Cell, Vol. 107, 567–578, November 30, 2001, Copyright ©2001 by Cell Press

Microarray Analysis and Organization of Circadian Gene Expression in *Drosophila*

Michael J. McDonald[1] and Michael Rosbash[1,2,3]
[1] Department of Biology
National Science Foundation Center
for Biological Timing
[2] Howard Hughes Medical Institute
Brandeis University
Waltham, Massachusetts 02454

Summary

We have used high-density oligonucleotide arrays to study global circadian gene expression in *Drosophila melanogaster*. Coupled with an analysis of clock mutant (*Clk*) flies, a cell line designed to identify direct targets of the CLOCK (CLK) transcription factor and differential display, we uncovered several striking features of circadian gene networks. These include the identification of 134 cycling genes, which contribute to a wide range of diverse processes. Many of these clock or clock-regulated genes are located in gene clusters, which appear subject to transcriptional coregulation. All oscillating gene expression is under *clk* control, indicating that *Drosophila* has no *clk*-independent circadian systems. An even larger number of genes is affected in *Clk* flies, suggesting that *clk* affects other genetic networks. As we identified a small number of direct target genes, the data suggest that most of the circadian gene network is indirectly regulated by *clk*.

Introduction

Circadian rhythms are endogenous, self-sustaining oscillations, which have been described in many eukaryotic organisms and some prokaryotes (Dunlap, 1999; Scully and Kay 2000; Young and Kay, 2001). Rhythms are manifest in such processes as locomotor activity and feeding behavior, sleep/wake patterns, and a variety of physiological and metabolic pathways. These circadian outputs are regulated by a central pacemaker, which receives environmental inputs and keeps circadian time. Integral to the circadian central pacemaker, in mammals as well as flies, is a transcriptional feedback loop. In flies, the positive limb of this loop is driven by a heterodimer of two bHLH-PAS transcription factors, CLOCK (CLK) and CYCLE (CYC); in mammals, the orthologs CLK and BMAL1 (or MOP3) perform a similar function (Allada et al., 1998; Bae et al., 1998; Darlington et al., 1998; Rutila et al., 1998; Gekakis et al., 1998; Bunger et al., 2000). The CLK:CYC heterodimer binds to E boxes and activates transcription of two key rhythm genes, *period* (*per*) and *timeless* (*tim)* (Darlington et al., 1998; Lee et al., 1999). Their protein products, PER and TIM, are proposed to inhibit directly CLK:CYC function, which is critical for circadian cycling of the transcriptional program (Hardin et al., 1990; Zeng et al., 1996; Allada et al., 1998; Darlington et al., 1998; Rutila et al., 1998; Lee et al., 1999). A similar feedback loop functions in mammals, but the cryptochrome gene products (mCRY1 and mCRY2) play a more important role in feedback inhibition (Griffin et al., 1999; Kume et al., 1999).

Genetic screens have been used as a powerful tool to identify clock genes. In addition, exon and enhancer trapping as well as reporter genes have been widely used to identify genes expressed in a circadian manner (Konopka and Benzer, 1971; McClung et al., 1989; Kondo et al., 1994; Liu et al., 1995; King et al., 1997). Molecular methods have also been used to identify clock and clock-controlled genes (Kondo et al., 1993; Millar et al., 1995; Blau and Young, 1999; So et al., 2000). However, these approaches are limited and do not provide a whole genome perspective.

Differential display analysis has recently been used to analyze circadian gene expression, e.g., in the mammalian liver and during sleep in *Drosophila melanogaster* (Kornmann et al., 2001; Shaw et al., 2001). This approach more closely approximates the genomic view but also carries the weakness that every interesting band must be individually cloned and sequenced. Microarrays are the ideal tool, but only in *Arabidopsis thaliana* has a global pattern of circadian gene expression been described (Harmer et al., 2000; Schaffer et al., 2001). In this paper, we use microarray technology to describe circadian transcripts in *Drosophila*. The data have a number of unexpected features, which suggest novel pathways and surprising mechanisms for the control of circadian gene expression.

Results

Cycling Circadian Genes

To isolate mRNA for analysis, we entrained wild-type Canton-S flies for 3 days in a standard 12:12 hr light dark (LD) cycle and then collected flies every 4 hr during the first full day in constant darkness (DD). This strategy was chosen to avoid light-regulated genes not under circadian control as well as the damping (e.g., a decreased cycling amplitude of circadian gene expression) that occurs during extended incubation in constant darkness (see Discussion). Fly head mRNA was harvested from the six time points, biotinylated cRNA prepared and Affymetrix *Drosophila* GeneChips used to probe the labeled cRNA. The final data set includes replicas of 4 chips for CT0, CT4, CT8 and CT12, 5 chips for CT16, and 3 chips for CT20. The GeneChip data were analyzed using a model-based expression approach with dCHIP software (Li and Hung Wong, 2001a, 2001b; for complete dataset, see Supplemental Table S3). To identify a set of circadian genes with confidence, we put the data through four sequential analyses. First, signals were averaged over the 6 time points, and those that did not have an average signal intensity greater than 20 were excluded. This step removed genes with very weak or dubious expression levels (~40% of the transcripts). Second, we required the difference between the highest

[3] Correspondence: rosbash@brandeis.edu

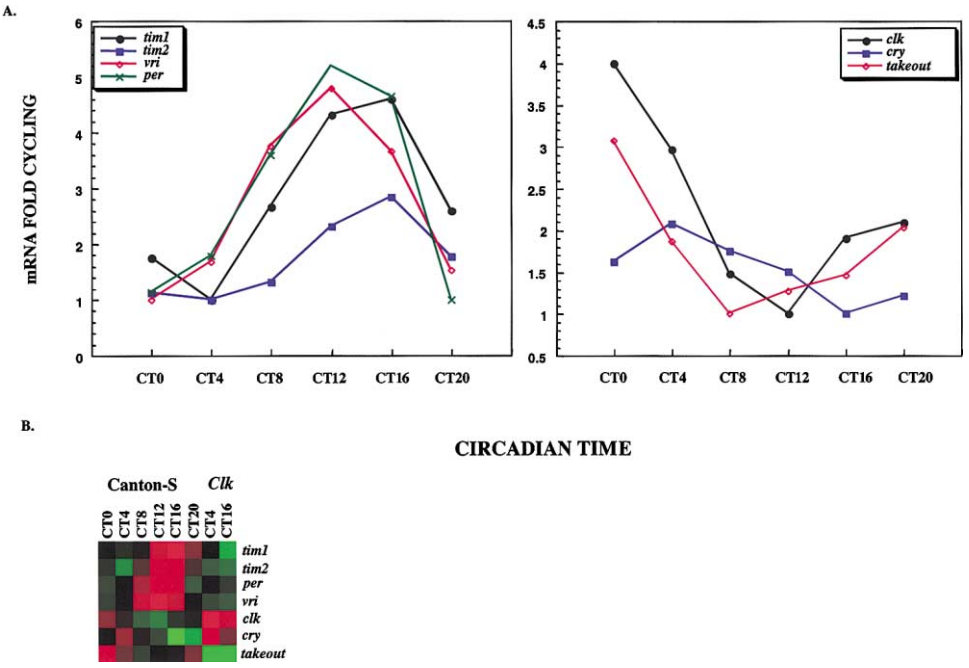

Figure 1. The mRNA Curves for Known Cycling Circadian Genes

tim1 refers to the 142154_at *tim* gene probe set, and *tim2* refers to the 154185_s_at *tim* probe set. (A) The left panel displays the mRNA curves in circadian time (CT0-CT20), generated from the Canton-S GeneChip for *per*, *tim1*, *tim2*, *vri*. The right panel displays the mRNA curves in circadian time (CT0-CT20), generated from the Canton-S GeneChip data for *clk*, *cry*, and CG11853 (*takeout*).

(B) Expanded view of the Treeview data for the known set of cycling genes. The first 6 lanes are circadian time (CT0-CT20) in Canton-S flies, and the remaining 2 lanes are CT4 and CT16 in *Clk* flies.

and lowest of the six time points (each time point was averaged over 3–5 experiments) to be statistically significant. In other words, the difference between the mean average difference values for the high and low time points (for each series of replicates) was larger than the standard error of the two time points. In practice, those genes that did not have a change factor quality above 1 were excluded. Third, a difference of at least 1.5-fold was required between the highest and lowest values of the 6 time points. This is almost equivalent to requiring a cycling amplitude of 50%. Fourth, we derived cross-correlation coefficients between the curves generated by the 6 time points and differently phased cosine waves with 24 hr periods. We excluded genes that did not have a correlation coefficient greater than 0.9. These criteria resulted in a total of 134 circadian genes with a variety of different predicted phases (Supplemental Figure S1, panel A; for a complete listing see Supplemental Table S1 at http://www.cell.com/cgi/content/full/107/5/567/DC1). Immediately validating this approach was the observation that all of the known major *Drosophila* cycling genes (*per, tim, vrille* (*vri*), *clk, cryptochrome* (*cry*), and *takeout*) were included in this list. The mRNA plots are consistent with previously published biochemical data (Figures 1A and 1B). Impressively, four of these six genes (*per*, *tim*, *vri* and *clk*) were among the top ten, when the list of 134 was sorted by fold-change, i.e., a measure approximately equivalent to cycling amplitude (Table 1). This suggests that central clock genes are enriched for high-amplitude oscillations.

Cycling Genes with Related Functions

Although approximately 25% of the remaining 128 genes possess no predictable function, two striking features are apparent for most of the list (Table 2 and Supplemental Table S2). First, there are several groups of genes with similar functions, many of which are reasonable from a circadian point of view. These include relatives of mammalian detoxification enzymes such as the cytochrome P450 genes, short chain dehydrogenases/reductases, and UDP-glucuronosyltransferases (Brierley and Burchell, 1993; Meyer, 1996; King et al., 2001) (Table 2). This is reminiscent of the large number of circadian gene products in mammalian liver (Kornmann et al., 2001). Much of this gene expression may originate in the fat body, which is present in "heads" as commonly prepared. Several of the genes have a proposed function in vertebrate olfaction (Lazard et al., 1990; Dear et al., 1991; Lazard et al., 1991) and in *Drosophila* odorant clearance (Hovemann and Sehlmeyer, 1997; Wang et al., 1999). Indeed, *Drosophila* antennae have been shown to harbor a circadian clock and the circadian cycling of these mRNAs might contribute to the reported oscillations of electrophysiological responses to olfactory stimuli (Krishnan et al., 1999). There are gene families involved in two aspects of ligand binding: these genes are proposed to convey nutritional information (*takeout* family) and to effect tocopherol (anti-oxidant) transport, which may be important in cell signaling or detoxification (Sarov-Blat et al., 2000; So et al., 2000; Zimmer et al., 2000). There are several cycling genes important in pro-

Table 1. Top 10 Highest Fold Cycling Genes

Flybase ID	Function	Fold Cycling	Map Position	mRNA Phase	Level in *Clk*
ldlr	scavenger receptor	40.8	89D8-89D8	23	high
CG11854	ligand binding or carrier	5.7	96C4-96C4	22	low
CG13856	unknown	5.6	94B1-94B1	10	low
per	PAS domain clock protein	5.3	3B4-3B4	12	low
vri	par domain clock protein	4.8	25D4-25D4	12	low
tim1	clock protein	4.6	23F3-23F5	14	low
CG5798	ubiquitin thiolesterase	4.4	93C1-93C1	11	low
CG2069	unknown	4.1	62B1-62B1	6	mid
clk	bHLH PAS clock protein	4.0	66A12-66A12	0	high
CG5156	unknown	3.8	21F2-21F2	2	low

The list of 134 cycling genes was sorted using the fold cycling value, which was obtained by dividing the timepoint with the highest average (over the series of replicates) by the timepoint with the lowest average. The top 10 genes are shown. Genes are annotated by their Flybase ID, probable function, map position, phase at which the mRNA is highest and the level of the transcript in *Clk* flies.

cessing neuropeptides such as CG8550, a neprilysin-like metalloendopeptidase and *Acer* (Angiotensin converting enzyme related; Taylor et al., 1996; Isaac et al., 2000; Turner et al., 2001). These mRNAs might contribute to oscillations in neuropeptide function (e.g., Renn et al., 1999; Park et al., 2000). There are three genes that function in cuticle formation, which is reminiscent of the circadian cycling of adult cuticle deposition observed in some insects (Saunders, 1982). Intriguingly, we may have found a connection between immunity/host defense and circadian rhythms, by identifying *immune deficiency* (*imd*), *u-shaped* (*ush*) and *acetylated low density lipoprotein receptor* (*ldlr*) as cycling genes (see Discussion). There are also genes of known or suspected function, not previously identified as under circadian control in any system. Although these are of great potential interest, we can only speculate about their contribution to circadian rhythms (see Discussion and Table 2).

Chromosomal Clustering of Cycling Genes with Related Functions

The second striking feature is that many of the genes with similar function cluster together within the same chromosomal region (Table 2 and Supplemental Table S2). This includes the 2 cytochrome P450 genes and UDP-glucuronosyltransferases mentioned above. The *takeout* ligand binding family is a second example of this phenomenon. These six genes are divided into a group of three adjacent genes, a pair of adjacent genes and one singlet; all six genes manifest circadian cycling (Figure 2, Table 2, and Supplemental Table S2). Another cluster is the contiguous family of three trypsin-like serine proteases at 88B3 (Supplemental Table S2). Of these three genes, CG9645 and CG9649 share the highest sequence similarity to each other and then to CG9631. PSI-BLAST was used to find significant members based on a position-specific matrix score (Altschul et al., 1997). After 5 iterations, the 3 genes each most closely resemble mammalian plasminogen. The third cluster is largely undescribed and comprises five members that all cycle with the same phase. They are grouped into two adjacent pairs and one single gene (Figure 2, and Supplemental Table S2). All five belong to a 22-member gene family cluster localized to 96C6-96C7 ($\sim$60 kb). These 22 genes belong to an even larger 34-gene family, all of which contain one or more conserved regions known as DUF227 domains. Although of unknown function, this domain is also present in a small number of *C. elegans* genes. The cycling genes within each cluster are similarly affected in *Clk* flies (Figure 2). In summary, these data strongly suggest that an important feature of circadian gene expression is transcriptional coregulation through common *cis*-acting elements (e.g., E boxes).

Analysis of *Clk* Mutant Flies

To examine the cycling mRNAs in a behaviorally arrhythmic genetic background, we entrained clock mutant *Clk* flies and isolated mRNA from two time points (CT4 and CT16) during the first day in DD. This experiment was repeated and the data averaged for the two trials. Remarkably, all of the cycling genes that show differences between these two time points failed to manifest circadian cycling in this mutant background. Most of the mRNA levels were high or low whereas a few were intermediate (Tables 1, 2, and Supplemental Tables S1 and S2). Consistent with previously published data, all of the suspected direct targets of CLK (*per, tim, vri*) had low mRNA levels in the *Clk* background. The two genes (*clk* and *cry*) previously reported to have high and constant mRNA levels by biochemical criteria in the *Clk* background have high levels in the microarray assay. The failure to maintain circadian cycling in this background is consistent with the notion that *clk* sits at the top of the circadian regulatory network; in the absence of functional CLK protein, the entire circadian program is shut down. We also conclude that there are no *clk*-independent circadian oscillators in the fly.

In addition to the lack of cycling transcripts in *Clk* flies, there are two additional, striking features of the data. First, there are an even larger number of noncycling genes with altered levels in this background. The application of stringent criteria indicates there are 267 *Clk*-regulated genes, and a more relaxed criterion indicates that there could be as many as 363 *Clk*-regulated genes (Supplemental Figure S1, panel B, and Supplemental Table S1; see Experimental Procedures). 80% of these noncycling mRNAs are more abundant in the *Clk* background, and 20% are less abundant. This list of genes should contain direct *clk*-targets with mRNA half-lives too long to support high-amplitude mRNA cycling. It should also contain genes affected only indirectly by CLK-CYC. Indirect targets would include genes that are

Table 2. Cycling Genes with Common Functions

Flybase ID	Function	Map Position	mRNA Phase	Level in *Clk*
Detoxification enzymes				
Cyp4d21	cytochrome P450, CYP4D21	28A6-28B1	19	mid
Cyp4p1	cytochrome P450, CYP4PI	45C1-45C1	4	high
Cyp6a21	cytochrome P450, CYP6A21	51D2-51D2	23	mid
Cyp6a20	cytochrome P450, CYP6A20	51D2-51D2	23	high
Cyp305a1	cytochrome P450, CYP305A1	76D3-76D3	3	low
CG15093	Short-chain dehydrogenase/reductase (SDR)	55F1-55F1	4	mid
CG3699	Short-chain dehydrogenase/reductase (SDR)	1D1-1D1	11	high
CG7322	Short-chain dehydrogenase/reductase (SDR)	17E4-17E4	9	high
Sodh-1	L-iditol 2-dehydrogenase	84C1-84C1	10	high
Cpr	ferrihemoprotein reductase-odorant clearance	26C3-26C4	22	low
CG18578	UDP-glucuronosyltransferase	86D4-86D4	3	mid
Ugt35b	UDP-glucuronosyltransferase	86D5-86D5	2	mid
Ligand binding or carrier				
CG10237	alpha-tocopherol transfer protein-like	37E3-37E3	18	mid
CG3823	tocopherol binding	5E1-5E1	2	low
CG10657	tocopherol binding	69C2-69C2	9	low
CG7079	*takeout* family	93B5-93B5	1	high
CG7096	*takeout* family	93B5-93B5	3	high
CG11852	*takeout* family	96C3-96C3	10	low
CG11853	*takeout* family	96C4-96C4	23	low
CG11854	*takeout* family	96C4-96C4	22	low
CG17189	*takeout* family	97E6-97E6	10	mid
Neuropeptide modulator				
Acer	angiotensis I-converting enzyme	29D1-29D2	10	high
CG8550	neprilysin-like metalloendopeptidase	49A7-49A7	14	high
Cuticle formation				
CG2555	larval cuticle protein-like	11B10-11B10	11	mid
CG3348	chitin binding domain	97F6-97F6	10	low
CG4784	cuticle protein	72F1-72F1	3	low
Genes involved in immune or host defense response				
imd	RIP1-like kinase	55C9-55C9	1	high
ldlr	acetylated LDL receptor-like	89D8-89D8	23	high
Known genes with unknown circadian function				
b	glutamate decarboxylase 2	34D2-34D3	0	mid
B4	expressed in neuroblast SOMC	34B11-34B12	23	mid
Ahcy89E	adenosylhomocysteinase	89E10-89E10	12	high
CG1147	neuropeptide Y receptor-like	83E1-83E1	1	low
crp	transcription activator protein	35F6-35F7	23	mid
Eip55E	cystathionine-gamma-lyase	55E5-55E5	7	high
KdelR	KDEL receptor	31E1-31E1	7	high
ATPCL	ATP-citrate (pro-S)-lyase	52E1-52E1	8	high
mew	integrin, alpha-subunit	11D10-11E1	0	mid
Nmdmc	methylenetetrahydrofolate dehydrogenase	85C5-85C5	5	low
ple	tyrosine 3-monooxygenase	65C3-65C3	2	high
s6kII	ribosomal protein S6 kinase	20A1-20A1	23	mid
slob	calcium binding protein	28B1-28B3	11	mid
ush	nuclear receptor involved in hemocyte diff.	21C5-21C5	7	low
Zw	glucose-6-phosphate 1-dehydrogenase	18D12-18D12	9	high

The 134 cycling genes were examined for probable function, using BLAST, PubMed literature searches, and predicted functions for conserved domains. Only the genes with predicted functions are shown; the remaining ~25% are shown in Supplemental Table S1. The Flybase ID, map position, phase of peak mRNA expression and levels in *Clk* flies are shown.

influenced by transcription factors like VRI, themselves direct targets of CLK-CYC, as well as genes that are even more indirectly affected by the clock system. The latter category might include genes affected by developmental abnormalities due to the *Clk* mutation.

Almost all the genes of obvious interest by sequence are in the first, more stringent set of *clk*-regulated genes (Table 3). These include two more neuropeptide-modulating enzymes, which belong to the neprilysin family of peptidases, as well as two additional neuropeptide hormone genes, *corazinin* and *capability*. There are five odorant binding proteins and four pheromone binding proteins, which also have members in two additional clusters. Lastly, there are a large number of immunity genes affected by the clock mutant (Table 3). Pathogen recognition molecules and almost all known members of the *imd* pathway of immunity genes constitute the majority of this list. Interestingly, all but one family of microbial resistance peptides (genes activated mainly through the *imd* pathway) are upregulated, including

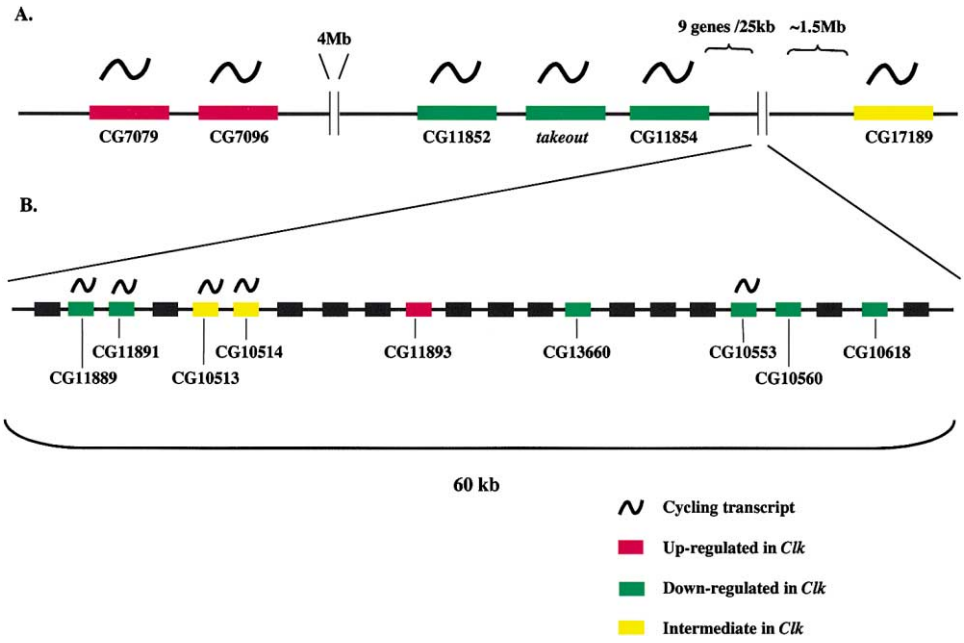

Figure 2. Homologous Cycling Gene Clusters of *takeout* and Unknown Families

Two examples of clustered genes are depicted. All doublets or triplets are similarly affected in *Clk* flies. Gross chromosomal distances are shown. (A) The *takeout* gene family. (B) The much larger unknown gene family.

members thought to be under control of the Toll receptor pathway (Khush et al., 2001). Although these genes do not display circadian oscillations, most manifest peculiar biphasic mRNA expression profiles (data not shown). There is evidence that serine proteases, like those present in the cycling and *Clk*-regulated gene lists, are involved in the initial proteolytic cascades of immunity (Stover et al., 2001; Vang Petersen et al., 2001). We speculate that this large set of *clk*-affected (but noncycling) mRNAs contributes to some of the phenotypes of *Clk* flies, e.g., cocaine sensitization (Andretic et al., 1999), and that a similar phenomenon will exist in *Clk* mutant mice (alteration of sleep homeostasis; Andretic et al., 1999; Naylor et al., 2000).

Numbers from Differential Display

To estimate the number of mRNAs under circadian clock control by an independent procedure, we subjected the same mRNA samples to a differential display procedure (ADDER) recently introduced by Schibler and colleagues (Kornmann et al., 2001). This procedure gave rise to very reproducible banding patterns, which showed no quantitative or qualitative differences when triplicate mRNA samples were assayed (data not shown). By assaying a majority of the 192 primer pairs, we identified ~19,200 bands of which 171 were cycling in a circadian manner. If we assume that these come from 13,500 genes, there would be 120 independent cycling mRNAs. Also in remarkable agreement with the microarray data, none of these cycling bands showed oscillations in mRNA from the *Clk* flies (assayed with approximately 1/3 of the primer pairs). Although we do not know how well these bands correspond to the cycling genes, this independent means of analysis reinforces the conclusion that the entire circadian program is under *clk* control.

Also in agreement with the microarray data, there are a much larger number of noncycling mRNAs with substantial level changes between *Clk* and wild-type flies. Extrapolating to 13,500 bands, there were 300 noncycling bands that showed increases and 250 bands that showed decreases compared to wild-type head mRNA (Figure 3; data not shown). This compares favorably with the number of transcripts from the microarray experiments that show increases in *Clk* flies (220). The only substantial point of disagreement between the ADDER and microarray data is that there are many fewer transcripts from the microarray data that are low in *Clk* flies (48). This inconsistency may result from unknown differences in sensitivity between the two techniques.

Direct Targets of CLK

To distinguish between direct and indirect targets of the CLK:CYC heterodimeric transcription factor, we established a stable S2 cell line expressing a CLOCK-glucocorticoid receptor hormone binding domain fusion protein (CLK-GR). Upon addition of dexamethasone (a glucocorticoid analog), CLK-GR is activated and drives the transcription of clock target genes (data not shown). This general strategy has been used previously for other transcription factors (Picard and Yamamoto, 1987; Godowski et al., 1988; Schena and Yamamoto, 1988) and relies on the almost complete absence of transcriptionally active CLK but the presence of endogenous CYC in the S2 cells (e.g., McDonald et al., 2001). Importantly, only primary targets are transcribed in the presence of the protein synthesis inhibitor cyclohexamide. With this approach and microarray screening, we identified a

Table 3. *Clk*-Regulated Genes

Immunity Genes			
Anti-microbial peptides		**Map Position**	**Levels *Clk*: CS**
CG4740	antibacterial response protein	50A8-50A8	high
AttA	antibacterial response protein	51C2-51C2	high
dro	antibacterial response protein	51C2-51C2	high
AttB	antibacterial response protein	51C2-51C2	high
mtk	antibacterial response protein	52A1-52A1	high
dpt	antibacterial response protein	55F5-55F5	high
dpt	antibacterial response protein	55F5-55F5	high
CG13422	antibacterial response protein	57A7-57A7	high
cecA1	antibacterial response protein	99E4-99E4	high
cecA2	antibacterial response protein	99E4-99E4	high
cecB	antibacterial response protein	99E4-99E4	high
cecC	antibacterial response protein	99E4-99E4	high
drs	antifungal response protein	63D1-63D1	high
LysX	lysozyme	61F3-61F3	high
Cg18107	immune response induced protein	55C9-55C9	high
CG15068	immune response induced protein	55C9-55C9	high
CG18108	immune response induced protein	55C9-55C9	high
Proteases and proteinase inhibitors			
CG9377	endopeptidase	34B6-34B6	high
CG4650	endopeptidase	35B6-35B6	high
CG12256	endopeptidase	87B6-87B6	high
TepIII	enzyme inhibitor	28C1-28C1	high
Immunity signal transduction			
Tak1	imd immunity pathway	19E1-19E1	high
TotM	immunity pathway	25D5-25D5	high
CG18589	immunity pathway	28C1-28C1	high
kenny	imd immunity pathway	60D15-60D15	high
Relish	imd immunity pathway	85C5-85C5	high
Non-self recognition			
CG11709	non-self recognition molecule	10C4-10C4	high
CG7106	non-self recognition molecule	28D3-28D3	high
CG11211	non-self recognition molecule	42A10-42A10	low
CG8343	non-self recognition molecule	42A10-42A10	low
CG14745	non-self recognition molecule	44D8-44D8	high
Hml	non-self recognition molecule	70C4-70C4	high
CG9681	non-self recognition molecule	73B5-73B5	high
Sr-CI	scavenger receptor	24D4-24D4	high
Neuropeptide or neuropeptide modulator			
Crz	neuropeptide hormone	88B7-88B7	high
capa	neuropeptide hormone	99D1-99D1	high
CG14527	neuropeptide peptidase	98F1-98F1	high
CG11501	neuropeptide peptidase	99B3-99B3	high
Pheromone or odorant binding			
CG8462	odorant binding protein	56E5-56E5	low
CG13421	odorant binding protein	57A6-57A6	high
CG9358	odorant binding protein	60E7-60E7	high
CG18111	odorant binding protein	99B9-99B9	low
CG7592	odorant binding protein	99C1-99C1	low
Pbprp5	pheromone binding protein	28A1-28A1	high
Pbprp3	pheromone binding protein	83C9-83D1	high
Os-E	pheromone binding protein	83D1-83D1	high
Os-C	pheromone binding protein	84E4-84E4	high

The 267-member stringent list of *Clk*-regulated genes was examined for genes that may be involved in similar physiological or functional pathways, using BLAST, PubMed literature searches, and predicted functions for conserved domains. Genes classified as immunity genes include: anti-microbial peptides, proteases and protease inhibitors (implicated in immunity), and non-self recognition molecules. The other classes of genes are neuropeptides or neuropeptide modulators, and pheromone or odorant binding proteins. The Flybase ID, map position, and corresponding levels in *Clk* flies relative to Canton-S flies are also shown.

small group of 9 candidate primary target genes (Table 4 and Figures 4A and 4B). Six cycle in CS flies and are low in *Clk* flies (Table 4). They fulfill all of the criteria to be direct CLK:CYC targets in the fly. They include three known or suspected direct targets (*tim, vri*, and *pdp1*; see below), which validate this approach. They also in-

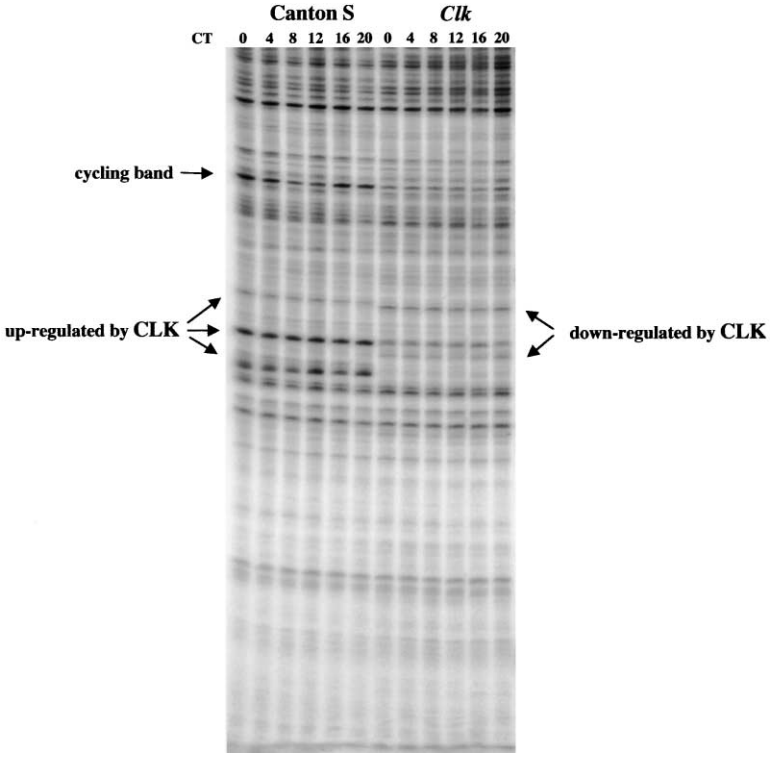

Figure 3. Example of Cycling and *Clk*-Regulated Bands from ADDER

Differential display of Canton-S and *Clk* fly head cDNAs over 6 time points (CT0-CT20). Shown is a primer set with a cycling band, 3 bands downregulated in *Clk* flies, and 2 bands upregulated in *Clk* flies. The cycling band is also low and constant in *Clk* flies, further suggesting it is a clock-controlled gene.

clude *sticky ch1*, which is a transcription factor and a new candidate CLK:CYC direct target gene (see Discussion). Of the three other genes, one is *Clk* regulated but does not cycle, and two fail to cycle or show *Clk* regulation. However, all 9 genes contain either canonical or noncanonical E boxes, suggesting that they are direct targets in S2 cells. RPA analysis in this system reveals that *per* mRNA is expressed in the absence of CLK and is only weakly further induced upon activation of CLK-GR (data not shown). Because so few of the cycling or *Clk*-regulated genes score positively in this assay, we suggest that only a small number are direct CLK:CYC targets. This hypothesis is strengthened by the observation that relatively few of the 134 cycling genes express peak mRNA levels in phase with the known direct CLK targets, *per, vri*, and *tim*. Therefore, the vast majority of *Clk*-regulated genes are probably indirect target genes.

Discussion

Our studies have identified 134 cycling genes, which contribute to a wide range of diverse processes. Many of these clock-regulated genes are located in gene clusters, which appear subject to transcriptional coregulation. All oscillating gene expression is absent in *Clk* mutant flies, consistent with the individual mRNAs previously examined in *Clk* mutant mice as well as flies (Allada et al., 1998; Darlington et al., 1998; Bae et al., 1998; Gekakis et al., 1998; Rutila et al., 1998; Glossop et al., 1999; Jin et al., 1999; Kume et al., 1999; Shearman et al., 2000). These results indicate that *Drosophila* has no *clk*-independent circadian system. A previous study of diapause in this organism suggested that a noncanonical circadian system may exist (Saunders et al., 1989; Saunders, 1990). If so, it is not apparent from this analy-

Table 4. CLK Direct Target Candidates

Flybase ID	Description	Function	Map Position	Fly Array Data
tim	clock protein	transcription factor	23F3-23F5	cycling in CS; low in *Clk*
vri	par domain clock protein	transcription factor	25D4-25D4	cycling in CS; low in *Clk*
CG8772	glutaminase	enzyme	49B8-49B8	no change in *Clk*
picot	Na/K symporter	transporter	53C12-53C14	low in *Clk*
CG15095	Na/K symporter	transporter	55F1-55F1	no change in *Clk*
pdp1	par domain protein	transcription factor	66A14-66A17	cycling in CS; low in *Clk*
CG8147	alkaline phosphatase-like	enzyme	85D21-85D21	cycling in CS; low in *Clk*
sticky ch1-1	bHLH transcription factor	transcription factor	86A5-86A5	not expressed
sticky ch1-2	bHLH transcription factor	transcription factor	86A5-86A5	cycling in CS; low in *Clk*
CG13610	ion transporter	transporter	95F2-95F2	cycling in CS; low in *Clk*

The putative direct targets are depicted by Flybase ID, a general description, probable function, map position, and corresponding data from the fly microarrays. *sticky ch1-1* and *sticky ch1-2* are different probe sets for the *sticky ch1* gene. One probe set for *sticky ch1* (*sticky ch1-1*) was not detected in head mRNA, reflecting perhaps some tissue-specific regulation.

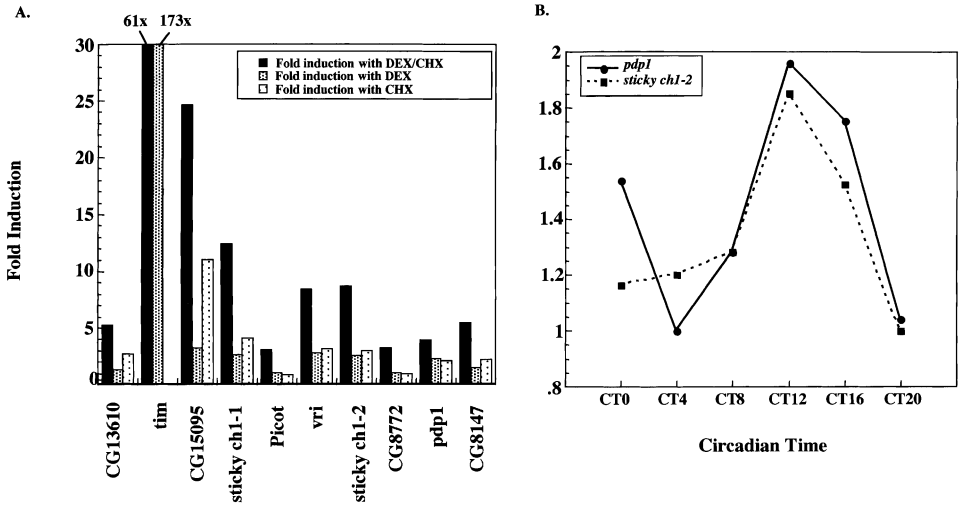

Figure 4. Candidate Direct Targets of CLK from the CLK-GR S2 Cell Assay

Results are shown as an average from duplicate GeneChip experiments. Briefly, stable, copper-inducible CLK-GR cell lines were induced to express CLK-GR with 500 μM copper sulfate. At time = 0, 48 hr after copper sulfate addition, 10 μM Dexamethasone (DEX) and/or 10 mg/ml Cycloheximide (CHX) were added. Another group of cultures were maintained in the copper induced state with no further additions. mRNA was harvested at 0, 2, and 4.5 hr after the additions. The data from the averaged replicates for the CLK-GR cells at 4.5 hr treated with DEX and/or CHX were divided by the average of the replicates for the copper only control. These values are shown as the fold change for those treatments. (A) The fold change values are shown for the 3 treatments (DEX, DEX/CHX, CHX) at 4.5 hr. (B) Graphs of the mRNA cycling curves from Canton-S flies of *pdp1* and *sticky ch1*. Values were obtained by dividing by the lowest of the 6 time points.

sis of head mRNA cycling. An even larger number of noncycling genes is also affected in the *Clk* mutant, consistent with reported phenotypes of this strain not obviously related to circadian phenomena (Allada et al., 1998; Andretic et al., 1999). Finally, our data indicate that most of the *clk*-regulated genes are only indirectly regulated by CLK, suggesting that a transcriptional cascade makes a major contribution to the genes and phenotypes under *clk*-control.

We chose to examine mRNA cycling during the first day in constant darkness, to avoid light-regulated genes and to minimize the impact of damping. It is known that many circadian features weaken as a function of time in constant darkness. In flies, only the lateral neurons appear to escape this phenomenon (Yang and Sehgal, 2001). In principle, the first day in DD should maximize the signal-to-noise ratio for circadian mRNA cycling. Coupled with our strict selection criteria, we are confident that most if not all of these are bona fide circadian clock-controlled genes. However, it is likely that this set of 134 genes, representing 1% of the current *Drosophila* genome estimate, is incomplete. For example, mRNAs that oscillate in tissues other than "heads" will not have been identified in this study. These transcripts may not be present in heads, have too low an abundance to be detectable or have a long half-life in head mRNA. The use of total RNA for probe preparation may have exacerbated this problem, but we obtained more reproducible results than with polyA$^+$ mRNA (data not shown), and all of the major cycling genes were easily detected. Consistent with the notion that there are a larger number of clock-regulated transcripts in *Drosophila*, the two *A. thaliana* microarray studies found 6% and 2% of genes under circadian control (Harmer et al., 2000; Schaffer et al., 2001, respectively).

It is also likely that our strict selection criteria led to the exclusion of some clock-controlled genes. Additional candidates include mRNAs with amplitudes (viz. fold cycling values) less than 1.5. In addition, small deviations from the other criteria have probably excluded some transcripts. For example, *pdp1* was excluded due to a small increase at CT0, which precluded an acceptable correlation with the cosine curve (∼.8 at a peak phase of 12). However, there are three indications that it is a bona fide cycling gene: *pdp1* scores as a direct target in the S2 cell assay; it shows significant 2-fold cycling (Figures 4A and 4B); *pdp1* mRNA levels fail to cycle and are very low in *Clk* flies. *sticky ch1*, another direct target, was also excluded from the cycling list, in this case because its standard error at CT16 was too large. It also meets all other criteria. Based on these two examples, we manually examined all of the genes with altered mRNA levels in the *Clk* background; although there were many additional genes with low (<1.5-fold) amplitude cycling, *pdp1* and *sticky ch1* were the only obvious head-expressed genes above this threshold (Figure 4B).

Some of the probe sequences could also preclude the identification of bona fide circadian genes. For example, there exist 3 *tim* probe sets on the GeneChip: 142154_at, 151484_i_at, and 151485_s_at. The results for each probe set are reproducible yet show substantial differences. Although probe sets 142154_at and 151485_s_at display similar mRNA accumulation profiles, the signal intensities are very different (data not shown). Moreover the signal intensity for the 151484_i_at probe set is well below the threshold for reliability. We do not know how many additional probe sets suffer from similar problems. In addition, there are almost certainly many *Drosophila* transcripts not represented on the GeneChip (e.g., Andrews et al., 2000).

Given these uncertainties, it is remarkable that the estimated number of cycling genes identified from ADDER (120) is so close to the number identified from the microarray (135). This assumes that the estimated ~19,200 total bands comes from 13,500 genes, consistent with the expectation that single genes will give rise to multiple bands (due for example to multiple polyadenylation sites or differential splicing). In addition, there are instances in which multiple primer pairs give rise to the same band (data not shown; Kornmann et al., 2001). In any case, these comparisons are tentative, as the relative sensitivity of the display and microarray methods is not known.

It is perhaps not surprising that some of the *Drosophila* CLK-controlled transcripts contribute to detoxification, as we have identified some of the same genes that are under clock control in the mouse liver (Kornmann et al., 2001). This may be due to the presence of fat body in the head preparations, and it will be interesting to see if the regulation of these genes is conserved between flies and mammals. Many of the identified detoxification enzymes are also hypothesized to be important in olfactory pathways involving odorant clearance. Alternatively or in addition, it is possible that some of these proteins are expressed in the fly brain and contribute to brain physiology and/or guard against xenobiotic-induced neurotoxicity (Hedlund et al., 2001).

There are several other cycling mRNAs of obvious interest for circadian phenomena. In addition to those described above, two are *slob* and *s6kII*. *slob* has been shown to interact with the K channel *slowpoke* and also the protein kinase *14-3-3* (Schopperle et al., 1998; Zhou et al., 1999). Cycling *slob* mRNA could give rise to circadian oscillations in K channel activity, which might then affect resting membrane potential. Cycling membrane potential has been previously reported (Michel et al., 1993) and should give rise to calcium oscillations, which might underlie the reported oscillations in neuropeptide staining in lateral neuron termini (Park et al., 2000). *s6kII* encodes a kinase which is the *Drosophila* ortholog of an important mammalian signal transduction protein. This kinase is downstream of the ERK and MAPK signaling pathways and affects a number of transcription factors in response to GPCR stimulation. These include CREB, reported to participate at several stages of the circadian program in both flies and mammals. The cycling of CREB activity (Belvin et al., 1999) and the circadian regulation of Ser131 phosphorylation (Ginty et al., 1993) could be downstream of *s6kII* cycling. Alternatively or in addition, the recent report of map kinase signaling in circadian output might involve the cycling of *s6kII* activity (Williams et al., 2001).

A major set of circadian genes is associated with the *imd* immunity pathway, revealed principally by analyzing *Clk* flies. *imd* itself is a cycling gene, which is expressed at high levels in the absence of *clk* function. Genes downstream of *imd*, such as *tak1*, *relish*, *kenny*, and the target genes of REL (which include the antimicrobial peptide gene families) are all upregulated in *Clk* mutant flies. Another cycling gene that could be involved in immune responses is *u-shaped* (*ush*). USH is a nuclear hormone transcription factor that is important in hemocyte development and limits hemocyte differentiation into crystal cells, which are involved in the defense-related melanization of encapsulated targets (Rizki, 1978). *ldlr* is a cycling gene that has been postulated to play a role as a scavenger receptor in the recognition of non-self proteins (Abrams et al., 1992). Additionally, *Clk* flies have increased mRNA levels of 5 non-self recognition and 3 c-type lectin genes, which are also believed to be involved in non-self recognition (Theopold et al., 1999). Finally, all but one of the regulated serine proteases (cycling or noncycling) increase in the absence of wild-type CLK, which could lead to the initiation of one or more immune system-relevant proteolytic cascades (Khush et al., 2001). Although we cannot exclude the possibility that the *Clk* gene induces a stress response, we propose that the fly immune system is under clock control. Supporting this hypothesis is the identification of cycling genes, *imd*, *ush*, *ldlr*, and a cluster of three serine proteases, which could contribute to the regulation of the immune network.

The remarkable clustering of clock-regulated genes suggests that genome organization plays a prominent role in circadian gene expression. We have discovered families of genes such as the serine proteases at 88B3 and the *takeout* family at 99C4-99C5; it appears that many members are transcriptionally coregulated. The structure of the *takeout* region suggests that the previously described *takeout* deletion removes more than just the original *takeout* gene (So et al., 2000). A large 22-gene region nearby at 99C6-99C7 (~60 kb) has clusters in which members are transcriptionally coregulated. Although we do not know the mechanisms that underlie this transcriptional regulation, it is possible that one or more clock-regulated enhancers govern the expression of a significant region of chromatin.

Indeed, our results indicate that CLK is a master transcriptional regulator of the circadian clock, relevant to detoxification, olfaction, neuropeptide signaling, nutritional state, and immunity. How does CLK exert this broad effect? The relatively few direct target genes identified from the S2 cell experiments suggest that it is in part by regulating a few key transcription factors. In support of this conclusion is the phase analysis of the 134 cycling genes, which have relatively few transcripts that peak between CT12-CT16. This early evening phase is characteristic of all known or suspected direct target genes, including *tim, per, vri, pdp1, and sticky ch1*. All of these mRNAs cycle and have low levels in the absence of CLK. VRI and PDP1 are especially attractive candidates for effecting a subsequent cascade of circadian transcription. It is known that the mammalian counterparts of these genes, E4BP4 (VRI) and DPB1 (PDP1), have regulatory input to many circadian genes (Doi et al., 2000; Ripperger et al., 2000; Mitsui et al., 2001). They act as repressors (E4BP4) and activators (DBP1) and can heterodimerize with many different proteins, features that amplify the range of transcriptional activities. We propose that *sticky ch1* also serves to extend the range of clock-controlled transcription.

This description of circadian transcripts more generally identifies a large number of candidate mechanisms as well as genes. More detailed studies are now required to verify their validity as well as their importance to the *Drosophila* circadian clock and to *Drosophila* physiology and behavior.

Experimental Procedures

Microarrays

Probe preparation. Total RNA was extracted from fly heads (collected every 4 hr for 24 hr starting with CT0) using Tri-Reagent (Molecular Research Center Inc.) according to the manufacturer's protocol. cDNA synthesis was carried out as described in the Expression Analysis Technical Manual (Affymetrix). The cRNA reactions were carried out using the BioArray High-Yield Transcript Labeling kit (Enzo). 30 μg of labeled cRNA was fragmented for 35 min at 94°C using fragmentation buffer (200 mM Tris-acetate, pH 8.1, 500 mM K acetate, 150 mM Mg acetate). Affymetrix high-density oligonucleotide arrays for *Drosophila melanogaster* were probed, hybridized, stained, and washed according to the manufacturer's protocol.

Creation of Stable S2 Cell CLK-GR Lines

Cloning: a Bgl II site was created to replace the stop codon in pBS-Clock (McDonald et al., 2001) and cloned into pAc 5.1 (Invitrogen) using a NotI-BglII digest, forming pAc-clk2. p6GRLS9 (a kind gift of Keith Yamamoto) was mutagenized so that an XbaI site replaced the stop codon. A NotI-BglII fragment from pAc-clk2, a BglII-XbaI fragment from p6GRLS9, and NotI-XbaI digested pMT (Invitrogen) were ligated to form pMT-CLK-GR (CLK-GR). S2 cells were maintained at high density in HyQ-SFX Insect media (Hyclone Laboratories Inc.) supplemented with 10% FBS (GIBCO) and 10 mg/ml Penicillin/Streptomycin (Sigma). S2 cells at a density of 3×10^6 per ml in HyQ-SFX lacking serum (SFM) were plated in 6-well tissue culture plates (Costar Corp.). After 24 hr, the medium was aspirated and replaced with 1 ml of transfection mix. This consisted of SFM containing 20 μl Geneporter reagent (Gene Therapy Systems), 5 μg of CLK-GR plasmid, and 0.5 μg of pCO-HYGRO (Invitrogen) hygromycin resistance plasmid. Transfection was stopped by adding 1 ml of serum-containing HyQ-SFX media. Cells were selected for Hygromycin resistance (300 μg/ml hygromycin) until stable cell lines formed.

Data Analysis

GeneChip .DAT files were analyzed by dCHIP software, using model-based expression analysis. Data output from dCHIP was analyzed first in Microsoft Excel. To identify genes cycling with a 24 hr period, genes surviving the first three criteria (see Results) were first normalized by subtracting the average value of the six time points (mx) from the expression value at a particular time point (x) and dividing by mx. A perl script was used to derive a cross correlation coefficient between the normalized data for each probe set and 24 hr period cosine curves with phases of 0–23. The probe sets with a coefficient greater than 0.9 were retained as circadian genes. 135 probe sets representing 134 genes (two probe sets for *tim*) passed this analysis. We required the absolute value of the difference between the average for the 6 Canton-S time points and the average of the *Clk* time points be 50 or greater. We required that the fold change difference (*Clk* average/CS average or CS average/*Clk* average) be greater than 1.5 and statistically have a change quality above 1. These genes are classified as the less stringent set, which contains 363 probe sets representing 361 genes. By demanding an even bigger difference between the CS and *Clk* values, we obtained the stringent set, containing 269 probe sets and 267 genes. This required that the *Clk* values are at least 1.5-fold higher than the highest CS value or 1.5-fold lower than the lowest CS value.

Supplemental Data

Supplemental Figure S1 and Supplemental Tables S1, S2, and S3 are available on the *Cell* website (http://www.cell.com/cgi/content/full/107/5/567/DC1).

Acknowledgments

We are very grateful to B. Kornmann and U. Schibler for their hospitality and instruction in the use of the ADDER differential display technique. We thank our colleagues Y. Peng, P. Nawathean, and J. Zhao for helpful suggestions and discussions, and P. Emery, R. Allada, V. Suri, and J. Hall for comments on the manuscript. We thank Ed Dougherty for help with figures, Steve Karel for assistance with statistics and algorithms, and Heather Felton for administrative assistance. The work was supported by a grant from the NIH (GM33205).

Received September 20, 2001; revised October 15, 2001.

References

Abrams, J.M., Lux, A., Steller, H., and Krieger, M. (1992). Macrophages in Drosophila embryos and L2 cells exhibit scavenger receptor-mediated endocytosis. Proc. Natl. Acad. Sci. USA *89*, 10375–10379.

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.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Andretic, R., Chaney, S., and Hirsh, J. (1999). Requirement of circadian genes for cocaine sensitization in Drosophila. Science *285*, 1066–1068.

Andrews, J., Bouffard, G.G., Cheadle, C., Lu, J., Becker, K.G., and Oliver, B. (2000). Gene discovery using computational and microarray analysis of transcription in the Drosophila melanogaster testis. Genome Res. *10*, 2030–2043.

Bae, K., Lee, C., Sidote, D., Chuang, K.Y., and Edery, I. (1998). Circadian regulation of a Drosophila homolog of the mammalian Clock gene: PER and TIM function as positive regulators. Mol. Cell. Biol. *18*, 6142–6151.

Belvin, M.P., Zhou, H., and Yin, J.C. (1999). The Drosophila dCREB2 gene affects the circadian clock. Neuron *22*, 777–787.

Blau, J., and Young, M.W. (1999). Cycling vrille expression is required for a functional Drosophila clock. Cell *99*, 661–671.

Brierley, C.H., and Burchell, B. (1993). Human UDP-glucuronosyl transferases: chemical defence, jaundice and gene therapy. Bioessays *15*, 749–754.

Bunger, M.K., Wilsbacher, L.D., Moran, S.M., Clendenin, C., Radcliffe, L.A., Hogenesch, J.B., Simon, M.C., Takahashi, J.S., and Bradfield, C.A. (2000). Mop3 is an essential component of the master circadian pacemaker in mammals. Cell *103*, 1009–1017.

Darlington, T.K., Wager-Smith, K., Ceriani, M.F., Staknis, D., Gekakis, N., Steeves, T.D.L., 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.

Dear, T.N., Boehm, T., Keverne, E.B., and Rabbitts, T.H. (1991). Novel genes for potential ligand-binding proteins in subregions of the olfactory mucosa. EMBO J. *10*, 2813–2819.

Doi, M., Okano, T., and Fukada, Y. (2000). Light-induced phase-delay of the chicken pineal circadian clock is associated with the induction of cE4bp4, a potential transcriptional repressor of cPer2 gene. Proc. Natl. Acad. Sci. USA *98*, 8089–8094.

Dunlap, J.C. (1999). Molecular bases for circadian clocks. Cell *96*, 271–290.

Gekakis, N., Staknis, D., Nguyen, H.B., Davis, C.F., Wilsbacher, L.D., King, D.P., Takahashi, J.S., and Weitz, C.J. (1998). Role of the CLOCK protein in the mammalian circadian mechanism. Science *280*, 1564–1569.

Ginty, D.D., Kornhauser, J.M., Thompson, M.A., Bading, H., Mayo, K.E., Takahashi, J.S., and Greenberg, M.E. (1993). Regulation of CREB phosphorylation in the suprachiasmatic nucleus by light and a circadian clock. Science *260*, 238–241.

Glossop, N.R., Lyons, L.C., and Hardin, P.E. (1999). Interlocked feedback loops within the Drosophila circadian pacemaker. Science *286*, 766–768.

Godowski, P.J., Picard, D., and Yamamoto, K.R. (1988). Signal transduction and transcriptional regulation by glucocorticoid receptor-LexA fusion proteins. Science *241*, 812–816.

Griffin, E.A., Jr., Staknis, D., and Weitz, C.J. (1999). Light-independent role of CRY1 and CRY2 in the mammalian circadian clock. Science *286*, 768–771.

Hardin, P.E., Hall, J.C., and Rosbash, M. (1990). Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. Nature *343*, 536–540.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A., and Kay, S.A. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science *290*, 2110–2113.

Hedlund, E., Gustafsson, J.A., and Warner, M. (2001). Cytochrome P450 in the brain; a review. Curr. Drug Metab. *2*, 245–263.

Hovemann, B.T., and Sehlmeyer, F. (1997). Drosophila melanogaster NADPH-cytochrome P450 oxidoreductase: pronounced expression in antennae may be related to odorant clearance. Gene *189*, 213–219.

Isaac, R.E., Siviter, R.J., Stancombe, P., Coates, D., and Shirras, A.D. (2000). Conserved roles for peptidases in the processing of invertebrate neuropeptides. Biochem. Soc. Trans. *28*, 460–464.

Jin, X., Shearman, L.P., Weaver, D.R., Zylka, M.J., de Vries, G.J., and Reppert, S.M. (1999). A molecular mechanism regulating rhythmic output from the suprachiasmatic circadian clock. Cell *96*, 57–68.

Khush, R.S., Leulier, F., and Lemaitre, B. (2001). Drosophila immunity: two paths to NF-kappaB. Trends Immunol. *22*, 260–264.

King, C.D., Rios, G.R., Green, M.D., and Tephly, T.R. (2001). UDP-glucuronosyltransferases. Curr. Drug Metab. *1*, 143–161.

King, D.P., Zhao, Y., Sangoram, A.M., Wilsbacher, L.D., Tanaka, M., Antoch, M.P., Steeves, T.D.L., Vitaterna, M.H., Kornhauser, J.M., Lowrey, P.L., et al. (1997). Positional cloning of the mouse circadian *clock* gene. Cell *89*, 641–653.

Kondo, T., Strayer, C.A., Kulkarni, R.D., Taylor, W., Ishiura, M., Golden, S.S., and Johnson, C.H. (1993). Circadian rhythms in prokaryotes: luciferase as a reporter of circadian gene expression in cyanobacteria. Proc. Natl. Acad. Sci. USA *90*, 5672–5676.

Kondo, T., Tsinoremas, N.F., Golden, S.S., Johnson, C.H., Kutsuna, S., and Isjiura, M. (1994). Circadian clock mutants of *Cyanobacteria*. Science *266*, 1233–1236.

Konopka, R.J., and Benzer, S. (1971). Clock mutants of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA *68*, 2112–2116.

Kornmann, B., Preitner, N., Rifat, D., Fleury-Olela, F., and Schibler, U. (2001). Analysis of circadian liver gene expression by ADDER, a highly sensitive method for the display of differentially expressed mRNAs. Nucleic Acids Res. *29*, E51–E61.

Krishnan, B., Dryer, S.E., and Hardin, P.E. (1999). Circadian rhythms in olfactory responses of Drosophila melanogaster. Nature *400*, 375–378.

Kume, K., Zylka, M.J., Sriram, S., Shearman, L.P., Weaver, D.R., Jin, X., Maywood, E.S., Hastings, M.H., and Reppert, S.M. (1999). mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. Cell *98*, 193–205.

Lazard, D., Tal, N., Rubinstein, M., Khen, M., Lancet, D., and Zupko, K. (1990). Identification and biochemical analysis of novel olfactory-specific cytochrome P-450IIA and UDP-glucuronosyl transferase. Biochemistry *29*, 7433–7440.

Lazard, D., Zupko, K., Poria, Y., Nef, P., Lazarovits, J., Horn, S., Khen, M., and Lancet, D. (1991). Odorant signal termination by olfactory UDP glucuronosyl transferase. Nature *349*, 790–793.

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.

Li, C., and Hung Wong, W. (2001a). Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol. *2*, RESEARCH0032.1–RESEARCH0032.11.

Li, C., and Hung Wong, W. (2001b). Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. Proc. Natl. Acad. Sci. USA *98*, 31–36.

Liu, Y., Tsinoremas, N.F., Johnson, C.H., Lebedeva, N.V., Golden, S.S., Ishiura, M., Kondo, T. (1995). Circadian orchestration of gene expression in cyanobacteria. Genes Dev. 9, 1469–1478.

McClung, C.R., Fox, B.A., and Dunlap, J.C. (1989). *frequency*, a clock gene in *Neurospora*, shares a sequence element with the *Drosophila* clock gene *period*. Nature *339*, 558–562.

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.

Meyer, U.A. (1996). Overview of enzymes of drug metabolism. J. Pharmacokinet. Biopharm. *24*, 449–459.

Michel, S., Geusz, M.E., Zaritsky, J.J., and Block, G.D. (1993). Circadian rhythm in membrane conductance expressed in isolated neurons. Science *259*, 239–241.

Millar, A.J., Carre, I.A., Strayer, C.A., Chua, N.H., and Kay, S.A. (1995). Circadian clock mutants in Arabidopsis identified by luciferase imaging. Science *267*, 1161–1162.

Mitsui, S., Yamaguchi, S., Matsuo, T., Ishida, Y., and Okamura, H. (2001). Antagonistic role of E4BP4 and PAR proteins in the circadian oscillatory mechanism. Genes Dev. *15*, 995–1006.

Naylor, E., Bergmann, B.M., Krauski, K., Zee, P.C., Takahashi, J.S., Vitaterna, M.H., and Turek, F.W. (2000). The circadian clock mutation alters sleep homeostasis in the mouse. J. Neurosci. *20*, 8138–8143.

Park, J.H., Helfrich-Forster, C., Lee, G., Liu, L., Rosbash, M., and Hall, J.C. (2000). Differential regulation of circadian pacemaker output by separate clock genes in Drosophila. Proc. Natl. Acad. Sci. USA *97*, 3608–3613.

Picard, D., and Yamamoto, K.R. (1987). Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. EMBO J. *6*, 3333–3340.

Renn, S.C., Park, J.H., Rosbash, M., Hall, J.C., and Taghert, P.H. (1999). A pdf neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in Drosophila. Cell *99*, 791–802.

Ripperger, J.A., Shearman, L.P., Reppert, S.M., and Schibler, U. (2000). CLOCK, an essential pacemaker component, controls expression of the circadian transcription factor DBP. Genes Dev. *14*, 679–689.

Rizki, T.M. (1978). Genetics and Biology of *Drosophila* (London: Academic), pp. 398–451.

Rutila, J.E., Suri, V., Le, M., So, W.V., Rosbash, M., and Hall, J.C. (1998). CYCLE is a second bHLH-PAS protein essential for circadian transcription of *Drosophila period* and *timeless*. Cell *93*, 805–814.

Sarov-Blat, L., So, W.V., Liu, L., and Rosbash, M. (2000). The *Drosophila takeout* gene is a novel link between circadian rhythms and feeding behavior. Cell *101*, 647–656.

Saunders, D.S. (1982). Insect Clocks (Oxford: Pergamon Press).

Saunders, D.S. (1990). The circadian basis of ovarian diapause regulation in *Drosophila melanogaster*: Is the *period* gene causally involved in photoperiodic time measurement? J. Biol. Rhythms *5*, 315–331.

Saunders, D.S., Henrich, V.C., and Gilbert, L.I. (1989). Induction of diapause in *Drosophila melanogaster*: Photoperiodic regulation and the impact of arrhythmic clock mutations on time measurement. Proc. Natl. Acad. Sci. USA *86*, 3748–3752.

Schaffer, R., Landgraf, J., Accerbi, M., Simon, V.V., Larson, M., and Wisman, E. (2001). Microarray analysis of diurnal and circadian-regulated genes in Arabidopsis. Plant Cell *13*, 113–123.

Schena, M., and Yamamoto, K.R. (1988). Mammalian glucocorticoid receptor derivatives enhance transcription in yeast. Science *241*, 965–967.

Schopperle, W.M., Holmqvist, M.H., Zhou, Y., Wang, J., Wang, Z., Griffith, L.C., Keselman, I., Kusinitz, F., Dagan, D., and Levitan, I.B. (1998). Slob, a novel protein that interacts with the Slowpoke calcium-dependent potassium channel. Neuron *20*, 565–573.

Scully, A.L., and Kay, S.A. (2000). Time flies for Drosophila. Cell *100*, 297–300.

Shaw, P.J., Cirelli, C., Greenspan, R.J., and Tononi, G. (2001). Corre-

lates of sleep and waking in Drosophila melanogaster. Science *287*, 1834–1837.

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., Hastings, M.H., and Reppert, S.M. (2000). Interacting molecular loops in the mammalian circadian clock. Science *288*, 1013–1019.

So, W.V., Sarov-Blat, L., Kotarski, C., McDonald, M.J., and Rosbash, M. (2000). *takeout*: a clock-regulated gene and the transcriptional regulation for novel cycling phase. Mol. Cell. Biol. *20*, 6935–6944.

Stover, C., Endo, Y., Lynch, N.J., Constantinescu, C., Vorup-Jensen, T., Thiel, S., Friedl, H., Hankeln, T., Hall, R., Gregory, S., et al. (2001). The human gene for mannan-binding lectin-associated serine protease-2 (MASP-2), the effector component of the lectin route of complement activation, is part of a tightly linked gene cluster on chromosome 1p36.2-3. Genes Immun. *2*, 119–127.

Taylor, C.A., Coates, D., and Shirras, A.D. (1996). The Acer gene of Drosophila codes for an angiotensin-converting enzyme homologue. Gene *181*, 191–197.

Theopold, U., Rissler, M., Fabbri, M., Schmidt, O., and Natori, S. (1999). Insect glycobiology: a lectin multigene family in Drosophila melanogaster. Biochem. Biophys. Res. Commun. *261*, 923–927.

Turner, A.J., Isaac, R.E., and Coates, D. (2001). The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. Bioessays *23*, 261–269.

Vang Petersen, S., Thiel, S., and Jensenius, J.C. (2001). The mannan-binding lectin pathway of complement activation: biology and disease association. Mol. Immunol. *38*, 133–149.

Wang, Q., Hasan, G., and Pikielny, C.W. (1999). Preferential expression of biotransformation enzymes in the olfactory organs of Drosophila melanogaster, the antennae. J. Biol. Chem. *274*, 10309–10315.

Williams, J.A., Su, H.S., Bernards, A., Field, J., and Sehgal, A. (2001). A circadian output in *Drosophila* mediated by *Neurofibromatosis-1* and Ras/MAPK. Science *293*, 2251–2256.

Yang, Z., and Sehgal, A. (2001). Role of molecular oscillations in generating behavioral rhythms in Drosophila. Neuron *29*, 453–467.

Young, M.W., and Kay, S.A. (2001). Time zones: a comparative genetics of circadian clocks. Nat. Rev. Genet. *2*, 702–715.

Zeng, H., Qian, Z., Myers, M.P., and Rosbash, M. (1996). A light-entrainment mechanism for the *Drosophila* circadian clock. Nature *380*, 129–135.

Zhou, Y., Schopperle, W.M., Murrey, H., Jaramillo, A., Dagan, D., Griffith, L.C., and Levitan, I.B. (1999). A dynamically regulated 14-3-3, Slob, and Slowpoke potassium channel complex in Drosophila presynaptic nerve terminals. Neuron *22*, 809–818.

Zimmer, S., Stocker, A., Sarbolouki, M.N., Spycher, S.E., Sassoon, J., and Azzi, A. (2000). A novel human tocopherol-associated protein: cloning, in vitro expression, and characterization. J. Biol. Chem. *275*, 25672–25680.