View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Current Biology 24, 1–10, January 6, 2014 ©2014 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2013.11.017

Article

cis-Regulatory Requirements for Tissue-Specific Programs of the Circadian Clock

Antonio C.A. Meireles-Filho,[1,2] Anaïs F. Bardet,[1,3] J. Omar Yáñez-Cuna,[1] Gerald Stampfel,[1] and Alexander Stark[1,*]
[1]Research Institute of Molecular Pathology (IMP), 1030 Vienna, Austria

Summary

Background: Broadly expressed transcriptions factors (TFs) control tissue-specific programs of gene expression through interactions with local TF networks. A prime example is the circadian clock: although the conserved TFs CLOCK (CLK) and CYCLE (CYC) control a transcriptional circuit throughout animal bodies, rhythms in behavior and physiology are generated tissue specifically. Yet, how CLK and CYC determine tissue-specific clock programs has remained unclear.
Results: Here, we use a functional genomics approach to determine the *cis*-regulatory requirements for clock specificity. We first determine CLK and CYC genome-wide binding targets in heads and bodies by ChIP-seq and show that they have distinct DNA targets in the two tissue contexts. Computational dissection of CLK/CYC context-specific binding sites reveals sequence motifs for putative partner factors, which are predictive for individual binding sites. Among them, we show that the opa and GATA motifs, differentially enriched in head and body binding sites respectively, can be bound by OPA and SERPENT (SRP). They act synergistically with CLK/CYC in the *Drosophila* feedback loop, suggesting that they help to determine their direct targets and therefore orchestrate tissue-specific clock outputs. In addition, using in vivo transgenic assays, we validate that GATA motifs are required for proper tissue-specific gene expression in the adult fat body, midgut, and Malpighian tubules, revealing a *cis*-regulatory signature for enhancers of the peripheral circadian clock.
Conclusions: Our results reveal how universal clock circuits can regulate tissue-specific rhythms and, more generally, provide insights into the mechanism by which universal TFs can be modulated to drive tissue-specific programs of gene expression.

Introduction

The basis for multicellular life is the ability to create functional specialization and distinct cell types through differentially controlled gene expression. This is achieved by gene regulatory networks controlled by transcription factors (TFs) [1]. TFs are regulatory proteins that bind to specific DNA *cis*-regulatory sequences (motifs) within enhancers of target genes and activate and repress expression. Typically, enhancer activity is determined by defined sets of TFs, and different TF combinations modulate the activity of individual factors that often function more broadly [2–11].

The circadian clock is an important example of a transcriptional circuit that functions broadly in animals but is modulated in a cell type-specific fashion [12]. Eukaryotic circadian clocks are governed by transcriptional negative feedback loops that control daily rhythms in gene expression, ultimately leading to cyclic output of behavior, metabolism, and physiology [13, 14]. In *Drosophila melanogaster*, circadian rhythms are controlled by transcriptional negative feedback loops interconnected by the TFs CLOCK (CLK) and CYCLE (CYC) [13, 15–17]. These form a heterodimer (CLK/CYC) that binds to E box sequences (CACGTG) upstream of the *period* (*per*) and *timeless* (*tim*) genes. PER and TIM dimerize and feed back on their own regulation by inhibiting CLK/CYC activity [17, 18]. Besides generating the pacemaker, this core is linked to downstream outputs in part by the CLK/CYC-mediated induction of downstream genes [19–23].

Interestingly, CLK and CYC are present almost ubiquitously in *Drosophila* [24, 25], yet they produce cell type-specific molecular and physiological outputs [26]. For example, the clock controls rhythms in locomotor activity behavior in the brain [27–29], odor receptivity in antennae [30, 31], and expression of metabolic enzymes in the fat body [32, 33]. Furthermore, circadian gene expression in fly heads versus bodies was found to be largely nonoverlapping [21]. Similarly, in mammals, only a small overlap exists between genes that cycle in the suprachiasmatic nucleus (SCN, the central clock in the brain), liver, and heart [14, 34–36]. Since circadian physiology is tissue specific and in part controlled by CLK/CYC-mediated gene expression, local TFs must modulate CLK/CYC and their targets in a context-specific manner. However, putative partner TFs involved in circadian clock tissue specification have remained unclear.

To identify sequence motifs and the corresponding TFs that define CLK/CYC context-specific binding and function, we epitope tagged the *Clk* and *cyc* genes in *Drosophila* by homologous recombination and identified CLK and CYC binding sites in fly heads and bodies by chromatin immunoprecipitation followed by massive parallel sequencing (ChIP-seq). Although CLK and CYC shared most binding sites, the sites differed between heads and bodies, consistent with tissue-specific functions. We identified motifs that were predictive for head- and body-specific CLK/CYC binding, including the motifs for the TFs *odd paired* (*opa*) and *serpent* (*srp*), respectively. We validated that both motifs were able to synergistically enhance transcription with CLK/CYC, and that SRP's GATA motif was required for enhancer activity in vivo. Our results suggest that tissue-specific circadian gene expression is achieved by the specific redirection of CLK/CYC to context-specific targets by motifs of partner TFs such as OPA and SRP.

Results

Clk^{tag} and *cyc*^{tag} Flies

To enable the identification of CLK and CYC binding sites by ChIP, we added a peptide tag including a V5 epitope to the 3′ terminus of the *Clk* and *cyc* coding sequences at their

[2]Present address: Laboratory of Systems Biology and Genetics, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
[3]Present address: Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland
*Correspondence: stark@starklab.org

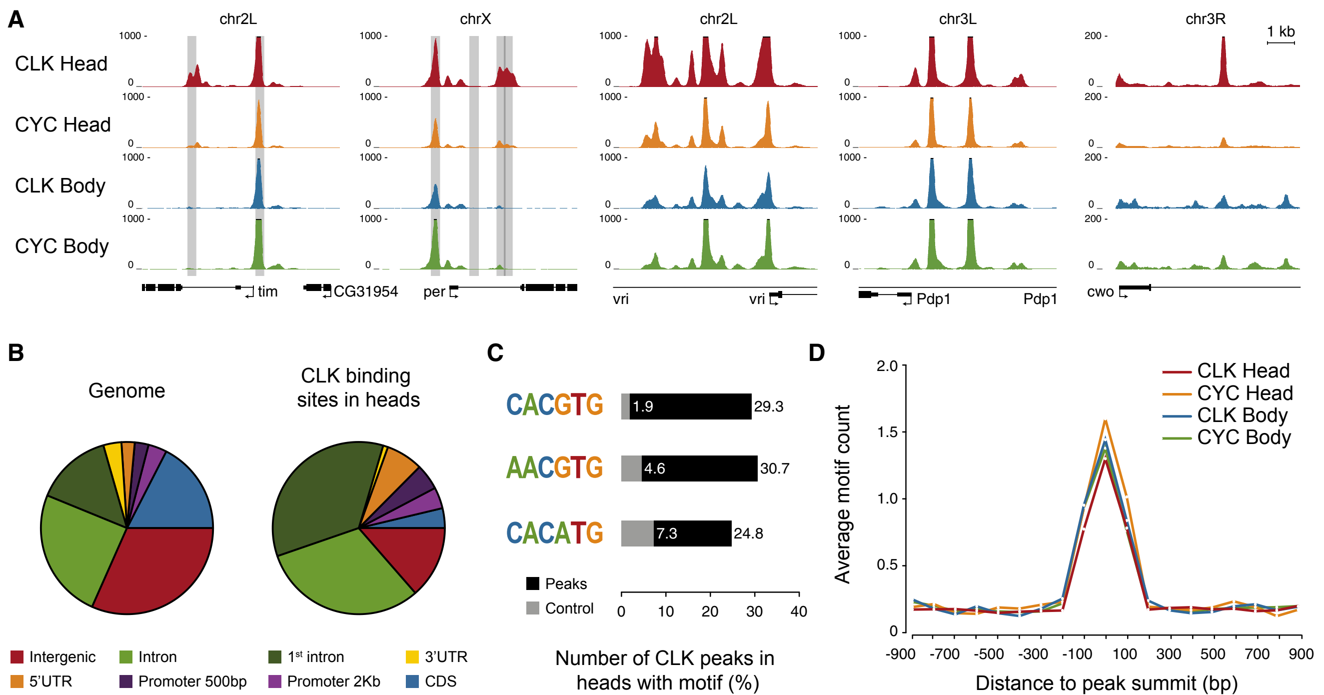

Figure 1. Genome-wide DNA Binding Profile of CLK and CYC

(A) UCSC Genome Browser screenshot with CLK and CYC fragment densities for head and body samples at known clock targets (gray shadings indicate known enhancers [44]).

(B) Genomic distribution of CLK peaks in heads (see Figure S2 for other samples).

(C) The three 6-mers most enriched in CLK head peaks are E boxes (see Figure S2 for other samples; gray shows the random occurrence of each motif).

(D) Distribution of E box motif instances across a 1.8 kb window (x axis) centered on the ChIP-seq peak summits for all CLK/CYC binding sites in heads or bodies.

endogenous genomic loci by ends-out homologous recombination [37–39] (see Figure S1 available online). This ensures that the tagged TFs are expressed at physiological levels, which is important, as increased *Clk* expression has been shown to shorten the locomotor period [40]. Additionally, it allows ChIP under identical conditions and facilitates comparisons (although tag accessibility might differ between proteins). We isolated one knockin line for each locus (*Clk*^{tag} and *cyc*^{tag}), for which we confirmed the correct integration by Southern blot and PCR analyses (Figures S1A and S1B) and the function of the circadian clock machinery by behavioral monitoring of the flies' locomotor activity rhythms (Figure S1C).

Genome-wide DNA Binding Profile of CLK and CYC

We performed ChIP-seq using a V5 antibody from heads and bodies of decapitated homozygous *Clk*^{tag} and *cyc*^{tag} flies at zeitgeber time (ZT) 13, at which CLK binding is close to its maximum [41] (ZT0 = lights-on; ZT12 = lights-off). We sequenced ChIP and input samples, as well as mock controls (V5-ChIP from wild-type flies without the V5 tag), for two independent biological replicates each. The replicates from independent fly collections were highly similar, with Pearson correlation coefficients (PCCs) between 0.83 and 0.88 (Table S1). We identified binding sites (peaks) using peakzilla [42] with stringent thresholds and required that both biological replicates concurred (as in [43]). This yielded 2,059 and 436 candidate peaks for CLK and CYC in heads versus 431 and 484 in bodies, respectively. In contrast, the mock ChIP resulted in only 62 versus 7 candidates, respectively, indicating false discovery rates (FDRs) between at most 1% and 14%. To exclude that antibody cross-reactivity influenced our results, we corrected the candidate peaks based on the mock results, yielding 1,959 final binding sites for CLK in heads, 369 for CYC in heads, 425 for CLK in bodies, and 481 for CYC in bodies.

Among the CLK binding sites reported previously [41], 58.2% and 59.1% overlapped with peaks in heads for CLK and CYC, respectively. We found multiple peaks near the core clock components *per*, *tim*, *vri*, *cwo*, and *Pdp1*, where they often coincided with previously described E boxes [44] (Figure 1A; Figure S2A). CLK and CYC had strong overlapping signals in both heads and bodies, in agreement with their role as heterodimers in the clock pacemaker [17, 45]. As expected, CLK and CYC peaks for head and body samples were enriched in 5′ UTRs, promoters, and introns, while they were depleted in coding sequences (Figure 1B; Figure S2B).

The most enriched 6-mer motifs in the CLK and CYC binding sites corresponded to the established CLK/CYC E box motif, but two variants were also found to be frequent (Figure 1C; Figure S2C). Collectively, these E boxes were found in more than 88% of all binding sites and showed a distinct enrichment at the ChIP peak summit (Figure 1D; Figure S2C). Taken together, these results suggest that we identified CLK and CYC binding sites at potential regulatory regions in heads and bodies genome-wide.

CLK and CYC Exhibit Context-Specific Binding

To study CLK/CYC in different tissues, we analyzed the overlap between their binding sites in heads versus bodies and defined shared and differential peaks across the different

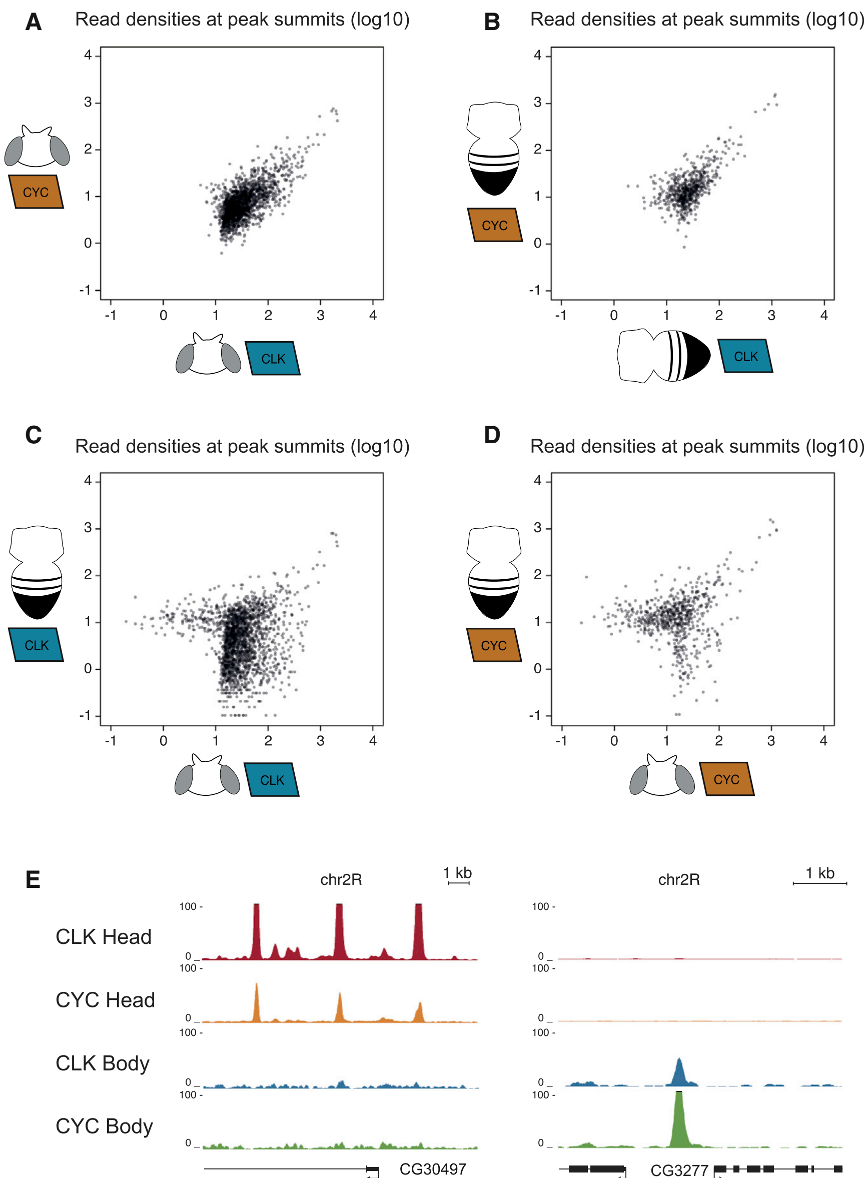

Figure 2. Context Specificity of CLK and CYC Binding Sites

(A–D) Scatterplot of fragment densities at peak summits of CLK versus CYC in heads (A), bodies (B), and heads versus bodies of CLK (C) and CYC (D).

(E) Examples of head- and body-specific binding sites (UCSC Genome Browser tracks).

conditions. CLK and CYC binding profiles were highly similar (Figures 2A and 2B), with PCCs above 0.88 (Table S1), suggesting that they work primarily as heterodimers to activate core clock components as well as downstream targets. In contrast, the binding profiles between heads and bodies differed substantially for both factors (Figures 2C–2E): more than 30% of the peaks found in heads were specific to heads, and more than 20% of peaks found in bodies were specific to bodies (e.g., Figure 2E); PCCs confirmed this disparity (PCCs between 0.38 and 0.72, below the PCC of the corresponding inputs [0.89]). This striking difference suggests that tissue-specific differences in circadian expression might be at least partly caused by differential CLK/CYC binding.

To test this hypothesis, we defined stringent classes of CLK and CYC binding sites that were shared, head-specific, or body-specific (Figure 3A; Table S2; leaving a large "gray zone" of sites not assigned to any class; see Supplemental Experimental Procedures); assigned each peak to the closest gene transcriptional start site (TSS); and assessed the genes' functions according to gene ontology (GO; [46]). Twenty-five percent of all peaks were shared and showed high enrichments for terms related to gene regulation, such as "regulation of transcription" and "transcription factor activity," suggesting that CLK/CYC are at the top of a gene regulatory hierarchy across tissues. Indeed, 19.8% of the shared peaks lie close to TFs, a 5-fold enrichment compared to all genes ($p = 2.36 \times 10^{-9}$). In contrast, we observed striking differences between the other sets: head-specific peaks were enriched in gene functions related to behavior and vision, while body-specific peaks were consistent with functions in metabolism (Figure 3B; Figure S3 for full list).

To test whether these differences were also reflected at the expression level, we first compared the CLK/CYC binding sites with gene expression according to FlyAtlas [25]. While shared CLK/CYC binding sites were close to genes expressed in head and body tissues, head- and body-specific peaks were located close to genes expressed either in heads or bodies (Figure S4). Next, we compared CLK/CYC binding sites with data sets of cycling mRNAs [21]. We found genes close to head-specific peaks to be 3.4-fold enriched in genes cycling in heads ($p < 0.05$), while genes close to body-specific peaks were 5.2-fold enriched in genes cycling in bodies ($p = 0.057$; note that the total of only 12 genes that cycled in heads and bodies precluded an analysis). Taken together, these results show that, in addition to their role as master regulators in a TF hierarchy, CLK/CYC also bind directly to downstream targets in a cell type-specific manner, suggesting that they drive tissue-specific programs.

Differential Motif Content is Predictive of Context-Specific CLK and CYC Binding

The presence of E boxes alone cannot explain how, in each cell type, CLK/CYC are recruited to distinct targets. As partner TFs can help to define context-specific binding [3, 5, 6, 10, 47], we searched for TF motifs that were differentially enriched between head and body binding sites. We found the "orphan" motif ME50 [48], opa, ME134/odd, and Adf1 to be enriched in head-specific binding sites, while the motifs bab1, TATA/Mef2, ME3, GATA, and Hox were enriched in body-specific binding sites (Figure 4A).

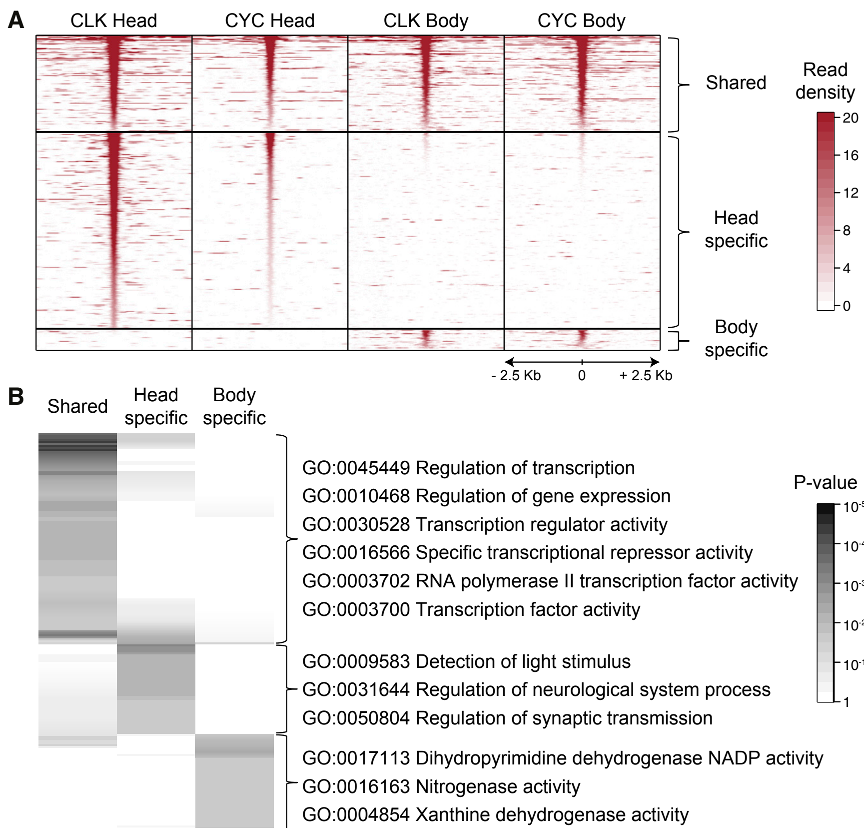

Figure 3. Context-Specific CLK/CYC Function

(A) Fragment density heatmaps for 5 kb windows centered on CLK/CYC peak summits (x axis).

(B) Genes next to shared, head-specific, or body-specific peaks are enriched in specific GO categories compared to all genes near peaks (unsupervised clustered heatmap of hypergeometric enrichment p values, highlighting a selection of GO categories; see Figure S3 for all categories).

To test whether combinations of differentially distributed motifs allow the discrimination of head versus body binding sites, we compared these sequences using a predictive binary classification framework [47]. Using sequence motif content alone, head and body peaks could be accurately distinguished using leave-one-out cross-validation (82% of peaks correctly classified; area under the receiver operating characteristic [ROC] curve [AUC] = 0.91; Figures 4B and 4C). This indicates that partner TF motifs surrounding CLK/CYC binding sites carry information indicative of binding and have the potential to determine context-specific clock target genes and function. It further suggests that the corresponding TFs could be novel tissue-specific CLK/CYC partner factors.

Identifying the *cis*-Regulatory Requirements of Individual Binding Sites

To assess the importance of each particular TF motif for each individual CLK/CYC binding site, we tested which one of them would affect the classification of the site toward the head or body classes after deleting them in silico [47]. Forty-four percent of the sites in heads could no longer be confidentially predicted after removal of opa motifs, whereas deletion of the schlank and Adf1 motifs affected 32% and 26% of the peaks, respectively (Figure 4D). For bodies, deletion of the GATA motif impaired the predictions of all sites, consistent with its enrichment at body peak summits (Figure 4E).

SRP and OPA Synergistically Activate CLK/CYC-Mediated Expression

Many aspects of both fly and mammalian circadian biochemistry can be simulated in *Drosophila* Schneider 2 (S2) cells [17, 49–54]. We therefore performed transactivation luciferase assays in S2 cells to test whether the CLK/CYC bound sequences can enhance transcription when activated by CLK/CYC and the predicted partner TFs.

We expected the GATA motif to be recognized by GATA factors, a family of TFs involved in different aspects of animal development and physiology [55, 56]. Indeed, a body-specific CLK/CYC binding site in the first intron of the CG34386 gene (hereafter "intCG34386"; Figure 5A) was activated by the GATA factor serpent (SRP) (Figure 5B). A second site showed similar results, while a third was negative for all tested GATA factors (Figures S5A–S5D). Interestingly, *srp* is indeed coexpressed with *Clk* and *cyc* in vivo in the *Drosophila* adult fat body [25].

Because S2 cells express *cyc*, *Clk* transfection is sufficient to activate E box-containing enhancers [17, 49, 52]. Cotransfection of *srp* and *Clk* induced expression 11-fold (Figure 5C), and this activation was reduced by 57% when we added *per*, the primary inhibitor of CLK/CYC activity in the circadian pacemaker in *Drosophila* [51]. Moreover, SRP-mediated CLK/CYC activation depended on wild-type CLK function and was not observed with the inactive CLK^{Jrk} mutant [15] (Figure 5C). Finally, activation by CLK and SRP was also abolished when either E box or GATA motifs were mutated (Figure 5D; Figures S5I and S5J).

We also tested three head-specific CLK/CYC binding sites for OPA responsiveness, including a site in the intron of *Slob* (*Slowpoke binding protein*), a putative regulator of the calcium-sensitive potassium channel *slowpoke* that is involved in the output of the molecular pacemaker in the brain [57]. This intronic site (hereafter "intSlob") showed a dose-dependent response to CLK only when OPA was present (Figure 5F; Figures S5E–S5H for the two other sites). As *opa*, *Clk*, and *cyc* are coexpressed in the *Drosophila* brain [25], OPA might indeed be a CLK/CYC partner for head-specific enhancers.

GATA-Containing CLK and CYC Body-Specific Binding Sites Act as Body-Specific Enhancers In Vivo

To test the function of body-specific CLK/CYC binding sites and the importance of the GATA motif in vivo, we selected nine sites based on the importance of the GATA motif for their correct classification toward bodies, but independent of neighboring genes. We cloned them upstream of a transcriptional reporter (Gal4), integrated the construct site specifically into the *Drosophila* genome, and analyzed Gal4 expression by crossing the transgenic fly lines to UAS-CD8:GFP flies. Eight

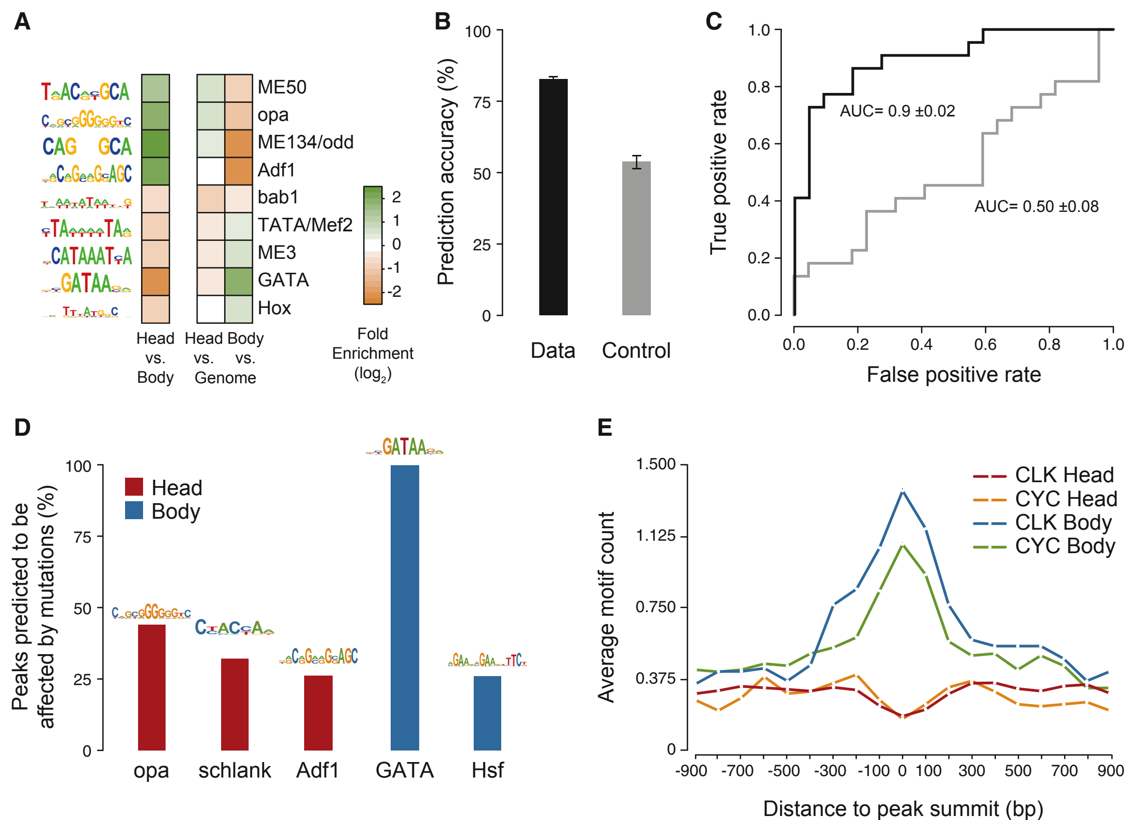

Figure 4. Context-Specific CLK and CYC Binding Sites Are Characterized by Specific TF Motif Content

(A) Heatmap showing motifs differentially enriched ($p \leq 0.01$) between head- and body-specific CLK/CYC binding sites (left) and the enrichment compared to the genomic average (right).

(B) Prediction accuracy (binary classification) for head- and body-specific CLK/CYC binding sites based on motif content (control: same after randomly assigning the sites to the head- or body-specific classes; error bars indicate standard deviations from three rounds of prediction).

(C) Receiver operating characteristic (ROC) curves and area under the curve (AUC) for the predictions (black) and controls (gray).

(D) Percent of head- (red) and body-specific (blue) binding sites that were affected by the in silico mutation of each of five motif types that were required for the predictions in (B) and (C).

(E) Distribution of GATA motif instances relative to the CLK and CYC peak summits.

out of the nine candidate *cis*-regulatory modules (CRMs) drove GFP expression in diverse tissues in the *Drosophila* body in which the circadian clock is important [26] (Table 1; Figure 6G; Figure S6A). Two lines were active in the fat body (CRMs 1 and 3), six in the midgut (CRMs 2, 4, 5, 6, 7 and 9), and two in Malpighian tubules (CRMs 4 and 9). Only two CRMs (3 and 9) also drove expression in the head, confirming that body-specific CLK/CYC binding sites predominantly correspond to body-specific enhancers.

To test whether the CRMs' activity depended on E box and GATA motifs as predicted, we also cloned variants with point mutations that disrupted both types of motifs. For two CRMs, one active in midgut (CRM5) and one in head and abdominal fat body (CRM3), GFP signals were not substantially different when E box motifs were mutated (probably due to the stability of GFP) but GAL4 reporter transcript levels were reduced as determined by quantitative PCR (Figures S6B and S6C). The mutation of GATA motifs severely reduced enhancer activity for both enhancers, even at the level of the GFP signal (Figures S6B and S6C). We therefore analyzed the requirement for GATA motifs in a tissue-specific manner in vivo for all eight active CRMs. While the activity of CRM1 and CRM2 was completely abolished, GATA mutations had more localized effects for CRMs active in more than one tissue: for CRM6 and CRM7, activity was impaired in midgut, but not in accessory glands (CRM6) or ovaries (CRM7). Similarly, the GATA mutant CRM4 was still active in midgut, but not in Malpighian tubules. Finally, mutating the GATA motifs of CRM9 abolished activity exclusively in the midgut (Figure S6D). These results suggest that the GATA motif is important for enhancer activity in the midgut, fat body, and Malpighian tubules and plays a fundamental role for tissue-specific enhancer activity of CLK/CYC bound regions in vivo.

Discussion

Although frequently not restricted to single cell types, individual TFs can control tissue-specific programs of gene expression through interactions with local TF networks [7, 10]. But despite substantial progress in identifying differential cell-specific circadian expression programs [21, 23, 33, 58–60], how CLK and CYC interact with local TF networks to generate output rhythms tissue specifically is still elusive.

Here, we used an integrative genomics approach to shed light on how the circadian clock drives tissue-specific gene expression. While shared CLK/CYC binding sites could not be explained by combinations of head- and body-specific motifs (yet were slightly more enriched in E box motifs

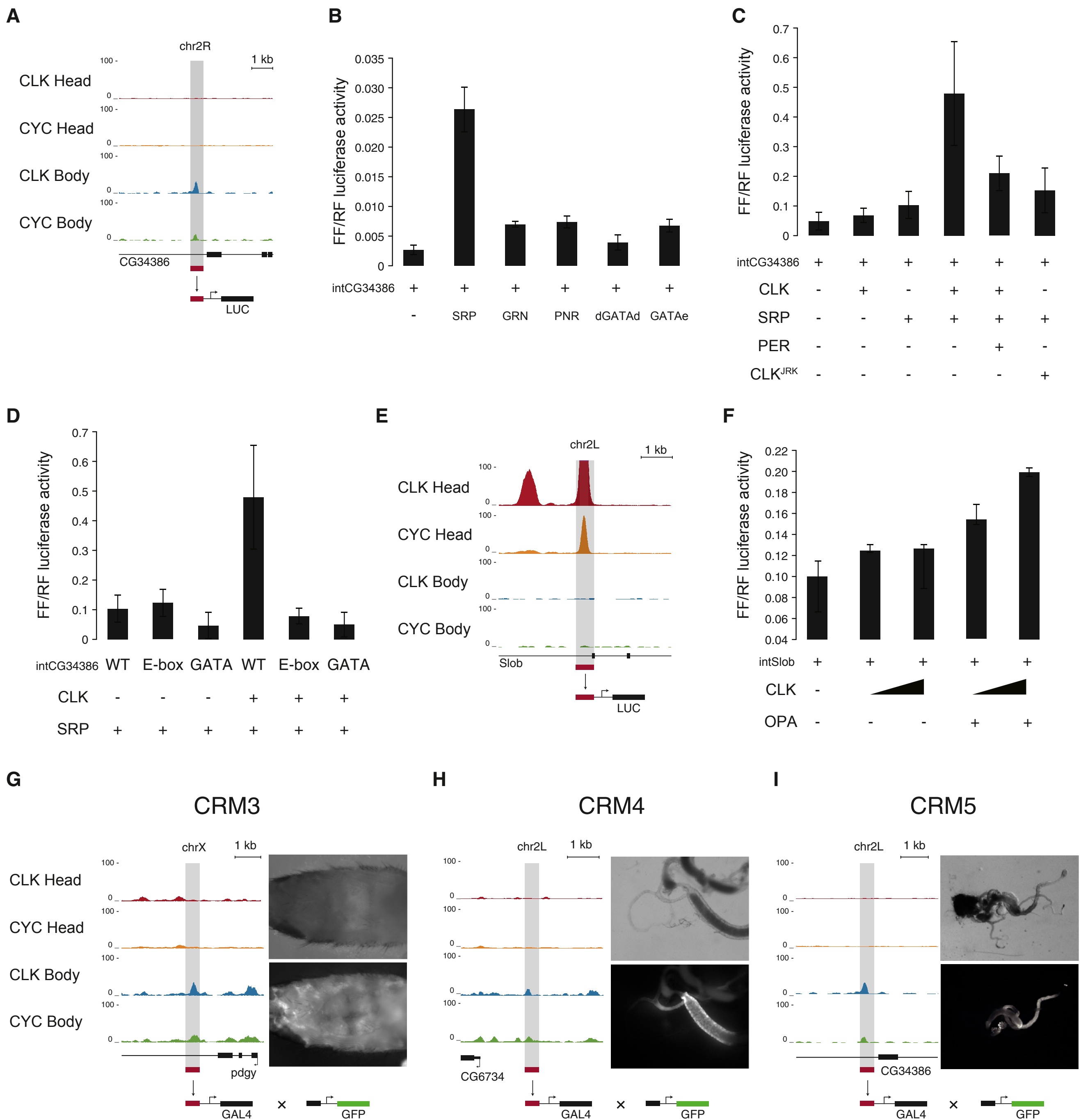

Figure 5. SRP and OPA Activate Transcription Synergistically with CLK/CYC

(A) UCSC Genome Browser screenshot for a CLK/CYC binding site in bodies (red bar indicates the tested region intCG34386).

(B) Normalized luciferase activity (Firefly versus Renilla, FF/FR) of extracts from S2 cells transiently transfected with 10 ng of intCG34386-LUC reporter gene and 50 ng of expression vectors for the indicated TFs or empty vector (–). Error bars in (B), (C), (D), and (F) show standard deviations of three independent transfections (B and F) or three independent experiments with three independent transfections each (C and D).

(C) Normalized activity for 10 ng of intCG34386-LUC with cotransfections of 2 ng of CLK or CLK^{Jrk} and 10 ng of SRP and/or PER vectors as indicated.

(D) Same as (C), but for E box or GATA mutant versions of intCG34386.

(E) Genome Browser screenshot for a CLK/CYC binding site in heads.

(F) Same as (C), but for 10 ng of int*Slob*-LUC reporter gene, 5 or 10 ng of CLK, 2.5 ng of OPA, or empty vector (–).

(G–I) UCSC Genome Browser screenshots for CLK/CYC-specific binding site in bodies (left; red bars indicate the tested regions CRM 3, 4, and 5). CRM activity was assayed by GFP fluorescence with transgenic reporters in adult flies (right). Images show bright-field (top) and GFP signals (bottom) for the fat body (CRM3), midgut and Malpighian tubules (CRM4), and midgut only (CRM5; see Figure S6 for all tested CRMs).

Table 1. GFP Spatial Expression in Transgenic Adult Flies

		GFP Localization					
Fly Line	Genomic Coordinates	Head	Midgut	Fat Body	Malpighian Tubules	MRT	FRT
CRM1	chr3L:4099109–4099771	–	–	+	–	–	–
CRM2	chr3R:14914965–14915499	–	+	–	–	–	–
CRM3	chrX:14843135–14843666	+[a]	–	+	–	–	–
CRM4	chr2L:12022549–12023056	–	+	–	+	–	–
CRM5	chr2R:13976794–13977305	–	+	–	–	–	–
CRM6	chrX:19134232–19134636	–	+	–	–	+	–
CRM7	chrX:16988629–16989153	–	+	–	–	–	+
CRM8	chr2R:19555521–19556095	–	–	–	–	–	–
CRM9	chr2R:16138228–16138731	+[b]	+	–	+	+	+

MRT, male reproductive tract; FRT, female reproductive tract.
[a]Head fat body expression.
[b]Antennae, labellum, and some parts of the brain.

[1.4-fold; $p < 0.004$] and—similar to highly occupied target (HOT) regions [61]—in Trithorax-like motifs [Trl/GAGA; 2-fold; $p < 1.8 \times 10^{-10}$]), a substantial number of CLK and CYC binding sites were specific to either heads or bodies and next to genes with different functional GO categories (Figure 3B; Figure S3). These binding sites differed substantially in their motif content, and this motif signature was predictive of context-specific CLK/CYC binding (Figure 4), suggesting that tissue-specific clock targets are determined by the binding site sequences.

GATA motifs were enriched in CLK/CYC binding sites in bodies and required for enhancer activity in the fat body, midgut, and Malpighian tubules. This suggests that GATA factors might play a key role for CLK/CYC-bound enhancers in bodies, potentially by helping to establish the chromatin landscape in tissues where they are specifically expressed (e.g., *srp* in the fat body and *GATAe* in the gut [25]). Interestingly, GATA motifs are also overrepresented in promoter regions of circadian genes in rodents [62], suggesting a conserved role for GATA factors in the circadian clock.

Here, we found that the GATA factor SRP could act synergistically with CLK, suggesting that it is an important determinant of clock function in peripheral tissues (Figures 5E and 5F). SRP has multiple functions in *Drosophila*, including the control of endodermal development and hematopoiesis in the embryo and the induction of immune response in the larval fat body [63–67]. Interestingly, *srp* is coexpressed with CLK and CYC in the fat body [25, 68], a tissue with roles in metabolic activity, innate immunity response, and detoxification [69–71]—all known to be controlled in a circadian manner [32, 72–74]. CLK body-specific peaks were 4.17-fold enriched ($p < 0.00001$) close to cycling fat body genes [1, 33], suggesting that *srp* might help determine the physiological outputs controlled by the fat body pacemaker (Figure 6). Interestingly, *srp* is also required for hormone-induced expression of the Fbp1 TF during fat body development, supporting the idea that it might be important for temporal or inducible regulation more generally [75].

Similarly, OPA, which belongs to the Zic family of mammalian TFs with conserved roles in head formation in flies and mammals [76, 77], is coexpressed with *Clk* and *cyc* in the adult brain [25]. In addition, an enhancer of *Slob*, an output gene of the clock pacemaker involved in the generation of locomotor activity rhythms [57, 78], responded to CLK and CYC in an OPA-dependent manner (Figure 5F), suggesting that OPA might be involved in the recruitment of CLK/CYC to regulate genes controlling fly behavior. Further studies on OPA and additional predicted partner TFs might provide new insights into the *Drosophila* clock in the head.

It is likely that different cofactors with functions equivalent to *srp* or *opa* exist in different cell types, which redirect CLK/CYC to tissue-specific binding sites and allow tissue-specific gene regulation. Indeed, we have identified several other motifs that are tissue-specifically enriched. This is reminiscent of studies showing that TFs downstream of signaling pathways are redirected in a tissue-specific manner by cell-specific master regulators [2–8, 10, 11, 79]. Our results might thus constitute an important example of how partner TFs adapt broadly active transcriptional regulators to achieve tissue-specific gene expression and function, contributing to a better understanding of gene regulatory networks more generally.

Our data on CLK/CYC binding in different contexts not only provide novel insights into clock regulatory networks and enhancer structure but also exemplify a new strategy to uncover cofactors of the circadian clock via their *cis*-regulatory motifs. Our approach is complementary to forward and reverse genetics or biochemistry, which have traditionally been used to reveal clock factors. It can also be applied more generally to identify factors that recruit broadly expressed TFs in different cell types or tissues. In addition, the

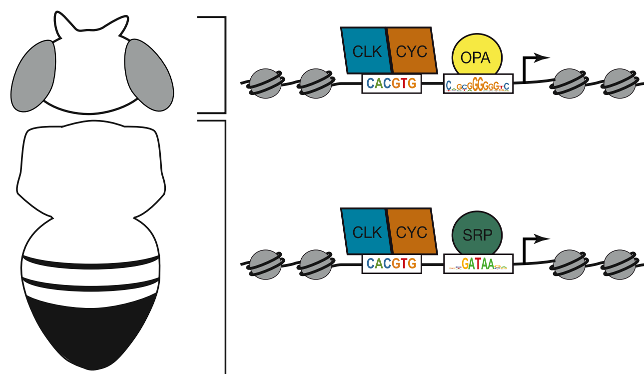

Figure 6. Model of Context-Specific CLK/CYC Binding and Tissue-Specific Transcriptional Programs

Schematic representation of the proposed model. CLK and CYC co-occupy the genome in heads and bodies, but CLK/CYC-specific binding in the head and body is determined with the help of OPA and SRP, respectively. Additional partner factors are likely also involved, and similar mechanisms likely apply to other broadly expressed TFs (not shown).

tagging of endogenous loci allows the study of TFs under physiological conditions in their endogenous expression domains, which is crucial especially for TFs that have large and complex regulatory regions and/or for which physiological expression levels are of fundamental importance. In summary, our results in the *Drosophila* circadian clock reveal how universal TF circuits can be modulated to generate transcriptional tissue-specific outputs and demonstrate a novel approach to determine regulatory partners more generally.

Accession Numbers

The NCBI GEO accession number for the deep sequencing data reported in this paper is GSE40467.

Supplemental Information

Supplemental Information includes two tables, six figures, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2013.11.017.

Acknowledgments

We would like to thank Michaela Pagani, Katharina Schernhuber, Martina Rath, Daria Shlyueva, Christian Machacek, and IMP/IMBA services for technical support and Johannes Tkadletz for help with Figure 6 and Figure S1. We are grateful to Korneel Hens, Bart Deplancke (EPFL, Lausanne), and Frank Weber (Biochemiezentrum der Universität Heidelberg) for plasmids and Carla Margulies and Andreas Ladurner (Ludwig Maximilian University) for sharing the chromatin preparation protocol prior to publication. Illumina sequencing was performed at the CSF NGS Unit (http://csf.ac.at/). A.C.A.M.-F. was supported by an EMBO long-term postdoctoral fellowship, and A.F.B and J.O.Y.-C. were supported by the Austrian Ministry for Science and Research through the GEN-AU Bioinformatics Integration Network III. Research in the Stark group is supported by a European Research Council (ERC) Starting Grant from the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement number 242922. Basic research at the IMP is supported by Boehringer Ingelheim GmbH.

Received: July 19, 2013
Revised: September 24, 2013
Accepted: November 6, 2013
Published: December 12, 2013

References

1. Stathopoulos, A., and Levine, M.S. (2005). Genomic regulatory networks and animal development. Dev. Cell *9*, 449–462.
2. Zeitlinger, J., Simon, I., Harbison, C.T., Hannett, N.M., Volkert, T.L., Fink, G.R., and Young, R.A. (2003). Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. Cell *113*, 395–404.
3. Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.-B., Reynolds, D.B., Yoo, J., et al. (2004). Transcriptional regulatory code of a eukaryotic genome. Nature *431*, 99–104.
4. Mullen, A.C., Orlando, D.A., Newman, J.J., Lovén, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF-β signaling. Cell *147*, 565–576.
5. Buck, M.J., and Lieb, J.D. (2006). A chromatin-mediated mechanism for specification of conditional transcription factor targets. Nat. Genet. *38*, 1446–1451.
6. Trompouki, E., Bowman, T.V., Lawton, L.N., Fan, Z.P., Wu, D.-C., DiBiase, A., Martin, C.S., Cech, J.N., Sessa, A.K., Leblanc, J.L., et al. (2011). Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. Cell *147*, 577–589.
7. Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell *132*, 958–970.
8. Zinzen, R.P., Girardot, C., Gagneur, J., Braun, M., and Furlong, E.E.M. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. Nature *462*, 65–70.
9. Wilczynski, B., and Furlong, E.E.M. (2010). Challenges for modeling global gene regulatory networks during development: insights from Drosophila. Dev. Biol. *340*, 161–169.
10. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.
11. Palii, C.G., Perez-Iratxeta, C., Yao, Z., Cao, Y., Dai, F., Davison, J., Atkins, H., Allan, D., Dilworth, F.J., Gentleman, R., et al. (2011). Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages. EMBO J. *30*, 494–509.
12. Zhang, E.E., and Kay, S.A. (2010). Clocks not winding down: unravelling circadian networks. Nat. Rev. Mol. Cell Biol. *11*, 764–776.
13. Hardin, P.E. (2011). Molecular genetic analysis of circadian timekeeping in Drosophila. Adv. Genet. *74*, 141–173.
14. Mohawk, J.A., Green, C.B., and Takahashi, J.S. (2012). Central and peripheral circadian clocks in mammals. Annu. Rev. Neurosci. *35*, 445–462.
15. 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.
16. 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.
17. 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.
18. Lee, C., Bae, K., and Edery, I. (1999). PER and TIM inhibit the DNA binding activity of a Drosophila CLOCK-CYC/dBMAL1 heterodimer without disrupting formation of the heterodimer: a basis for circadian transcription. Mol. Cell. Biol. *19*, 5316–5325.
19. McDonald, M.J., and Rosbash, M. (2001). Microarray analysis and organization of circadian gene expression in Drosophila. Cell *107*, 567–578.
20. 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.
21. 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. *22*, 9305–9319.
22. Ueda, H.R., Matsumoto, A., Kawamura, M., Iino, M., Tanimura, T., and Hashimoto, S. (2002). Genome-wide transcriptional orchestration of circadian rhythms in Drosophila. J. Biol. Chem. *277*, 14048–14052.
23. Keegan, K.P., Pradhan, S., Wang, J.-P., and Allada, R. (2007). Meta-analysis of Drosophila circadian microarray studies identifies a novel set of rhythmically expressed genes. PLoS Comput. Biol. *3*, e208.
24. Plautz, J.D., Kaneko, M., Hall, J.C., and Kay, S.A. (1997). Independent photoreceptive circadian clocks throughout Drosophila. Science *278*, 1632–1635.
25. Chintapalli, V.R., Wang, J., and Dow, J.A.T. (2007). Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat. Genet. *39*, 715–720.
26. Tomioka, K., Uryu, O., Kamae, Y., Umezaki, Y., and Yoshii, T. (2012). Peripheral circadian rhythms and their regulatory mechanism in insects and some other arthropods: a review. J Comp. Physiol. B *182*, 729–740.
27. Grima, B., Chélot, E., Xia, R., and Rouyer, F. (2004). Morning and evening peaks of activity rely on different clock neurons of the Drosophila brain. Nature *431*, 869–873.
28. Stoleru, D., Peng, Y., Agosto, J., and Rosbash, M. (2004). Coupled oscillators control morning and evening locomotor behaviour of Drosophila. Nature *431*, 862–868.
29. Stoleru, D., Peng, Y., Nawathean, P., and Rosbash, M. (2005). A resetting signal between Drosophila pacemakers synchronizes morning and evening activity. Nature *438*, 238–242.
30. Krishnan, B., Dryer, S.E., and Hardin, P.E. (1999). Circadian rhythms in olfactory responses of Drosophila melanogaster. Nature *400*, 375–378.

31. Tanoue, S., Krishnan, P., Krishnan, B., Dryer, S.E., and Hardin, P.E. (2004). Circadian clocks in antennal neurons are necessary and sufficient for olfaction rhythms in Drosophila. Curr. Biol. *14*, 638–649.
32. Xu, K., Zheng, X., and Sehgal, A. (2008). Regulation of feeding and metabolism by neuronal and peripheral clocks in Drosophila. Cell Metab. *8*, 289–300.
33. Xu, K., DiAngelo, J.R., Hughes, M.E., Hogenesch, J.B., and Sehgal, A. (2011). The circadian clock interacts with metabolic physiology to influence reproductive fitness. Cell Metab. *13*, 639–654.
34. Akhtar, R.A., Reddy, A.B., Maywood, E.S., Clayton, J.D., King, V.M., Smith, A.G., Gant, T.W., Hastings, M.H., and Kyriacou, C.P. (2002). Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. Curr. Biol. *12*, 540–550.
35. Panda, S., Antoch, M.P., Miller, B.H., Su, A.I., Schook, A.B., Straume, M., Schultz, P.G., Kay, S.A., Takahashi, J.S., and Hogenesch, J.B. (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. Cell *109*, 307–320.
36. Storch, K.-F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F.C., Wong, W.H., and Weitz, C.J. (2002). Extensive and divergent circadian gene expression in liver and heart. Nature *417*, 78–83.
37. Gong, W.J., and Golic, K.G. (2003). Ends-out, or replacement, gene targeting in Drosophila. Proc. Natl. Acad. Sci. USA *100*, 2556–2561.
38. Venken, K.J.T., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747–1751.
39. Huang, J., Zhou, W., Watson, A.M., Jan, Y.-N., and Hong, Y. (2008). Efficient ends-out gene targeting in Drosophila. Genetics *180*, 703–707.
40. Kadener, S., Menet, J.S., Schoer, R., and Rosbash, M. (2008). Circadian transcription contributes to core period determination in Drosophila. PLoS Biol. *6*, e119.
41. 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.
42. Bardet, A.F., Steinmann, J., Bafna, S., Knoblich, J.A., Zeitlinger, J., and Stark, A. (2013). Identification of transcription factor binding sites from ChIP-seq data at high resolution. Bioinformatics *29*, 2705–2713.
43. Bardet, A.F., He, Q., Zeitlinger, J., and Stark, A. (2012). A computational pipeline for comparative ChIP-seq analyses. Nat. Protoc. *7*, 45–61.
44. Taylor, P., and Hardin, P.E. (2008). Rhythmic E-box binding by CLK-CYC controls daily cycles in per and tim transcription and chromatin modifications. Mol. Cell. Biol. *28*, 4642–4652.
45. Cyran, S.A., Buchsbaum, A.M., Reddy, K.L., Lin, M.-C., Glossop, N.R.J., Hardin, P.E., Young, M.W., Storti, R.V., and Blau, J. (2003). vrille, Pdp1, and dClock form a second feedback loop in the Drosophila circadian clock. Cell *112*, 329–341.
46. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al.; The Gene Ontology Consortium (2000). Gene ontology: tool for the unification of biology. Nat. Genet. *25*, 25–29.
47. Yáñez-Cuna, J.O., Dinh, H.Q., Kvon, E.Z., Shlyueva, D., and Stark, A. (2012). Uncovering cis-regulatory sequence requirements for context-specific transcription factor binding. Genome Res. *22*, 2018–2030.
48. Stark, A., Lin, M.F., Kheradpour, P., Pedersen, J.S., Parts, L., Carlson, J.W., Crosby, M.A., Rasmussen, M.D., Roy, S., Deoras, A.N., et al.; Harvard FlyBase curators; Berkeley Drosophila Genome Project (2007). Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. Nature *450*, 219–232.
49. Ceriani, M.F., Darlington, T.K., Staknis, D., Más, P., Petti, A.A., Weitz, C.J., and Kay, S.A. (1999). Light-dependent sequestration of TIMELESS by CRYPTOCHROME. Science *285*, 553–556.
50. Shearman, L.P., Sriram, S., Weaver, D.R., Maywood, E.S., Chaves, I., Zheng, B., Kume, K., Lee, C.C., van der Horst, G.T.J., Hastings, M.H., and Reppert, S.M. (2000). Interacting molecular loops in the mammalian circadian clock. Science *288*, 1013–1019.
51. Chang, D.C., and Reppert, S.M. (2003). A novel C-terminal domain of drosophila PERIOD inhibits dCLOCK:CYCLE-mediated transcription. Curr. Biol. *13*, 758–762.
52. Weber, F., Hung, H.-C., Maurer, C., and Kay, S.A. (2006). Second messenger and Ras/MAPK signalling pathways regulate CLOCK/CYCLE-dependent transcription. J. Neurochem. *98*, 248–257.
53. Maurer, C., Hung, H.-C., and Weber, F. (2009). Cytoplasmic interaction with CYCLE promotes the post-translational processing of the circadian CLOCK protein. FEBS Lett. *583*, 1561–1566.
54. Hung, H.-C., Maurer, C., Zorn, D., Chang, W.-L., and Weber, F. (2009). Sequential and compartment-specific phosphorylation controls the life cycle of the circadian CLOCK protein. J. Biol. Chem. *284*, 23734–23742.
55. Patient, R.K., and McGhee, J.D. (2002). The GATA family (vertebrates and invertebrates). Curr. Opin. Genet. Dev. *12*, 416–422.
56. Gillis, W.Q., Bowerman, B.A., and Schneider, S.Q. (2008). The evolution of protostome GATA factors: molecular phylogenetics, synteny, and intron/exon structure reveal orthologous relationships. BMC Evol. Biol. *8*, 112.
57. 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 Neurosci. *5*, 3.
58. Nagoshi, E., Sugino, K., Kula, E., Okazaki, E., Tachibana, T., Nelson, S., and Rosbash, M. (2010). Dissecting differential gene expression within the circadian neuronal circuit of Drosophila. Nat. Neurosci. *13*, 60–68.
59. Kula-Eversole, E., Nagoshi, E., Shang, Y., Rodriguez, J., Allada, R., and Rosbash, M. (2010). Surprising gene expression patterns within and between PDF-containing circadian neurons in Drosophila. Proc. Natl. Acad. Sci. USA *107*, 13497–13502.
60. Hughes, M.E., Grant, G.R., Paquin, C., Qian, J., and Nitabach, M.N. (2012). Deep sequencing the circadian and diurnal transcriptome of Drosophila brain. Genome Res. *22*, 1266–1281.
61. Kvon, E.Z., Stampfel, G., Yáñez-Cuna, J.O., Dickson, B.J., and Stark, A. (2012). HOT regions function as patterned developmental enhancers and have a distinct cis-regulatory signature. Genes Dev. *26*, 908–913.
62. Bozek, K., Relógio, A., Kielbasa, S.M., Heine, M., Dame, C., Kramer, A., and Herzel, H. (2009). Regulation of clock-controlled genes in mammals. PLoS ONE *4*, e4882.
63. Rehorn, K.P., Thelen, H., Michelson, A.M., and Reuter, R. (1996). A molecular aspect of hematopoiesis and endoderm development common to vertebrates and Drosophila. Development *122*, 4023–4031.
64. Sam, S., Leise, W., and Hoshizaki, D.K. (1996). The serpent gene is necessary for progression through the early stages of fat-body development. Mech. Dev. *60*, 197–205.
65. Petersen, U.M., Kadalayil, L., Rehorn, K.P., Hoshizaki, D.K., Reuter, R., and Engström, Y. (1999). Serpent regulates Drosophila immunity genes in the larval fat body through an essential GATA motif. EMBO J. *18*, 4013–4022.
66. Lebestky, T., Chang, T., Hartenstein, V., and Banerjee, U. (2000). Specification of Drosophila hematopoietic lineage by conserved transcription factors. Science *288*, 146–149.
67. Hayes, S.A., Miller, J.M., and Hoshizaki, D.K. (2001). serpent, a GATA-like transcription factor gene, induces fat-cell development in Drosophila melanogaster. Development *128*, 1193–1200.
68. Senger, K., Harris, K., and Levine, M. (2006). GATA factors participate in tissue-specific immune responses in Drosophila larvae. Proc. Natl. Acad. Sci. USA *103*, 15957–15962.
69. Arrese, E.L., and Soulages, J.L. (2010). Insect fat body: energy, metabolism, and regulation. Annu. Rev. Entomol. *55*, 207–225.
70. Leclerc, V., and Reichhart, J.-M. (2004). The immune response of Drosophila melanogaster. Immunol. Rev. *198*, 59–71.
71. Lazareva, A.A., Roman, G., Mattox, W., Hardin, P.E., and Dauwalder, B. (2007). A role for the adult fat body in Drosophila male courtship behavior. PLoS Genet. *3*, e16.
72. Shirasu-Hiza, M.M., Dionne, M.S., Pham, L.N., Ayres, J.S., and Schneider, D.S. (2007). Interactions between circadian rhythm and immunity in Drosophila melanogaster. Curr. Biol. *17*, R353–R355.
73. Krishnan, N., Davis, A.J., and Giebultowicz, J.M. (2008). Circadian regulation of response to oxidative stress in Drosophila melanogaster. Biochem. Biophys. Res. Commun. *374*, 299–303.
74. Lee, J.-E., and Edery, I. (2008). Circadian regulation in the ability of Drosophila to combat pathogenic infections. Curr. Biol. *18*, 195–199.
75. Brodu, V., Mugat, B., Roignant, J.Y., Lepesant, J.A., and Antoniewski, C. (1999). Dual requirement for the EcR/USP nuclear receptor and the dGATAb factor in an ecdysone response in Drosophila melanogaster. Mol. Cell. Biol. *19*, 5732–5742.
76. Aruga, J. (2004). The role of Zic genes in neural development. Mol. Cell. Neurosci. *26*, 205–221.

77. Lee, H., Stultz, B.G., and Hursh, D.A. (2007). The Zic family member, odd-paired, regulates the Drosophila BMP, decapentaplegic, during adult head development. Development *134*, 1301–1310.
78. Ma, H., Zhang, J., and Levitan, I.B. (2011). Slob, a Slowpoke channel-binding protein, modulates synaptic transmission. J. Gen. Physiol. *137*, 225–238.
79. Wilczyński, B., and Furlong, E.E.M. (2010). Dynamic CRM occupancy reflects a temporal map of developmental progression. Mol. Syst. Biol. *6*, 383.