Supporting Information

Rund et al. 10.1073/pnas.1100584108

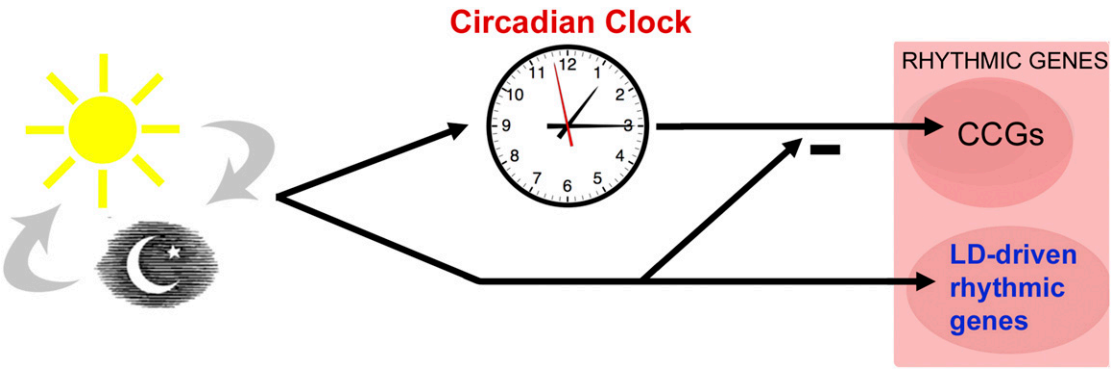

Fig. S1. Model for the regulation of 24-h rhythmic gene expression based on the current study in *Anopheles gambiae* and on studies in other organisms (1–8). Clock-controlled genes (CCGs) are a subset of rhythmic genes. CCG expression is driven by the endogenous circadian clock. Activity of the clock and rhythms in CCGs persist under constant environmental conditions. Light can synchronize or reset the clock. Under light/dark (LD) cycle conditions, additional genes are expressed rhythmically; these genes are defined as "LD-driven rhythmic genes." This direct action of the LD cycle also inhibits oscillations in a proportion of CCGs.

1. Ceriani MF, et al. (2002) Genome-wide expression analysis in *Drosophila* reveals genes controlling circadian behavior. *J Neurosci* 22:9305–9319.
2. Lin Y, et al. (2002) Influence of the period-dependent circadian clock on diurnal, circadian, and aperiodic gene expression in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 99: 9562–9567.
3. Albers HE, Liou SY, Ferris CF, Stopa EG, Zoeller RT (1991) Neurochemistry of circadian timing. In *Suprachiasmatic Nucleus: The Mind's Clock*, eds Klein DC, Moore RY, Reppert SM (Oxford Univ Press, New York).
4. Inouye ST (1996) Circadian rhythms of neuropeptides in the suprachiasmatic nucleus. *Prog Brain Res* 111:75–90.
5. Chen CH, Ringelberg CS, Gross RH, Dunlap JC, Loros JJ (2009) Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in *Neurospora*. *EMBO J* 28: 1029–1042.
6. Michael TP, et al. (2008) Network discovery pipeline elucidates conserved time-of-day-specific *cis*-regulatory modules. *PLoS Genet* 4:e14.
7. Hastings JW (1960) Biochemical aspects of rhythms: Phase shifting by chemicals. *Cold Spring Harb Symp Quant Biol* 25:131–143.
8. Wijnen H, Naef F, Boothroyd C, Claridge-Chang A, Young MW (2006) Control of daily transcript oscillations in *Drosophila* by light and the circadian clock. *PLoS Genet* 2:e39.

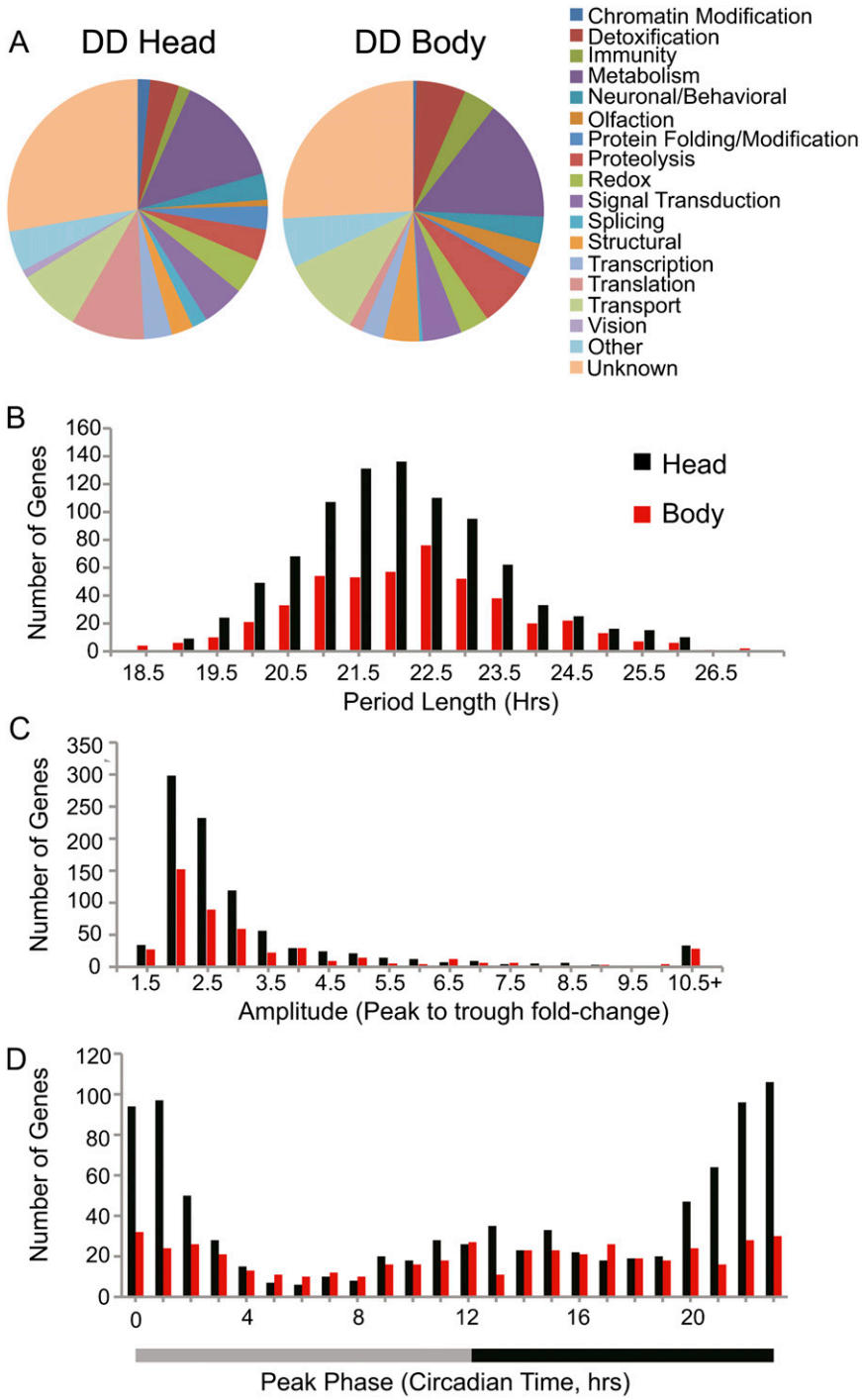

Fig. S2. *An. gambiae* circadian transcription analysis reveals a great number and diversity of rhythmic genes under constant dark (DD) conditions. (*A*) Manual annotation of rhythmic genes in both heads and bodies reveals a wide range of biological functions among rhythmic genes. (*B*) Histogram of numbers of rhythmic genes in head and body under DD conditions distributed according to the period length, as determined by the COSOPT algorithm. Under DD conditions, mean ± SEM period length of rhythmic genes was 21.83 ± 0.05 h in the head and was 22.02 ± 0.07 h in the body. (*C*) Rhythmic gene amplitude determined as peak-to-trough fold change. Under DD conditions, rhythms had a median 2.21-fold change. However, as in LD conditions, a significant number of rhythmic genes had >10-fold amplitude in expression. (*D*) Phase histogram of numbers of rhythmic genes in head and body under DD conditions distributed according to the phase of peak expression (as determined by the COSOPT algorithm). Circadian day and night are indicated by the horizontal gray/black (subjective day/subjective night) bar below the histogram. Peaks of transcriptional expression occur at the subjective dawn and dusk transitions, circadian time (CT) 0 and CT12, respectively. In *B–D* data are binned according to their value up to and including the values provided on the *x* axes.

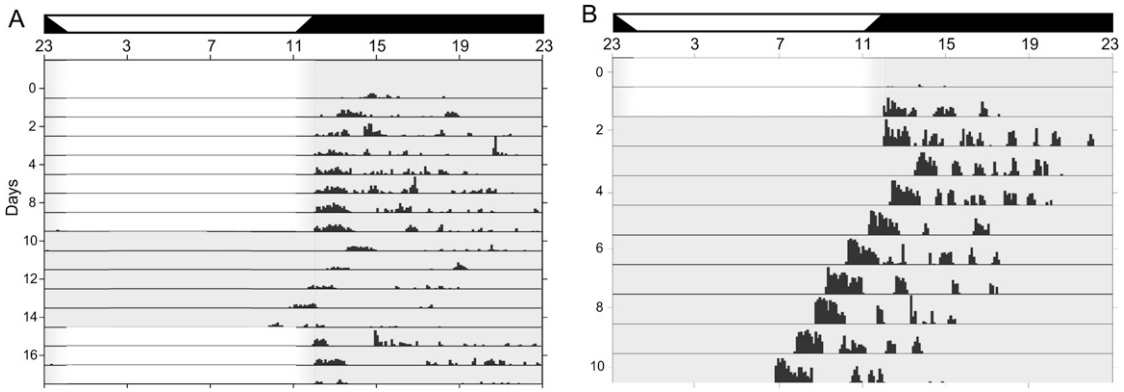

Fig. S3. *An. gambiae* locomotor/flight activity and the determination of free-running period length in *An. gambiae*. (*A*) Representative actogram of a female mosquito maintained under 12-h/12-h LD conditions (with 1-h dawn and dusk transitions), followed by 5 d of free running in constant dark (DD) conditions and then returned to LD. Individual mosquito locomotor/flight activity was monitored with the Locomotor Activity Monitor 25 (L.A.M. 25) system in which mosquitoes are placed in individual recording tubes. Data were analyzed using the ClockLab analysis program. Each horizontal line represents a 24-h period, and vertical bars represent periods of movement activity across an infrared beam. Numbers on the left indicate the number of days in the study. Day and night are indicated by the horizontal white and black bar above the actogram, and the numbers designate Zeitgeber time (ZT) in hours; ZT12 is the time of lights off, and ZT0 marks the end of the dawn transition. (*B*) Representative actogram of a female mosquito maintained under DD conditions. Mated and unmated female mosquitoes first were entrained to a 12-h/12-h LD cycle with 1-h dawn and dusk transitions for at least 7 d, followed by constant DD conditions during which activity was recorded. Mosquitoes (n = 66) displayed a free-running period length of 23.27 ± 0.03 h (mean ± SEM) under DD conditions, determined by manually fitting a line to the onset of activity in DD conditions. Only mosquitoes that persisted in free-running conditions for at least 7 d were included in the analysis of period length, and the first day in DD was excluded from analysis. Note that the small delay of 0.93 ± 0.09 h (mean ± SEM, n = 18) in onset of locomotor/flight activity during the first 24 h of transition from LD to DD has been reported previously (1). No significant differences in period lengths were found between mated and unmated females (two-tailed Student's t-test).

1. Jones MD (1973) Delayed effect of light on the mosquito "clock". *Nature* 245:384–385.

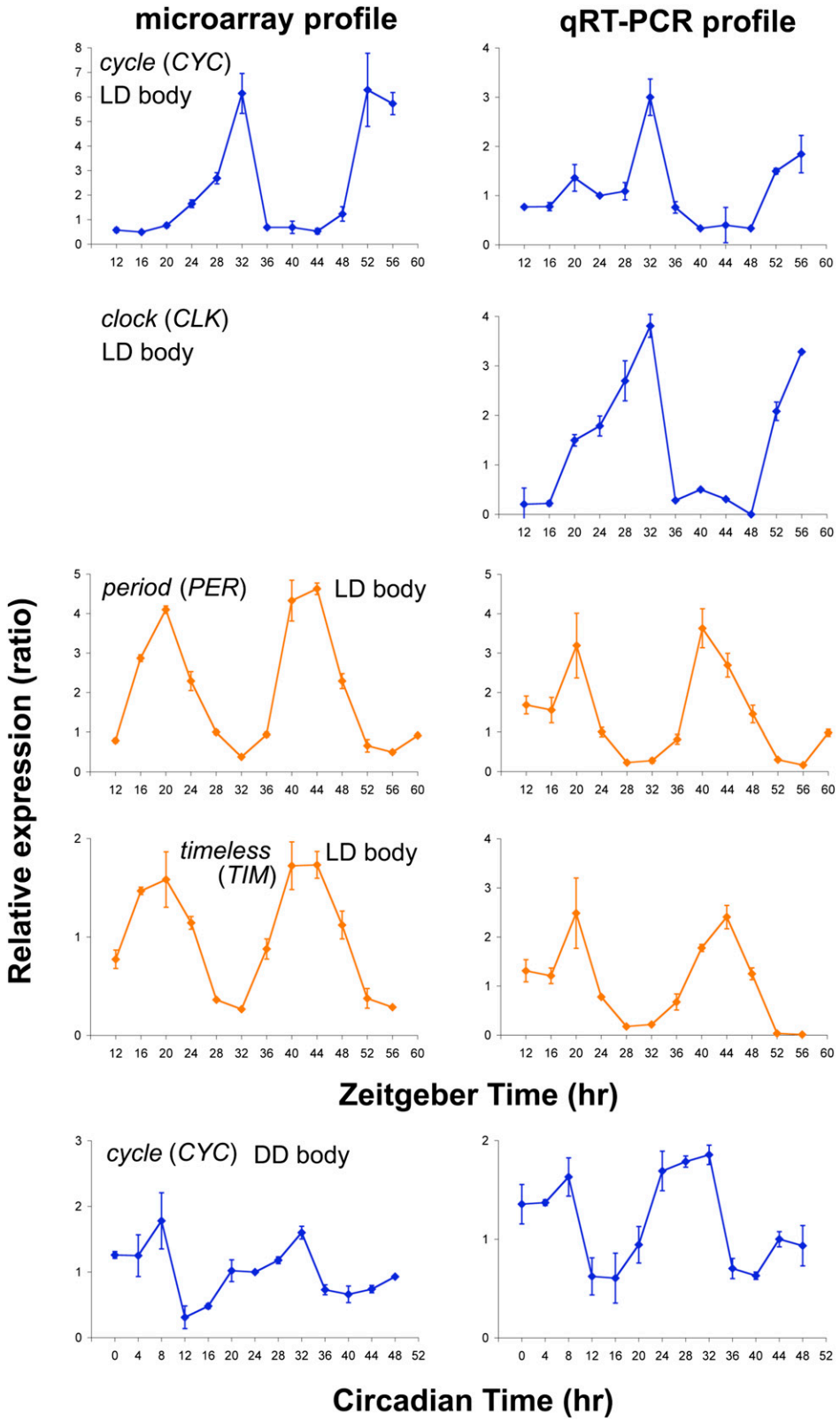

Fig. S4. (Continued)

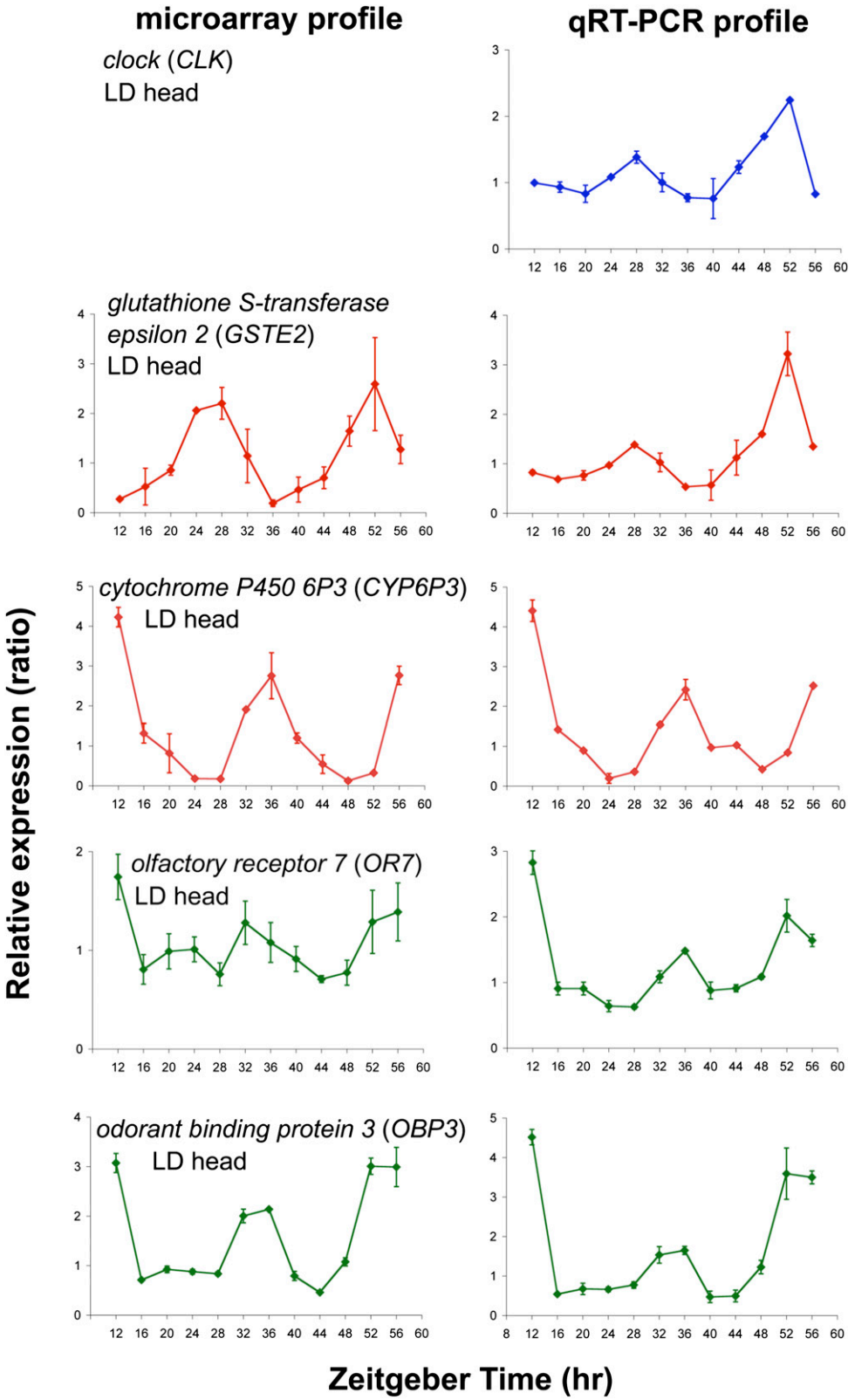

Fig. S4. Quantitative real-time RT-PCR (qRT-PCR) validates microarray analysis gene-expression profiles of rhythmic genes. (*Right*) qRT-PCR using SYBR green reagents and (*Left*) expression profiles from microarray analysis using matching RNA samples from head under LD conditions and body under both LD and DD conditions. *40S ribosomal protein S7* (*RPS7*) was used as the internal control for qRT-PCR. Both qRT-PCR and microarray data are mean ± SD fold difference compared with the median expression value (median expression value is 1.0). Clock genes are in blue and orange, representing the positive and negative elements of the transcriptional–translational feedback loop, respectively. Metabolic detoxification genes are in red, and olfaction-related genes are in green. Real-time quantitative RT-PCR primer sequences (5′→3′): *cycle* (*CYC*) (AGAP005655) forward: CATTCGAACGATGCAATCAC, reverse: GTTGATGCGCGTAGTCACTG; *clock*

Legend continued on following page

(*CLK*) (AGAP005711) forward: GTGTACGAGGACGACCAGAA, reverse: CGCTTGATGTAGCAGGAGAA; *period* (*PER*) (AGAP001856) forward: CGATAGCGGACGTGTTTGC, reverse: AGCGTGTCGTCGGAGAATTG; *timeless* (*TIM*) (AGAP008288) forward: ACATAGTGACGCTCGTGCAGTAC, reverse: GTCCTCTGGTTTGGGCGAAT; *GSTE2* (AGAP009194) forward: ATCACCGAGAGCCACGCAATCAT, reverse: GCCACCGTTCGCTTCCTCGTAGT (1); cytochrome P450 *CYP6P3* (AGAP002865) forward: AGCTAATTAACGCGGTGCTG, reverse: AAGTGTGGATTCGGAGCGTA (2); *odorant receptor 7* (*OR7*) (AGAP002560) forward: TGCTGCTACACATGCTGAC, reverse: TAGGTGACAACGGCTCCAA (3); *odorant-binding protein 3* (*OBP3*) (AGAP001409) forward: GATTCGTGCTGGAGCTCGAG, reverse: GTAAAAAGTAGTGCACCGGGTCC (4); *ribosomal protein S7* (*RPS7*) (AGAP010592) forward: CATTCTGCCCAAACCGATG, reverse: AACGCGGTCTCTTCTGCTTG (5).

1. Ding Y, Ortelli F, Rossiter LC, Hemingway J, Ranson H (2003) The *Anopheles gambiae* glutathione transferase supergene family: Annotation, phylogeny and expression profiles. *BMC Genomics* 4:35.
2. Müller P, et al. (2008) Field-caught permethrin-resistant *Anopheles gambiae* overexpress CYP6P3, a P450 that metabolises pyrethroids. *PLoS Genet* 4:e1000286.
3. Pitts RJ, Fox AN, Zwiebel LJ (2004) A highly conserved candidate chemoreceptor expressed in both olfactory and gustatory tissues in the malaria vector *Anopheles gambiae*. *Proc Natl Acad Sci USA* 101:5058–5063.
4. Justice RW, Dimitratos S, Walter MF, Woods DF, Biessmann H (2003) Sexual dimorphic expression of putative antennal carrier protein genes in the malaria vector *Anopheles gambiae*. *Insect Mol Biol* 12:581–594.
5. Dana AN, et al. (2005) Gene expression patterns associated with blood-feeding in the malaria mosquito *Anopheles gambiae*. *BMC Genomics* 6:5.

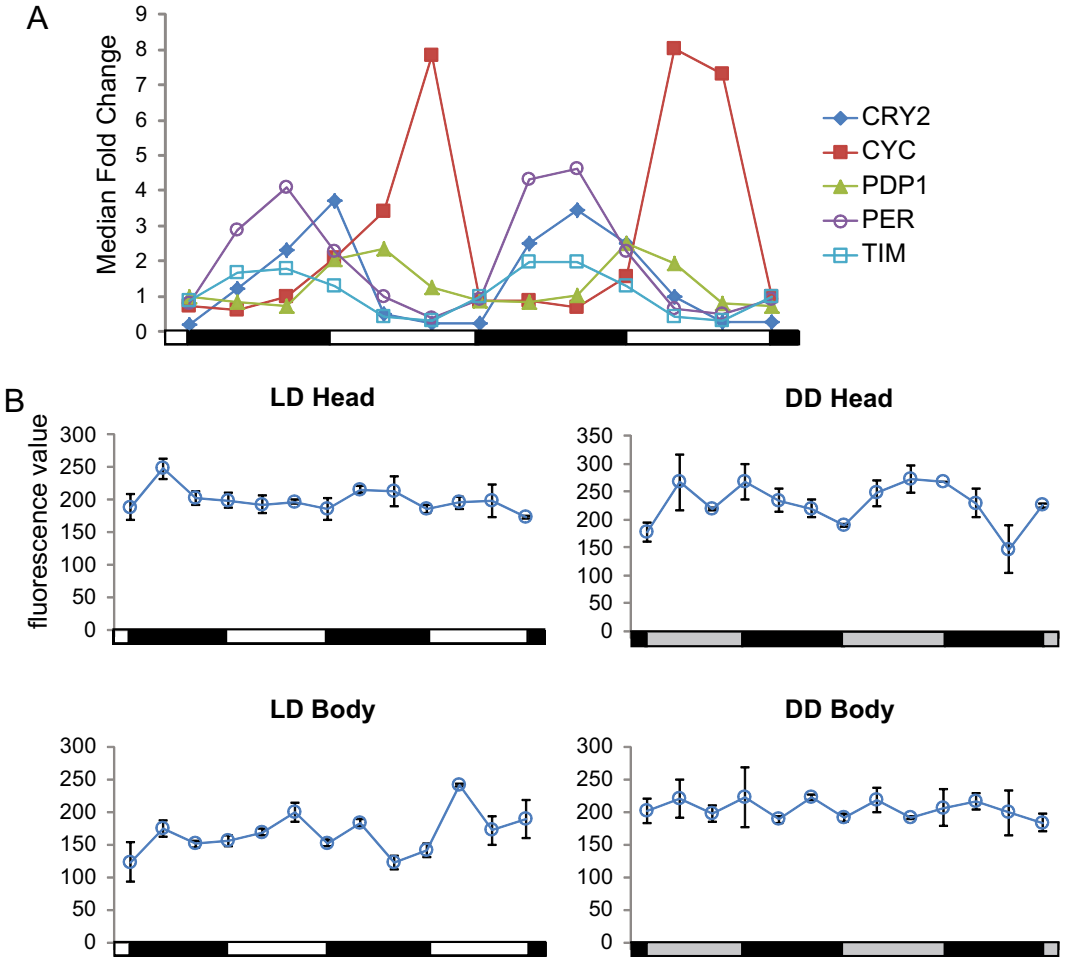

Fig. S5. Expression profiles of *An. gambiae* clock genes. (*A*) Expression of *cryptochrome 2* (*CRY2*), *cycle* (*CYC*), *PAR-domain protein 1* (*PDP1*), *period* (*PER*), and *timeless* (*TIM*) genes in mosquito bodies under LD conditions. Data have been normalized to median fold change, and SD error bars have been omitted for viewing purposes. See Fig. 2 for clock gene profiles under LD conditions in mosquito heads. (*B*) The expression profile of *cryptochrome 1* (*CRY1*) in all four tissue/conditions revealed a constitutive pattern of expression, i.e., no or limited rhythmic expression. Data are normalized microarray probe set mean ± SD fluorescence values.

A

Gene	Ensembl ID	*An. gambiae* Identification
timeless, TIM	AGAP008288	Here identification reassigned from AGAP010787 (by Das and Dimopoulos 2008 (1)) reflecting additions to the *An. gambiae* gene assembly beginning with assembly 45. Phylogenetic analysis using VectorBase (2) now reveal AGAP008288 to share higher percent identity to *Drosophila timeless* than the previously identified candidate gene (see Figure S6B). *Timeless* has been previously identified in *Cx. quinquefasciatus* (CPIJ007082) and *Ae. aegypti* (AAEL006411) by Gentile *et al.* 2009 (3).
cryptochrome 1, CRY1	AGAP001958	VectorBase annotated gene. Previously identified as the *Drosophila*-like *cryptochrome 1, CRY1* by Zhu *et al.* 2005 (4)
cryptochrome 2, CRY2	AGAP004261	Previously identified as the mouse-like *cryptochrome 2, CRY2* by Zhu *et al.* 2005 (4)
period, PER	AGAP001856	Previously identified by Das and Dimopoulos 2008 (1)
cycle, CYC	AGAP005655	Previously identified by Das and Dimopoulos 2008 (1)
clock, CLK	AGAP005711	Previously identified by Das and Dimopoulos 2008 (1)
PAR-domain protein 1, PDP1	AGAP006376	Here identified based on phylogenetic analysis using VectorBase Figure S6B (2). *PAR-domain protein 1* has been previously identified in *Cx. quinquefasciatus* (CPIJ014920) and *Ae. aegypti* (AAEL005255) by Gentile *et al.* 2009 (3).

B

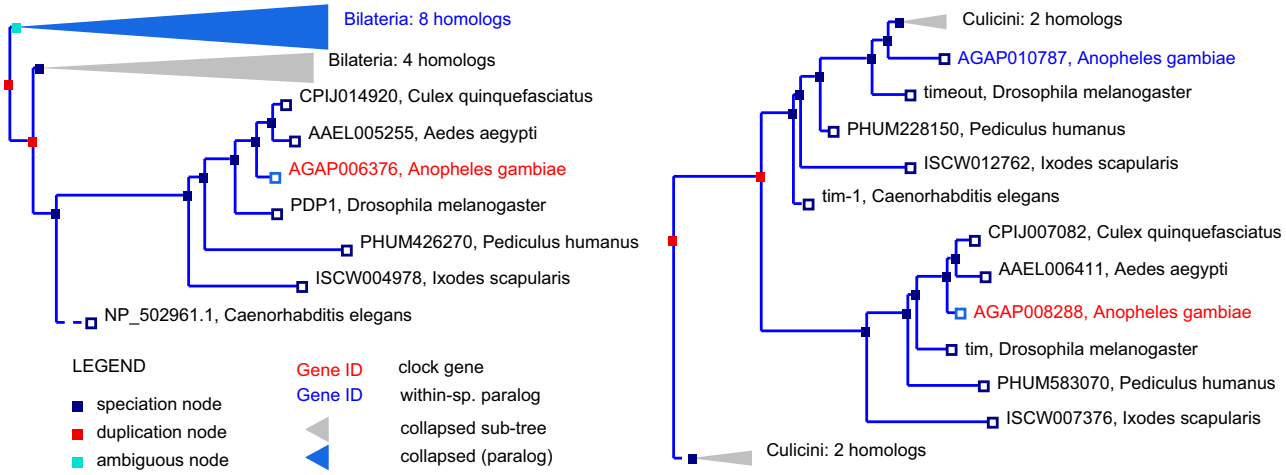

Fig. S6. (*A*) Identification and annotation of *An. gambiae* clock genes. (*B*) Identification of *An. gambiae timeless* (*TIM*) and *PAR-domain protein 1* (*PDP1*). The two clock genes were identified in the current study based on gene tree analysis in VectorBase (2). *Timeless* has been identified previously in *Culex quinquefasciatus* (CPIJ007082) and *Aedes aegypti* (AAEL006411) (3). Similarly, *PDP1* has been identified previously in *Cx. quinquefasciatus* (CPIJ014920) and *Ae. aegypti* (AAEL005255) (3). VectorBase gene trees were constructed using Tree Building guided by Species Tree (TreeBeST, http://treesoft.sourceforge.net/treebest.shtml) and utilization of protein alignments generated with the multiple sequence comparison by log-expectation (MUSCLE) algorithm (5).

1. Das S, Dimopoulos G (2008) Molecular analysis of photic inhibition of blood-feeding in Anopheles gambiae. *BMC Physiol* 8:23.
2. Lawson D, et al. (2009) VectorBase: A data resource for invertebrate vector genomics. *Nucleic Acids Res* 37(Database issue):D583–D587.
3. Gentile C, Rivas GBS, Meireles-Filho ACA, Lima JBP, Peixoto AA (2009) Circadian expression of clock genes in two mosquito disease vectors: *Cry2* is different. *J Biol Rhythms* 24:444–451.
4. Zhu H, Yuan Q, Froy O, Casselman A Reppert SM (2005) The two CRYs of the butterfly. *Curr Biol* 15:R953–R954.
5. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.

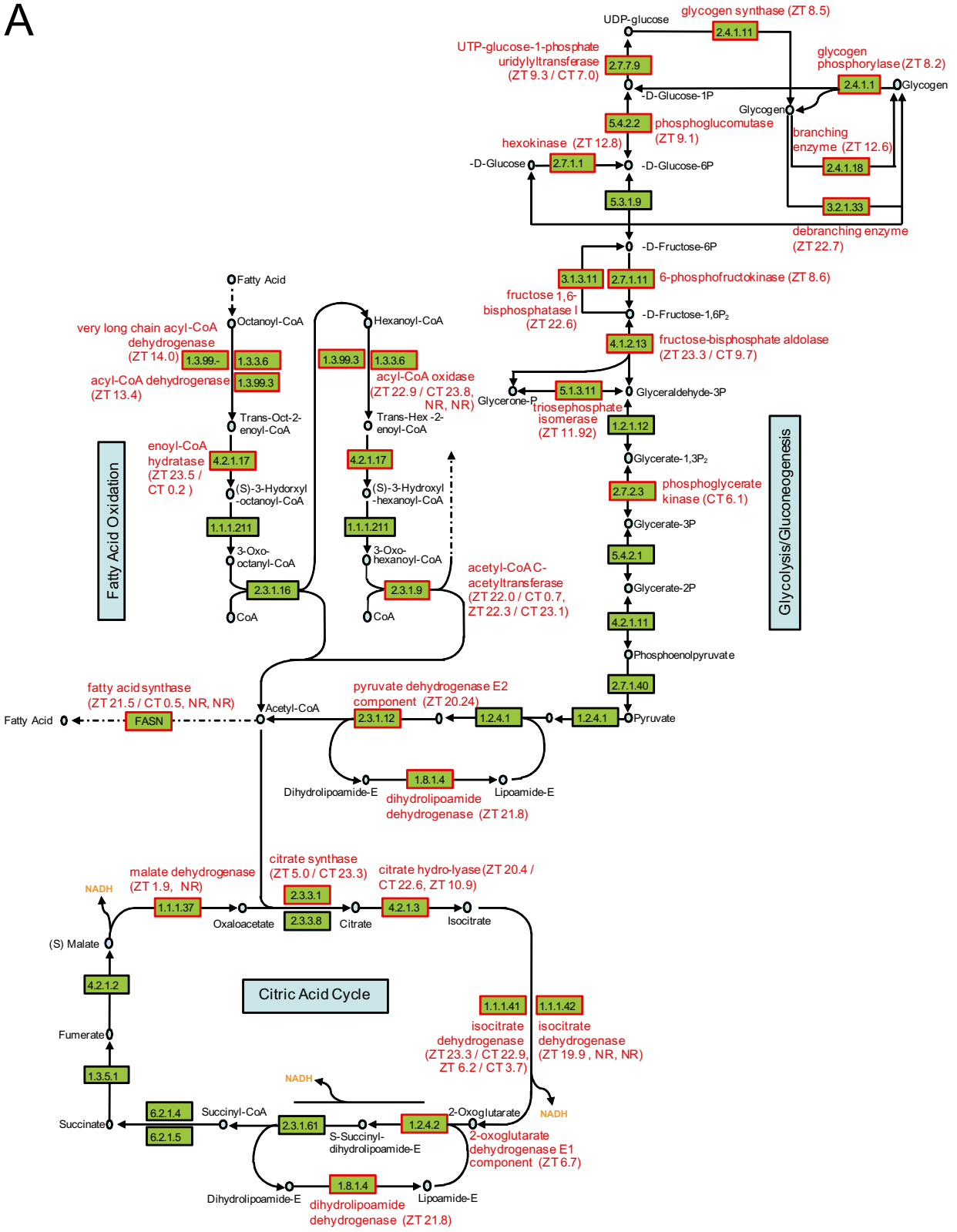

Fig. S7. (Continued)

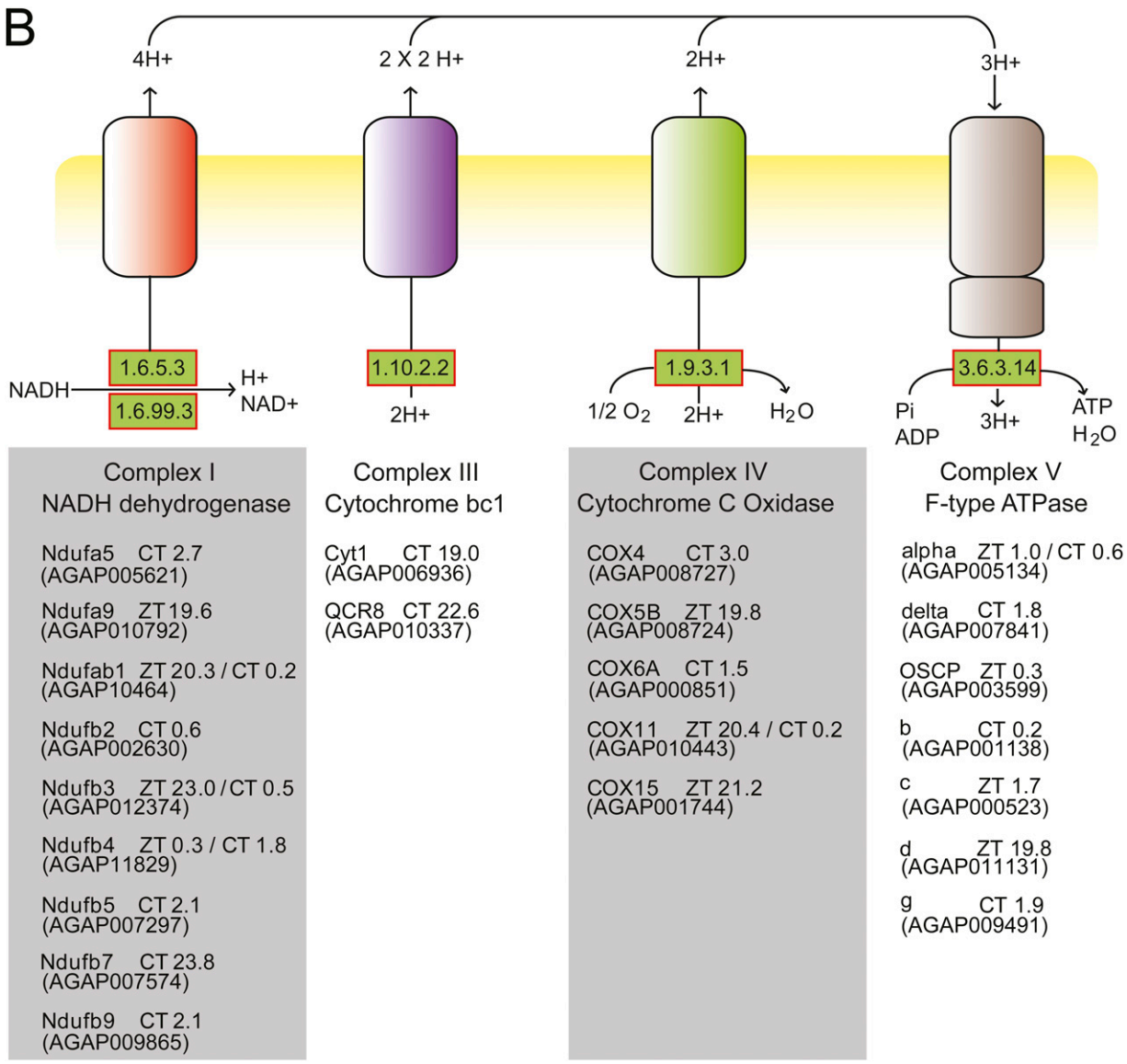

Fig. S7. Genes associated with fatty acid oxidation, glycolysis/gluconeogenesis, the citric acid cycle, and oxidative phosphorylation are rhythmically expressed in *An. gambiae* heads under LD and DD conditions. (*A*) Green boxes represent enzymes that have been identified in *An. gambiae*; numbers in boxes are the Enzyme Commission classification number for that enzyme. Red outlined boxes indicate genes found to be rhythmic in heads under LD or DD conditions. Peak phase under diel (ZT) and circadian (CT) conditions for rhythmic genes is indicated in red text along with the inferred gene name. Some enzymes have more than one gene associated with them and thus have multiple peak phases reported or may not be scored as rhythmic in our analysis; these enzymes are noted as "NR." Note: Many enzymes in fatty acid oxidation are present in more than one reaction; this information has been condensed in the figure. (*B*) Rhythmic components of oxidative phosphorylation and their peak phases in LD and DD heads. Metabolic pathways and genes were predicted using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (1, 2).

1. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.
2. Dennis G, Jr., et al. (2003) DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol* 4:3.

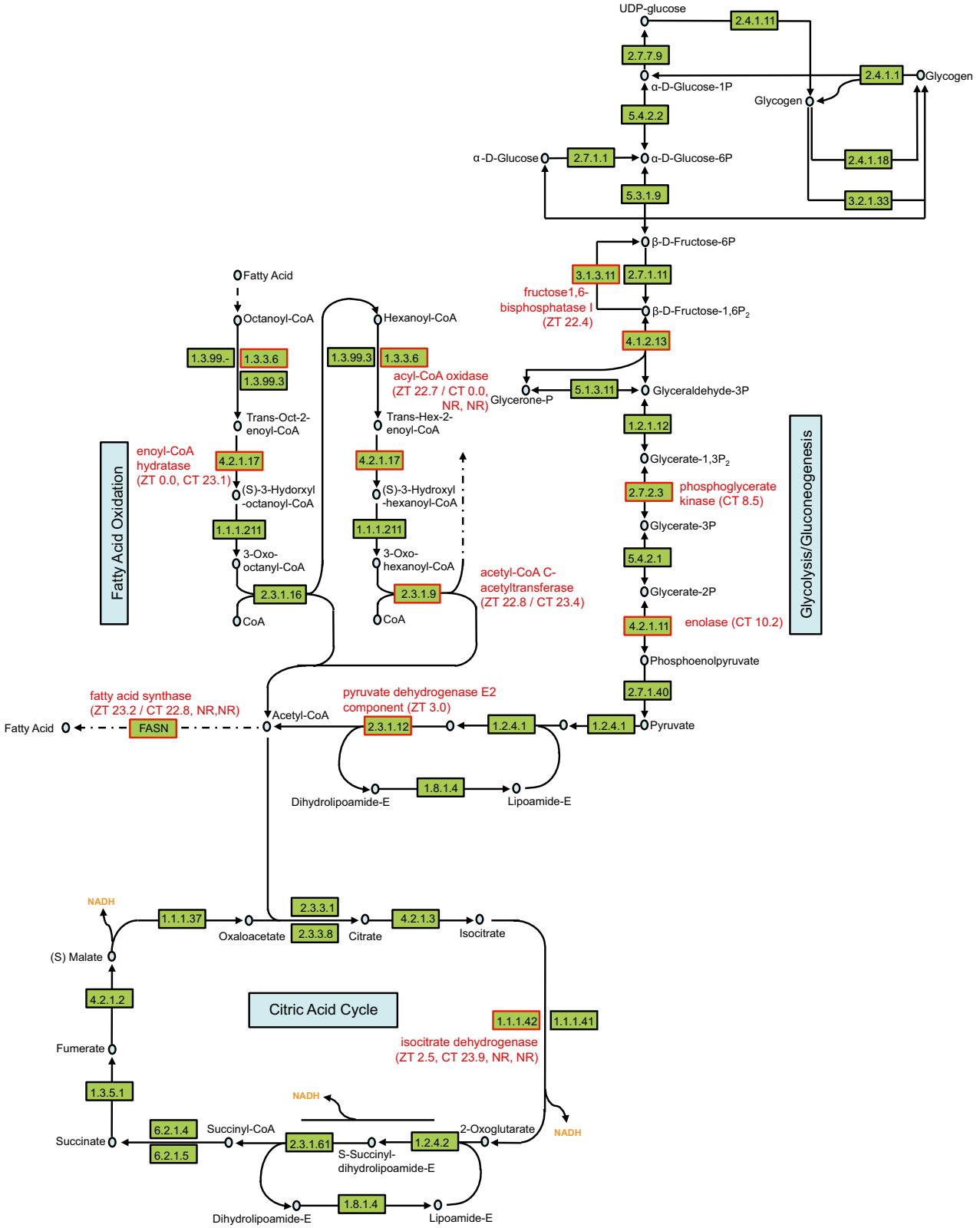

Fig. S8. Genes associated with fatty acid oxidation, glycolysis/gluconeogenesis, and the citric acid cycle are rhythmically expressed in *An. gambiae* bodies under LD and DD conditions. (*A*) Green boxes represent enzymes that have been identified in *An. gambiae*; numbers in boxes are the Enzyme Commission classification number for that enzyme. Red outlined boxes indicate genes found rhythmic in heads under LD or DD conditions. Peak phase under diel (ZT) and circadian (CT) conditions for rhythmic genes is indicated in red text along with the inferred gene name. Some enzymes have more than one gene associated with them and thus have multiple peak phases reported or may not be scored as rhythmic in our analysis; these enzymes are noted as "NR." Note: Many enzymes in fatty acid oxidation are present in more than one reaction; this information has been condensed in the figure.

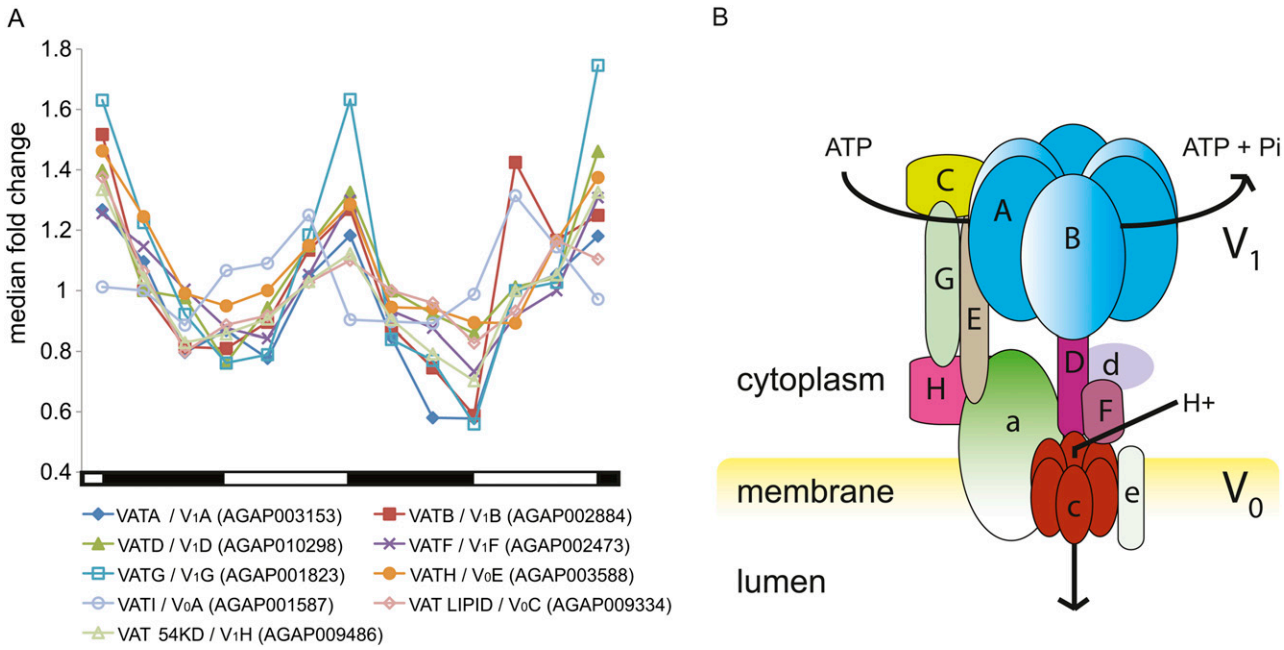

Fig. S9. Multiple subunits of vesicular type H^+ ATPase (V-ATPase) are rhythmically expressed in LD bodies. (*A*) Nine of the 12 V-ATPase subunits are rhythmically expressed and are phase concordant. Apart from the *An. gambiae* gene *V-type proton ATPase catalytic subunit A* (*VATA*) (AGAP003153), all other genes shown are orthologs predicted using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (1, 2). Data have been normalized to median fold change. (*B*) Model of V- ATPase showing the V_1 and V_0 complexes. V_1 subunits are represented by capital letters, and V_0 subunits are represented lowercase letters.

1. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.
2. Dennis G, Jr., et al. (2003) DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol* 4:3.

Dataset S1. ***Anopheles gambiae*** **genes that are rhythmic in both tissues and both environmental conditions, genes that are rhythmic in heads under LD and/or DD conditions, and genes that are rhythmic in bodies under LD and/or DD conditions**

Dataset S1

(Tab 1) Genes rhythmic in both tissues and both environmental conditions. The COSOPT-derived probability, multiple means corrected β (pMMCβ) value and peak phase data are provided for each gene from each of the four individual experiments. (Tabs 2 and 3) COSOPT-derived pMMCβ value and peak phase data are provided for each gene found rhythmic in the head (Tab 2) or body (Tab 3) under LD and/or DD conditions. Gene lists are sorted first by category and then by Ensembl number (VectorBase identity). Where a gene is represented by multiple Affymetrix probe sets on the GeneChip, the data showing the lowest COSOPT pMMCβ value (i.e., highest significance value) and associated phase determination are provided. If more than one Affymetrix probe set number is provided in the data file, the second probe set number refers specifically to the COSOPT pMMCβ and phase data for genes examined under DD conditions.