

## Supplementary Information

### Methods

#### Behavioral analysis

Rhythmic flies were defined by  $X^2$  periodogram analysis with the following criteria (filter on): power  $\geq 20$ , width  $\geq 2$  h, with no selection on period value. Power and width are the height and width of the periodogram peak, respectively, and give the significance of the calculated period. Mean daily activity (number of events per 0.5 h) is calculated over the whole period in LD or DD conditions.

#### Immunolabelings

Fluorescence signals were analyzed with a Zeiss Z1 epifluorescence microscope equipped with a AxioCam MRm digital camera. A Zeiss apotome confocal module was used for some experiments (indicated in legends). For apotome images, stacks of 5-15 confocal sections were used generate a maximum intensity projection. Fluorescence intensity was quantified from digital images with the ImageJ software. We calculated the fluorescence intensity by applying the formula  $I = 100n(S - B)/N*B$ , which gives the fluorescence percentage above background (where n is the number of labeled cells among the N cells of the considered group (PDF-expressing sLNvs: 4, larval DN1s: 2), S is fluorescence intensity and B is average intensity of the region adjacent to the positive cell).

#### Primers for quantitative RT-PCR

TCCTTGTCGCGTGTGAAACA (5') and CCGAACGAGTGGAAGATGAG (3') for *tubulin* (464 bp), CCAAATAAATCCACTGAAAACAT (5') and GAGAGGCACCATTTTCGGAGTAC (3') for *clk* (367 bp), ACCCGCATCCTTCGCTTTTCTACA (5') and GGGCAGGAGTGGTGACCGAGTGGA

(3') for *per* (219 bp), AGTTGGTCATGCGCAGCAAATG (5') and TCTTTTCGTACACAGATGCCA (3') for *tim* (448 bp), CCGCCGGGCAAACCTCTATC (5') and GCGACTGGGCCATCATTCC (3') for *vri* (298 bp), GCGGCAACTGGTAATG (5') and ATTCCTGCCTGAGCT (3') for *pdp1e* (217 bp), GTACGTCCCGGAGTTGATGAATGT (5') and CACGTCCGGCCAGCCAGAAGAACT (3') for *cry* (252 bp).

### Degradation assays

dsRNA constructs were generated as described previously (Hung *et al.*, 2007), using primers that contained a 5' T7 RNA-polymerase binding site followed by gene specific sequences for *egfp* (Hung *et al.*, 2007), for exon 6 of *ctrip* in OBS 66948: TAATACGACTCACTATAGGGACTCAGCATGGTAGTGGATCG (5') and TAATACGACTCACTATAGGGCATGTTTAGGCAGCAATTAGCC (3'), and for exon 3 of *ctrip* in OBS 71876: TAATACGACTCACTATAGGGATCACACCAGACTTGCTACTGG (5') and TAATACGACTCACTATAGGGGGAGCGGTAGAAATCAACTTAGG (3'). DNA fragments of the coding region of *egfp* and *ctrip* target genes were generated by PCR, *in vitro* transcribed using the T7 RiboMAX Express Large Scale RNA Production System (Promega) and RNA products were purified using the SV Total RNA Isolation System (Promega). Purified dsRNAs were annealed by incubation at 65°C for 15 min and slow cooling to room temperature. To analyze for effects of dsRNAs on CLK degradation kinetics, 1 ml of *Drosophila* S2 cells ( $3 \times 10^6$  cells/ml) were incubated in serum-free Schneider's insect medium (Sigma) with 15  $\mu$ g of dsRNA for 2 hours prior to the addition of 1 ml of Schneider's insect medium containing 20 % FBS. Cells were incubated for 48 hours prior to the transfection of 3 mg of CLK-expression plasmid (pAc-*Clk*) with 10 ml of Lipofectin (Invitrogen) as described previously (Hung *et al.*, 2009), with the difference that another 15  $\mu$ g of dsRNA were added 45 min prior to the addition of 20 % FBS. Two days after

transfection, cells were split into 24 well plates and incubated for 24 hours. Subsequently, new protein synthesis was inhibited by the addition of cycloheximide at a final concentration of 0.58 mM. Cells were harvested over a time-course as indicated in the figure after the addition of cycloheximide and CLK protein levels were determined by western blot analysis, using the anti-CLK antibody described in Hung *et al.*, (2007). CLK protein levels were quantified for different time-points by densitometry and normalized towards an unspecific control band. Amounts of CLK proteins are shown relative to the amount of CLK prior to the addition of cycloheximide (time-point 0 set to 1).

## Legends to the Supplementary Figures

**Fig S1: *ctrip* cDNA and predicted protein sequences.** The cDNA sequence of the longest isolated head transcript is shown (GenBank JF775582). Exons 1, 3, 5, 7, 9 and 11 are in black, exons 2, 4, 6, 8, 10 and 12 are in blue, with exon 4 split in two parts, defined by a splicing acceptor site that was found in some transcripts (GenBank JF775584). The first 30 nucleotides (in red) represent a 5' extension (exon 1) of the FlyBase sequence that we identified by 5' RACE analysis. Polymorphic nucleotides changes (compared to the Flybase sequence) are shown in red (14 sites). The CAG triplet at the E4-E5 junction (indicated in green) was present in half of the E5-containing cDNAs, reflecting the use of two different splicing acceptor sites. Two protein sequences are shown, which correspond to the longest cDNA (first) and the most abundant (second, lacking exons E4-E5, Genbank JF775583) predicted proteins. The 12 polymorphic nucleotides of the coding region generate three amino-acid changes (C197G, Q852H, V1069M) indicated in red, in addition to the optional A913 that is shown in green (numbering according to the longest protein sequence). The Armadillo, WWE and HECT domains are shown on the sequence of the most abundant protein as shaded in yellow, green and grey, respectively.

**Fig S2 *ctrip* transcripts do not cycle in the head.** Quantitative RT-PCR for three different *ctrip* cDNA regions in LD: E3-containing transcripts (all transcripts) (blue), E3-E4b -containing transcripts (red) and E3-E4a -containing transcripts (green). mRNA abundance values of E3-E4a and E3-E4b were multiplied by 10 for easier comparison. Black bars represent s.e.m., which is calculated for time point values averaged from three independent experiments. Other values are averaged from two independent experiments. Primers: E3 (377 bp):

5': TCGAATTCAGAAATCAACTTAGGTGCAGC, 3':

CGGTAGTGTCCCTCCGCTGGTAT; E3-E4b (517 bp): 5':  
TGCTGCTGAGACTGGAGGTGAA, 3': GCTGCTGTTGTGGGTTGTAGAC; E3-E4a (550  
bp): 5': TGCTGCTGAGACTGGAGGTGAA, 3': GAGGTGGCTGCTGTTGTTGGTGT.

**Fig S3 Immunofluorescence triple-labelings of larval (top) and adult (bottom) sLNvs at CT0.** CT0 Images are shown to illustrate the labelings that were used for the quantifications shown in Figure 2. Only one focal plan is shown that contains either two or four neurons. CLK is green, PER is red and TIM is blue. Although GP47 recognizes non CLK specific epitopes in non clock cells (Houl *et al.*, 2006, 2008), no GP47 signal could be observed in the clock neurons of *clk<sup>irk</sup>* larval and adult brains (not shown). Scale bar represents 10µm.

**Fig S4 : *ctrip* but not *Hus1-like* affect CLK and PER levels.** Quantification of CLK and PER immunolabelings in the larval DN1s (A), larval sLNvs (B) and adult sLNvs (C) at CT0. Images were taken with an apotome module. *w* flies are used as controls. Df6142 is for the Df(3R)Exel6142 deficiency, which deletes the 82D2-82D6 region of chromosome 3R chromosome and thus includes *ctrip* (Parks *et al.*, 2004). Df6243 is for Df(3R)Exel6143, which deletes the 82E4-82E7 adjacent region that does not include *ctrip*. Error bars indicate the s.e.m. (n=20-25 brain hemispheres).

**Fig S5: Clock protein oscillations in flies expressing *ctrip* RNAi under *pdf-gal4* control**

CLK, PER and TIM immunofluorescence in the adult sLNvs. Flies were entrained in 12:12 LD cycles and transferred to DD, at 25°C. Brains were dissected every four hours during the first and second days of DD. Labelings and quantifications of the signal were done as

described in the Methods section. Error bars indicate the s.e.m. (n=20-25 brain hemispheres). Grey and black bars represent subjective day and night, respectively. Circadian Time is indicated in hours.

**Fig S6 CLK levels in head extracts of *ctrip* RNAi flies and controls.** Quantification of anti-CLK western blots in LD and DD1 (as illustrated in Fig. 3A). Autoradiograms of anti-CLK and anti-PER (after stripping the membrane and reprobing with anti-PER antibody) western blots were scanned. Files were analyzed with ImageJ software to quantify the CLK band and a PER unspecific band that was used as a loading control (see Fig. 3A). A CLK/control ratio was calculated and values were normalized to the *tim-gal4/+* ZT0 or CT0 value for each gel. Curves represent mean values from two (top) or three (bottom) independent gels. Top: Triangles (*tim-gal4/+*) and squares (*tim-gal4/ctrip RNAi*) bracketing each mean value represent the two independent values that were used to calculate the mean. Bottom: bars represent the s.e.m.

**Fig S7 PER and TIM oscillations in head extract of *ctrip* RNAi flies.** Western blot analysis of PER and TIM protein in head extracts during LD and the first day of DD. Flies were collected every 1.5h during day or subjective day and every 3h during night or subjective night. Control: *w;tim-gal4/+*, *ctrip* RNAi: *w;tim-gal4/+;UAS-ctrip RNAi/+*.

## References

Houl JH, Yu W, Dudek SM, Hardin PE (2006) Drosophila CLOCK is constitutively expressed in circadian oscillator and non-oscillator cells. *J Biol Rhythms*, **21**: 93–103

Houl JH, Ng F, Taylor P, Hardin PE (2008) CLOCK expression identifies developing circadian oscillator neurons in the brains of Drosophila embryos. *BMC Neurosci*, **9**: 119

Hung HC, Maurer C, Kay SA, Weber F (2007) Circadian transcription depends on limiting amounts of the transcription co-activator nejd/CBP. *J Biol Chem*, **282**: 31349–31357

Hung HC, Maurer C, Zorn D, Chang WL, Weber F (2009) Sequential and Compartment-specific Phosphorylation Controls the Life Cycle of the Circadian CLOCK Protein. *J Biol Chem*, **284**: 23734–23742

Parks AL, Cook KR, Belvin M, Dompe NA, Fawcett R, Huppert K, Tan LR, Winter CG, Bogart KP, Deal JE, Deal-Herr ME, Grant D, Marcinko M, Miyazaki WY, Robertson S, Shaw KJ, Tabios M, Vysotskaia V, Zhao L, Andrade RS, Edgar KA, Howie E, Killpack K, Milash B, Norton A, Thao D, Whittaker K, Winner MA, Friedman L, Margolis J, Singer MA, Kopczynski C, Curtis D, Kaufman TC, Plowman GD, Duyk G, Francis-Lang HL (2004) Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. *Nat Genet*, **36**: 288–292

Genotype	Total flies	%R	Tau (s.e.m)	Power (s.e.m)
<i>w;tim-gal4,UAS-gfp/+</i>	16	100	24.3 (0.1)	105 (8)
<i>w;;17R2,17R3/UAS-ctrip</i>	10	80	24.0 (0.1)	113 (14)
<i>w;;17R3,14R3/UAS-ctrip</i>	14	100	24.2 (0.2)	125 (7)
<i>w;tim-gal4/UAS-gfp;UAS-ctrip</i>	22	97	23.6 (0.1)	126 (6)
<i>w;tim-gal4,UAS-gfp/+;17R3,14R3/+</i>	28	93	26.2 (0.1)	123 (8)
<i>w;tim-gal4/+;17R3,14R3/UAS-ctrip</i>	31	100	24.5 (0.1)	119 (6)
<i>w;tim-gal4,UAS-gfp/+ ;17R2,17R3/+</i>	17	60	25.7 (0.1)	126 (10)
<i>w;tim-gal4/+;17R2,17R3/UAS-ctrip</i>	22	91	24.1 (0.1)	128 (6)

**Supplementary Table 1:** Activity rhythms of *ctrip* RNAi rescued flies in DD.

R: Rhythmic flies; Tau: period value in hours; s.e.m: standard error of the mean; *14R3* and *17R3* are *UAS-ctrip* RNAi transgenes targeting *ctrip* exons 2 and 6, respectively. *17R2* (RNAi line *17735R2*) and *17R3* are different insertions of the same *CG17735* RNAi transgene targeting the *ctrip* E6 exon. *tim-gal4* -driven *UAS-ctrip* induces a 4 fold increase of *ctrip* mRNA levels in heads, which were reduced by two folds in the presence of *ctrip* RNAi (not shown). *UAS-gfp* is added to mimick putative titration of the GAL4 protein by UAS sequences that may occur when adding *UAS-ctrip* to *ctrip* RNAi flies. As opposed to *UAS-ctrip*, *UAS-gfp* does not change the circadian period of the *ctrip* RNAi flies.



Figure S1

ctrip longest cDNA (E1-E12, 10 587 bp)

GTTTCGTAGCAAGACGCGGTCTCACTGGTGCACATTGTTTTCCGTCGCGTGTGAAAAATTGCTG  
AGTGTATAAAAACTGTGCATACTTAGAGAAAACCTTCTGCATTTGGTTGAAAGCACTACCAAT  
CTGTCCAACgAAGCACGCGAATGGGAAATCAGAACGGAAAACCTGTTAAAACTTGAAGTGAAA  
GAGTCAAAATCACAGCTTATGAAATAGTGTGTTAATAAGAAAGCGACAGTAAATTGAATCC  
ACAGCTGCTTAGAGACTATTACCATCCTTTAGGCAACGGACCGATTACGAGCAAACATTTTCG  
AGATCAATATTCCATTATTGGTTGTGTTGGCATGTTAAGTGTGTTTCGATGCCGGTGCGGTTA  
AGTGTGCTCTATGGCCGAATCCGTTAAAAGTCAATCGCTCTCTGCATTGACGGAGGGGCAGC  
ACGACGACGGTGGTTCAACAGCAACATCAGCGACATTGGCTAATAAAAACCAGTGGGACCAAC  
ACGACAAGTGCATCCAATCACTCGTCCCAGAGGCGACGCAACAACAACACAAACAGCAACAG  
AAACAAGAGTCACCACAAAACGAGAAACGCAACGCACCGGCTACATCAGGCCCCGCTGTAA  
GCTCCTCCGTCGACGTTGTACCCGCCACCTCGACTGCATCTGCAACCTTACCCGTCGGTCT  
CGCAGTCAAGGTGCCACTCCTCCGCCGAGCCCTCAGCAGCAAGGAGTCACAAATCCCCAA  
GCGCTCCGTGGGAGCCATCCGCTCCCATCTTACCAGCATTCAACTCCATACCGGTCGTCA  
GTGATCAATCGCTATCGCCCGGCAGTCGTAAGCGCCAGCTTAACCAGCACAACAACAGTAAC  
AACAAAGCCAGCAGTAGCCACCAATCGGCGCCCGGGGGAGATCAGCTGGTTCGCGCGCCCT  
TAAAAAACGCCGCTTGCAGCAGTCTTTGGGATCCGGTATCGGCGAGGGAGCGGTAGAAATCA  
ACTTAGGTGCAGCAAGCTTTGACCACACTCCTCGTCAGGCATCCGGAGACTGTGTTGCACAG  
CGCACTCGATCTAAGACCATTTCCGCCGAGGAGTTGCCAAGCACAAGCAGCGCTGCGGCAGC  
ACGCCATCACCAGATTCGCCCCAGTGCCAGCTTACCAGCTTCTCGAGCAGCCATCGAAACA  
AACGAAAGGCCAGTATTGGAGGATCAATAGCTGCGGTGAGACCACTCCGCATAGAGGAACA  
GCAGCTGCTCGCGCCGACGACGCGGTAACCTGCTAAACTATTACAGAAAACGCGTAAGGT  
TAGCCACACACGCTCCTCCCAAAGTCGAAAAGCAGTCTATAACCAGCGGAGGACACTACGG  
CCAACCGAAACGGATCAGCCGGATCAGGATCCAGTAGCAAGTCTGGTGTGATACCCAAGAAC  
AGGAAGAGCCTGCGTCATTCGCAGCAGAACCTGCTGGACACAGAGGCGGCAGAAGTAGAGAG  
TGCTGCTGCTGAGACTGGAGGTGAACAGCAGTCCGAGCACGGAAATCTGGACGTGATCGACC  
AACTGCCGACCCCTGAAGCTGGTCTTAATCAATCAGAGGCATCAGGAATCCACAACCAGAGT  
CAATTGGAGGCGGATCGCCAGCCCGTGGTGCAGGACGAAGACGACGAGGAGGACGACGAGGA  
AGAGGAAGAAGAGGAGGAAGAGGAGGTGGGCTTCTACGGGATTGTAAACTCTGCTGGCTCGT  
CATACGAGGAAGACACACAAATCGTTGCCGAGGAGGACGAGATCACAACGGAGGAGGAAGTT  
GACGACGAGGACGACGACGAAATCGAGGAAGAAGACCTCTCGGAGAGCGAATTCGCGCAGCA  
ACTGATTGGTGAACCTGGTGAGAAACGACAACAACCTCAACGTGCCACCAACAACGCAAAAA  
AAACCAAGAACAACACACCGGCGGTAGCAGCAGCCAGCACTATCAAGGCGACAACAACAGCA  
ACAACGTCGTCCGGTTATCAACAACGTGCCGCCCCCCCATGGACAGGGGGCGGCGCCAGGCGC  
TGGAGGCGGCGGTCCGCGTGCCACGCGCAGCAGCAACAGCAACTCCAGCAACAACGTTGTTG  
ACTACAGCATTATGCCGCACCTCACAACGGTCGCAACAACCACACCACATGCCTTGACACCG  
CAACAGCACCAACAACAACGACAACCGCCACAACCTCCGTCGCATCAGCAACAACAAGTACA  
CCAACAACAGCAGCCACCTCAATTGTTGCCACACATCAGTTTGCACACTTGTGCTCCTTTG  
TGACACCGACAGCAGCGGCAGCAGCAGCAGCAGCGGCGGCCACTATCCCCAGCTTCTGCT  
GCTGCCGCTTACTTTGCACAGCAGCAACACCATCAGCATCAACATCAGCATCAGCAACAACA  
CGTGCAGCAGGCACAGGCACCGCAGCCGCACTATTACCATAGTTCAGTGGCTGTGTTGCATC  
AACCTCCACCGCACCATTTTACCTCCACAGGAGCAGGACCACCGCCAGCACTCTTTTCAGCAG  
CAGGCACCGCCGAGCTAACCCGCTACACGCCAGCAACAGCTACAACAGCGGCTACGTTTGT  
GCCTCCGAGCAGCAACAGGTTGTCTACAACCACAACAGCAGCAGCACTCATCACTGCGAC  
GCAGTTCAGTGGCAAGACGGTTCTTGTGTGAGTTCGGCAGCGGCAGCACAGCAGCAGCAG  
CACCATCAGCAGCAGCAACATAGCTCCGCTGCTGCTGCTGTACAGCAGCAATTGCTGCCACC  
GCCTGGTACCTATCAATACCAACAGGTGGGCGGCAATACCGTTGTAGTTGCCGTGCGTCATC  
AGCAACAGTTGCAACAGCAGCAGCAGCAACAGCAGCAACTGGCTTTACATCATCAGCATCAG  
CAGCAGCAGCATCAGCAGCAGCATCAGCAGCAGCAGCATCAGCAACAGCAGCG  
TGCCACTCTCGTGCAGCCAGGATTTGTTTCAGCTATCGCAGCAATCAACCCAGCAGCAAC

AGCAACCAGCGCAGCAAATACATCAGGGCCCCAAAGTGACGCATAGCAGCGCAGCTTCGGAT  
GCTTTAACATATAGTTTGGATGGCACAACAACCGCCAAGTGGACCGCCGCATGCAGGTGGACA  
GCAAATAACGCCAGGTGCAAACCTCTGCAAATTTAAGCATTTGTAGCGGCCGCTTTGAGCGCCG  
CTCGTGACGTTCGGGGGAGGATCCGACGGAGGAGGATCAGCGGGAGGCGCAACGCCAGCTACT  
GGAGCATCGGCTTCATCCGTAGGAAACACCAGCGCCGTGGGCGCGTCGAGCAGCAGCAACAG  
CAGTGCCGGCCAAGCAGCCAGCAGCAACAGCAACAATGTGACTGCAACAGGATCGGGGTCGG  
CTCCTGGCGGAGGACCCACGAGTACCGGGACGACCAGTGGCACTCAGCATGGTAGTGGATCG  
GGGGCAGCTGCAGCCGTTGATTCGAAAGCGATGACAGTGAGGTGGTTCGACTACAGGCTTT  
GCTCGAAGCCCAGGACTGCCGCCCATTTATTCGGTGCCTTGGCCCAAGGATGACACATA  
TTCTCCACCGAACCATTTGTAACAGCAGCAGTTCGAAGCCAATCAACTTCTGCAAGGATTA  
CAATCCCACGACGAATCCCAGCAGCTGCAGGCTGCCATTGAAATGTGCCAGATGCTCGTGAT  
GGGCAATGAGGACACCCTAGCCGGTTTTTCCAATCAAGCAAGTGGTACCTGCTCTAATAACAAC  
TCCTTCGTATGGAGCAACCTTTGATATAATGAACAATGCATGTAGGGCATTTGGCATAACATG  
CTAGAAGCTTTGCCCGCTCGTCAGGCACCGTGGTGGAGGCGGTTCAGTCTTTCTAGAGAA  
ACTTCAGGTCATCCAATGCATGGATGTGGCTGAGCAAAGTTTATCGGCCTTGGAGATTCTGT  
CACGGCGTCACAACAAGGCCATCCTGCAGGCAAACGGAATTTCCGGCTGCCTCACCTATTTG  
GACTTCTTTTCGATTGTGGCTCAGCGGGCAGCATTGGCGATTGCGGCTAATTGCTGCCTAAA  
CATGCATCCGGAGGAATTCACCTTTGTGGCGGAAAGCTTGCCTCTACTAGCTCGCTTGCTCT  
CCCAGCAAGATAAAAAATGCATTGAAAGCGTTTGTCTGCATTCTGTCTAGTAGAGAGT  
TTCCAGCAGCAGGTCAGCGCTTGCAGCAAATCGCTAGTCCGGACCTGTTGAAGAACTGCCA  
GCAATTACTGCTGGTAAACACCAGCCATCCTGAATACAGGAACCTTCACAGCCGTCGTTCCGA  
TGCTAAGTCTAATGTGTTGCAGCTGTCCAGACTTGGCAATTTCTTTACTAAGGAACGATATA  
GCGGCCACCCTGCTTTATTTGCTCACCGGAAATGCAGAGCCGGCTGCTGCCAGTGTACCCA  
CGTGGAGCTGATCTCACGCTCGCCTTCAGAGCTGTACGAACTTACCTGCCTCATTGGGGAGT  
TAATGCCGCGCCTACCTTTGGATGGAATCTTTGCTGTGGACTCACTGCTGGACAGACCAACG  
CTTAATACCCAGGACCAGGTGCATTGGCAGTGGCGGATGATCGCGGCTCCTGGCACAACCTA  
TTCTACAATTGATTCGCGCTGATCGAGGCCGCAACCAGAGCAGCGAGGACGAGATAAGCC  
TAAGCACCTTTGGACGCACCTACACTGTGATTTTCATGCCATGCAGCAGATAAATGAGGAC  
ACGGGCAC'TACACGACCCGTTCAACGTGCTCTTAACCATAACTATGTGGCTCCCATGTCTGC  
GGCCAGGACTTGACTACAACATCGGCTGGATCGGCGGCCGCGGGAGGAGCTTCCACGTCCG  
CTGCAGCAGCGGCAGCGTCCCTCCAACAACAATAACAATAATAATAATCCACCAGGCAAT  
AGTGTCAACCTCAATCAAGTAAAGCGCCGTCATCTCTGGATGCTAGAATCGCATGCCTTAA  
GGAGGAGCGGGTCTGGCTGCGGACTTTATAAAACACATCTTTAACGTACTGTACGAAGTAT  
ACAGCTCTTCTGCCGGACCAATGTGCGTTACAAATGCCTGCGCGCCCTTCTGCGGATGGTG  
TACTATGCCACGCCGGAGTTGTTGCGTCAGGTGCTGAAGTATCAGCTCGTCTCTAGCCACAT  
TGCTGGCATGCTAGGAAGCAATGATTTGCGCATCGTTGTGGGAGCTCTTCAAATGGCTGAGA  
TCTTAATGCGCCAACCTTCCGATGTCTTTGGCAGCAGCTTTTCGTCGTGAGGGAGTGATCTAC  
CAGTTTACTCAGCTCACTGATCCAAATAATCCGATATGCGCGAATCCCTCGCCGAAACCGCT  
TAGCGCAACAGCAACGCCACCAGCAAATGCAGGCGGATCACAATCGGCTCCTGCCTCCGCCA  
ACAGCCTGCAAGTAAATCCATTTTTTCATGGATAGTGCTCCGGGTTTGTGAGTGCTTCCACG  
ACTCCCAGCAGTAGCAAGCATCAATCGTACAGTGTGAAGAGCTTCTCACATGCCATGAACGC  
CCTAACAGCCAGCGCCAAAGGAACTCCTTCTGGAGCGTTGGACGCCACAAGCTCATCTACCA  
CAGCTGGTGGTTATAATTACAGCAGCTCGGCACCGTCTTCTTCCCTCCGGAGACCCGCTGCC  
TACTTTGTGACCCAACAGGGAGATCCACGCCAGTACGTCCACTTTCAGCAGCCGGCCGTTCC  
AGCACCACCACCGCAGCAGGAACTGTTACCATCAGGAGTTCAGCAACAAGGGCAGCAGGTGC  
CTCAGGTGATCTACCAGCCACATCATCAACAGCCTGCACACTTAGTGCTGGCCTCCACCAGC  
AGTGGCGCTGCCTCGTCTGCTTTCGCTTTCGTCCTTTCGTCCTCGGCTCTGCACTGCAGCA  
CAAGATGACAGACATGCTAAAGCGAAAAGCGCCTCCGAAGCGTAAATCGCAAAGCGGTGGCA  
GGGCAAATCTAGACAAGAAGATGCAGCAGTCGACCAGCGGGATCTGGACCAGGGGGAGCG  
CCTCCTTCTTCATCCGGCTCAGCTATGCATGAACTTCTTAGCCGTGCTACAAGTCTTGGTAG  
CGGGAACGGTGGCAGGAGTACGCCAAATTCGGGAGGTGGCTCTGGAAGCTCCAAGTCCCGGT  
TTAACGCCGGAACCTCGTCCAATGCCGGATCGAGCAAATCTTCGTTCTTGGCATCGCTTAAT

CCAGCTCGCTGGGGCCGTCAAACATCAAAAATCATCACCATCACCACCAGCAGTCCCAGCA  
GCAGCATCATGGTCTTTCTAAAGATTCGGGCAACTCGAACTCAACAGGTTCCGGCTCCGGAG  
CTGGTTTAGCATAACGGTGAGCCAACATGGCGCTGGTAGTGGAGCCGGAGGACTGAACGCT  
GCAGCCGTTGCGGCGAGTATAAGTAAGAGCATTTCCCACGCCAATCTCCTAGCGGGGCCAA  
CAGAGAGAGGGCGCGTCACTGGGTCCGCGAACAGGCTGTGGATTTTGTAAGCGTTACACTG  
AGCAGGAGGCCAAAAGGAGCAAGGCGCATCCGAGAGTGGTGCCACTCAGAGCGGGAGTAGT  
GGAGTGGGATTATCGTCGACGGGCAATACTCCATTGTCCACAGCGGGCAGTACCAATGTCTT  
AGAGCGTTTGTCCAGCATTTCTGTTAAGCTTAACGGGAGCTACCACGACTGTTTAGATGCTC  
TCCTCGAGCTAAAGACCATTTCTTTGGAGAGTGACATCTCTCCCTTTGAGGTAATCATTCCG  
GGCCTGATCAAAGCCATGCTCAACTATATGACTAGCGAGACTGGACTGGTTGAGCGCGATGC  
ACGTCTGCGCAGCTTTATGCACGTATTCGCTGGCCTTCCATTAGAGCCTCTGCTGCAAAACG  
TCGGCCAGATGCCAACTATCGAGCCGATAGCCTTCCGGAGCCTTTGTGGCCAAGCTAAATGGT  
TGTGTCACGCAACTAGAACAGTTTCCCTGTTAAGGTACACGACTTTCCCTGCCGGTCTGGTGG  
AAGGTCCAATCAGAGTGCTCTGCGTTTCTTCAACACCCATCAGCTTAAGTGCAACCTGCAAC  
GCCATCCGCAGTGCAATAACCTACGTCACTGGAAGGGCGGAACTGTAAAGATCGACCCCTTG  
GCCATGGTCCAGGCCATCGAGCGTTACCTGGTGGTGC GCGGCTATGGTGGCATTCTGTGCAGA  
CTCCGACGACGACAGTGAAGAGGATATGGATGACAATGTGGCGGCAGTGGTGTGTCCAGG  
CGAGCTTTAGACACAAGTTGCAGTTTACGATTTGGGGATCACGTGCTGCCCTATAACATGACG  
GTTTATCAGGCGGTCAAGCAGTTCTCGCCACTGGTCACTGAACAACCTGAGACGGACAACGA  
GTCGGAGACTCTTTTAGGAAACGCTAGTATATGGGTGCAGCAGCATACTATTCACTATCGTC  
CTGTAGAGGAGGAAGTGACTTCGGGAGCGGCTGCGGGAGCAGCATCCAGCAGCAGTTCTGTGT  
AGCAGTGGCGTGCAAAAACAGCAGAGCTCTTCAAGCTCCGCATCCAGCTGTGTAAACGCTAC  
CTCTTCGTGCTCCTCATCTTCTGGAGTGGCAAGCGGTGGCGGGTCTTAAACAAAGAAGGCTC  
ACAAATCTAGCAGCAAGTTCATGCGCAAGAAGACTGAGCTATGGCACGAAGGAATAGCTCCT  
GTAGTGATTTAGCTCTAAAGCCTTTCCTTAGCAGCTCTCTTCCAGCTGATGTTGTGACCGT  
ACAAGACGCCTCATTGGATGCACTGTGTATGCTGCGAGTTATTTCATGCTCTAAATCGTCACT  
GGGACCATCTCTACGGTTGTGTGGTGCGCCAAAACATTAATCCCCAGTCAGACTTTATACAT  
CCTAAGATTATGGCCAAGGCGAATCGTCAGCTGCAGGATCCGCTTGTTCATCATGACAGGCAA  
TCTACCACAGTGGCTTCCCCAGATAGGCATGGCCTGTCCCTTCTCTTCCCATTTCGAGACTC  
GTCACCTGCTTTTCTACGCAACCAGCTTTGATCGGGATCGCGCTTTGCAGCGTCTGTTGGAC  
ACTACCCCGATCTTAATGCAGCGGAGTCTTCCAGAGCGCGTGGCTCCTCGCCTTGATCGCCG  
CAAGCGAGCGATTTCCCGGACAGAGATACTCAAACAGGCAGAGCACATCCTACAGGACTTTG  
GGCACTCGAAGGCACTGCTGGAATTCAGTACGAGAATGAGGTCGGAACGGGTCTGGGACCT  
ACATTGGAATTCTACGCTTTGGTTTCTGCGGAGCTGCAGCGTACCGACCTTGGCCTATGGAA  
TGGCAGTGATAGCTACAAACAGAACTCTGTGACCATTGTAGACGTAGTCAAGGCCAACAGCG  
CTGTGCTGCACATCGAGGATGCTCTTGGAGGCAACCACCACGGACCAGAACACTCCAGCAGTA  
GCAGGAGCGTCGCTTGTGAGCAGCAGCACCACCCTACAACAACAACCGCACAACAACACCA  
GCATCCGCCACACGCTCTAGCAGTCGTTTCGCACGCTTTCGCTAGCGGAGCTGGCCAACAGC  
CGGTGGAGCACAGCTCCTCTAGCGCCGGCGCCAATGAGAATGCTTTAAATATGGTTATAGCA  
CAGCAATTTAGTGATACAACTCTGCTAACCAGCAGCGATTGATAATCCAGCAGCACCAC  
TACCGCAACAACGGTTGTACAGCATAATACCACTACGAATAACTCAAGCATTATCACGACTA  
CAACCACGACGAGTTATGTGCATGCGGTGCATGGCCTGTTCCCTTTGCCTCTGGGTAAATCT  
TCGAAGCTTCCGCAGATGACTAAGGCCAAGGCCAAATCAAATCTTGGGAAAGTTCATGGC  
AAAGGCGGTGATGGACAGTCGCATGTTGGATTTGCCATTTTCTTTGCCATTTCTATCGCTGGT  
TAGTCAGCGAAGAGCATTCAATCGGCCTGGCCGATTTGATGCGTGTGGCCCCGGAGGTCCAA  
AACACTCTAGTGCGACTCCAAGACTTGGTTCGCCAGCGGGAGTACATTCTATCCGATCCAAA  
CATCGATGCCATGGAAAAGACTGAAAAGATTGAACAGCTAGACTTGGATGGCTGTCCCATTG  
CTGACCTAGGCTTGGACTTTGTATTGCCCGGACATGCCAATATAGAGTTGTGTGGGGCGGT  
CGCGATACACCTGTCACTGTGCACAACCTTACCAGTACATCTCCCTGGTCACATACTGGTT  
CCTTATCGAGGGTGTCCAGAAACAGTTTGGAGCTTTGCGTGAAGGTTTTGATTCCGTTTTTC  
CAATCCAACGGTTGCGAATGTTCTATCCAGAGGAGTTGGAGTGCCTTTTCTGTGGATCGGGC  
AGCGAGCAGCAGCATTCGCGATGGGAAATCAAGATGCTGCAGGAGAGCTGCCGCACTGATCA

TGGATTCCACCAGGACTCGCAAGCTATCCAATATCTGTATGAGATCCTGGCCTCTTACAACC  
GCGACGAACAGCGCGCCTTCTTACAGTTTGTGACTGGATCACCACGCCTTCCGACTGGAGGA  
TTCAAGGCCCTTACGCCACCCTGACTATTGTACGCAAGACGTTGGATGAGAACCAAAACCC  
TAACGATTACTTACCATCTGTGATGACCTGTGTCAACTATCTAAAGTTGCCCGACTACTCTA  
GTCGCGAGGTGATGAGGCAGAAGCTGAAAGTGGCTGCTAACGAAGGCAGCATGTCTTTCCAC  
CTCTCATAAGCAAGCAACAAACCAAAACCAACCGAACCCCGCTTTCACAAACAACCAGAT  
TTACAGCAACACAAATCGGAAACTACAAAATACTACATGGAATATATTAAGGAGGGTATAT  
TCCGGATGGATTAATTGGATGGGGGTCCTATTTGCATTTTTTCTTTTCGATCGTTAGTTTCGC  
CGCGCGGTTTTCCCAAGCTCCTTAAATTTTTTCATTTTGGCAAGCGCGTGTACATACTAAATT  
TCAATATAAATATATAAGTATAATATATTTTTTATTACCAAGTTTTCGAAAAAAGTTAAACG  
GAAACACGGATAACATCGAGAAAGAAAATAATGATGGACCTCTCACACAAAATGTTACCTTT  
AAAAAATTTGAATAGTAAACATCGATTATTCAATAAAAAGACCATTTGTCCACCATAAACGC  
ACAAAGCGGTATATTTGTGCGATTAGGGCTTTGCATTTCGTTCTAGTTAAATTTGTTTAGTTTG  
CTTTTCCCATTCTTATTATCCTCAAGCAAATCTAACTCATATTATTACATAATCAAACCATT  
ATTTTATTAAGTTTTTATTTGTTATGTTGAACTATTTCTCTGGCGTTTAACTATGAGTTTAA  
TTATTTGTAAGCATACAATGTGCATAAACGCTTATTAGAAAGAAACGTGAATGAAAATACAA  
TGAAGGGGAAACACTGAAGCATTGGAGGTGAGGAGTAACACTTATATGGAAAAGAAAGAATT  
CATAATGTTTTACATATAACTCCCAAATAAATGATCTGAATAAATT

**Longest putative CTRIP protein (3140 aa, encoded by E2-E12)**

MAESVKSQSLSALTEGQHDDGGSTATSATLANKTSGTNTTSASNHSSQRRRNNTNSN  
RNKSHHKNEKRNAPATSGPAVSSVVDVVPATSTASATSTRRSRSQGRHSSAAALSSKE  
SQIPKRSVGAIRSPSSPAFNSIPVVSQSLSPGSRKRQLNQHNNSNNKASSSHQSAPG  
GDQLVCAPLKKRRLQQSLGSGI GEGAVEINLGAASFDHTPRQASGDCVAQRTRSKTIS  
PEELPSTSSAAAARHHQIRPSASSTSFSSSHRNKRKASIGGSIAAVETTPHRGTAAAR  
AGRSGNLLNYYRKRTRKVSHTRSSQKSEKQSI PAEDTTANRNGSAGSGSSSKSGVIPKN  
RKSLRHSQQNLLDTEAAEVESAAAETGGEQQSEHGNLDVIDQLPTPEAGPNQSEASGI  
HNQSQLEADRQPVVQDEDEDEDEEEEEEEEEVEGFYGI VNSAGSSYEEDTQIVA EED  
EITTEEEVDEDDDEIEEEDLSESEFAQQLIGELG GERQQPQRATNNAKKTKNNTPAV  
AAASTIKATTTATTSSGYQQRAAPHGQGAAPGAGGGGPRATRSSNSNSNNVVDYSIM  
PHLTTVATTTTPHALTPQQHQOQRQPPQPPSHQQQQVHQQQQPPQLLPTHQFAHLSSFV  
TPTAAAAAHHYPPASAAAAYFAQQQHHQHQQHQQQHVQQAQAPQPHYHSSVA  
VLHQPPPHFTSTGAGPPALFQQQAPPQLTRYTPATATTAATFVPPQQQQVVYNPQQ  
QQHSSLRRSSRGKTGSCVSSAAAQQQQHQQQHSSAAAQVQQQLLPPPGTYQYQQV  
GGNTVVAVRHQQQLQQQQQQQQQLALHHQHQQQQHQQQHQQQHQQQHQQQORATLVQ  
PGFLFSYRSNQPQQQQQPAQQIHQGPKVTHSS AASDALTYSLMAQQPPSGPPHAGGQQ  
ITPGANSANLSIVAAALSAARDVGGGSDGGGSAGGATPATGASASSVGNTSAVGASSS  
SNSSAGQAASSNSNNVTATGSGSAPGGGPTSTGTTSGTQHGS GSGAAA AVDSESDSE

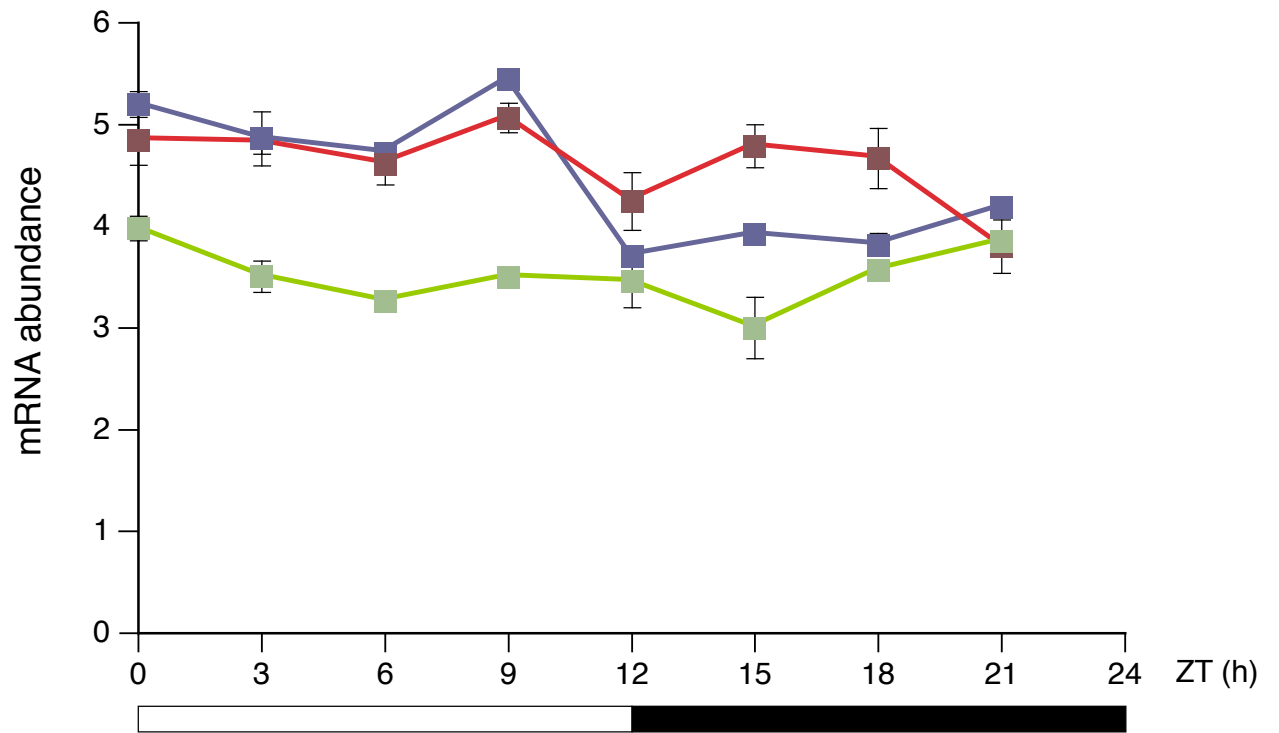
VGRLQALLEARGLPPhLFGALGPRMTHILHRTIGNSSSSKANQLLQGLQSHDESQOLQ  
AAIEMCQMLVMGNEDTLAGFPIKQVVPALIQLLRMEHNFDIMNNACRALAYMLEALPR  
SSGTVVEAVPVFLEKLQVIQCMDVAEQSLSALEILSRRHNKAILQANGISACLTYLDF  
FSIVAQRAALAIANCCCLNMHPEEFHFVAESLPLLARLLSQDKKCIESVCSAFCLV  
ESFQHDGQRLQOQIASPDLLKNCQQLLLVTPAILNTGTFTAVVRMLSLMCCSCPDLAIS  
LLRNDIAATLLYLLTGNAEPAAASATHVELISRSPSELYELTCLIGELMPRLPLDGIF  
AVDSLLDRPTLNTQDQVHWQWRDDRGSWHNYSTIDSRLIEAANQSSSEDEISLSTFGRT  
YTVDFHAMQOINEDTGTTRPVQRRLNHNYVAPMSAGQDLTTTSAGSAAAGGASTSAAA  
AAASSNNNNNNNNPPGNSVNLNQVKRRPSLDARIAACLKEERGLAADFIKHIFNVLYE  
VYSSSAGPNVRYKCLRALLRMVYYATPELLRQVLKYQLVSSHIAGMLGSNDLRIVVGA  
LQMAEILMRQLPDVFGTHFRREGVIYQFTQLTDPNNPICANPSPKPLSATATPTANAG  
GSQSAPASANSLOVNPFFMDSAPGLSSASTTPSSSKHQSYSVKFSHAMNALTASAKG  
TPSGALDATSSSTTAGGYNYSSSAPSSSSGAPAAVFVTQQGDPRQYVHFQQPAVPAPP  
PQQEELLPSGVQQQGOQVPQVIYQPHHQPPAHLVLASTSSGAASSSSSSSSSSASALQ  
HKMTDMLKRKAPPKRKSQSGGRAKSRQEDAAVAPAGSGPGGAPPSSSSGSAMHELLSRA  
TSLGSGNGGRSTPNSGGGSGSSKSRFNAGNSSNAGSSKSSFLASLNPARGRQTHQNH  
HHHHQOSQQQHGLSKDSGNSNSTGSGSGAGLAYTVSQHGAGSGAGGLNAAVAASIS  
KSI SHANLLAAANRERARQVWREQAVDFVKRYTEQEAKRSKAASESGATQSGSSGVGL  
SSTGNTPLSTAGSTNVLERLSSILFKLNGSYHDCLDALLELKTILLESDISPFEVNHS  
GLIKAMLYMTSETGLVERDARLRSFMHVFAGLPLEPLLQNVGQMPTIEPIAFGAFVA  
KLNGCVTQLEQFPVKVHDFPAGPGGRSNQSALRFFNTHQLKCNLQRHPQCNNLRQWKG  
GTVKIDPLAMVQAIERYLVVRGYGGIRADSDDDSEEDMDDNVAAVVLSQASFRHKLQF  
TIGDHVLPYNMTVYQAVKQFSPLVSEQPETDNESETLLGNASIWVQOHTIHYRPVEEE  
VTSGAAAGAASSSSSCSSGVQKQOSSSSSASSCVNATSSCASSSSGVASGGGSLTKKAH  
KSSSKFMRKKTTELWHEGIAPVVISALKPFLSSSLPADVVTVQDASLDALCMLRVIHAL  
NRHWDHLYGCVVRQNIIPQSDFIHPKIMAKANRQLQDPLVIMTGNLPQWLPQIGMACP  
FLFPFETRHLLEYATSFDRDRALQRLLDTTPDLNAAESSERVAPRLDRRKRAISRTEI  
LKQAEHILQDFGHSKALLEIQYENEVGTGLGPTLEFYALVSAELQRTDLGLWNGSDSY  
KQNSVTIVDVVKANSVAVLHIEDALEATTTDQNTPAVAGASLVSSSTTTTTTTAQQHQH  
PPTRSSSRSHVLRSGAGQOPVEHSSSSAGANENALNMVIAQQFSDTNSANPAAIDNPS  
STTTATTVVQHNTTTNNSIIITTTTTTSYVHAVHGLFPLPLGKSSKLPQMTKAKAKFK  
FLGKFMAKAVMDSRMLDLPLPFLPYRVLVSEEHSIGLADLMRVAPQNTLVRLQDLV

RQREYILSDPNIDAMEKTEKIEQLDLDGCP IADLGLDFVLPGHANIELCRGGRDTPVT  
VHNLHQYISLVTYWFLIEGVQKQFEALREGFDSVFP IQRLRMFYPEEELECVFCGSGSE  
QQHSRWEIKMLQESCRTDHGFHQDSQAIQYLYEILASYNRDEQRAFLQFVTGSPRLPT  
GGFKALTPPLTIVRKTLDENQNPNDYLP SVMTCVNYLKLDPDYSSREVMRQKLKVAANE  
GSMSFHLS

**Most abundant putative CTRIP protein (2707 aa, lacking E4-E5)**

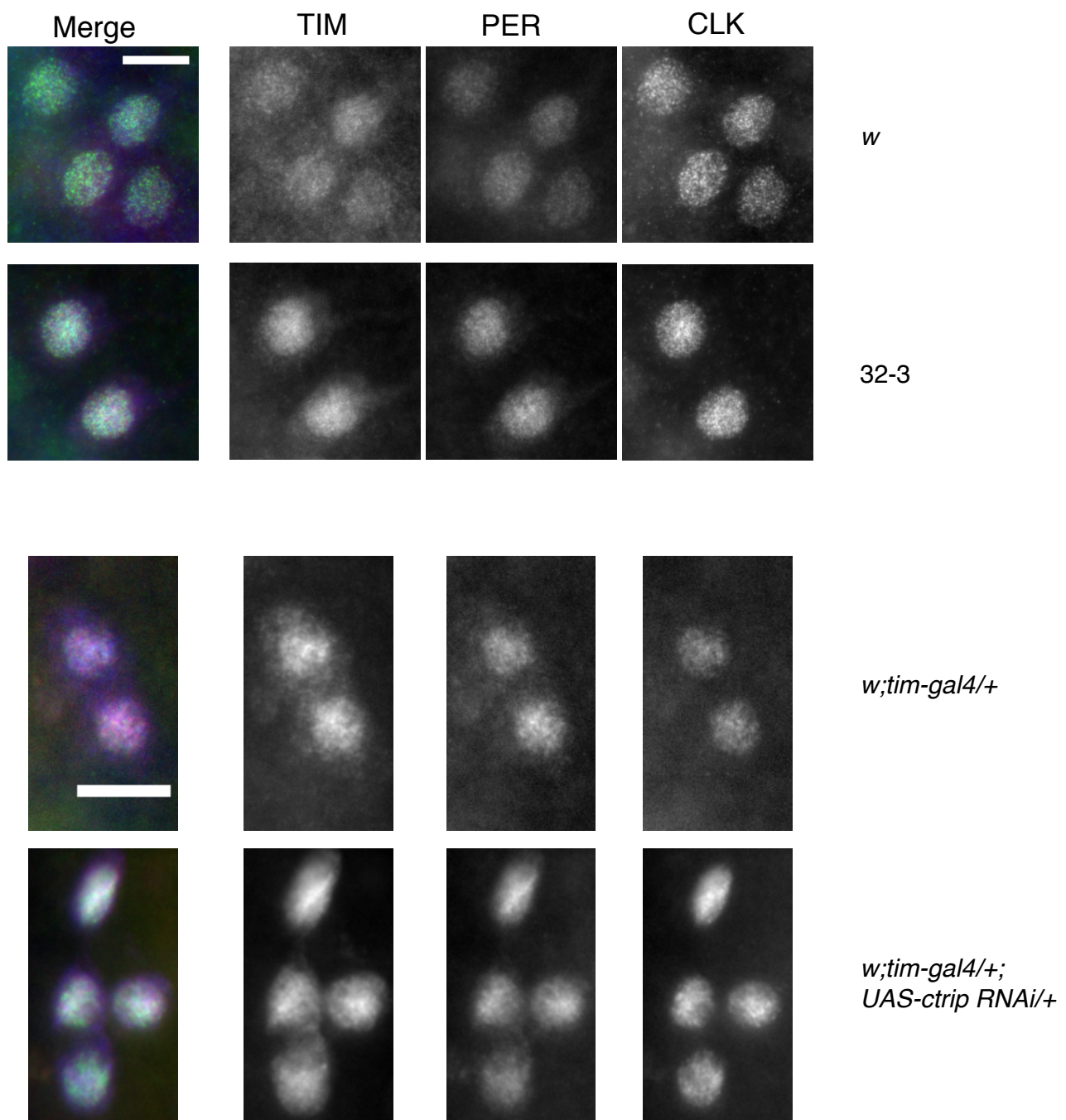
MAESVKSQSL SALTEGQHDDGGSTATSATLANKTSGTNTTSASNHSSQRRRNNNTNSN  
RNKSHHKNEKRNAPATSGPAVSSSVDVVPATSTASATSTRRSRSQGRHSSAAALSSKE  
SQIPKRSVGAIRSPSSPAFNSIPVVSQSLSPGSRKRQLNQHNNSNNKASSSHQSAPG  
GDQLVCAPLKKRRLQQLGSGI GEGAVEINLGAASFDHTPRQASGDCVAQRTRSKTIS  
PEELPSTSSAAAARHHQIRPSASSTSFSSSHRNKRKASIGGSI AAVETTPHRGTAAAR  
AGRSGNLLNYRKRTRKVSHTRSSQKSEKQSI PAEDTTANRNGSAGSGSSSKSGVIPKN  
RKSLRHSQQNLLDTEAAEVESAAAETGGEQQSEHGNDVIDQLPTPEAGPNQSEASGI  
HNQSQLEADRQPVVQDE  
EITTEEEVDDEDDDEIEEEDLSESEFAQQLIGELGANSANLSI VAAALSAARDVGGGS  
DGGGSAGGATPATGASASSVGNTSAVGASSSSNSSAGQAASSNSNNVTATGSGSAPGG  
GPTSTGTTSGTQHGS GSGAAA AVDSDDSEVGR LQALLEARGLPPHLFGALGPRMTH  
ILHRTIGNSSSSKANQLLQGLQSHDESQQLQAAIEMCQMLVMGNE DTLAGFPIKQVVP  
ALIQLLRMEHNF DIMNNACRALAYMLEALPRSSGT VVEAVPVFLEKLQVIQCMDVAEQ  
SLSALEILSRRHNKAILQANGISACLT YLDFFSIVAQRAALAI AANCCLNMHP EEFHF  
VAESPLLRLLS QODKKCIESVCSAF CRLVESFQHDGQRLOQ IASPDLLKNCQQLLL  
VTPAILNTGTFTAVVRMLSLMCCSCPD LAISLLRNDIAATLLYLLTGNAEPAAASATH  
VELISRSPSELYELTCLIGELMPRLPLDGI FAVD SLLDRPTLNTQDQVHWQRDDRGS  
WHNYSTIDSRLIEAANQSS EDEISLSTFGRTYTVDFHAMQQIN EDTGTRPVQRRLNH  
NYVAPMSAGQDLTTTSAGSAAAGGASTSAAAAAASSNNNNNNNNNNPPGNSVNLNQVKR  
RPSLDARIA CLKEERGLAADFIKHIFNVLYEVYSSSAGPNVRYKCLRALLRMVYYATP  
ELLRQVLKYQLVSSH IAGMLGSNDLRIVVGALQMAEILMRQLPDVFGTHFRREGVIYQ  
FTQLTDPNNPICANPSPKPLSATATPTANAGGSQSAPASANS LQVNPFFMDSAPGLSS  
ASTTPSSSKHQSYSVKSF SHAMNALTASAKGTPSGALDATSSSTTAGGYNYS SSSAPSS  
SSGAPAAAYFVTQ QGDPRQYVHFQ QPAVPAPPPQO ELLPSGVQOQGOQVPQVIYQPHHQ

QPAHLVLASTSSGAASSSSSSSSSSASALQHKMTDMLKRKAPPKRKSQSGGRAKSRQ  
EDAAVAPAGSGPGGAPPSSSSGSAMHELLSRATSLGSGNGGRSTPNSSGGSGSSKSRFN  
AGNSSNAGSSKSSFLASLNPARWGRQTHQNHSHHHHQSSQQQHHGLSKDSGNSNSTGSG  
SGAGLAYTVSQHGAGSGAGGLNAAVAASISKSI SHANLLAAANRERARQWVREQAVD  
FVKRYTEQEAKRSKAASESGATQSGSSGVGLSSTGNTPLSTAGSTNVLERLSSILFKL  
NGSYHDCLDALLELK TILLES DISPFVNHSGLIKAMLNMTSETGLVERDARLRSFM  
HVFAGLPLEPLLQNVGQMPTIEPIAFGAFVAKLNGCVTQLEQFPVKVHDFPAGPGGRS  
NQSALRFFNTHQLKCNLQRHPQCNNLRQWKGGTVKIDPLAMVQAIERYLVVRGYGGIR  
ADSDDDSEEDMDDNVAAVVLSQASFRHKLQFTIGDHVLPYNMTVYQAVKQFSPLVSEQ  
PETDNESETLLGNASIWVQQHTIHYRPVEEEVTSGAAAGAASSSSSSCSSGVQKQSSS  
SSASSCVNATSSCSSSSGVASGGSLTKKAHKSSSKFMRKKTTELWHEGIAPVVISALK  
PFLSSSLPADVVTVQDASLDALCMLRVIHALNRHWDHLYGCVVRQNIIPQSDFIHPKI  
MAKANRQLQDPLVIMTGNLPQWLPQIGMACPFLFPFETRHLLEYATSFD RDALQRLD  
DTTPDLNAAESSERVAPRLDRRKRAISRTEILKQAEHILQDFGH SKALLEIQYENEVG  
TGLGPTLEFYALVSAELQRTDLGLWNGSDSYKQNSVTIVDVVKANS AVLHIEDALEAT  
TTDQNTPAVAGASLVSSSTTTTTTTTAQQHQHPTRSSSRSHVLRSGAGQQPVEHSSSS  
AGANENALNMVIAQQFSDTNSANPAAIDNPSSTTTATTVVQHNTTTNNSII TTTTTT  
SYVHAVHGLFPLPLGKSSKLPQMTKAKAKFKFLGKFMAKAVMDSRMLDL PFSLPFYRW  
LVSEEHSIGLADLMRVAPVQNTLVRLQDLVRQREYILSDPNIDAMEKTEKIEQLDLD  
GCPIADLGLDFVLPGHANIELCRGGRDTPVTVHNLHQYISLVTYWFLIEGVQKQFEAL  
REGFDSVFPIQRLRMFYPEELECVCVFCGSGSEQQHSRWEIKMLQESCRTDHGFHQDSQA  
IQYLYEILASYNRDEQRAFLQFVTGSPRLPTGGFKALTPPLTIVRKTLDENQNPNDYL  
PSVMTCVNYLKLDPDYSSREVMRQKLVKVAANEGSMSFHLS

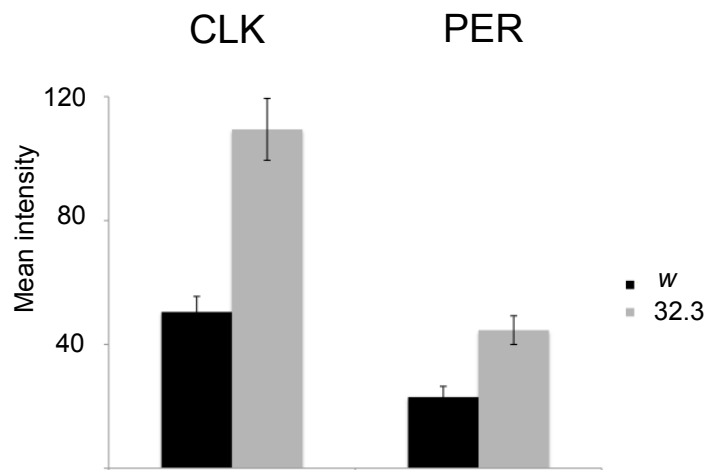
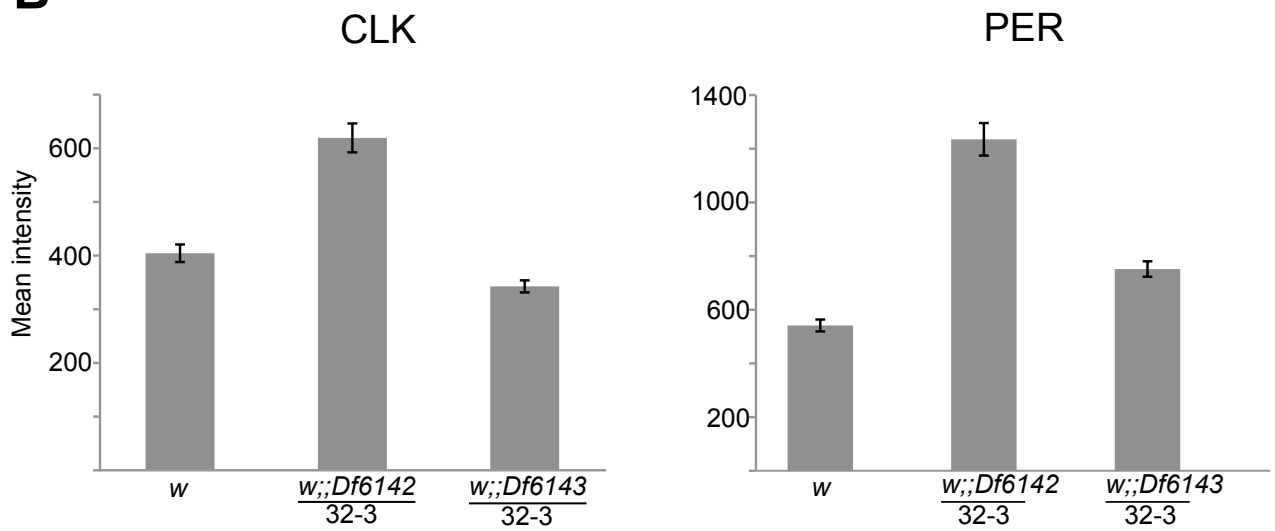
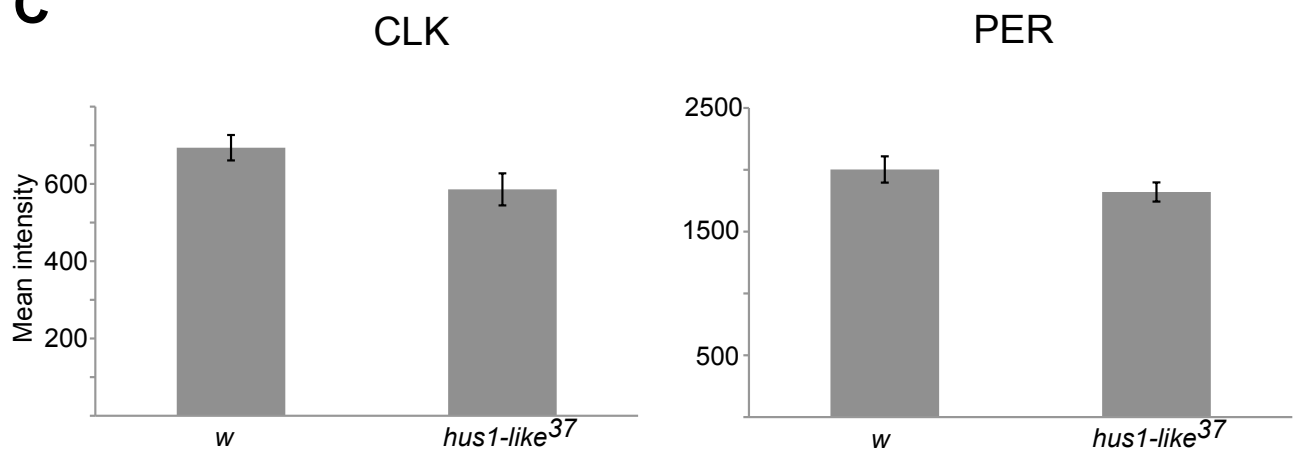


**Figure S2**

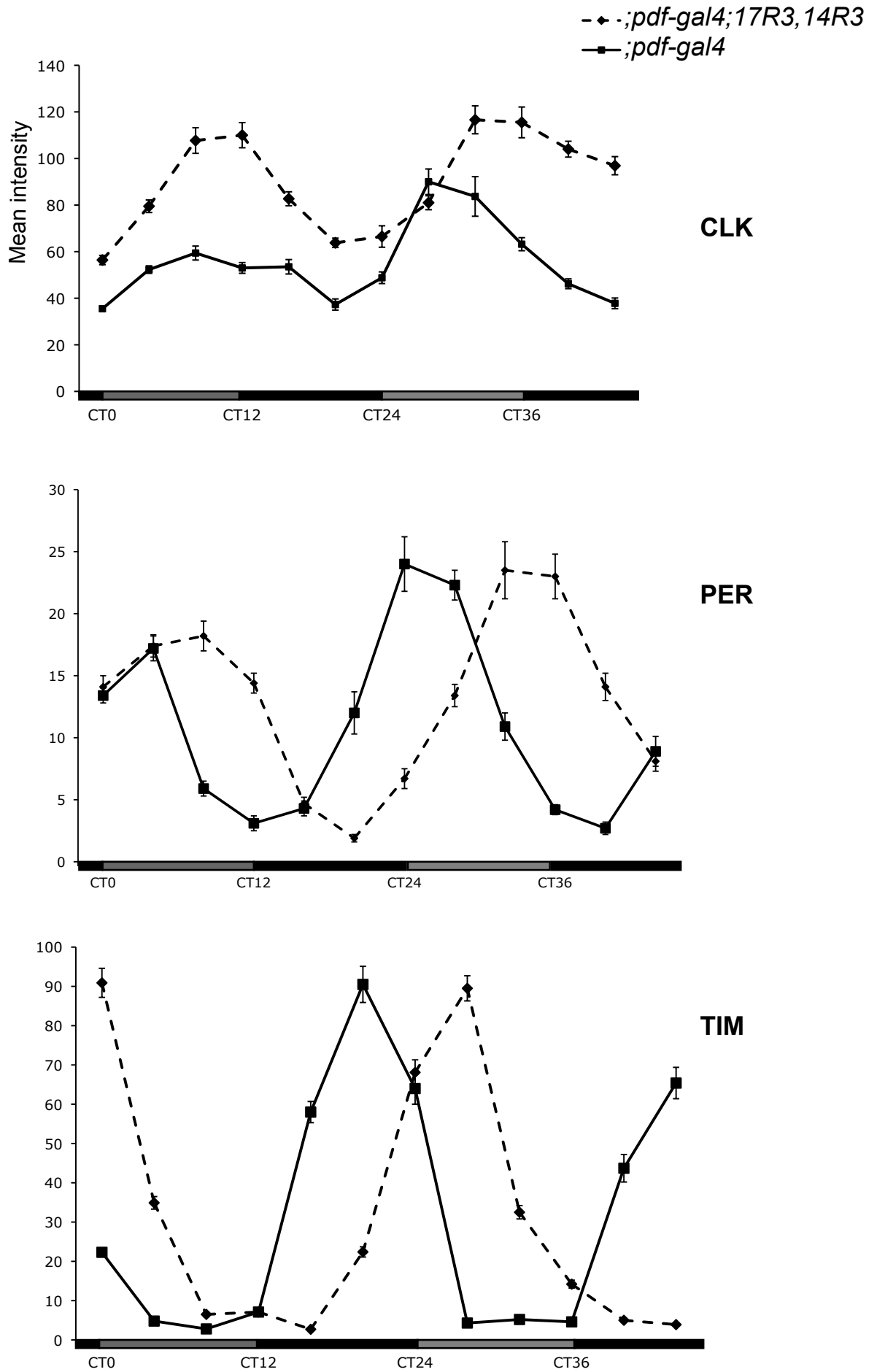


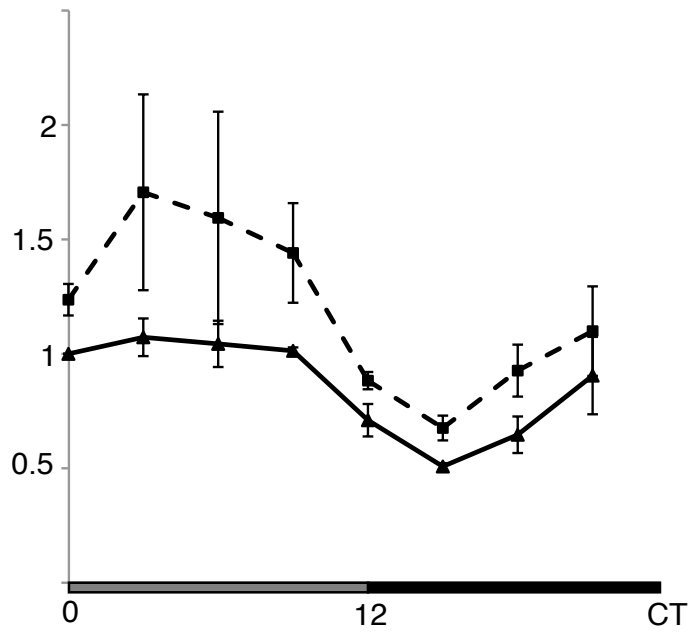
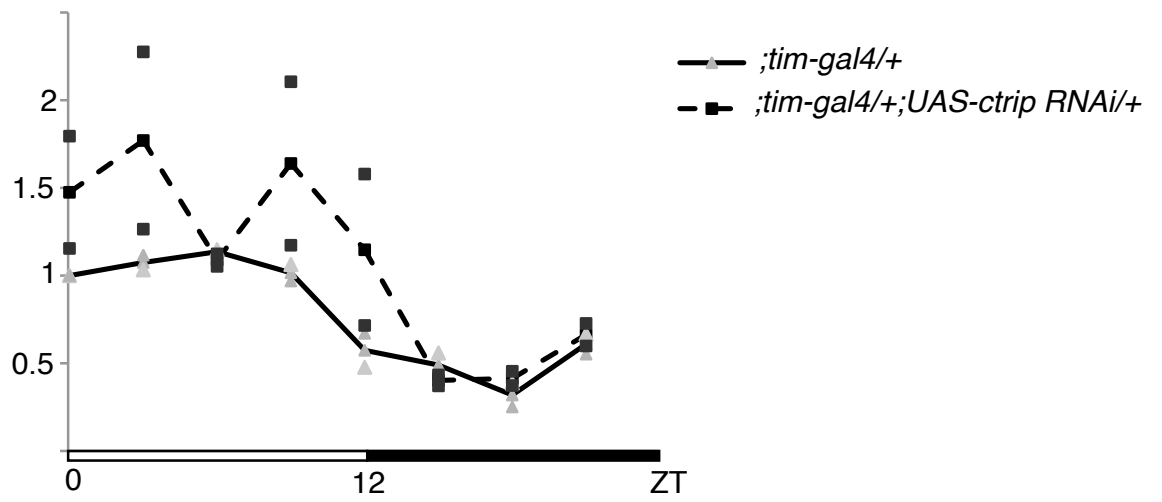


**Figure S3**

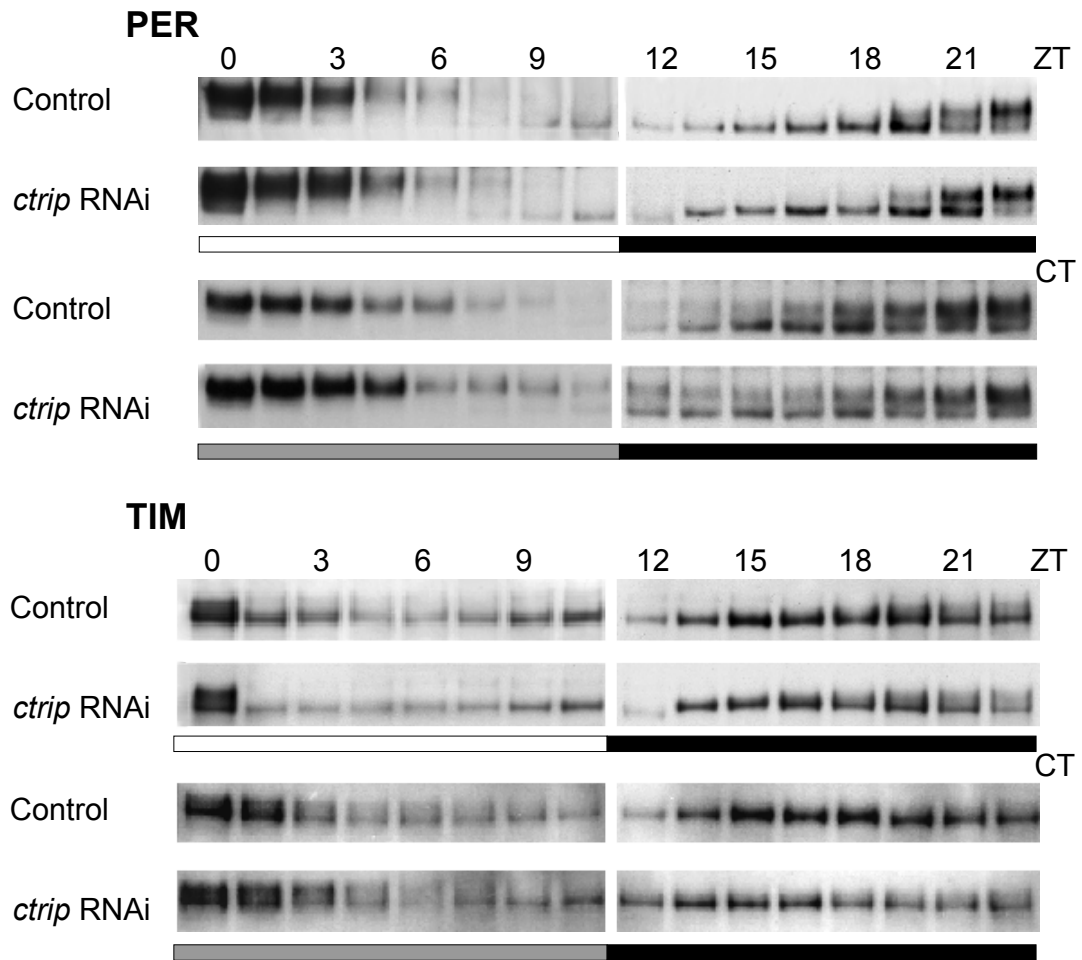
**A****B****C****Figure S4**

**Figure S5**





**Figure S6**



**Figure S7**