructs (pPT<sub>luc</sub>) were created by blunt-end cloning to replace a 3193 bp fragment containing *Gal4* flanked by *PstI* sites with a 1955 bp *HindIII*–*BamHI* fragment containing the *luciferase* gene from pGL3-Basic (Promega) (15). Copies of pPT<sub>luc</sub> can be requested online ([http://www.addgene.org/Herman\\_Wijnen/](http://www.addgene.org/Herman_Wijnen/)). For each reporter construct, several independent lines were generated.

### *Drosophila* stocks

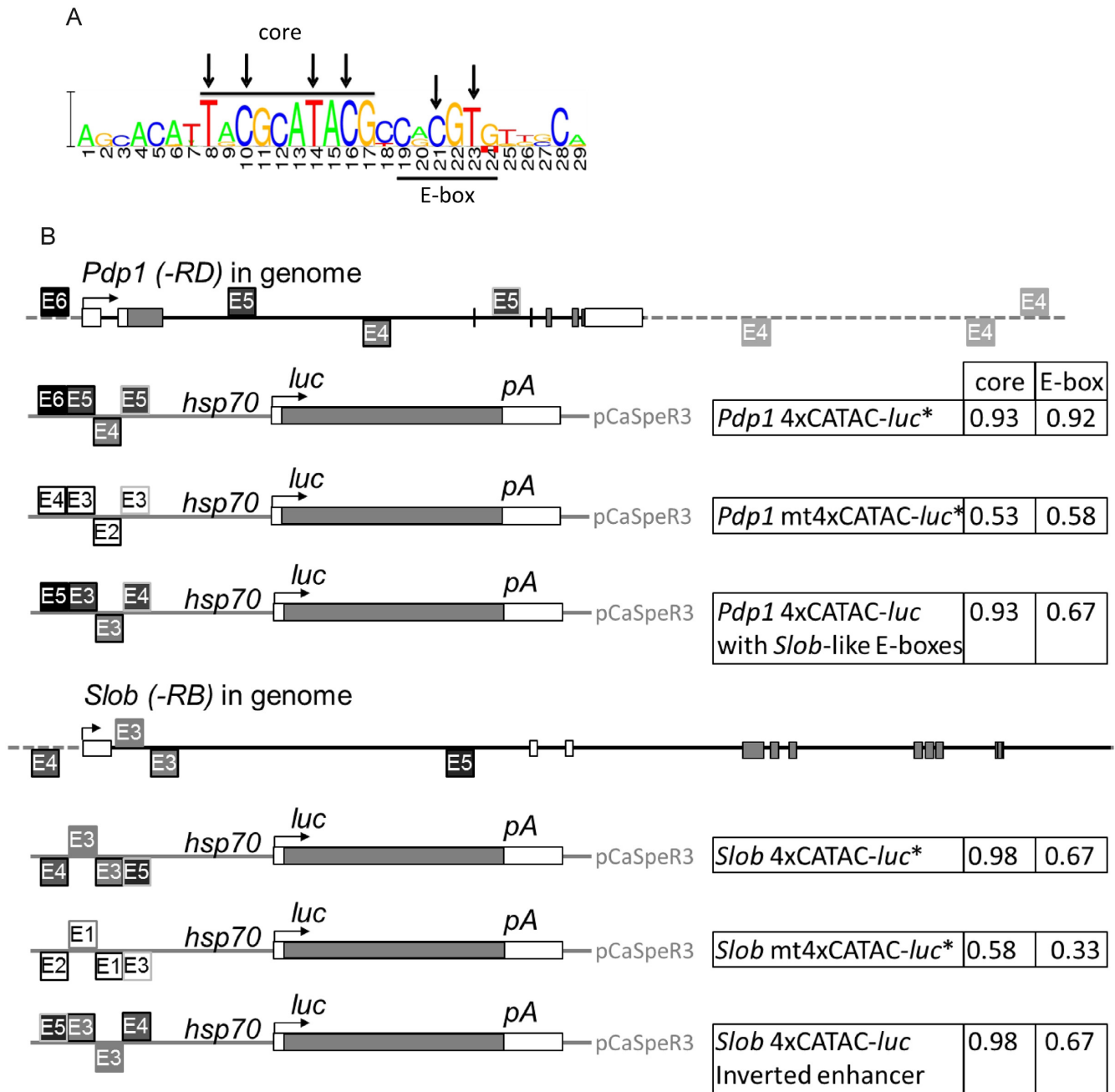
Fly stocks bearing the *w<sup>1118</sup>*, *UAS-CD8::GFP*, and *eya<sup>2</sup>* mutations were obtained from the Bloomington *Drosophila* Stock Center. The *cyc<sup>01</sup>*, *per<sup>01</sup>* and *tim-luc* alleles have been described previously (3,16,17). All stocks were maintained on standard yeast cornmeal agar food.

### *In vivo* luciferase monitoring

Bioluminescence monitoring of flies was carried out as described previously (17–19). 100  $\mu$ l of a 5% sucrose 1% agar solution containing 15 mM luciferin (GOLDBIO) was added to every other well of a white 96-well microtiter plate (Optiplate, Perkin Elmer). Flies were entrained for 3 days in a 12-h light, 12-h dark cycle (LD 12:12). On day 3 during the light phase, flies were anesthetized, added to separate wells and covered with clear plastic domes to reduce noise caused by fly movement closer to and farther away from the photodetector. Plates were then placed into a Top-Count Scintillation Counter and subjected to the remaining portion of the LD phase and subsequent constant darkness (DD). Luminescence from each fly was monitored for 7–17 s per time point, thereby allowing data for each fly to be collected roughly once every hour.

### Quantitative data analysis

Luciferase assay data were analyzed for period in the circadian range (15–35 h) and relative amplitude error (RAE) by an iterative, coupled fast Fourier transform-non-linear



**Figure 1.** Clock-associated transcriptional activation cassette (CATAC) and reporter constructs. **(A)** Consensus sequence for the CATAC element derived from five clock genes (*tim*, *per*, *cwo*, *vri* and *Pdp1*) from 12 *Drosophila* species. The highly conserved core region is overlined and the E-box sequence match is underlined. Residues mutated in this study are marked with arrows. **(B)** CATAC-*luciferase* reporter constructs used in this research. Reporters were generated by inserting CATAC sequences from the *Pdp1* or *Slob* genes upstream of the basal promoter in the pP<sub>T</sub>*luc* vector. Individual CATAC elements are indicated as E1/E2/E3/E4/E5/E6 depending on the match of their E-box region to the CACGTG consensus (1–6 matching residues). The native CATAC elements used for the *Pdp1* and *Slob* CATAC constructs are also shown in their genomic context with the *Pdp1*-RD and *Slob*-RB transcripts, respectively (the patterns of shading and outline color allow matching between transgenic elements and their genomic origin). The mt4xCATAC constructs contain six (A ↔ T or C ↔ G) transversion mutations in each of the four elements at the positions indicated in (A). ‘*Pdp1* 4xCATAC-*luc* with *Slob*-like E-boxes’ (PSE) was created by changing the E-box consensus matches of the original elements as indicated. The ‘core’ and ‘E-box’ scores for each construct represent the fraction of residues in each of the four elements matching the respective TRCGCATACG and CRCGTG consensus sequences from (A). For constructs marked with an asterisk, additional versions carrying *Gal4* instead of *luc* as the reporter were generated. Sequences are specified in Materials and Methods.



least squares (FFT-NLLS) multicomponent cosine analysis (19). RAE is the ratio of the 95% confidence interval by the amplitude estimate—the ratio of amplitude error to most probable amplitude. Lower RAE values suggest stronger rhythms. In our analysis, individual flies with RAE values  $<0.7$  were considered rhythmic,  $\geq 0.7$  were considered weakly rhythmic, and those for which the program returned no data (RAE  $> 1$ ), arrhythmic. The first 24-h of data were excluded from the plots and quantitative analysis since flies placed on fresh luciferin media require time to inactivate previously synthesized luciferase. For each fly, only DD data were analyzed for statistical analysis.

### Imaging of GFP and luciferase reporters

Larvae and adult flies producing 4xCATAC-*Gal4* driven expression of membrane-tethered green fluorescent protein were dissected in Ringer's solution (20) under a dissecting scope equipped with UV light. Tissues expressing GFP were wet or dry mounted and immediately imaged with a fluorescence microscope equipped with a CCD camera. Luciferase expressing flies were dehydrated for 24-h then fed on a cotton plug soaked in 200  $\mu$ l of a 1% sucrose solution containing 15 mM luciferin (GOLDBIO). Feeding occurred for 1–2 h prior to imaging of male heads and for 25–26 h prior to imaging of male bodies. Luciferase-expressing adult fly heads or bodies were mounted on a sterile filter insert (Millicell-CM; Millipore Inc.) and immobilized under a 13-mm coverglass with the help of sterile vacuum grease (21). The insert was then placed in a sterile glass-bottom dish (FluoroDish FD35PDL) containing insect tissue culture media with 0.1 mM luciferin. Sterile vacuum grease was then applied to seal the dish. Luminescence imaging was conducted as previously described by Sellix *et al.* (21), with an inverted epifluorescence microscope (Olympus CKX-41 equipped with a cooled intensified CCD camera (Mega10Z; Stanford Photonics Inc., Palo Alto, CA, USA) housed in a light-tight wooden dark box. Luminescence images were collected using Piper image analysis software (Stanford Photonics).

### Quantitative reverse transcriptase PCR (qRT-PCR) analyses

qRT-PCR expression analyses were carried out as described previously (22,23). *Slob* or *Pdp1* 4xCATAC-*luciferase* flies were entrained to 12 h light/12 h dark cycles at 25°C prior to release into constant conditions (DD 25°C). Flies were harvested onto dry ice at time points CT0, CT6, CT12, CT18 and CT24, and adult heads were dissected on a chilled platform and transferred to guanidinium thiocyanate buffer. Four separate groups of flies (~50 each) were used for each experimental condition. Total RNA was obtained from the heads using the RNAqueous4PCR kit (Ambion). Sample concentration and purity was analysed using a NanoDrop spectrophotometer. Samples exhibited OD 260/280 ratios between 1.8 and 2.1. Concentrations were adjusted to 25 ng/ $\mu$ l in 10 mM Tris-HCl 0.1 mM EDTA pH 8.0 buffer and samples were frozen at  $-80^{\circ}\text{C}$  in aliquots until further use. The RNA samples were then analyzed with the SensiFAST SYBR No-ROX One-Step qPCR Kit (Bioline) using experimental primer pairs designed to specifically amplify fragments of the circadian *Pdp1* or *Slob* transcripts, the

transgenic *luciferase* transcript or the *EF1 $\beta$*  control transcript (see Supplementary Table S1). All primers and amplicons used in this study have been described before (22,23). Oligonucleotides were sourced from Integrated DNA Technologies (25 nmol scale, standard desalting purification). No-template and RNase-treated controls were included to avoid false positive results. The following thermocycling protocol was used for the *Pdp1* and *Slob* amplicons: 60°C 180 s, 95°C 300 s, [95°C 15 s, 62°C 30 s]  $\times 45$ , 40°C 60 s, melting curve 60°C to 95°C 0.2°C/s. Thermocycling for *luciferase* and *EF1 $\beta$*  made use of an adjusted annealing temperature of 60°C instead of 62°C during the 45 amplification phase. Expression levels measured on a SmartCycler system (Cepheid) relative to *EF1 $\beta$*  were determined using the comparative Cycle threshold (Ct) method (24) and analysis of timed gene expression was carried out after normalization to the time course average. Amplicon sizes and specificity were verified by 2% agarose gel electrophoresis of all experimental and control PCR samples. Statistical analyses were conducted using SPSS.

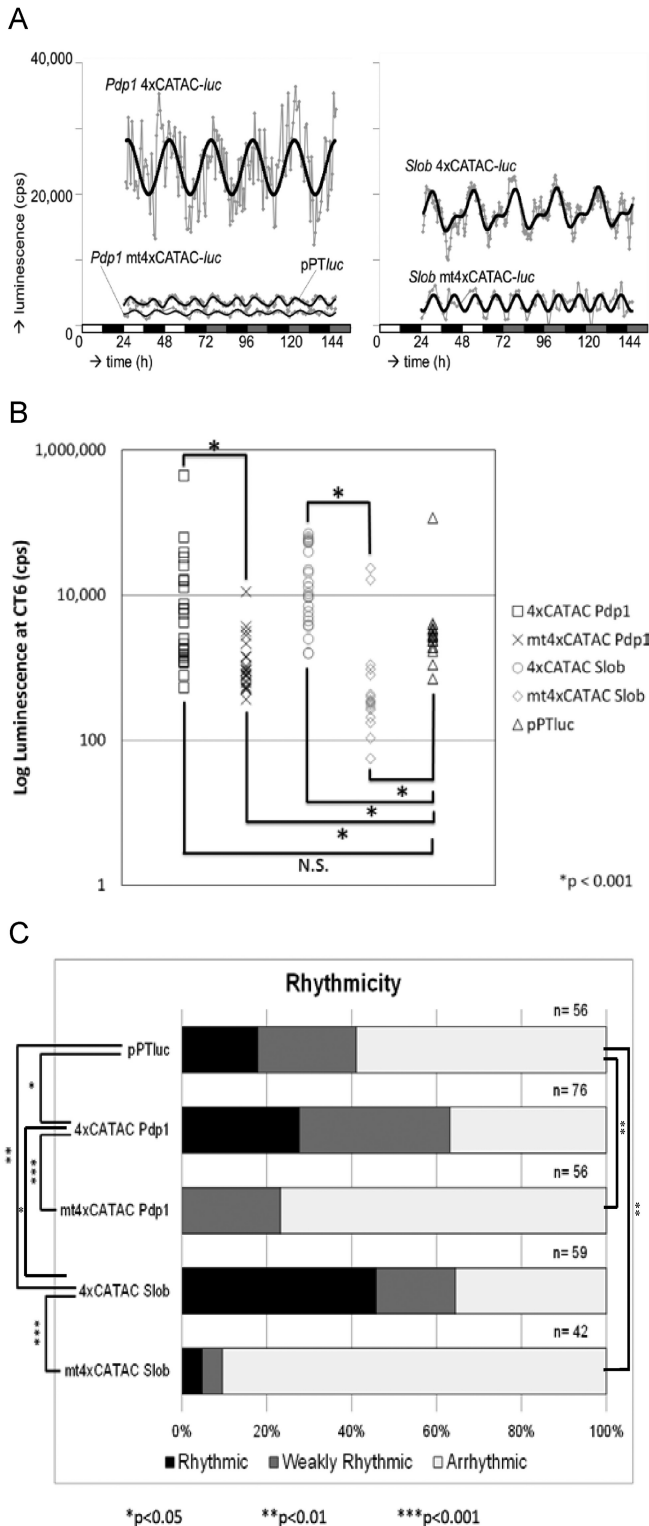
## RESULTS

### Identification of a conserved promoter element in *Drosophila*

In order to detect over-represented *cis*-acting elements associated with clock gene expression, we combined MEME (13) with a Hidden Markov Model ((12) and Materials and Methods) to analyse an alignment of promoter sequences of the core clock genes (*per*, *tim*, *Pdp1*, *cwo* and *vri*) for 12 species of *Drosophila*. As a result, a 29-bp motif designated the Clock-Associated Transcriptional Activation Cassette or 'CATAC,' was identified as over-represented in these promoters (Figure 1A, Supplementary Figure S1). The consensus bears a well-conserved core' motif (nucleotides 8–17) and, although the model was not seeded with an E-box, it bears an E-box-like motif as well (nucleotides 19–24) (Figure 1A). To our knowledge, no functional analyses of individual CATAC elements or corresponding genomic sites have been reported to date. An illustration of the relationship between the CATAC element and the major known circadian enhancer element in the *per* gene (69-bp enhancer (5)) is provided in supplemental Supplementary Figure S1B.

A possible association of CATAC with clock-controlled and/or circadian transcription was explored further by determining the relative frequencies at which the CATAC element occurred in relevant promoter sequences outside of the five training genes. In particular, enrichment of CATAC was found in the top 0.5% genes ( $n = 61$  genes) predicted to be induced upon activation of the circadian regulator CLK (25) as well as in the top 0.7% genes ( $n = 97$  genes) encoding transcripts predicted to exhibit circadian oscillations (25–28) (Supplementary Figure S2). A prominent circadian transcript associated with CATAC promoter elements is *Slob* (*Slowpoke binding protein*) (27). With the exception of the element in the first intron of *per*, all CATAC elements associated with these six genes coincide with conserved sequence elements predicted by PhastCons analysis from an alignment of 27 insect genomes (29).

Based upon the sequence alignment, conservation, and subsequent CLK/CYC target gene enrichment analyses,



**Figure 2.** Regulation of 4xCATAC-*luc* reporter constructs. (A) Comparison of detrended, average traces (gray) and their corresponding FFT theoretical overlays (black) demonstrated a strong rhythm in wild-type CATAc-*luc* reporter activity. Average traces were derived from 26 *Pdp1* 4xCATAC-*luc*, 23 *Pdp1* mt4xCATAC-*luc* and 19 pPT*luc* lines (left panel) as well as 23 *Slob* 4xCATAC-*luc* and 17 *Slob* mt4xCATAC-*luc* lines (right panel). The luminescence profiles of reporters with mutant CATAc elements showed weaker, ultradian rhythms that resembled those observed

CATAc was hypothesized to be a regulatory element involved in mediating spatiotemporal expression of clock-regulated genes.

**CATAc imparts rhythmicity to a luciferase reporter**

To test our hypothesis, we studied CATAc elements found in two genes, *Pdp1* and *Slob*, that not only exhibit strong CLK/CYC-associated circadian oscillations, but also have an unusually high number of CATAc motifs. The *Pdp1* gene encodes a clock component (PDP1-ε) that impacts molecular circadian oscillations, particularly in the clock neurons, as well as circadian behaviour (30,31). Remarkably, *Pdp1* has seven CATAc elements in close proximity, one of which bears a canonical E-box (Figure 1B). The *Slob* gene, which encodes proteins that regulate the slowpoke channel, exhibits strong circadian regulation of its expression. The *Slob* promoter has four CATAc elements which all possess noncanonical E-boxes. To address the spatiotemporal expression information associated with CATAc, enhancer elements were generated, bearing four separate native CATAc elements from the *Slob* or *Pdp1* promoters (Figure 1B; Supplementary Figure S3). Each enhancer was assembled such that the four CATAc elements matched their order and orientation in the native *Pdp1* and *Slob* promoters (Figure 1B; Supplementary Figure S3). Moreover, like their promoters, the *Pdp1* 4xCATAC enhancer possesses one CATAc element with a canonical E-box while the *Slob* 4xCATAC enhancer has four noncanonical E-box-bearing CATAc elements. More specifically, in comparison with the *Pdp1* 4xCATAC enhancer the *Slob* 4xCATAC enhancer has more mismatches relative to both the CACGTG canonical E-box consensus (9/24 versus 4/24 mismatches) as well as the CRCGTG consensus for the E-box-like element within CATAc (Figure 1A; 8/24 versus 2/24 mismatches).

Mutated CATAc constructs were also generated by making transversion mutations to each of the four native elements. Mutations were made to the well-conserved core motif at residues 8, 10, 14 and 16, as well as to the E-box-like motif at residues 21 and 23 (Figure 1A; Supplementary Figure S3). Native and mutated 4xCATAC enhancers were then inserted upstream of a basal promoter driving expression of the firefly *luciferase* reporter gene (Figure 1B). Multiple independent transgenic lines were generated by P element

for the enhancer-less control (pPT*luc*). (B) 4xCATAC-induced reporter expression was determined after 6 h in DD from non-detrended, averaged raw data. The expression levels of *Pdp1* and *Slob* 4xCATAC were higher than their mutant counterparts ( $P < 0.001$ , Mann-Whitney U/Wilcoxon W rank sum test). The pPT*luc* expression levels were significantly higher than *Pdp1* and *Slob* mt4xCATAC but lower than *Slob* 4xCATAC. (C) 4xCATAC-*luc* reporter rhythmicity was determined from TopCount *in vivo* luminescence data by the FFT-NLLS utility of the BRASS software program. Individual flies with relative amplitude error (RAE) < 0.7, between 0.7 and 1, or > 1 were classified as rhythmic, weakly rhythmic and arrhythmic, respectively. Pairwise Chi-square test comparisons were used to generate the indicated  $P$  value estimates. *Pdp1* and *Slob* 4xCATAC generated greater rhythmicity than their mutated counterparts. Residual rhythmicity was observed for empty pPT*luc* vector, and to a lesser degree in mutant 4xCATAC constructs, but this was always significantly reduced relative to the wild-type *Pdp1* and *Slob* 4xCATAC reporters.

**Table 1.** Comparison of in vivo luminescence rhythms of different 4xCATAC-*luc* reporters

Genotype	#lines	#flies	%R	%WR	%AR	Mean RAE ± SEM	Mean period ± SEM
<i>Pdp1</i> 4xCATAC- <i>luc</i>	26	76	28	36	37	0.69 ± 0.02 <sup>a,b</sup>	24.09 ± 0.26 <sup>g</sup>
<i>Pdp1</i> with <i>Slob</i> -like E-boxes 4xCATAC- <i>luc</i>	18	105	44	25	31	0.62 ± 0.02 <sup>c</sup>	23.01 ± 0.30
<i>Pdp1</i> mt4xCATAC- <i>luc</i>	23	56	0	23	77	0.90 ± 0.02 <sup>a,c,d,e</sup>	20.29 ± 0.80 <sup>g,h</sup>
<i>Slob</i> 4xCATAC- <i>luc</i>	23	59	46	19	36	0.57 ± 0.02 <sup>b,f</sup>	24.72 ± 0.53 <sup>h</sup>
Flipped <i>Slob</i> 4xCATAC- <i>luc</i>	8	71	32	31	37	0.68 ± 0.03 <sup>d</sup>	23.17 ± 0.53
<i>Slob</i> mt4xCATAC- <i>luc</i>	17	42	5	5	90	0.71 ± 0.07	24.15 ± 0.83
pPT <i>luc</i>	19	56	18	23	59	0.71 ± 0.04 <sup>e,f</sup>	23.29 ± 0.71

RAE: Relative Amplitude Error, only provided for weakly rhythmic and rhythmic flies. %R/WR/AR: percentage of single fly luminescence traces that are rhythmic, weakly rhythmic or arrhythmic.

ANOVA with post-hoc Tamhane's T2  $P < 0.001^{a,d,e}$ ;  $P < 0.01^{b,c,g,h}$ ;  $P < 0.05^f$ .

transformation of native and mutated 4xCATAC *Pdp1* or *Slob* reporter constructs. Transformants were tested in an automated bioluminescence assay (TopCount) which allowed measurement of the transcriptional reporter constructs at relatively high frequency in vivo. Native *Pdp1* and *Slob* 4xCATAC both showed strong luciferase expression (Figure 2, Supplementary Figure S4) and rhythms (Figure 2, Supplementary Figure S5; Table 1), and their average profiles showed a single peak per 24 h (Figure 2). Thus, the presence of a canonical E-box, which was lacking from the *Slob* 4xCATAC reporter, did not seem to be required for either rhythms or induction. A similar conclusion was reached in a previous study of the *tim* promoter, which showed that noncanonical E-boxes can significantly contribute to circadian transcriptional activity (6). If anything, native *Pdp1* 4xCATAC transgenes, which included better matches to both the canonical and CATAC E-box consensus sequences tended to exhibit somewhat lower levels of rhythmicity than native *Slob* 4xCATAC transgenes with poorer quality E-boxes (Figure 2, Supplementary Figure S5; Table 1).

When we mutated all elements in the *Pdp1* and *Slob* 4xCATAC enhancers, the luciferase expression levels decreased (Figure 2, Supplementary Figure S4) and the rhythms were disrupted (Figure 2, Supplementary Figure S5; Table 1). Thus, the CATAC consensus sequence contributes to reporter induction and rhythmicity. Empty vector (pPT*luc*) and both mt4xCATAC constructs shared characteristically low level expression (Figure 2, Supplementary Figure S4) and comparable oscillation patterns, with two-peaks occurring per 24-hrs (Figure 2, Supplementary Figure S6). Earlier studies also found this 12-hr rhythm component in mutant E-box and basal promoter controls (4) and it has been reported to be a component of all expression patterns in such bioluminescence assays (4,19,32).

Although pPT*luc* and both of the mt4xCATAC constructs exhibited residual rhythmicity, neither exhibited the proportion of rhythmic flies observed for native *Pdp1* or *Slob* 4xCATAC (Figure 2). The residual rhythm for these constructs may have arisen due to cryptic regulatory element(s) present in the reporter vector or, alternatively, as a result of rhythmic post-transcriptional modulation of luciferase reporter activity in the tissues expressing these constructs. We also generated two additional constructs, 'flipped *Slob*' (FS) in which the orientation of *Slob* 4xCATAC in the vector was inverted and '*Pdp1* 4xCATAC with *Slob*-like E-boxes' (PSE) in which the four E-boxes of

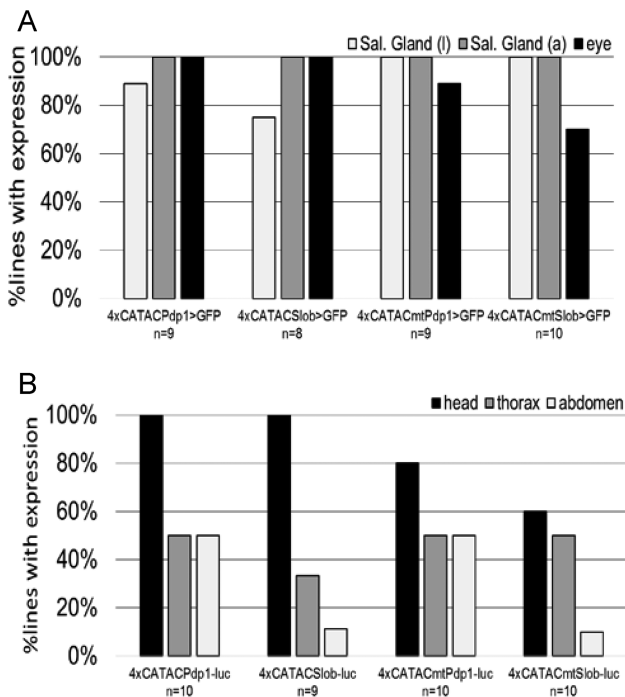
*Pdp1* 4xCATAC were mutated such that their sequence fidelity matched that of the four E-boxes in *Slob* 4xCATAC (Figure 1B; Supplementary Figure S3). These constructs also exhibited better rhythmicity than pPT*luc* and both mt4xCATACs (Supplementary Figure S5; Table 1). Moreover, FS and PSE induced significantly more luciferase activity than the respective *Slob* and *Pdp1* mt4xCATAC controls (Supplementary Figure S4). Thus, in sum, in vivo luciferase reporter analyses uncovered preferential activity induction and rhythmicity for constructs containing wild-type CATAC elements.

We performed ANOVA analyses of luminescence data using only rhythmic and weakly rhythmic individual flies as determined by FFT-NLLS. The coherence of rhythms observed for *Pdp1* and *Slob* 4xCATAC reporters, as measured by the inversely correlated Relative Amplitude Error (RAE), was significantly stronger than those for mt4xCATAC and empty vector controls, respectively (Table 1). Period measurements, did not differ significantly between reporter constructs, except for the *Pdp1* 4xCATAC versus *Pdp1* mt4xCATAC comparison (Table 1). However, the unexpected short period length observed for *Pdp1* mt4xCATAC rhythms may be attributable to reduced accuracy in period estimation due to weak rhythms for this reporter.

### CATAC-driven reporters are expressed in the eye and other tissues with peak transcript phases during (subjective) day

To further examine the spatiotemporal expression of the CATAC reporter, we crossed our 4xCATAC-*Gal4* lines with *UAS-CD8::GFP* flies to generate offspring that report CATAC activity by expression of membrane-tethered green fluorescent protein. In larvae and adult flies, 4xCATAC expression primarily included the salivary glands in addition to the photoreceptor cells of the compound eye in the adult (Figure 3A, Supplementary Figure S7). We also examined brains of the 4xCATAC flies but found no evidence of GFP expression. Little is known about the salivary glands as a circadianly rhythmic tissue and, although the observed signal in salivary glands was not due to autofluorescence or leaky expression of the *UAS-CD8::GFP* transgene (Supplementary Figure S7B), it is possible that the reporter gene expression in this tissue reflects the presence of a suspected salivary gland enhancer in the *hsp70* sequences (33) included in the pPTGal vector. Conversely, the photorecep-





**Figure 3.** Spatial expression patterns of transgenic CATAc reporters. Larvae (l) and adult (a) male flies from multiple (n) lines expressing either 4xCATAC-*Gal4*-mediated GFP (A) or 4xCATAC-luciferase (B) were dissected to determine the spatial expression pattern of CATAc by fluorescence or luminescence imaging, respectively. The proportion of lines with tissues positive for a given reporter is indicated.

tors are known to possess an autonomous, circadian oscillator (34). Furthermore, both the native *Slob* and *Pdp1* transcripts are expressed in the photoreceptors (30,35).

To follow-up, we imaged dissected heads and bodies of male *Pdp1* and *Slob* 4xCATAC luciferase flies that were fed a sucrose solution containing 15mM of luciferin, as in the TopCount assays. These tissues were then cultured in luciferin-containing media and imaged as close to their LD peak expression time as possible. The peak times were determined from the previously described TopCount assays (Figure 2) – *Pdp1* (ZT 2–3) and *Slob* (ZT 3–5). Heads showed luciferase expression in not only the compound eyes, but also the proboscis and antennae; each of which has been shown to possess a circadian oscillator (17,18) (Figure 3B, Supplementary Figure S8A). The bodies (including appendages) demonstrated signal in the wing, another tissue with an oscillator, and in other tissues within the thorax and abdomen (Figure 3B, Supplementary Figure S8B).

Given the status of the adult compound eye as a tissue with not only native expression of *Pdp1* and *Slob* (30,35) but also >80% of the circadian oscillators in fly heads (36), follow-up experiments were conducted to determine the contribution of the compound eyes to overall 4xCATAC reporter signal. We genetically removed the eyes via the eyes absent mutation, *eya*<sup>2</sup> (37). While the luciferase expression levels in these flies did decrease, there was still significant luciferase activity (Supplementary Figure S9, Table 2). Thus, while the eyes are a source of 4xCATAC driven luciferase signal, they are not its only relevant source.

Therefore, spatial expression analyses of the wild-type and mutant 4xCATAC reporters uncovered expression in a number of tissues including the known oscillators of the compound eye. However, despite some trends suggesting a role for eye clocks in CATAc-mediated rhythmicity, no statistically significant differences between the spatial expression patterns of wild-type and mutant 4xCATAC reporters were uncovered (Figure 3, Supplementary Figure S8). As described above, this stands in contrast to clear differences in rhythmicities (Figure 2, Supplementary Figure S5, Table 1) and expression levels (Figure 2, Supplementary Figure S4, Table 1) produced by wild-type versus mutant 4xCATAC-*luc* constructs. Taken together, these findings suggest that CATAc primarily specifies temporal rather than spatial information.

The native *Pdp1* and *Slob* transcripts oscillate robustly in fly heads (25–28). While we had found that *Pdp1* and *Slob* 4xCATAC-*luc* oscillated in whole flies, it remained unclear whether the 4xCATAC reporter transcripts oscillated in the same phase as their respective native genes. To address this question, we performed a qRT-PCR time course analysis on native and reporter gene mRNA in whole fly heads. We found that both *Pdp1* and *Slob* 4xCATAC-*luc* oscillated in phase with each other, peaking at CT6 (Figure 4A and B). This is consistent with observed in-phase luminescence patterns for reporter activity of these constructs (Figure 4C). Native *Slob* transcript also peaked at CT6, while the native *Pdp1* transcript peaked slightly out of phase with *Pdp1* 4xCATAC at CT12. The different phases of CLK/CYC-regulated *Pdp1* and *Slob* transcripts may reflect subtly different transcriptional activities of their promoters or differing mRNA half-lives (28,30).

### CATAc-*luc* reporters require *cyc* and *per* for rhythmicity, but not induction

Based on the circadian rhythmicity of 4xCATAC-*luc* reporter genes and the presence of a canonical CACGTG E-box in the CATAc consensus sequence, we hypothesized that the circadian E-box transcription factor CLK/CYC might be responsible for the observed rhythmic regulation of the CATAc element. To address this, 4xCATAC-*luc* reporter genes were introduced into a *cycle* null (*cyc*<sup>01</sup>) genetic background (3). As expected, *Pdp1* and *Slob* 4xCATAC lost rhythmicity in the *cyc*<sup>01</sup> homozygous background, as did our *tim*-luciferase positive control (Figure 5, Supplementary Figure S10, Table 3). Multimerized E-box-*luc* reporters also have been reported to show this in *Clk*<sup>Jrk</sup> flies (4). However, unlike *tim*-*luc*, and multi-E-box-*luc* constructs, 4xCATAC-*luc* transgenes did not exhibit a decrease in reporter activity in *cyc*<sup>01</sup> homozygotes (Figure 5).

To better understand this result, we performed further analysis of the association between CATAc and observed CLK/CYC binding using published CLK antibody ChIP-chip data (38). CATAc elements in the core clock genes, *Pdp1*, *tim*, *per* and *cwo* as well as the clock-controlled gene, *Slob*, tended to coincide with CLK or PER ChIP signals only in the presence of either an internal canonical E-box or closely-linked external E1E2 motif (Supplementary Figure S11A and B). Moreover, an extended analysis across the entire genome exposed significant enrichment of CLK

**Table 2.** Impact of the eyeless mutant *eya<sup>2</sup>* on 4xCATAC-*luc* in vivo luminescence rhythms

genotype	#flies	%R	%WR	%AR	mean RAE±SEM	mean period±SEM
<i>eya<sup>2</sup>/CyO; Pdp1 4xCATAC-luc</i>	18	44	33	22	0.67±0.03	23.03±0.68
<i>eya<sup>2</sup>; Pdp1 4xCATAC-luc</i>	19	16	26	58	0.68±0.06	24.71±2.28
<i>eya<sup>2</sup>/CyO; Slob 4xCATAC-luc</i>	21	38	33	29	0.71±0.03	25.03±1.09
<i>eya<sup>2</sup>; Slob 4xCATAC-luc</i>	9	44	44	11	0.67±0.05	22.15±1.17

**Table 3.** Impact of the arrhythmic mutant *cyc<sup>01</sup>* on 4xCATAC-*luc* in vivo luminescence rhythms

Genotype	#flies	%R	%WR	%AR	Mean RAE ± SEM	Mean period ± SEM
<i>Pdp1 4xCATAC-luc; cyc<sup>01</sup>/+</i>	7	29	57	14	0.70 ± 0.06	25.74 ± 1.21
<i>Pdp1 4xCATAC-luc; cyc<sup>01</sup></i>	14	0	14	86	0.86 ± 0.09	19.56 ± 2.59
<i>Slob 4xCATAC-luc; cyc<sup>01</sup>/+</i>	24	33	46	21	0.70 ± 0.04	21.98 ± 0.64
<i>Slob 4xCATAC-luc; cyc<sup>01</sup></i>	23	4	35	61	0.80 ± 0.05 <sup>a</sup>	22.68 ± 1.98
<i>Pdp1 mt4xCATAC-luc; cyc<sup>01</sup>/+</i>	22	9	18	82	0.73 ± 0.04 <sup>b</sup>	23.72 ± 0.41
<i>Pdp1 mt4xCATAC-luc; cyc<sup>01</sup></i>	18	0	6	94	(0.73)	(16.97)
<i>Slob mt4xCATAC-luc; cyc<sup>01</sup>/+</i>	24	17	29	54	0.76 ± 0.05	24.03 ± 0.68
<i>Slob mt4xCATAC-luc; cyc<sup>01</sup></i>	22	0	18	82	0.85 ± 0.05 <sup>c</sup>	18.55 ± 1.17
<i>pPThuc; cyc<sup>01</sup>/+</i>	24	8	38	54	0.81 ± 0.05 <sup>d</sup>	23.66 ± 0.93
<i>pPThuc; cyc<sup>01</sup></i>	20	10	25	65	0.79 ± 0.08	22.10 ± 2.19
<i>tim-luc; cyc<sup>01</sup>/+</i>	23	83	13	4	0.51 ± 0.04 <sup>a,b,c,d,e</sup>	24.35 ± 0.46
<i>tim-luc; cyc<sup>01</sup></i>	23	4	22	74	0.82 ± 0.05 <sup>e</sup>	19.59 ± 1.75

ANOVA with post-hoc Tamhane's T2  $P < 0.01^{a,d}$ ;  $P < 0.05^{b,c,e}$ .**Table 4.** Impact of the arrhythmic mutant *per<sup>01</sup>* on 4xCATAC-*luc* in vivo luminescence rhythms

Genotype	#flies	%R	%WR	%AR	Mean RAE ± SEM	Mean period ± SEM
<i>Pdp1 4xCATAC-luc; per<sup>01</sup>/+</i>	24	71	25	4	0.62 ± 0.03 <sup>a</sup>	24.92 ± 0.65
<i>Pdp1 4xCATAC-luc; per<sup>01</sup></i>	23	17	13	70	0.73 ± 0.05	23.89 ± 3.00
<i>Slob 4xCATAC-luc; per<sup>01</sup>/+</i>	20	45	45	10	0.66 ± 0.03 <sup>b</sup>	24.19 ± 1.25
<i>Slob 4xCATAC-luc; per<sup>01</sup></i>	22	9	18	73	0.76 ± 0.05	25.18 ± 2.47
<i>Pdp1 with Slob-like E-boxes 4xCATAC-luc; per<sup>01</sup>/+</i>	21	29	48	24	0.69 ± 0.04 <sup>c</sup>	25.37 ± 0.77
<i>Pdp1 with Slob-like E-boxes 4xCATAC-luc; per<sup>01</sup></i>	21	5	5	90	0.63 ± 0.17	21.92 ± 4.19
<i>Pdp1 mt4xCATAC-luc; per<sup>01</sup>/+</i>	20	30	30	40	0.69 ± 0.04 <sup>d</sup>	24.44 ± 1.03
<i>Pdp1 mt4xCATAC-luc; per<sup>01</sup></i>	25	8	20	72	0.80 ± 0.05 <sup>e</sup>	22.91 ± 2.63
<i>Slob mt4xCATAC-luc; per<sup>01</sup>/+</i>	24	63	17	21	0.59 ± 0.03 <sup>f</sup>	23.73 ± 0.75
<i>Slob mt4xCATAC-luc; per<sup>01</sup></i>	24	0	17	83	0.89 ± 0.03 <sup>a,f,g,h</sup>	22.53 ± 2.67
<i>pPThuc; per<sup>01</sup>/+</i>	21	62	24	14	0.60 ± 0.04 <sup>h</sup>	26.75 ± 0.85
<i>pPThuc; per<sup>01</sup></i>	21	0	19	81	0.85 ± 0.05 <sup>i</sup>	20.27 ± 1.51
<i>tim-luc; per<sup>01</sup>/+</i>	24	88	13	0	0.46 ± 0.04 <sup>b,c,d,e,g,i</sup>	23.87 ± 0.44
<i>tim-luc; per<sup>01</sup></i>	24	17	4	79	0.65 ± 0.05	18.62 ± 2.58

ANOVA with post-hoc Tamhane's T2  $P < 0.001^g$ ;  $P < 0.01^e$ ;  $P < 0.05^{a,b,c,d,f,h,i}$ .

binding at E1E2 sites, but not CATAC sites (Supplementary Figure S11C; Supplementary Table S2). Taken together, although CATAC-mediated circadian rhythms depended on CLK/CYC, and it is possible that CATAC and CLK/CYC-regulated elements act cooperatively in native clock gene promoters, our results indicated that CLK/CYC is not the (only) direct regulator of CATAC. This conclusion is supported by our observation that 4xCATAC-*luc* differed from multimerized E-box-*luc* or *tim-luc* in its ability to maintain high expression levels in the absence of CLK/CYC activity and the fact that *in vivo* CLK-binding near individual CATAC elements did not appear predictive of their ability to mediate circadian rhythms, but rather of their proximity to canonical E-boxes.

In *Drosophila*, CLK/CYC is rhythmically repressed by the PER/TIM heterodimer to produce oscillations of clock-regulated gene transcripts such as *vri* and *Pdp1* (30). A previous study showed that in a *Clk<sup>Jrk</sup>* or *cyc<sup>01</sup>* genetic background *tim*, *per* and *vri* transcription is low (39)—which we confirmed with *tim* in our bioluminescence assay (Figure 5). However, 4xCATAC does not act in a similar fashion and instead remains at high to intermediate expression levels—a phenotype expected of VRI/PDP1-regulated transcripts such as *Clk* and *cry* (30). If CATAC were regulated in a fashion similar to the *Clk* and *cry* promoters, we would expect its activity to be not only arrhythmic, but also strongly reduced in a *per<sup>01</sup>* background, where increased VRI activity is thought to result in suppression of these promoters. In contrast, CLK/CYC-regulated promoters such



as *tim* are expressed at intermediate levels in this background (27) (Figure 6). Although there was some variation among individual wild-type 4xCATAC-*luc* lines (for example, the *Slob* 4xCATAC line that was tested showed somewhat reduced expression levels in the *per*<sup>01</sup> background), by and large there was little effect of the *per*<sup>01</sup> mutation on their expression levels (Figure 6). CATAC, therefore, does not appear to be co-regulated with VRI/PDP1 elements.

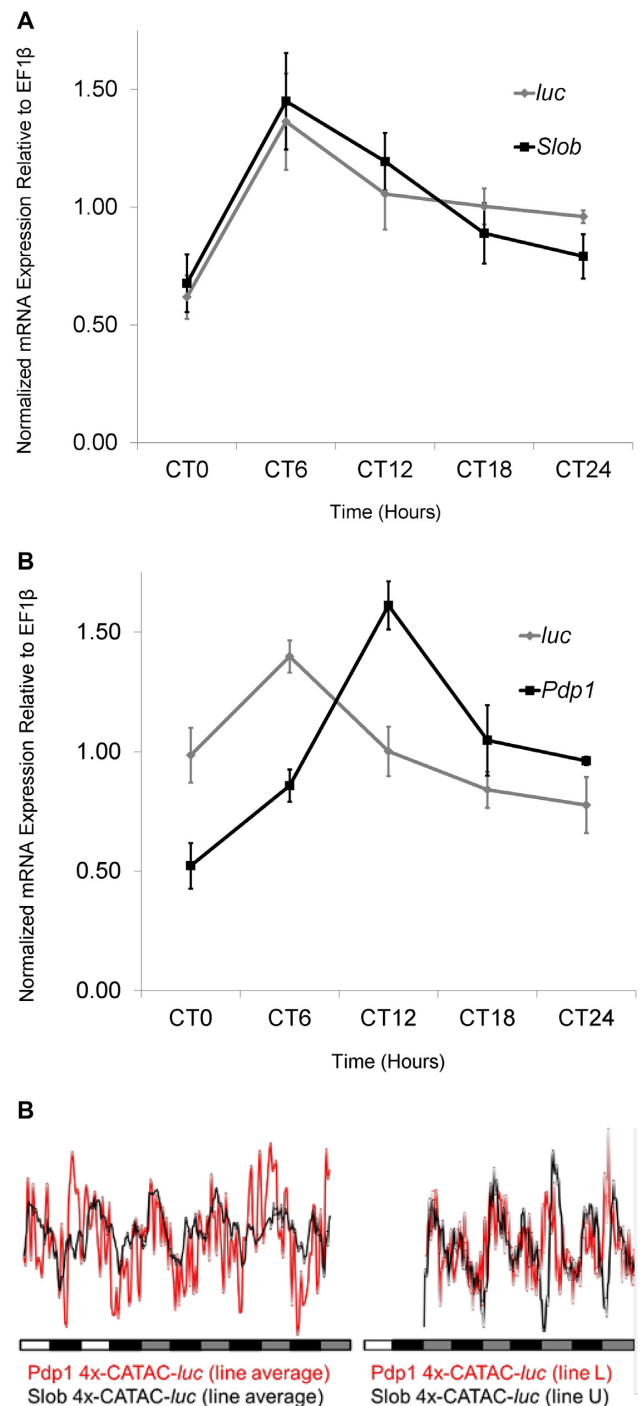
Additional experiments were performed to examine whether the somewhat variable response of 4xCATAC-*luc* expression levels to the *per*<sup>01</sup> mutation could be explained by the presence or absence of a canonical E-box. To this aim, we tested our *Pdp1* 4xCATAC with *Slob*-like E-boxes (PSE) construct. PSE is a modified version of *Pdp1* 4xCATAC, which has E-box sequences that were mutated such that their deviation from the CACGTG consensus matched that of the four E-boxes in the *Slob* 4xCATAC enhancer (Figure 1A). PSE still showed high expression levels in *per*<sup>01</sup> flies, which suggests that the E-box content is not responsible for the observed differences. Nevertheless, in all cases 4xCATAC lost rhythmicity in the *per*<sup>01</sup> homozygotes, confirming that PER was required for CATAC-mediated rhythms (Figure 6, Supplementary Figure S12, Table 4).

## DISCUSSION

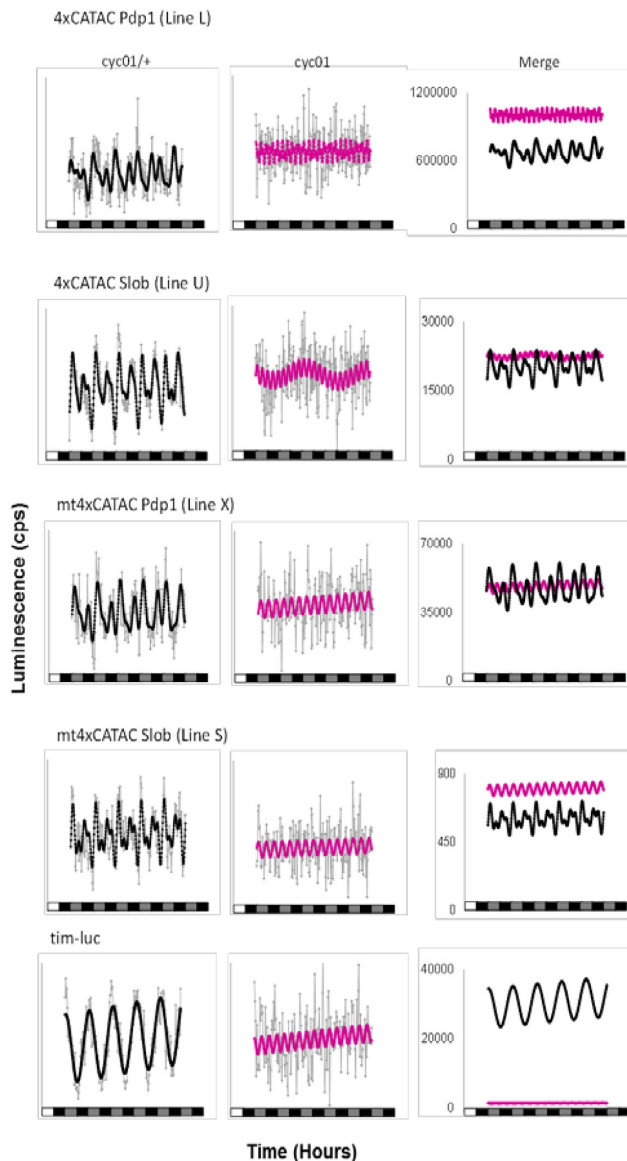
Regulation of circadian gene expression relies on a number of transcriptional elements found in the promoters and introns of core clock and clock-regulated genes. We now know of at least 5 reported functional transcriptional elements that are represented in circadianly regulated genes in *Drosophila*—the E-box, PERR element, VRI/PDP1-box, CRE element and TER box (27). Among them, the E-box is the best studied. The E-box is a rather versatile regulatory element, capable of functioning with deviations from its canonical form and utilizing flanking sequences and/or nearby E-boxes and regulatory elements to modulate its functionality (7,8).

Noncanonical E-box regulation of circadian genes has previously been identified in *Drosophila tim* and mammalian *dbp* (6,40). The CATAC element possesses an E-box sequence which may be either canonical or noncanonical in nature. Regardless, CATAC is able to generate robust, sustainable rhythms as observed with the *Pdp1* and *Slob* 4xCATAC, as well as the *Pdp1* with *Slob*-like E-boxes (PSE), constructs. Thus, the considerable variation in quality of E-box sequences between these constructs did not obviously impact reporter gene rhythms. The sequence immediately flanking an E-box element has also been implicated in affecting regulation (9). Simultaneous mutation of residues in the conserved core region of CATAC as well as its E-box, resulted in disrupted rhythms and decreased reporter expression. This suggests that the conserved core sequences are important to CATAC regulation, although it is unclear as to whether this corresponds to rhythmicity, expression levels or both. This question may be addressed in future studies by mutational analyses of only the CATAC core sequence.

As previously observed by McDonald *et al.* (6), transcriptional elements nearby an E-box can influence circadian gene regulation. The *tim* promoter has a canonical E-



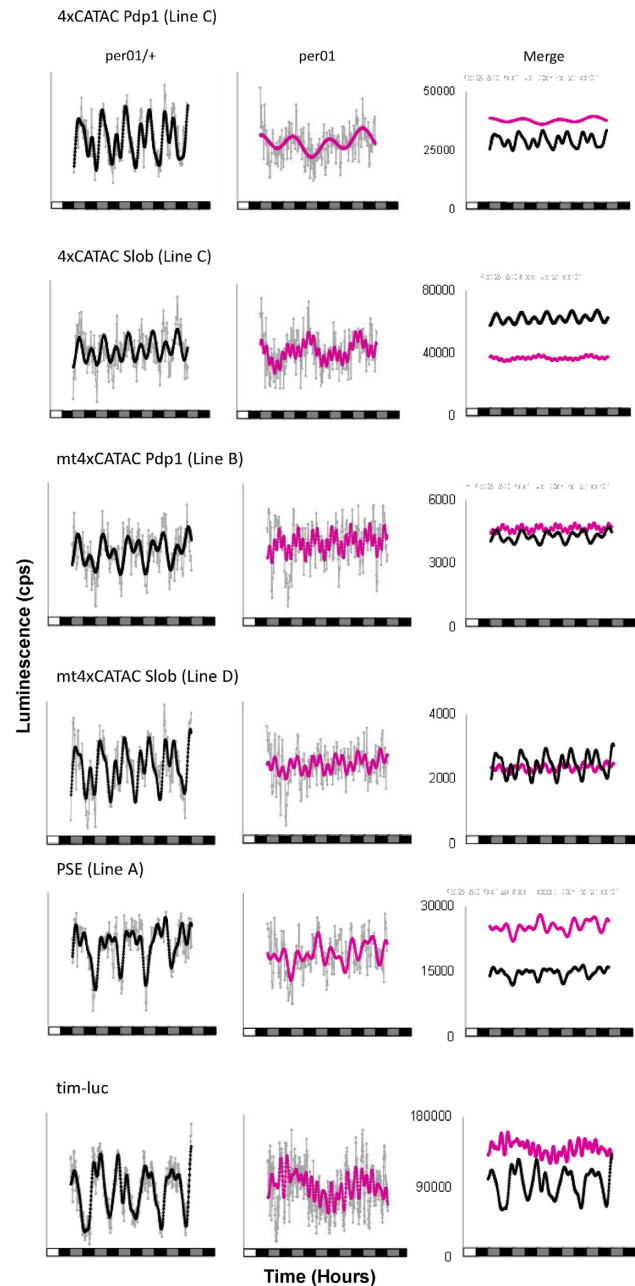
**Figure 4.** Time course of mRNA expression of *Pdp1* and *Slob* 4xCATAC-*luc* relative to their respective native gene transcripts in fly heads. (A, B) Transcript profiles were determined relative to *EF1β* by qRT-PCR for (A) *Slob* 4xCATAC-*luc* (Line U) and (B) *Pdp1* 4xCATAC-*luc* (Line L) homozygotes (gray diamonds with gray lines) and compared to the respective native transcripts, *Pdp1* and *Slob* (black squares with black lines). Each time point represents the mean data from four independent experiments ( $\pm$ SEM). Each profile exhibited significant regulation (Kruskal–Wallis [*P*-value;peak]: *Pdp1*[0.007;CT12] *Pdp1* 4xCATAC-*luc*(0.028;CT6) *Slob*[0.023;CT6] *Slob* 4xCATAC-*luc*(0.026;CT6)). (C) The matching mRNA phases for *Slob* 4xCATAC-*luc* (A) versus *Pdp1* 4xCATAC-*luc* (B) are consistent with in-phase luciferase activity rhythms observed for the across-the-board line average (left; data from Figure 2A) or for the individual lines used in (A) and (B) (right; data from control flies in Figure 5, below).



**Figure 5.** Impact of the *cyc<sup>01</sup>* mutation on 4xCATAC-*luc* rhythmicity and expression. Comparison of detrended, average traces (gray) and their corresponding FFT theoretical overlays (black or magenta) for a single fly line per genotype. Heterozygote controls (black) and the corresponding *cyc<sup>01</sup>* homozygotes (magenta) show that without a functioning clock, CATAc fails to yield circadian reporter oscillations. Residual *Pdp1* and *Slob* mt4xCATAC rhythmicity is further diminished in the homozygous *cyc<sup>01</sup>* genetic background. Noticeably, in the *cyc<sup>01</sup>* homozygote background *tim-luciferase* expression levels drop considerably, a phenomenon not observed with 4xCATAC. For statistical measures of the RAE, period and overall rhythmicity refer to Supplementary Figure S10 and Table 3.

box that relies on two proximal cis-regulatory elements that can also interact with CLK/CYC to produce rhythmic transcriptional activity. Genomic analysis of CATAc showed that a high likelihood of CLK/CYC binding coincided with the presence of either a canonical E-box internal to CATAc or the presence of E1E2 motifs in close proximity (Supplementary Figure S11).

For all the similarities that CATAc shares with E-box motifs, CATAc is not simply another E-box element.



**Figure 6.** Impact of the *per<sup>01</sup>* mutation on 4xCATAC-*luc* rhythmicity and expression. Comparison of detrended, average traces (gray) and their corresponding FFT theoretical overlays (black or magenta) for a single fly line per genotype. PSE as well as *Pdp1* and *Slob* 4xCATAC heterozygote controls (black) and the corresponding *per<sup>01</sup>* homozygotes (magenta) show that without a functioning clock, CATAc fails to yield reporter oscillations. Residual *Pdp1* and *Slob* mt4xCATAC rhythmicity is further diminished in the homozygous *per<sup>01</sup>* genetic background. Most reporter constructs, including *tim-luc* and various CATAc-*luc* insertions, maintain similar or somewhat increased expression levels in the *per<sup>01</sup>* homozygous background; only *Slob* 4xCATAC expression levels exhibited a mild decrease. For statistical measures of the RAE, period and overall rhythmicity refer to Supplementary Figure S10 and Table 4.

Normally, an E-box regulated transcriptional element exhibits low reporter expression in a *cyc<sup>01</sup>* genetic background and intermediate-to-high expression in *per<sup>01</sup>* (3,41). CATAc shows high reporter expression in a *cyc<sup>01</sup>* background and generally intermediate-to-high expression in *per<sup>01</sup>*. While CATAc rhythmicity relies on a functioning core clock, CLK/CYC does not appear to be the only potential regulator of CATAc. There is likely another regulator, or regulators, that act on the CATAc to generate increased transcriptional activity in the null backgrounds. Other E-box-binding transcription factors may contribute to the regulation of CATAc along with transcription complexes binding to the conserved core sequence. In particular, the basic helix-loop-helix transcription factor CLOCKWORK ORANGE (CWO) is known to bind E-boxes in clock-controlled genes (42,43) and may, therefore, modulate CATAc activity. However, since CWO exhibits an inhibitory effect on E-box-mediated transcription it is unlikely to be directly responsible for the enhancer activity of CATAc elements.

Thus, we have discovered a novel circadian regulatory element that, although it possesses an E-box-like motif, exhibits non-E-box-like responses. Through use of multimerized CATAc element, we concluded that CATAc is capable of contributing to the rhythmicity of clock genes, such as *Pdp1* and *Slob*. Of course, this conclusion comes with the caveat that the multimerized constructs do not represent the complexity of the native genomic environment. Mutational analysis of CATAc sequences in their natural context will constitute an important next step in the functional analysis of this *cis*-regulatory element. Nevertheless, the identification of CATAc contributes to our knowledge of circadian transcriptional elements, and may be used to further characterize the regulatory regions of clock and clock-regulated genes.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors would like to thank Mallory Kuchem for assistance with the generation of constructs, Raisa Saab and Chris Uyehara for assistance with some of the genetic and luminescence assays, and Genetic Services, Inc. for assistance with the generation of transgenic flies. We thank the Bloomington Drosophila Stock Center, Drosophila Genomics Resource Center, Michael Young, Michael Rosbash, Jeffrey Hall, Ralf Stanewsky, Liqun Luo and Dan Eberl for the generation or provision of reagents used in this study.

## FUNDING

National Institutes of Health [GM78339 and RR025537 to H.W.]; National Science Foundation [0923366 to H.W.]; Biotechnology and Biological Science Research Council [BB/L023067/1 to H.W.]. Funding for open access charge: RCUK funds held at University of Southampton.

*Conflict of interest statement.* None declared.

## REFERENCES

- Partch,C.L., Green,C.B. and Takahashi,J.S. (2014) Molecular architecture of the mammalian circadian clock. *Trends Cell Biol.*, **24**, 90–99.
- Allada,R., White,N.E., So,W.V., Hall,J.C. and Rosbash,M. (1998) A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless. *Cell*, **93**, 791–804.
- Rutila,J.E., Suri,V., Le,M., So,W.V., Rosbash,M. and Hall,J.C. (1998) CYCLE is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of Drosophila period and timeless. *Cell*, **93**, 805–814.
- Darlington,T.K., Lyons,L.C., Hardin,P.E. and Kay,S.A. (2000) The period E-box is sufficient to drive circadian oscillation of transcription in vivo. *J. Biol. Rhythms*, **15**, 462–471.
- Hao,H., Allen,D.L. and Hardin,P.E. (1997) A circadian enhancer mediates PER-dependent mRNA cycling in Drosophila melanogaster. *Mol. Cell Biol.*, **17**, 3687–3693.
- McDonald,M.J., Rosbash,M. and Emery,P. (2001) Wild-type circadian rhythmicity is dependent on closely spaced E boxes in the Drosophila timeless promoter. *Mol. Cell Biol.*, **21**, 1207–1217.
- Kyriacou,C.P. and Rosato,E. (2000) Squaring up the E-box. *J. Biol. Rhythms*, **15**, 483–490.
- Munoz,E., Brewer,M. and Baler,R. (2002) Circadian transcription. Thinking outside the E-Box. *J. Biol. Chem.*, **277**, 36009–36017.
- Lyons,L.C., Darlington,T.K., Hao,H., Houl,J., Kay,S.A. and Hardin,P.E. (2000) Specific sequences outside the E-box are required for proper per expression and behavioral rescue. *J. Biol. Rhythms*, **15**, 472–482.
- Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F. *et al.* (2000) The genome sequence of Drosophila melanogaster. *Science*, **287**, 2185–2195.
- Stapleton,M., Carlson,J., Brokstein,P., Yu,C., Champe,M., George,R., Guarin,H., Kronmiller,B., Pacleb,J., Park,S. *et al.* (2002) A Drosophila full-length cDNA resource. *Genome Biol.*, **3**, RESEARCH0080.
- Paquet,E.R., Rey,G. and Naef,F. (2008) Modeling an evolutionary conserved circadian cis-element. *PLoS Comput. Biol.*, **4**, e38.
- Bailey,T.L. and Elkan,C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **2**, 28–36.
- Sharma,Y., Cheung,U., Larsen,E.W. and Eberl,D.F. (2002) PPTGAL, a convenient Gal4 P-element vector for testing expression of enhancer fragments in drosophila. *Genesis*, **34**, 115–118.
- Darlington,T.K., Wager-Smith,K., Ceriani,M.F., Staknis,D., Gekakis,N., Steeves,T.D., Weitz,C.J., Takahashi,J.S. and Kay,S.A. (1998) Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. *Science*, **280**, 1599–1603.
- Konopka,R.J. and Benzer,S. (1971) Clock mutants of Drosophila melanogaster. *Proc. Natl. Acad. Sci. U.S.A.*, **68**, 2112–2116.
- Stanewsky,R., Jamison,C.F., Plautz,J.D., Kay,S.A. and Hall,J.C. (1997) Multiple circadian-regulated elements contribute to cycling period gene expression in Drosophila. *EMBO J.*, **16**, 5006–5018.
- Plautz,J.D., Kaneko,M., Hall,J.C. and Kay,S.A. (1997) Independent photoreceptive circadian clocks throughout Drosophila. *Science*, **278**, 1632–1635.
- Plautz,J.D., Straume,M., Stanewsky,R., Jamison,C.F., Brandes,C., Dowse,H.B., Hall,J.C. and Kay,S.A. (1997) Quantitative analysis of Drosophila period gene transcription in living animals. *J. Biol. Rhythms*, **12**, 204–217.
- Sullivan,W., Ashburner,M. and Hawley,R.S. (2000) *Drosophila Protocols*. Cold Spring Harbor Laboratory Press, NY.
- Sellix,M.T., Currie,J., Menaker,M. and Wijnen,H. (2010) Fluorescence/luminescence circadian imaging of complex tissues at single-cell resolution. *J. Biol. Rhythms*, **25**, 228–232.
- Goda,T., Mirowska,K., Currie,J., Kim,M.H., Rao,N.V., Bonilla,G. and Wijnen,H. (2011) Adult circadian behavior in Drosophila requires developmental expression of cycle, but not period. *PLoS Genet.*, **7**, e1002167.
- Goda,T., Sharp,B. and Wijnen,H. (2014) Temperature-dependent resetting of the molecular circadian oscillator in Drosophila. *Proc. Biol. Sci./R. Soc.*, **281**.



24. Schmittgen, T.D. and Livak, K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.*, **3**, 1101–1108.
25. McDonald, M.J. and Rosbash, M. (2001) Microarray analysis and organization of circadian gene expression in *Drosophila*. *Cell*, **107**, 567–578.
26. Ceriani, M.F., Hogenesch, J.B., Yanovsky, M., Panda, S., Straume, M. and Kay, S.A. (2002) Genome-wide expression analysis in *Drosophila* reveals genes controlling circadian behavior. *J. Neurosci. Off. J. Soc. Neurosci.*, **22**, 9305–9319.
27. Claridge-Chang, A., Wijnen, H., Naef, F., Boothroyd, C., Rajewsky, N. and Young, M.W. (2001) Circadian regulation of gene expression systems in the *Drosophila* head. *Neuron*, **32**, 657–671.
28. Wijnen, H., Naef, F., Boothroyd, C., Claridge-Chang, A. and Young, M.W. (2006) Control of daily transcript oscillations in *Drosophila* by light and the circadian clock. *PLoS Genet.*, **2**, e39.
29. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.*, **15**, 1034–1050.
30. Cyran, S.A., Buchsbaum, A.M., Reddy, K.L., Lin, M.C., Glossop, N.R., Hardin, P.E., Young, M.W., Storti, R.V. and Blau, J. (2003) *vriille*, *Pdp1*, and *dClock* form a second feedback loop in the *Drosophila* circadian clock. *Cell*, **112**, 329–341.
31. Zheng, X., Koh, K., Sowcik, M., Smith, C.J., Chen, D., Wu, M.N. and Sehgal, A. (2009) An isoform-specific mutant reveals a role of PDP1 epsilon in the circadian oscillator. *J. Neurosci. Off. J. Soc. Neurosci.*, **29**, 10920–10927.
32. Brandes, C., Plautz, J.D., Stanewsky, R., Jamison, C.F., Straume, M., Wood, K.V., Kay, S.A. and Hall, J.C. (1996) Novel features of *Drosophila* period transcription revealed by real-time luciferase reporting. *Neuron*, **16**, 687–692.
33. Gerlitz, O., Nellen, D., Ottiger, M. and Basler, K. (2002) A screen for genes expressed in *Drosophila* imaginal discs. *Int. J. Dev. Biol.*, **46**, 173–176.
34. Cheng, Y. and Hardin, P.E. (1998) *Drosophila* photoreceptors contain an autonomous circadian oscillator that can function without period mRNA cycling. *J. Neurosci.*, **18**, 741–750.
35. Jaramillo, A.M., Zheng, X., Zhou, Y., Amado, D.A., Sheldon, A., Sehgal, A. and Levitan, I.B. (2004) Pattern of distribution and cycling of SLOB, Slowpoke channel binding protein, in *Drosophila*. *BMC Neuroscience*, **5**, 3.
36. Glossop, N.R. and Hardin, P.E. (2002) Central and peripheral circadian oscillator mechanisms in flies and mammals. *J. Cell Sci.*, **115**, 3369–3377.
37. Bonini, N.M., Leiserson, W.M. and Benzer, S. (1993) The eyes absent gene: genetic control of cell survival and differentiation in the developing *Drosophila* eye. *Cell*, **72**, 379–395.
38. Abruzzi, K.C., Rodriguez, J., Menet, J.S., Desrochers, J., Zadina, A., Luo, W., Tkachev, S. and Rosbash, M. (2011) *Drosophila* CLOCK target gene characterization: implications for circadian tissue-specific gene expression. *Genes Dev.*, **25**, 2374–2386.
39. Blau, J. and Young, M.W. (1999) Cycling *vriille* expression is required for a functional *Drosophila* clock. *Cell*, **99**, 661–671.
40. Schibler, U. and Naef, F. (2005) Cellular oscillators: rhythmic gene expression and metabolism. *Curr. Opin. Cell Biol.*, **17**, 223–229.
41. Glossop, N.R., Lyons, L.C. and Hardin, P.E. (1999) Interlocked feedback loops within the *Drosophila* circadian oscillator. *Science*, **286**, 766–768.
42. Kadener, S., Stoleru, D., McDonald, M., Nawathean, P. and Rosbash, M. (2007) Clockwork Orange is a transcriptional repressor and a new *Drosophila* circadian pacemaker component. *Genes Dev.*, **21**, 1675–1686.
43. Matsumoto, A., Ukai-Tadenuma, M., Yamada, R.G., Houl, J., Uno, K.D., Kasukawa, T., Dauwalder, B., Itoh, T.Q., Takahashi, K., Ueda, R. *et al.* (2007) A functional genomics strategy reveals clockwork orange as a transcriptional regulator in the *Drosophila* circadian clock. *Genes Dev.*, **21**, 1687–1700.