



Aberystwyth University

Identification and temporal expression of putative circadian clock transcripts in the amphipod crustacean Talitrus saltator

O'Grady, Joseph Francis; Hoelters, Laura Sophie; Swain, Martin; Wilcockson, David

Published in:

PeerJ

DOI:

[10.7717/peerj.2555](https://doi.org/10.7717/peerj.2555)

Publication date:

2016

Citation for published version (APA):

O'Grady, J. F., Hoelters, L. S., Swain, M., & Wilcockson, D. (2016). Identification and temporal expression of putative circadian clock transcripts in the amphipod crustacean *Talitrus saltator*. *PeerJ*, [e2555]. <https://doi.org/10.7717/peerj.2555>

Document License

CC BY

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400

email: is@aber.ac.uk

Identification and temporal expression of putative circadian clock transcripts in the amphipod crustacean *Talitrus saltator*

Joseph F. O'Grady, Laura S. Hoelters, Martin T. Swain
and David C. Wilcockson

Institute of Biological, Environmental and Rural Sciences, University of Wales, Aberystwyth, Ceredigion, United Kingdom

ABSTRACT

Background. *Talitrus saltator* is an amphipod crustacean that inhabits the supralittoral zone on sandy beaches in the Northeast Atlantic and Mediterranean. *T. saltator* exhibits endogenous locomotor activity rhythms and time-compensated sun and moon orientation, both of which necessitate at least one chronometric mechanism. Whilst their behaviour is well studied, currently there are no descriptions of the underlying molecular components of a biological clock in this animal, and very few in other crustacean species.

Methods. We harvested brain tissue from animals expressing robust circadian activity rhythms and used homology cloning and Illumina RNAseq approaches to sequence and identify the core circadian clock and clock-related genes in these samples. We assessed the temporal expression of these genes in time-course samples from rhythmic animals using RNAseq.

Results. We identified a comprehensive suite of circadian clock gene homologues in *T. saltator* including the 'core' clock genes *period* (*Talper*), *cryptochrome 2* (*Talcry2*), *timeless* (*Taltim*), *clock* (*Talclk*), and *bmal1* (*Talbm1*). In addition we describe the sequence and putative structures of 23 clock-associated genes including two unusual, extended isoforms of pigment dispersing hormone (*Talpdh*). We examined time-course RNAseq expression data, derived from tissues harvested from behaviourally rhythmic animals, to reveal rhythmic expression of these genes with approximately circadian period in *Talper* and *Talbm1*. Of the clock-related genes, *casein kinase IIβ* (*TalckIIβ*), *ebony* (*Talebony*), *jetlag* (*Taljetlag*), *pigment dispersing hormone* (*Talpdh*), *protein phosphatase 1* (*Talpp1*), *shaggy* (*Talshaggy*), *sirt1* (*Talsirt1*), *sirt7* (*Talsirt7*) and *supernumerary limbs* (*Talslimb*) show temporal changes in expression.

Discussion. We report the sequences of principle genes that comprise the circadian clock of *T. saltator* and highlight the conserved structural and functional domains of their deduced cognate proteins. Our sequencing data contribute to the growing inventory of described comparative clocks. Expression profiling of the identified clock genes illuminates tantalising targets for experimental manipulation to elucidate the molecular and cellular control of clock-driven phenotypes in this crustacean.

Submitted 8 July 2016
Accepted 11 September 2016
Published 5 October 2016

Corresponding author
David C. Wilcockson,
dqw@aber.ac.uk

Academic editor
Juan Riesgo-Escovar

Additional Information and
Declarations can be found on
page 22

DOI 10.7717/peerj.2555

© Copyright
2016 O'Grady et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Animal Behavior, Genomics, Marine Biology, Molecular Biology, Zoology

Keywords Circadian, Crustacean, *Talitrus saltator*, Transcriptome, Rhythms

INTRODUCTION

Nearly all organisms are exposed to monotonous cyclic alterations in their environment. Competence to gate behaviour and physiology in tune with these changes is orchestrated by so-called circadian clocks, the cardinal features of which are: the clock mechanism is autonomous and endogenous i.e., it persists in the absence of external cues; the period of the rhythm is temperature compensated and is entrained by relevant cyclic environmental cues. For example, in the terrestrial realm, organisms invariably exhibit daily activity rhythms in temporal correspondence to the light and dark of day and night. Our understanding of the molecular and cellular underpinnings of circadian clocks has advanced tremendously in recent years and comparative studies have benefitted from the advent of next generation sequencing strategies to reveal clock genes in non-model species, including crustaceans ([Christie et al., 2013](#); [De Pitta et al., 2013](#); [Nesbit & Christie, 2014](#); [Toullec et al., 2013](#)).

In the fruit fly brain about 150 cells, primarily in the protocerebrum and optic lobe, contribute to the central clock oscillatory system, although this number varies throughout arthropods. Indeed, in some Lepidoptera ([Sauman et al., 2005](#)) and Crustacea ([Beckwith et al., 2011](#); [Zhang et al., 2013](#)) this number seems to be much lower. The consensus arthropod model of the central oscillatory mechanism is based on transcriptional/translational feedback loops (TTFLs), with positive loops driving the expression of negative factors that feed-back to inhibit the positive loops and hence their own transcription. Despite interspecific variation, the principle mechanisms and components are conserved across diverse taxa (For reviews, see [Allada & Chung, 2010](#); [Hardin, 2009](#); [Sandrelli et al., 2008](#)). Briefly, in *Drosophila* the transcription factors CLOCK and CYCLE (CLK and CYC) drive the expression of *period* (*per*) and *timeless* (*tim*). The translation of *per* and *tim* throughout the evening and early night results in the cytoplasmic accumulation of their cognate proteins that subsequently form stable heterodimers and translocate to the nucleus where they interfere with the action of CLK and CYC to inhibit their own transcription. Under constant conditions such as DD, this autoregulatory feedback loop takes circa 24 h to complete due to the opposing phosphorylation and dephosphorylation actions of various casein kinases and phosphatases that alter circadian protein stability. For example, the stability of PER is facilitated by casein kinase 1 ϵ (CK1 ϵ) or DOUBLETIME (DBT) phosphorylation and translocation of the PER/TIM dimer is mediated by phosphorylation via glycogen synthase kinase β or “SHAGGY” (SGG). In fruit flies maintained under LD conditions the blue light activated flavoprotein cryptochrome (dCRY1) acts as a transduction pathway to relay photic stimuli to the TTFL. When photo-activated at dawn dCRY1 facilitates the degradation of TIM, via the CRYPTOCHROME/JETLAG (JET) complex. This consequently leads to destabilisation of PER, which is ubiquitinated by E3 ubiquitin ligase SUPERNUMERARY LIMBS (SLIMB), and degraded via the proteasome. The resultant de-repression of dCLOCK and dCYCLE ‘resets’ the clock at dawn synchronising (or entraining) the system to the ambient conditions. Protein phosphatases (PP) including PP2A and PP1 regulate these phosphorylation events and so their action contributes to the temporal control of the feedback loops. A secondary negative loop also governs precisely timed oscillations; the CLK-CYC heterodimer initiates the transcription of the basic zipper (bZip) activator

PAR domain protein 1 ϵ (PDP1 ϵ) and bZip transcriptional repressor *vri* (*vri*) (Blau & Young, 1999) by binding E-boxes in the *vri* promoter. VRI accumulation culminates in the repression of *clk* transcription by binding VRI/PAR domain protein 1 ϵ (PDP1 ϵ) promoters (Cyrán et al., 2003). VRI dependent repression also mediates transcription of output mRNAs that cycle in phase with *clk*, such as *cryptochrome* (Glossop et al., 2003). In mammals and some arthropods, CRYPTOCHROME 2 (CRY2) acts as a negative repressor of the CLK/BMAL positive loop. In flies the actions of CLK/CYC E-boxes are also targeted by a basic-helix-loop-helix repressor, CLOCKWORK ORANGE (CWO), which is thought to sustain rhythmicity (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007).

In insects, the octadecapeptide pigment dispersing factor (PDF) that occurs in subsets of clock neurons plays pivotal roles in synchronising oscillator neurons in the clock network. Perturbation of PDF function leads to arrhythmicity or incompetence to adjust phase and period of rhythmic phenotypes to changing environmental cues (Shafer & Yao, 2014). PDF, or rather the orthologous pigment dispersing hormone (PDH) was first discovered in crustaceans where they play a neurohormonal role as well as a neuromodulatory function, effecting dispersion of retinal pigments and integumental chromatophores, including the diurnal rhythm of this action. In crustaceans, PDHs occur in several isoforms and the neuroarchitecture of PDH cells is more complex than for PDF, suggesting additional functions to pigment dispersion. However, the defined role(s) for PDH in rhythmic phenotypes of crustaceans remains poorly described (Beckwith et al., 2011; Strauss & Dirksen, 2010; Wilcockson et al., 2011)

Talitrus saltator lives in the sandy substratum of the supralittoral zone of beaches on the European Atlantic coast and in the Mediterranean. During the day it remains buried in the sand, emerging after dark to make foraging excursions along the land-beach axis. Before dawn it relocates to its burrowing zone to seek refuge, reinterred in the sand. This circadian locomotor activity persists in the absence of external cues and is thus endogenously driven (Bregazzi & Naylor, 1972; Edwards & Naylor, 1987). Remarkably, *T. saltator* also maintains its preferred position on the shore by orientating using visual cues such as the sun and moon as a compass guide (Pardi & Papi, 1952; Ugolini, 2003; Ugolini et al., 1999; Ugolini, Melis & Innocenti, 1999; Ugolini et al., 2007). This capability necessitates compensation for azimuthal changes over time (Ugolini, Tiribilli & Boddi, 2002). Thus, these clock-controlled phenotypes contribute to the fitness of *T. saltator*.

The well-defined behavioural phenotypes of *T. saltator* enable a range of comparisons with other organisms to be drawn. For example, another peracarid crustacean, the isopod *Eurydice pulchra* possesses independent circatidal and circadian clocks (Zhang et al., 2013) and the staggering time-compensated navigational mechanisms of the monarch butterfly (Reppert, 2006) are comparable to those of *T. saltator*. This brings the relevance of comparative clock biology sharply into focus; detailed analysis of a diverse non-model species may reveal commonalities or differences that give insight into how each has evolved and functions. Thus, *T. saltator* represents an excellent, tractable model for time-compensated orientation.

Despite decades of elaborate behavioural analyses, the molecular basis of clock-driven behavioural phenotypes in *T. saltator* is not understood, even at the most fundamental

level. Therefore, we sought to elucidate the sequences of canonical clock genes and their temporal expression dynamics in *T. saltator* to provide a platform from which to explore the neuromolecular mechanisms of circadian clock mediated phenotypes in this animal. This work contributes also to the steadily growing inventory of crustacean transcriptomes and is applicable to further exploration of comparative chronobiology and animal orientation in a tractable and ecologically important model.

MATERIALS AND METHODS

Animal husbandry, tissue sampling and RNA extraction

Animals were caught by hand from Ynyslas beach, Wales, UK at night during their active period and returned to glass tanks containing damp sand at ambient temperature (17 °C) and a 12:12 LD regime. Lighting at dawn and dusk was ramped/dimmed over 30 min to approximate natural conditions. Animals were fed fish food flakes *ad libitum*. After seven days acclimation, the rhythmic emergence and locomotor activity of a sub-sample of 60 animals was established. *T. saltator* has been shown to express more robust and less variable locomotor rhythms in small groups (Bregazzi & Naylor, 1972). Therefore, animals were housed in groups of five in a glass tank containing 10 cm-deep damp sand and compartmentalized with Plexiglas dividers. Across each compartment, infrared beams were passed via bespoke recording apparatus fabricated by Trikinetics (Waltham, MA, USA). All activity, registered as interruptions to infrared beams was recorded via proprietary software on a stand-alone PC. Activity data were recorded in one minute bins and was analysed and plotted using ClockLab software (Actimetrics, Wilmette, IL, USA) run via Matlab[®] v6.2. *T. saltator* brains were rapidly dissected from animals drawn from the population shown to be behaviourally rhythmic (see above). Dissections were done in ice-chilled DEPC treated physiological saline. Ten brains were pooled to form each replicate at 3-hour intervals and snap frozen in liquid nitrogen before storing at −80 °C until use.

Total RNA was extracted using Trizol[®] (Invitrogen[™], Thermo Fisher, UK) according to the manufacturer's instructions except that an additional wash step in 75% ethanol of the RNA pellet was introduced prior to drying and rehydration in 30µl DEPC-treated water. All RNA samples were treated with Turbo DNA-free[™] (Ambion[®], Thermo Fisher, Hampshire, UK) to remove contaminating DNA. RNA was then pooled for degenerate PCR, RACE PCR and Illumina HiSeq2500 sequencing, whilst RNA time-course samples were prepared separately for temporal expression analysis.

Degenerate PCR

Initially we used a strategy of degenerate PCR and 5' and 3' Rapid amplification of cDNA ends (RACE) to identify full-length core canonical clock genes. Full details of the PCR conditions used can be found in [File S1](#) and primers used are tabulated in [Table S1](#).

Illumina RNAseq protocol

The TruSeq cDNA library preparation protocol (Illumina, Cambridge, UK) was carried out on eight time-point samples according to manufacturer's instructions. Amplified cDNA was run on a 1% agarose gel to validate correct library size range. The cDNA libraries were

sequenced on an Illumina RNA TruSeq sequencer (Illumina Inc.) and quality checked at Aberystwyth University's Core Genomics Facility. Read library quality was investigated using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), thereafter all samples were pooled and *de novo* transcriptome assembly was performed with Trinity software, version 2012-10-05 (*Grabherr et al., 2011*). Transcripts smaller than 300 bases were removed from the pooled assembly. This consisted of 186,495 contigs with a mean length of 1,215 bp pairs and an assembly N50 statistic of 2,371 bp.

Transcriptome mining and sequence analysis

Assembled contigs were initially interrogated for canonical clock gene transcripts using the programme BioEdit (*Hall, 1999*), Basic local alignment (BLAST) searches (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) were done using *Drosophila melanogaster* protein sequences as search terms and tblastn search functions set with default parameters (*e*-value threshold at $1e^{-100}$ except for PDH, where stringency was relaxed to 1000). Where no sequence information was available for *Drosophila* (e.g., cryptochrome 2) alternative arthropod sequences were used, details of which are given in [Table 1](#). Contigs were translated using EXPASY Translate tool (<http://web.expasy.org/translate/>) and the coding sequences checked manually using blastP for homology to known circadian proteins.

EMBL SMART (*Letunic, Doerks & Bork, 2009*) servers were used to detect and analyse the conserved functional domains and motifs of clock proteins. Deduced *T. saltator* protein sequences were used to query the FlyBase (<http://flybase.org>) and NCBI non-redundant protein databases using the blastp algorithm. The identity and similarity between protein sequences were calculated using EMBOSS Pairwise Alignment Algorithms (<http://www.ebi.ac.uk>). The server SignalP v.4 (*Petersen et al., 2011*) was used to predict the presence and location of signal peptide cleavage sites.

Quantitation of transcript abundance

Using the Trinity downstream analysis tools, reads from each time-point were mapped to the pooled assembly and transcript abundance estimation was carried out using RSEM (*Li & Dewey, 2011*). In this way the FPKM (Fragments Per Kilobase of exon per Million fragments mapped) values were calculated for each time point. Following this TPM values (transcripts per million) were calculated following *Wagner, Kin & Lynch (2012)*. Quantitative PCR methods for TPM validation are given in [File S1](#) and primers in [Table S1](#). All data files for time-course RNAseq can be found on the NCBI SRA data base, bioproject Accession No. [297565](#).

Blast2GO analysis- transcriptome annotation

Before running Blast2GO, the number of transcripts was further reduced by clustering them using CD-HIT-EST (version 4.5.4) with default options except for the sequence identity threshold, which was set to 95% (*Li & Godzik, 2006*; *Li, Jaroszewski & Godzik, 2001*). The clustering step removes very similar splice variants, which may include some misassemblies, and it significantly reduces the number of transcripts to be included in the time-consuming BLAST analysis. In total 156,766 transcripts with a mean length of 1,534 bp and an assembly N50 statistic of 968 bp were run through BLAST. Using the

Table 1 Identified putative *Talitrus saltator* circadian protein-encoding transcripts.

Clock gene query protein	Transcriptome search sequence Accession No. (all <i>D. melanogaster</i>)	<i>Talitrus</i> transcript/protein identifications					
		Transcript			Protein		
		Trinity contig ID	Length ^a	RACE	Name	CDS	Length ^b
<i>Core clock proteins</i>							
CRY2	n/a ^c	comp100937_c0_seq1 ^c comp102609_c0_seq3 ^c	1,843	5' + 3'	Tal-CRY2	Full	565
CLOCK	AAC62234	comp100688_c1_seq1	5,723	3' ^d	Tal-CLK	Partial	1,907
PERIOD	AAF45804	comp102279_c0_seq7	8,001		Tal-PER	Full	1,557
TIMELESS	AAC46920	comp849619_c0_seq1	1,209		Tal-TIM	Partial	402
CYCLE	AAF49107	comp12103_c0_seq1 comp939723_c0_seq1	1,807 ^e		Tal-BMAL1	Full	602
<i>Clock-associated proteins</i>							
PDH	n/a ^f	comp92607_c0_seq2 comp97165_c0_seq3	2,471 ^g 3,392	5' + 3' ^d	Tal-PDH I Tal-PDH II	Full Full	129 89
CASEIN KINASE 2 α	AAN11415	comp102480_c0_seq1	5,147		Tal-CK2 α	Full	353
CASEIN KINASE 2 β	AAF48093	comp99101_c0_seq3	1,567		Tal-CK2 β	Full	220
CLOCKWORK ORANGE	AAF54527	comp1009591_c0_seq1	503		Tal-CWO	Partial	167
DOUBLETIME (or CK1 ϵ)	AAF57110	comp87763_c0_seq1	1,092		Tal-DBT	Full	310
PDP1 ϵ	AAF04509	comp98345_c0_seq1	3,423		Tal-PDP1 ϵ	Full	508
PP1	CAA39820	comp97405_c0_seq1	1,725		Tal-PP1	Full	357
PP2A –subunit MICROTUBULE STAR	AAF52567	comp98380_c0_seq1	2,981		Tal-MTS	Full	309
PP2A –subunit WIDERBORST	AAF56720	comp102157_c1_seq1	2,474		Tal-WBT	Full	458
PP2A –subunit TWINS	AAF54498	comp99704_c0_seq3	1,633		Tal-TWS	Full	445
SHAGGY	AAN09084	comp99811_c0_seq7	4,413		Tal-SGG	Full	418

(continued on next page)

Table 1 (continued)

Clock gene query protein	Transcriptome search sequence Accession No. (all <i>D. melanogaster</i>)	<i>Talitrus</i> transcript/protein identifications					
		Transcript			Protein		
		Trinity contig ID	Length ^a	RACE	Name	CDS	Length ^b
SUPERNUMERARY LIMBS	AAF55853	comp98870_c0_seq1	2,121		Tal-SLIMB	Full	588
VRILLE	AAF52237	comp100474_c0_seq6	3,949		Tal-VRI	Full	509
EBONY	AAF55870	comp99283_c0_seq2	4,380		Tal-EBONY	Full	974
RORA	NP_788301	comp99654_c0_seq3	2,217		Tal-RORA	Partial	599
REVERB	NP_730321	comp101252_c0_seq2	5,385		Tal-REVERB	Full	1,110
SIRT1	NP_477351	comp101818_c1_seq1	4,033		Tal-SIRT1	Partial	955
SIRT2	NP_650880	comp97450_c0_seq4	2,275		Tal-SIRT2	Partial	376
SIRT4	NP_572241	comp92313_c0_seq2	2,977		Tal-SIRT4	Partial	354
SIRT6	NP_649990	comp69157_c0_seq1	1,209		Tal-SIRT6	Partial	402
SIRT7	NP_651664	comp95761_c0_seq1	5,180		Tal-SIRT7	Partial	948
JETLAG	NP_608880	comp100423_c0_seq4	2,454		Tal-JET	Partial	458

Notes.^alength in nucleotides.^blength in amino acids.^cHomology cloning sequence used to search transcriptome. Trinity contigs likely from misassembly, other gene details stated taken from homology cloning and RACE PCR.^dTrinity contig extended 3' sequence past RACE-derived sequence, other gene details stated taken from contig sequence.^eOverlapping contigs combined.^f*Uca pugnator* octadecapeptide β -PDH consensus sequence used.^gIncludes combined contig sequence and overlapping 5' RACE sequence.

BLAST output the transcripts were annotated using BLAST2GO software (Conesa *et al.*, 2005; Gotz *et al.*, 2008) including reports for Gene Ontology (GO) terms and EC numbers for the KEGG pathway. The BLAST2GO cut-off parameters used to filter out poor quality BLAST hits for the annotation were as follows: Annotation rule cut-off = 55; *E*-value = 1e-6; Hit-HSP overlap = 0; and the GO weight = 5.

Determining cycling transcripts

Rhythmicity in clock gene expression was determined using the JTK_CYCLE software (http://openwetware.org/wiki/HughesLab:JTK_Cycle) developed by Professors Michael E. Hughes, Karl Kornacker and John Hogenesch (Hughes, Hogenesch & Kornacker, 2010; Miyazaki *et al.*, 2011) following the JTK_CYCLE Users guide. Changes in gene expression values (TPM) were also analysed by one-way ANOVA.

Data availability

Raw data files have been deposited in public sequencing databases as indicated in the text or at <https://figshare.com/s/a2513243c63bf557b720>.

RESULTS

Animals entrained under 12:12LD regimes showed robust activity rhythms when released into constant darkness (DD) with peak activity occurring in the middle of the subjective night (Fig. 1A). Periodogram analysis of representative animals revealed a period (τ) of 24.15 h (Fig. 1B). The demonstration of rhythmicity in these animals was essential to our sampling strategy; tissue for gene discovery was taken across one complete daily cycle and pooled to ensure capture of transiently expressed transcripts. Additionally, expression profiling of transcriptome contigs required samples harvested from rhythmic animals in-phase with respect to each other.

Identification of putative *Talitrus saltator* circadian proteins

Our principle objective was to describe the cDNAs encoding the core elements of the circadian clock system in *T. saltator*, initially adopting homology cloning and standard sequencing approach but, subsequently superseded by RNAseq strategies. The Illumina RNAseq generated 141,769,456 reads, 128,386,193 of these were assembled into 156,766 clustered contigs (minimum length 300 bp). This *T. saltator* Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the Accession No. [GDUJ00000000](https://www.ncbi.nlm.nih.gov/nuclink/GDUJ00000000). The version described in this paper is the first version, [GDUJ01000000](https://www.ncbi.nlm.nih.gov/nuclink/GDUJ01000000). From this transcriptome we identified contigs encoding putative circadian clock genes (Table 1). The coding regions and conserved functional domains of identified 'core' clock genes *period* (*Talper*), *timeless* (*Taltim*), *cryptochrome2* (*Talcry2*) and *bmal1* (*Talbm1*) and *clock* (*Talclk*) are diagrammatically represented in Fig. 2. In the interests of space, alignments of all genes (core and clock-related) are shown in Figs. S1–S26 and all BLAST results are tabulated in Tables S2–S4.

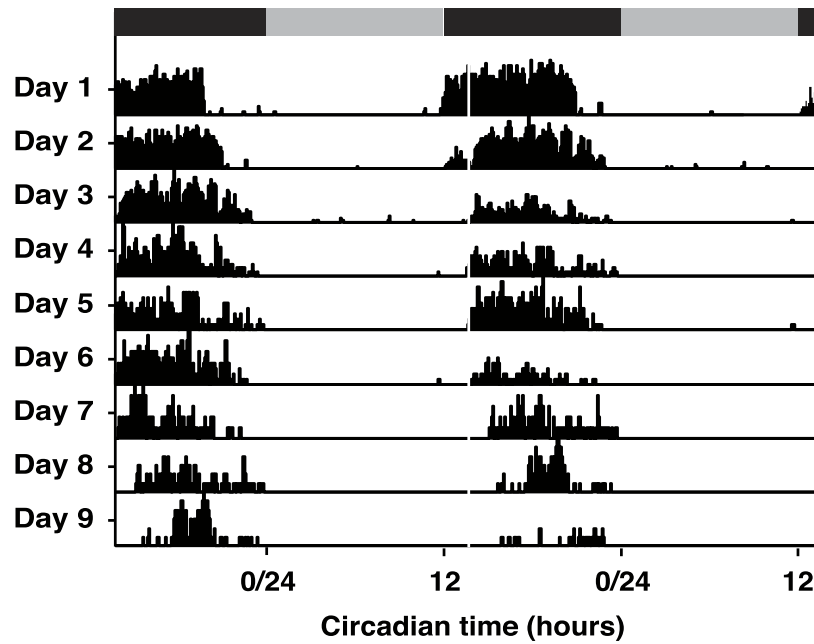
Cryptochrome 2

Degenerate PCR coupled with RACE PCR yielded a 1,843 bp cDNA sequence coding for a 565 amino acid protein; a putative cryptochrome 2 assigned TalCRY2. An identical sequence was revealed in the transcriptome data but assembled as two separate contigs (Table 1). High fidelity PCR amplification and sequencing confirmed that *Talcry2* is expressed as one contiguous transcript. The deduced protein sequence contains two SMART identified domains, a DNA photolyase domain and a FAD binding 7 domain (Figs. 2 and S1) and shared an identity of 46.4% and a similarity of 56.0% with *Danaus plexippus* CRY2 (Accession No. [ABA62409](https://www.ncbi.nlm.nih.gov/nuclink/ABA62409); Zhu et al., 2005). The highest sequence identity protein in the NCBI non-redundant database is that for the closely related isopod crustacean, *Eurydice pulchra* CRY2 sequence (Accession No. [AGV28717](https://www.ncbi.nlm.nih.gov/nuclink/AGV28717); Zhang et al., 2013).

Period and timeless

For *Talper* we mined an 8,001 bp contig encoding a putative full-length protein of 1557 amino acids aligning to dPER with an identity of 29% and a similarity of 50%. The deduced

A



B

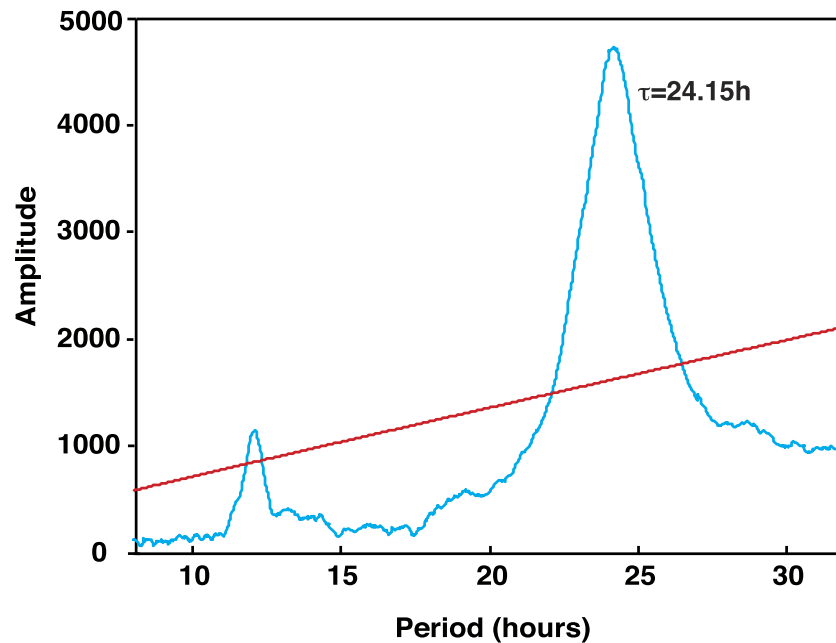


Figure 1 *Talitrus saltator* exhibits free-running circadian locomotor activity rhythms. (A) Plots of activity in five representative animals removed from the shore and held in DD over 9 days. Grey and black bars show time of subjective night and day, respectively. (B) Chi Square Periodogram analysis of activity data of five representative animals recorded in DD over nine days. The period of activity (τ) is shown inside the plot. Red line represents significance at $P < 0.001$.

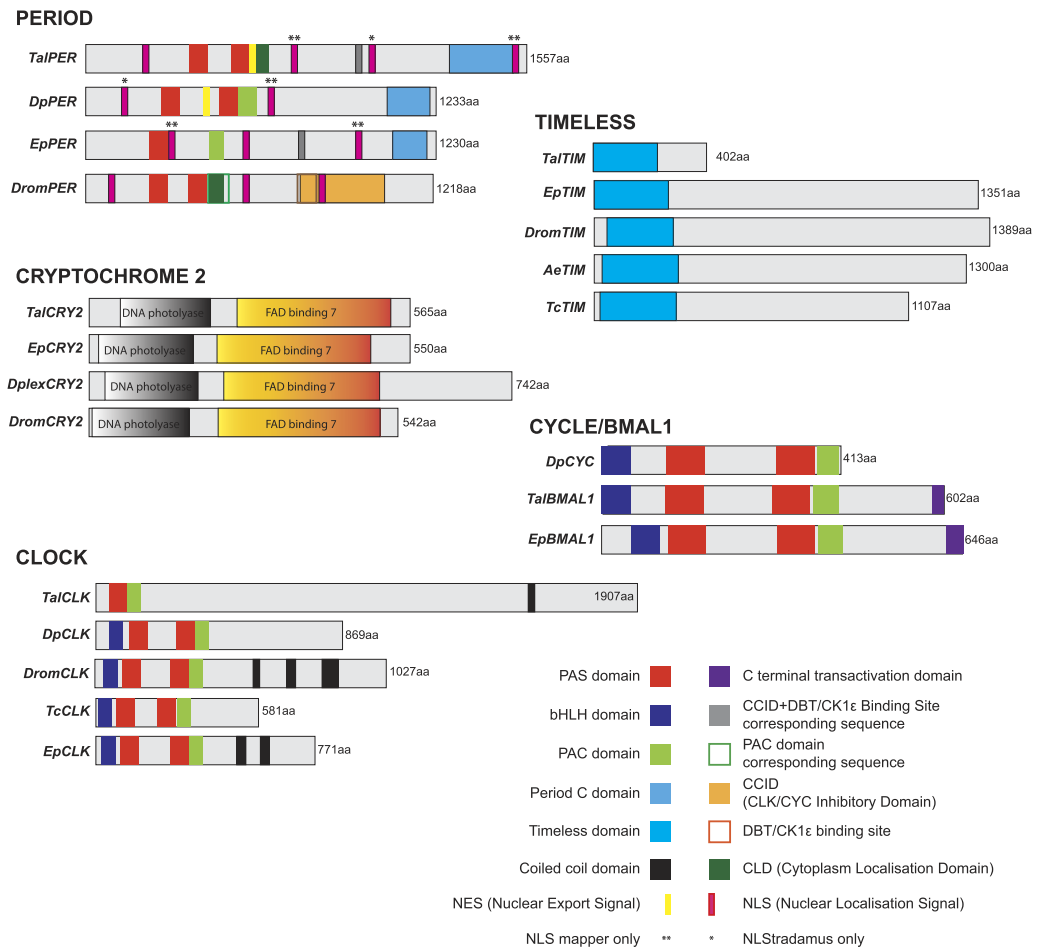


Figure 2 Cartoon representing conserved domains of core clock gene proteins. Putative *Talitrus saltator* core clock protein amino acid sequences aligned with other arthropod circadian clock protein sequences. SMART identified domains are coloured (see key for colours corresponding to domains). *Tal*, *Talitrus saltator*; *Dp*, *Daphnia pulex*; *Ep*, *Eurydice pulchra*; *Drom*, *Drosophila melanogaster*; *Dplex*, *Danaus plexippus*; *Tc*, *Tribolium castaneum*; *Ae*, *Aedes aegypti*.

TalPER contains characteristic features of PER proteins including two PAS domains and a PAC domain as well as nuclear export signal, cytoplasmic localization and nuclear localization signal motifs. In common with the isopod *E. pulchra*, TalPER also has a CK1ε binding site sequence and Period C domain (Figs. 2 and S2). The putative TalPER sequence aligned most closely to dPER-PB in FlyBase (FlyBase No. FBpp0304590) and in the NCBI database to *Eurydice pulchra* PER (Accession No. AGV28714; Zhang et al., 2013). We also used the *Drosophila melanogaster* TIM (dTIM) protein (Accession No. AAC46920; Myers et al., 1995) as a search term to identify a 1,209 bp contig with a coding region for a 402 amino acid protein sharing sequence similarity to dTIM (identity of 24.3% and similarity of 41.0%, Fig. S3). SMART analysis of this sequence identified a 228 amino acid Timeless domain (Figs. 2 and S3). The partial protein sequence, assigned TalTIM aligned most closely in the NCBI database to the partial Timeless-like protein of the termite *Zootermopsis nevadensis*

(Accession No. [KDR17447](#); *Terrapon et al., 2014*) and in FlyBase to Timeout-PA (FlyBase No. [FBpp0082180](#)).

Clock

Using the *D. melanogaster* CLK (dCLK) sequence (Accession No. [AAC62234](#); *Bae et al., 1998*) as search term, a single 5,723 bp contig was identified as a candidate TalCLK encoding sequence. However, this sequence lacked a stop codon and we were unable to deduce the C-terminal end of the protein, which is missing a C-terminal transactivation domain (Figs. 2 and S4). Nevertheless, our partial sequence contains SMART identified domains including one PAS domain, one PAC domain and coiled coil domains, whilst missing a bHLH domain and a second PAS domain present in dCLK. The TalCLK PAS domain is likely homologous to the PAS-2 domain of dCLK as it aligned better (identity of 64.2% and a similarity of 80.6%) than it did to the PAS-1 domain (identity of 17.3% and a similarity of 36.0%). Moreover, the PAC domains align closely with an identity of 75.0% and a similarity of 90.9% to that of dCLK. The most significant hit for TalCLK in the NCBI database was CLK of the prawn *Macrobrachium rosenbergii* (Accession No. [AAX44045](#); *Yang et al., 2006*). In FlyBase the most similar protein is the *D. melanogaster* CLK orthologue (FlyBase No. [FBpp0306710](#)).

Cycle/Bmal1

By searching the transcriptome using the *Drosophila melanogaster* CYC (dCYC) sequence (Accession No. [AAF49107](#); *Adams et al., 2000*), we identified two contigs as potential TalCYC/BMAL encoding sequences each of which aligned closely to separate regions of dCYC (Figs. 2 and S5). In fact these two contigs, when correctly oriented, overlap to form one contiguous sequence encoding a characteristic bHLH domain, two PAS domains a PAC and a C terminal transactivation domain. This sequence was confirmed by standard, Hi-fidelity PCR, cloning and sequencing. TalBMAL1 aligned with identity of 43% and similarity of 57% to BMAL1 of signal crayfish *Pacifastacus leniusculus* (Accession No. [AFV39705](#)). TalBMAL1 aligned with 44% identity 58% similarity to *E. pulchra* BMAL1.

In addition to the 'core' clock elements, we also isolated a comprehensive suite of 'clock-associated' transcripts from contigs that conceptually translate to proteins known to contribute to circadian clock functioning (Table 1). In the interests of space, we refer the reader to the [Supplemental Information](#); herein the majority of detail is omitted from the text. Sequence alignments are displayed in Figs. S6–S26 whilst BLAST search outputs of identified contig sequences to nucleotide, protein and FlyBase repositories are given in Tables S2–S4.

Casein Kinase II (CKII) α and β subunits

We identified single contigs coding for putative CKII α and CKII β . For CKII α we found a 5,147 bp contig enveloping a coding sequence for a 353 amino acid protein including a highly conserved serine/threonine protein kinase catalytic domain (Table 1; Fig S6). This putative TalCKII α sequence aligned in the NCBI database most closely to human CK2 Chain A (Accession No. [1NA7_A](#); *Niefind et al., 1998*) and CKII α from the copepod crustacean, *Paracyclopina nana* (Accession No. [AII16523](#)). For CKII β we identified a

1,567 bp coding sequence for a 220 amino acid protein, including a casein kinase regulatory subunit domain (Table 1; Fig. S6). This sequence aligned most closely in the NCBI database to parasitic wood wasp *Orussus abietinus* CKII β (Accession No. XP_012287730).

Clockwork orange (CWO)

A candidate *cwo* contig was mined that encodes a partial protein sequence of 167 amino acids (Table 1; Fig. S7) and included a SMART-identified bHLH domain and the Orange of the Hairy/E (SPL) family domain. TalCWO matched most closely in the NCBI database to an uncharacterized protein from the mite *Metaseiulus occidentalis* (Accession No. XP_003744690); however, the same sequence aligned with CLOCKWORK ORANGE from the firebug *Pyrrhocoris apterus* (Accession No. AGI17571; Bajgar, Jindra & Dolezel, 2013) with an *e* value of $8e^{-42}$.

Doubletime (DBT)

A putative 1,092 bp *dbt* contig coding for a deduced a full-length 310 amino acid protein was identified containing a serine/threonine protein kinase catalytic domain (Table 1; Fig. S8). The deduced sequence aligned most closely in the NCBI database to DBT of the isopod crustacean *Eurydice pulchra* (Accession No. AGV28719; Zhang et al., 2013) and in FlyBase to the *Drosophila melanogaster* orthologue DCO (FlyBase No. FBpp0306615).

Par domain protein 1 ϵ (PDP1 ϵ)

A single putative 3,423 bp *pdp1 ϵ* contig with a coding sequence for a 508 amino acid protein was identified containing a basic leucine zipper domain in the C-terminal region (Table 1; Fig. S9). The deduced TalPDP1 ϵ protein sequence aligned most closely in the NCBI database to the hepatic leukaemia factor of the clonal raider ant *Cerapachys biroi* (Accession No. EZA50108; Oxley et al., 2014); however PDP1 ϵ from the mosquito *Culex quinquefasciatus* (Accession No. XP_001865130) was also a top ranking BLAST hit aligned with an *e* value of $1e^{-35}$. In FlyBase the sequence aligned most closely to *Drosophila melanogaster* PDP1 (FlyBase No. FBpp0289727).

Protein phosphatase 1 (PP1)

A 1,725 bp contig was identified coding for a 357 putative PP1 containing a protein phosphatase 2Ac catalytic domain of the serine/threonine phosphatase family (Table 1; Fig. S10). The translated sequence aligned most closely to PP1 of the jumping ant *Harpegnathos saltator* (Accession No. EFN86649; Bonasio et al., 2010) in the NCBI database. In FlyBase the sequence aligned most closely to *Drosophila melanogaster* PP1 (FlyBase No. FBpp0306442).

Protein phosphatase 2A (PP2A)

PP2A catalytic subunit “MICROTUBULE STAR” (MTS)

The *Drosophila melanogaster* MTS sequence (dMTS, Accession No. AAF52567; Adams et al., 2000), was used as a search term to identify a contig 2,981 bp in length and coding for a putative 309 orthologue to dMTS with an identity of 91.9% and a similarity of 96.8% (Table 1; Fig. S11). The putative TalPP2A includes a protein phosphatase 2Ac catalytic domain of the serine/threonine phosphatase family with an identity of 93.4% and

a similarity of 98.2% to the fruit fly orthologue. The closest match in the NCBI database was the putative serine/threonine protein phosphatase PP-V from the body louse *Pediculus humanus corporis* (Accession No. [XP_002426726](#)). The closest match in FlyBase was the *Drosophila melanogaster* MTS homologue (FlyBase No. [FBpp0310063](#)).

PP2A catalytic subunit “WIDERBORST” (WBT)

We identified a 2,474 bp sequence that codes for a putative 458 amino acid protein. The deduced protein sequence, although 66 amino acids shorter than that of *Drosophila* WBT (the TalWBT protein lacks a sequence at the C terminal end) aligned with the fruit fly orthologue (identity of 73.8%) (Table 1; Fig. S12). In common with other species, the *T. saltator* candidate contains B56 domains (serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit epsilon). The deduced TalWBT protein aligned most closely in the NCBI database to serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit epsilon of the red flour beetle *Tribolium castaneum* (Accession No. [XP_971164](#)) and in FlyBase to *Drosophila melanogaster* WBT (FlyBase No. [FBpp0084575](#)).

PP2A catalytic subunit TWINS (TWS)

A unique 1,633 bp contig coding for a putative 445 amino acid protein with high identity to *Drosophila* TWS, (Table 1; Fig. S13) was identified. The deduced sequence contained seven WD40 domains identified using SMART. The protein aligned most closely in the NCBI database to PP2A subunit B of the mud crab *Scylla paramamosain* (Accession No. [AFK24473](#)) and to *Drosophila melanogaster* TWINS (FlyBase No. [FBpp0081671](#)).

Shaggy (SGG)

A 4,413 bp transcript incorporating a coding region for a 418 amino acid protein that contained a SMART identified serine/threonine protein kinase catalytic domain (Table 1; Fig. S14). In the NCBI database, the sequence aligned most closely to glycogen synthase kinase-3 of the turnip sawfly *Athalia rosae* (Accession No. [XP_012256017](#)) and in FlyBase to *Drosophila melanogaster* SGG (FlyBase No. [FBpp0070450](#)).

Supernumerary limbs (SLIMB)

A candidate 588 amino acid putative SLIMB protein encoded within a 2,121 bp contig was mined. This sequence included one D domain of beta-TrCP, one F box and seven WD40 domains (Table 1; Fig. S15). The deduced sequence aligned in the NCBI database most closely to F-box/WD repeat-containing protein 1A of the termite *Zootermopsis nevadensis* (Accession No. [KDR19729](#); [Terrapon et al., 2014](#)) and in FlyBase to *Drosophila melanogaster* SLIMB (FlyBase No. [FBpp0303082](#)).

Vrille (VRI)

A single 3,949 bp contig incorporating coding sequence for a 509 amino acid protein was identified as a putative *vrille* transcript. The putative TalVRI protein contains a SMART identified basic region leucine zipper domain (Table 1; Fig. S16) and aligned most closely in the NCBI database to nuclear factor interleukin-3-regulated protein of the termite *Zootermopsis nevadensis* (Accession No. [KDR19729](#); [Terrapon et al., 2014](#)) and in FlyBase to *Drosophila melanogaster* VRI ([FBpp0312171](#)).

Ebony

A single 4,380 bp contig coding for a putative 974 amino acid protein with one AMP binding domain, one AMP binding C domain and one PP binding domain was mined (Table 1; Fig. S17). The deduced protein most closely aligned in the NCBI database to the β -alanyl conjugating enzyme of the cockroach *Periplaneta americana*. *P. americana* EBONY (Accession No. CAI26307; Blenau & Baumann, 2005) has been shown to have β -alanyl-dopamine (DA) synthase (BAS) enzymatic activity and the *Drosophila melanogaster* EBONY sequence (Accession No. ABO27280) aligned with the TalEBONY candidate sequence with an e value of $8e^{-90}$, supported by EBONY (FlyBase No. FBpp0083505) being the most closely aligned protein to the TalEBONY query in FlyBase.

Pigment dispersing hormone 1 and II (PDHI and PDHII)

In the Trinity assembly, a 2,430 bp contig was identified by searching for the conserved NSELINS domain. The identified contig included a coding region for a putative 129 amino acid TalPDH prepropeptide (TalPDH-I, Table 1, Fig. S18). This was subsequently extended to 2,471 bp by the addition of a 5' UTR by RACE PCR. The TalPDH-I contains a signal peptide between residues 20–21, a 77 amino acid PDH-precursor-related peptide (PPRP) sequence ending in a K-R dibasic cleavage site. Unusually however, the deduced mature peptide sequence is 32 amino acids long and lacks an amidation signal. SMART identified PDH domains are present in the TalPDH. The deduced 32 amino acid Tal-PDH-I mature peptide has an identity of 40.6% and a similarity of 53.1% with the *Uca pugilator* β -PDH consensus sequence.

A second putative TalPDH encoding transcript was found in the transcriptome of 3,392 bp in length that included a coding region for an 89 amino acid prepropeptide that we assigned TalPDH-II (Table 1; Fig. S18). This prepropeptide sequence was predicted to include a 23 residue signal peptide. The 43 amino acid PPRP thus extends from residue 24 to a K-R dibasic cleavage site at residue 66. The 23 amino acid mature peptide terminates in an amidation signal and has an identity of 47.8% and a similarity of 69.6% with *Uca pugilator* β -PDH consensus sequence. No domains were identified by SMART but a corresponding PDH domain region was identified.

RORA

One 2,217 bp partial contig, coding a 599 amino acid containing a C4 zinc finger domain was identified (Table 1; Fig. S19) as a predicted candidate for a *T. saltator* homologue of the hormone receptor RORA. The 3' HOLI ligand binding domain present in other RORA proteins is not identified in TalRORA. It is possible that the HOLI ligand binding domain is present in the unsequenced 3' section. The most closely aligned protein in the NCBI database is the house fly *Musca domestica* nuclear hormone receptor HR3 (Accession No. XP_011290218). The *Drosophila* hormone receptor-like in 46 protein (FlyBase No. FBpp0297438) was the closest protein to TalRORA in FlyBase.

Reverb

A 5,385 bp full length contig was identified, coding for a putative 1110 amino acid protein containing both a C4 zinc finger domain and a HOLI ligand binding domain (Table 1; Fig. S20). The sequence aligned most closely in the NCBI database to nuclear hormone receptor E75, the non-mammalian REVERB homologue in the carpenter ant *Camponotus floridanus* (Accession No. XP_011259848). The closest aligning protein in FlyBase was the *Drosophila melanogaster* E75 protein (FBpp0297726).

Sirt 1, 2, 4, 6 and 7

Five contig sequences of lengths 4,033 bp, 2,275 bp, 2,977 bp, 1,209 bp and 5,180 bp all coding for partial proteins of length 955, 376, 354, 402 and 948 amino acids, respectively were identified in the transcriptome as homologues for SIRTUIN proteins 1, 2, 4, 6, and 7. Each sequence contained one SMART identified SIR2 domain (Table 1 and Figs. S21–S25). TalSIRT1 aligned most closely to the trematode *Schistosoma mansoni* SIRT1 (Accession No. ABG78545) in the NCBI database and *Drosophila melanogaster* SIRT1 (FBpp0080015) in FlyBase. TalSIRT2 was most closely aligned to the red flour beetle *Tribolium castaneum* hypothetical protein (Accession No. EFA06770) in the NCBI database and *D. melanogaster* SIRT2 (FBpp0310647) in FlyBase. The putative TalSIRT4 protein is most closely aligned in the NCBI database to SIRT4 of the Asian citrus psyllid *Diaphorina citri* (Accession No. XP_008480918) and in FlyBase to the *D. melanogaster* SIRT4 (FBpp0070817). The TalSIRT6 sequence most closely aligns to the water flea *Daphnia pulex* hypothetical protein (Accession No. EFX74386) in the NCBI database and in the FlyBase database to *D. melanogaster* SIRT6 (FBpp0293897). TalSIRT7 most closely aligns in the NCBI database to the SIRT7 protein of the leafcutter bee *Megachile rotundata* (Accession No. XP_012143211) and aligns most closely in the FlyBase database to the *D. melanogaster* SIRT7 protein (FBpp0084733).

Jetlag

One 2,454 bp contig coding a partial putative 458 amino acid protein was identified. This candidate putative protein sequence contains one F-box domain and multiple leucine-rich repeat domains (Table 1; Fig. S26). The sequence was most closely aligned in the NCBI database to the red flour beetle *Tribolium castaneum* F-box/LRR-repeat protein (Accession No. XP_008193983). In the FlyBase database the TalJET sequence aligned most closely to a *D. melanogaster* protein from the F-box and leucine rich repeat region group (FBpp0111980).

Blast2Go analysis

Gene ontology (GO) annotations were categorised into the three groupings: ‘Biological Process,’ ‘Molecular Function’ and ‘Cellular Component’. These categories of annotation were most easily displayed graphically at ontology level 2 (Fig. 3) where the most frequent Molecular Functions were “Binding” and “Catalytic activity”; the most common Biological Processes were “Cellular and Metabolic”; and the most common Cellular Components were “Cell Components” and “Organelle components.” These GO term outputs were subsequently compared with results from transcriptomes previously generated and analysed

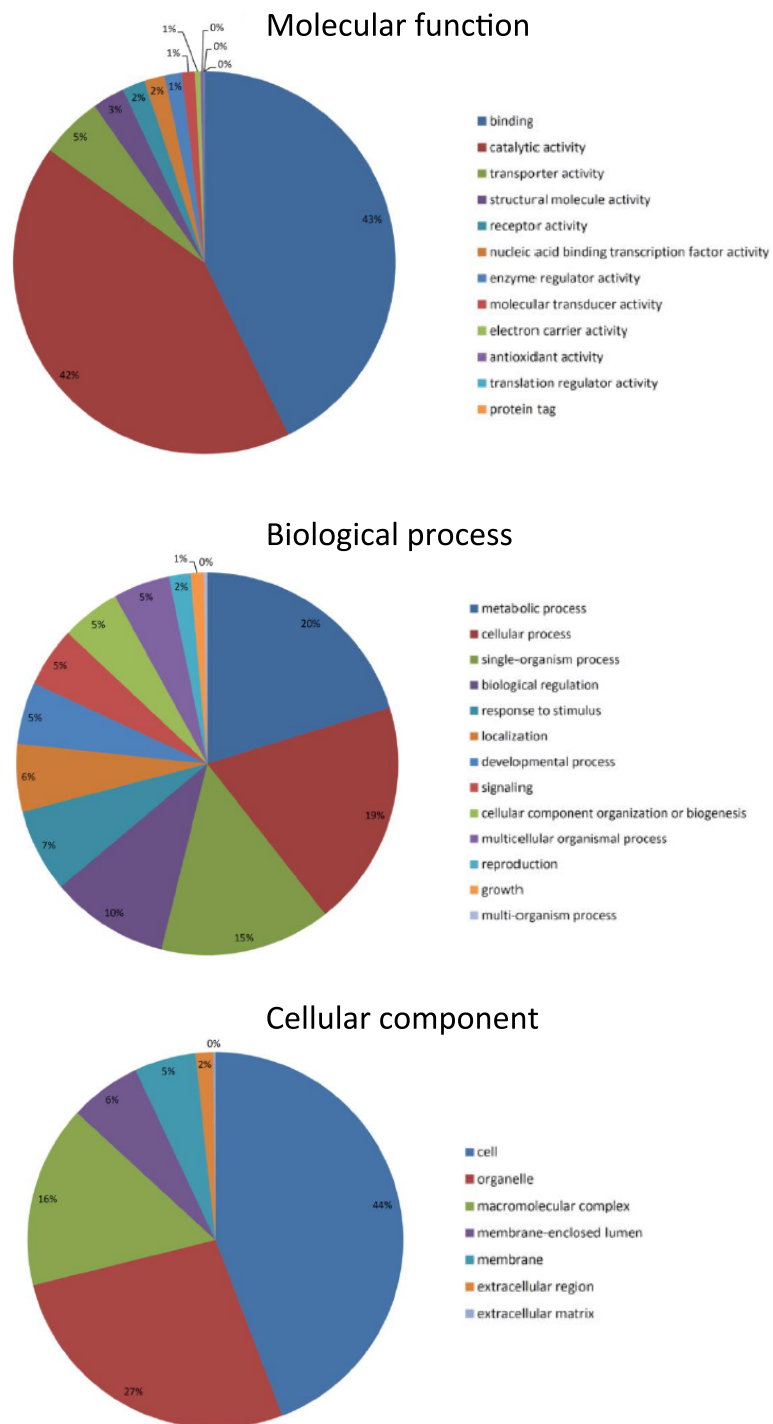


Figure 3 *Talitrus saltator* brain tissue transcriptome BLAST2GO analysis. Ontology level 2 data showing functional classification of the brain transcripts from *T. saltator* for three main gene ontology categories.

including the larva of the moth *Plutella xylostella* (Xie et al., 2012), the silkworm *Bombyx mori*, the parasitic nematode *Teladorsagia circumcincta* (Menon et al., 2012) and the North Atlantic copepod *Calanus finmarchicus* (Lenz et al., 2014). Proportions of GO terms were all found to be at broadly similar levels in all organisms transcriptomes and for each category analysed; for example, in the category Molecular Function ontology level 2 'binding' terms for *T. saltator*, *P. xylostella*, *B. mori* and *T. circumcincta* were 51%, 48%, 41% and 40% respectively whilst level 2 'catalytic activity' terms were 31%, 36%, 38% and 40% respectively.

Temporal expression of clock genes

JTK_CYCLE analysis applied to TPM values indicated oscillations (limited to a period of between 21–27h) in the core clock genes *Talper* ($q = 1.69^{-07}$) and *Talbm1* ($q = 0.001$) (Fig. 4). Alternative analyses by ANOVA showed significant differences in mean values at each interval for *Talcry2* ($F_{7,30} = 2.745$, $P = 0.032$), *Talper* ($F_{7,30} = 4.86$, $P = 0.002$) and *Talbm1* ($F_{7,30} = 2.445$, $P = 0.049$). Of the clock-related genes *Talpdh-II* ($q = 5.43^{-04}$), *TalckII β* ($q = 0.0017$), *Talpp1* ($q = 0.023$), *Talshaggy* ($q = 0.021$), *Talebony* ($q = 0.028$), *Talsirt1* ($q = 0.001$), *Talsirt7* ($q = 0.037$) and *Taljetlag* ($q = 0.027$) exhibited oscillatory expression profiles by JTK_CYCLE analysis. Interrogation of these data by ANOVA showed *TalCK11β* ($F_{7,30} = 3.284$, $P = 0.014$), *Talpp1* ($F_{7,30} = 3.99$, $P = 0.005$), *Talsirt1* ($F_{7,30} = 4.91$, $P = 0.002$) and *Talslimb* ($F_{7,30} = 3.19$, $P = 0.016$) to be differentially expressed over time. *Talper* mRNA accumulation increased approximately 6-fold during early night and peaked towards the middle of the subjective night (CT18-21) before diminishing at CT24 reaching a nadir at CT6. This profile is approximately antiphasic to *TalckII β*, *Talpp1*, *Talsirt1* and *Talslimb* that all showed peaks in expected daytime. *Talcry2* and *Talbm1* respective contigs also showed antiphasic relationships to *Talper*. However, JTK_CYCLE analysis did not reveal significant rhythmicity in the *Talcry2* transcript.

Given the robust rhythm of *Talper* expression we chose this transcript with which to validate our RNAseq expression data by quantitative PCR. RNA levels of *Talper* in samples taken from an independent time-course experiment show peak expression (~3-fold change from CT3) at CT19 ($F_{7,32} = 4.54$, $P = 0.01$, see Fig. S27). These data support those revealed by RNAseq with peak expression in the early to mid night albeit with a slightly lower amplitude.

Reads for *Talcwo* were below levels acceptable for analysis and excluded from further investigation.

DISCUSSION

Talitrus saltator exhibits intriguing clock-driven behavioural phenotypes, including circadian locomotor rhythms (Bregazzi & Naylor, 1972; Ugolini et al., 2007) and time-compensated solar and lunar navigation (Ugolini et al., 1999; Ugolini, Melis & Innocenti, 1999; Ugolini, 2003; Ugolini et al., 2007). Their abundance and suitability for behavioural experimentation make them an excellent model for comparative clock analysis but to make the species more genetically tractable we set out to define the neuromolecular components of its circadian clock. Initially, our strategy included homology cloning and RACE PCR to

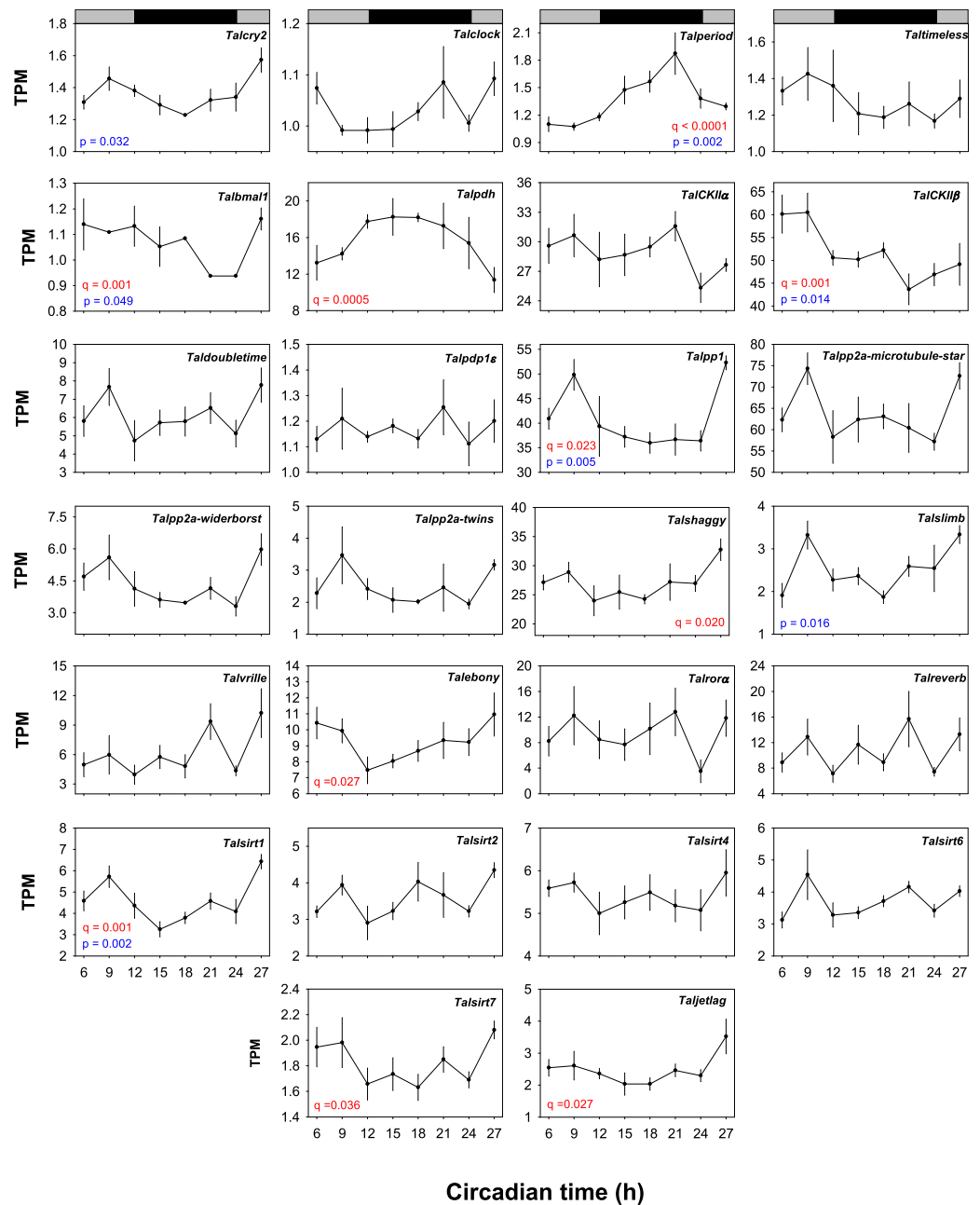


Figure 4 Temporal expression profiles of gene transcripts in the brains of free-running *Talitrus saltator*. Plots show TPM values of each identified core clock and clock-associated genes over a 24 h period. Data points represent mean TPM values \pm SEM ($N = 4$ biological replicates, except CT18 where $N = 3$). Significance values are shown in red for JTK_YCLE and in blue for ANOVA. Grey and black bars indicate expected daytime and nighttime, respectively.

sequence full-length *Talcry2*, including the 3' and 5' UTRs, but we subsequently exploited RNAseq technologies to sequence the cerebral ganglia transcriptome and identify 'core' clock genes and clock related genes, known in other species to play a role in normal clock functioning.

The quality of the Trinity assembled transcriptome was evaluated in a number of ways. First, we set out to maximise coverage of temporally regulated transcripts by harvesting brain tissue across a complete day-night cycle, and from behaviourally rhythmic animals. The clock gene transcripts identified from our assemblies and the proportion of assembled contigs for that encoded full-length proteins (17 out of 28) suggest we satisfied this aim. Second, standard cloning and sequencing of several genes, including *Talcry2*, *Talper* and *Talbm1* and *Talpdh-1* resulted in 100% nucleotide alignment with the corresponding contig coding region. Third, confidence in the identity of our contigs as clock gene transcripts was provided by the identification of conserved functional motifs in each conceptual translation. This was reflected also in the BLAST searching and high confidence levels of the search returns (expressed as *e*-values) from the NCBI nucleotide and protein databases as well as FlyBase. Taken together, these attributes indicate that the transcriptome is a faithful representation of all genes expressed in this tissue across a day-night cycle. Furthermore, Blast2GO analysis of the assembled contigs resulted in broadly similar numbers of GO terms and categorisation pattern to previous studies (Lenz *et al.*, 2014; Menon *et al.*, 2012; Xie *et al.*, 2012). The few differences observed between GO term numbers of *T. saltator* to other non-model species can be attributed to transcriptomes being compared across different tissues of diverse species at different stages of development and subjected to various environmental conditions.

Clock genes are well conserved across diverse taxa (Tauber & Kyriacou, 2005; Tauber *et al.*, 2004) and the discovery of structurally similar transcripts in *T. saltator* was expected. Nevertheless, our data yielded some interesting comparative insights. In *T. saltator* we failed to find a CRY1 but the CRY2 showed very high sequence similarity to that of the isopod *Eurydice pulchra*, also lacking a CRY1 but possessing a mammalian-like CRY2 that has been demonstrated *in vivo* to function as the principle negative repressor BMAL1:CLK transcriptional activity (Zhang *et al.*, 2013). Given the close relationship of *T. saltator* to *E. pulchra* it is tempting to speculate that the CRY2 we describe is also a negative regulator in the core oscillator system.

Whilst *Drosophila* do not have a CRY2, other insects, such as the monarch butterfly, *Danaus plexippus* have both CRY1 and CRY2 (Zhu *et al.*, 2005). This is considered to be the ancestral state- Cry2 being lost in fruit flies (Sandrelli *et al.*, 2008). However, peracarid crustaceans, that presumably predate insects, have only a CRY2, offering alternative evolutionary perspectives on the central clock mechanism. However, the evolutionary status of cryptochrome is likely highly complex; for example, and in contrast to *E. pulchra* and *T. saltator*, the copepod crustaceans, *Calanus finmarchicus* and *Tigriopus californicus* have been shown to express a putative CRY1 transcript (Lenz *et al.*, 2014; Nesbit & Christie, 2014) whilst some insects such as the flour beetle, *Tribolium castaneum* and the honey bee, *Apis mellifera* have only a mammalian CRY2 (Rubin *et al.*, 2006).

Using the *Drosophila* CYC as a search term we elucidated a CYC-like sequence in *T. saltator* containing a bHLH and 2 PAS domains, but differing from *Drosophila* in that it also contained a C-terminal transactivation domain, on the basis of which we assigned the *T. saltator* sequence as a vertebrate-like BMAL1. A very similar structural scenario was described for BMAL1 in *E. pulchra* and the water flea *Daphnia pulex* (Zhang *et al.*, 2013).

In *E. pulchra* deletion of the transactivation domain resulted in severely attenuated functionality (of the EpBMAL:EpCLK dimer) in a cell-based *in vitro* luciferase assay (Zhang *et al.*, 2013). Again, whilst we have no experimental evidence to support the notion, the close structural and phylogenetic relationship between *T. saltator* and *E. pulchra* hint at a similar role for TalBMAL1.

Although our data reveal most of the clock gene candidates we had anticipated, some such as *Taltim* and *Talclk* were incomplete. However, given the presence of the conserved domains identified within the transcripts, we are confident that these partial sequences are the products of the *Taltim* and *Talclk* genes in *T. saltator*. Manual cloning and sequencing strategies are necessary to reveal the full-length sequences to these. We also uncovered two unusual sequences identified as *Talpdh*. For TalPDH-I the N-terminal end of the mature hormone contains the conserved NSE/ALINSSLLG signature but, the remaining sequence extends beyond the expected 18 residues that define PDHs, including a further 14 residues, and lacking an amidation signal. A second contig discovered with relaxed search stringency also exhibited features of PDH-like peptides with a 23 amino acid signal peptide, a 43 residue PPRP and a 23 amino acid mature PDH with C-terminal amidation. Concentrated efforts to locate other PDH candidates in our transcriptome were unsuccessful. Interestingly, a combined transcriptomic and mass spectrometry approach on the transcriptome/neuropeptidome of the krill *Euphasia crystallorophias* (Toullec *et al.*, 2013) revealed three PDH isoforms PDH-L β 1, PDH-L β 2 and PDH-L α . In this species PDH-L β 1 expresses the characteristic N-terminal 12 residues but has an extended C-terminus, making the mature, amidated peptide 24 amino acids in length. The PDH-L2 β is almost identical except for Ser² being substituted for Ala², reminiscent of the case in *E. pulchra* (Wilcockson *et al.*, 2011). TalPDH-II also expresses the conserved N-terminus with a conserved substitution of Leu⁴ for Ile⁴. In accord with the structure of PDH prepropeptides, we observed the signal peptide and PPRP, separated from the mature peptide by a dibasic cleavage site (K-R). These PDH preprohormone-like features and confirmation of the contig sequences by standard and RACE PCR and cloning, together with the unusual PDH isoforms reported in euphausiids, allay our initial fears that our sequence data were mis-assemblies or anomalies. The structural and functional significance of PDHs in *T. saltator* are worthy of further investigation.

Many organisms show cyclic changes in gene expression of core and clock-related genes as a function of the TTFL. A feature of biological clocks is that they free-run in constant conditions i.e., the transcriptional/translational activity of the central oscillator persists with a circadian period in DD. We determined the transcriptional dynamics of the *T. saltator* clock genes in behaviourally rhythmic animals by mapping RNAseq sequencing reads for each time-course sample back to the assembled transcriptome and analysing temporal changes in these data by JTK_CYCLE analysis and ANOVA. Of the canonical clock genes only *Talper* and *Talbm* were indicated as rhythmically expressed by JTK_CYCLE analysis and ANOVA with *Talcry 2* abundance varying in time (ANOVA). In addition, the clock related genes, *TalckII β* , *Talpp1*, *Talsirt1*, *Talsirt7* *Talebony*, *Taljetlag*, *Talshaggy* and *Talpdh-II* and showed oscillating expression using the same

analytical parameters. *TalCK11 β* , *Talpp1*, *Talsirt1* and *Talslimb* abundance differences were also noted over time (ANOVA).

We chose the JTK_CYCLE algorithm because it is reported to be very robust to outliers and returns low numbers of false-positives (*Hughes, Hogenesch & Kornacker, 2010*). In our study we sacrificed an extended sampling period for increased biological replicates and three-hour sampling intervals. We are confident that this approach, coupled with the JTK_CYCLE data treatment yielded genuinely cycling genes that offer suitable targets for further exploration.

We caution that, it would be entirely speculative to draw conclusions on the functional significance of the rhythmicity or phase relationships of cyclic expression identified in the current study. Although circadian genes and proteins are conserved and the central tenet of the TTFL serves to describe the basis of rhythmicity across taxa, simple interspecific comparisons are confounded by the fact that clocks do show nuances in their organisation of the TTFL, e.g., in *E. pulchra* only *Eptim* shows robust cycling in animals expressing circadian and circatidal behaviour and yet *Epper*, which is non-rhythmically expressed, appears essential to circadian phenotypes (*Zhang et al., 2013*). Indeed, rhythmicity in gene expression itself is not a prerequisite for a functional clock (*Lakin-Thomas, 2006*). The tremendous scope for complex interplay of gene products and their cognate proteins will only be revealed by more targeted experimental approaches in non-model species. Nevertheless, our data indicate that sub-sets of clock and clock-related genes in *T. saltator* continue to cycle in free-running conditions and thus represent tantalising targets for investigation to uncover the role of these genes in either circadian or circalunidian (24.8 h) rhythms. For example, *T. saltator* has the capability to orientate by the sun and moon suggesting they may have separate daily and lunidian clocks (*Ugolini et al., 1999*). Indeed a precedence for independent clock mechanisms in marine organisms has been set, e.g., *E. pulchra* has been shown to possess a dual clock system with separate clocks orchestrating circatidal and circadian phenotypes and the marine worm *Platyneries dumerilli* employs separate mechanisms to keep circadian as well as lunar time (*Zantke et al., 2013*). The sequence data and expression analysis described here provide a foot-hold in a behaviourally tractable system from which we can address these questions.

In conclusion, we have sequenced the entire brain transcriptome of an ecologically important beach crustacean that exhibits intriguing clock-controlled phenotypes and report the sequences of key clock and clock-related genes that are likely key players in these phenotypes. Our dataset is one of only very few to describe the putative clock mechanism in a non-model crustacean and contribute to the growing inventory of crustacean and non-model transcriptomes that may have wide-ranging utility in the research community.

ACKNOWLEDGEMENTS

Thanks to Justin Pachebat, (Aberystwyth University) for his advice and help with initial transcriptome sequencing strategies.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by a Natural Environment Research Council, UK grant NE/K000594/1 awarded to D.C.W. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Natural Environment Research Council, UK: NE/K000594/1.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Joseph F. O'Grady conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Laura S. Hoelters performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Martin T. Swain analyzed the data, reviewed drafts of the paper.
- David C. Wilcockson conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:
NCBI SRA and GenBank. Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession [GDUJ00000000](#).

Data Availability

The following information was supplied regarding data availability:
Figshare: <https://figshare.com/s/a2513243c63bf557b720>.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.2555#supplemental-information>.

REFERENCES

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Abril JF, Agbayani A, An HJ,

- Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, De Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, Mcintosh TC, Mcleod MP, Mcpherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RD, Scheeler F, Shen H, Shue BC, Siden-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, Woodage T, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM, Venter JC. 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287:2185–2195 DOI 10.1126/science.287.5461.2185.
- Allada R, Chung BY. 2010. Circadian organization of behavior and physiology in *Drosophila*. *Annual Review of Physiology* 72:605–624 DOI 10.1146/annurev-physiol-021909-135815.
- Bae K, Lee C, Sidote D, Chuang KY, Ederly I. 1998. Circadian regulation of a *Drosophila* homolog of the mammalian Clock gene: PER and TIM function as positive regulators. *Molecular and Cellular Biology* 18:6142–6151 DOI 10.1128/MCB.18.10.6142.
- Bajgar A, Jindra M, Dolezel D. 2013. Autonomous regulation of the insect gut by circadian genes acting downstream of juvenile hormone signaling. *Proceedings of the National Academy of Sciences of the United States of America* 110:4416–4421 DOI 10.1073/pnas.1217060110.
- Beckwith EJ, Lelito KR, Hsu YW, Medina BM, Shafer O, Ceriani MF, De la Iglesia HO. 2011. Functional conservation of clock output signaling between flies and intertidal crabs. *Journal of Biological Rhythms* 26:518–529 DOI 10.1177/0748730411420242.
- Blau J, Young MW. 1999. Cycling vrille expression is required for a functional *Drosophila* clock. *Cell* 99:661–671 DOI 10.1016/S0092-8674(00)81554-8.
- Blenau W, Baumann A. 2005. Molecular characterization of the ebony gene from the American cockroach, *Periplaneta americana*. *Archives of Insect Biochemistry and Physiology* 59:184–195 DOI 10.1002/arch.20064.

- Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, Qin N, Donahue G, Yang P, Li Q, Li C, Zhang P, Huang Z, Berger SL, Reinberg D, Wang J, Liebig J. 2010.** Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* 329:1068–1071 DOI [10.1126/science.1192428](https://doi.org/10.1126/science.1192428).
- Bregazzi PK, Naylor E. 1972.** Locomotor activity rhythm of *Talitrus saltator* (Montagu) (Crustacean, Amphipoda). *Journal of Experimental Biology* 57:375–391.
- Christie AE, Fontanilla TM, Nesbit KT, Lenz PH. 2013.** Prediction of the protein components of a putative *Calanus finmarchicus* (Crustacea, Copepoda) circadian signaling system using a *de novo* assembled transcriptome. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 8:165–193 DOI [10.1016/j.cbd.2013.04.002](https://doi.org/10.1016/j.cbd.2013.04.002).
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005.** Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676 DOI [10.1093/bioinformatics/bti610](https://doi.org/10.1093/bioinformatics/bti610).
- Cyran SA, Buchsbaum AM, Reddy KL, Lin MC, Glossop NR, Hardin PE, Young MW, Storti RV, Blau J. 2003.** vrille, Pdp1, and dClock form a second feedback loop in the *Drosophila* circadian clock. *Cell* 112:329–341 DOI [10.1016/S0092-8674\(03\)00074-6](https://doi.org/10.1016/S0092-8674(03)00074-6).
- De Pitta C, Biscontin A, Albiero A, Sales G, Millino C, Mazzotta GM, Bertolucci C, Costa R. 2013.** The Antarctic Krill *Euphausia superba* shows diurnal cycles of transcription under natural conditions. *PLoS ONE* 8: Article e68652 DOI [10.1371/journal.pone.0068652](https://doi.org/10.1371/journal.pone.0068652).
- Edwards JM, Naylor E. 1987.** Endogenous circadian changes in orientational behavior of *Talitrus saltator*. *Journal of the Marine Biological Association of the United Kingdom* 67:17–26 DOI [10.1017/S0025315400026333](https://doi.org/10.1017/S0025315400026333).
- Glossop NR, Houl JH, Zheng H, Ng FS, Dudek SM, Hardin PE. 2003.** VRILLE feeds back to control circadian transcription of Clock in the *Drosophila* circadian oscillator. *Neuron* 37:249–261 DOI [10.1016/S0896-6273\(03\)00002-3](https://doi.org/10.1016/S0896-6273(03)00002-3).
- Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. 2008.** High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research* 36:3420–3435 DOI [10.1093/nar/gkn176](https://doi.org/10.1093/nar/gkn176).
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, Di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011.** Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644–652 DOI [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883).
- Hall TA. 1999.** BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: *Nucleic Acid Symposium*. p 95–98.
- Hardin PE. 2009.** Molecular mechanisms of circadian timekeeping in *Drosophila*. *Sleep and Biological Rhythms* 7:235–242 DOI [10.1111/j.1479-8425.2009.00412.x](https://doi.org/10.1111/j.1479-8425.2009.00412.x).
- Hughes ME, Hogenesch JB, Kornacker K. 2010.** JTK-CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *Journal of Biological Rhythms* 25:372–380 DOI [10.1177/0748730410379711](https://doi.org/10.1177/0748730410379711).

- Kadener S, Stoleru D, McDonald M, Nawathean P, Rosbash M. 2007.** Clockwork Orange is a transcriptional repressor and a new *Drosophila* circadian pacemaker component. *Genes and Development* **21**:1675–1686 DOI [10.1101/gad.1552607](https://doi.org/10.1101/gad.1552607).
- Lakin-Thomas PL. 2006.** Transcriptional feedback oscillators: maybe, maybe not. *Journal of Biological Rhythms* **21**:83–92 DOI [10.1177/0748730405286102](https://doi.org/10.1177/0748730405286102).
- Lenz PH, Roncalli V, Hassett RP, Wu LS, Cieslak MC, Hartline DK, Christie AE. 2014.** De Novo Assembly of a transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) - the dominant zooplankter of the North Atlantic Ocean. *PLoS ONE* **9**:ARTNe88589 DOI [10.1371/journal.pone.0088589](https://doi.org/10.1371/journal.pone.0088589).
- Letunic I, Doerks T, Bork P. 2009.** SMART 6: recent updates and new developments. *Nucleic Acids Research* **37**:D229–D232 DOI [10.1093/nar/gkn808](https://doi.org/10.1093/nar/gkn808).
- Li B, Dewey CN. 2011.** RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**:323 DOI [10.1186/1471-2105-12-323](https://doi.org/10.1186/1471-2105-12-323).
- Li W, Godzik A. 2006.** Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**:1658–1659 DOI [10.1093/bioinformatics/btl158](https://doi.org/10.1093/bioinformatics/btl158).
- Li W, Jaroszewski L, Godzik A. 2001.** Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**:282–283 DOI [10.1093/bioinformatics/17.3.282](https://doi.org/10.1093/bioinformatics/17.3.282).
- Lim C, Chung BY, Pitman JL, McGill JJ, Pradhan S, Lee J, Keegan KP, Choe J, Allada R. 2007.** Clockwork orange encodes a transcriptional repressor important for circadian-clock amplitude in *Drosophila*. *Current Biology* **17**:1082–1089 DOI [10.1016/j.cub.2007.05.039](https://doi.org/10.1016/j.cub.2007.05.039).
- Matsumoto A, Ukai-Tadenuma M, Yamada RG, Houl J, Uno KD, Kasukawa T, Dauwalder B, Itoh TQ, Takahashi K, Ueda R, Hardin PE, Tanimura T, Ueda HR. 2007.** A functional genomics strategy reveals clockwork orange as a transcriptional regulator in the *Drosophila* circadian clock. *Genes and Development* **21**:1687–1700 DOI [10.1101/gad.1552207](https://doi.org/10.1101/gad.1552207).
- Menon R, Gasser RB, Mitreva M, Ranganathan S. 2012.** An analysis of the transcriptome of *Teladorsagia circumcincta*: its biological and biotechnological implications. *BMC Genomics* **13**: Article S10 DOI [10.1186/1471-2164-13-S7-S10](https://doi.org/10.1186/1471-2164-13-S7-S10).
- Miyazaki M, Schroder E, Edelmann SE, Hughes ME, Kornacker K, Balke CW, Esser KA. 2011.** Age-associated disruption of molecular clock expression in skeletal muscle of the spontaneously hypertensive rat. *PLoS ONE* **6**:e27168 DOI [10.1371/journal.pone.0027168](https://doi.org/10.1371/journal.pone.0027168).
- Myers MP, Wagersmith K, Wesley CS, Young MW, Sehgal A. 1995.** Positional cloning and sequence analysis of the *Drosophila* Clock gene, Timeless. *Science* **270**:805–808 DOI [10.1126/science.270.5237.805](https://doi.org/10.1126/science.270.5237.805).
- Nesbit KT, Christie AE. 2014.** Identification of the molecular components of a *Tigriopus californicus* (Crustacea, Copepoda) circadian clock. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **12**:16–44 DOI [10.1016/j.cbd.2014.09.002](https://doi.org/10.1016/j.cbd.2014.09.002).

- Niefind K, Guerra B, Pinna LA, Issinger OG, Schomburg D. 1998. Crystal structure of the catalytic subunit of protein kinase CK2 from *Zea mays* at 2.1 Å resolution. *EMBO Journal* 17:2451–2462 DOI 10.1093/emboj/17.9.2451.
- Oxley PR, Ji L, Fetter-Pruneda I, McKenzie SK, Li C, Hu HF, Zhang GJ, Kronauer DJC. 2014. The genome of the clonal raider ant *Cerapachys biroi*. *Current Biology* 24:451–458 DOI 10.1016/j.cub.2014.01.018.
- Pardi L, Papi F. 1952. Die Sonne Als kompass bei *Talitrus saltator* (Montagu), (Amphipoda, Talitridae). *Naturwissenschaften* 39:262–263.
- Petersen TN, Brunak S, Von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8:785–786 DOI 10.1038/nmeth.1701.
- Reppert SM. 2006. A colorful model of the circadian clock. *Cell* 124:233–236 DOI 10.1016/j.cell.2006.01.009.
- Rubin EB, Shemesh Y, Cohen M, Elgavish S, Robertson HM, Bloch G. 2006. Molecular and phylogenetic analyses reveal mammalian-like clockwork in the honey bee (*Apis mellifera*) and shed new light on the molecular evolution of the circadian clock. *Genome Research* 16:1352–1365 DOI 10.1101/gr.5094806.
- Sandrelli F, Costa R, Kyriacou CP, Rosato E. 2008. Comparative analysis of circadian clock genes in insects. *Insect Molecular Biology* 17:447–463 DOI 10.1111/j.1365-2583.2008.00832.x.
- Sauman I, Briscoe AD, Zhu H, Shi D, Froy O, Stalleicken J, Yuan Q, Casselman A, Reppert SM. 2005. Connecting the navigational clock to sun compass input in monarch butterfly brain. *Neuron* 46:457–467 DOI 10.1016/j.neuron.2005.03.014.
- Shafer OT, Yao Z. 2014. Pigment-dispersing factor signaling and circadian rhythms in insect locomotor activity. *Current Opinion in Insect Science* 1:73–80 DOI 10.1016/j.cois.2014.05.002.
- Strauss J, Dirksen H. 2010. Circadian clocks in crustaceans: identified neuronal and cellular systems. *Front Biosci (Landmark Ed)* 15:1040–1074 DOI 10.2741/3661.
- Tauber E, Kyriacou CP. 2005. Molecular evolution and population genetics of circadian clock genes. *Methods in Enzymology* 393:797–817 DOI 10.1016/S0076-6879(05)93042-5.
- Tauber E, Last KS, Olive PJ, Kyriacou CP. 2004. Clock gene evolution and functional divergence. *Journal of Biological Rhythms* 19:445–458 DOI 10.1177/0748730404268775.
- Terrapon N, Li C, Robertson HM, Ji L, Meng XH, Booth W, Chen ZS, Childers CP, Glastad KM, Gokhale K, Gowin J, Gronenberg W, Hermansen RA, Hu HF, Hunt BG, Huylmans AK, Khalil SMS, Mitchell RD, Munoz-Torres MC, Mustard JA, Pan HL, Reese JT, Scharf ME, Sun FM, Vogel H, Xiao J, Yang W, Yang ZK, Yang ZQ, Zhou JJ, Zhu JW, Brent CS, Elsik CG, Goodisman MAD, Liberles DA, Roe RM, Vargo EL, Vilcinskis A, Wang J, Bornberg-Bauer E, Korb J, Zhang GJ, Liebig J. 2014. Molecular traces of alternative social organization in a termite genome. *Nature Communications* 5: Article 3636 DOI 10.1038/ncomms4636.
- Toullec JY, Corre E, Bernay B, Thorne MA, Cascella K, Ollivaux C, Henry J, Clark MS. 2013. Transcriptome and peptidome characterisation of the main neuropeptides and

- peptidic hormones of a euphausiid: the Ice Krill, *Euphausia crystallorophias*. *PLoS ONE* **8**:e71609 DOI 10.1371/journal.pone.0071609.
- Ugolini A. 2003.** Activity rhythms and orientation in sandhoppers (Crustacea, Amphipoda). *Frontiers in Bioscience* **8**:s722–s732 DOI 10.2741/1041.
- Ugolini A, Melis C, Innocenti R. 1999.** Moon orientation in adult and young sandhoppers. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology* **184**:9–12 DOI 10.1007/s003590050301.
- Ugolini A, Melis C, Innocenti R, Tiribilli B, Castellini C. 1999.** Moon and sun compasses in sandhoppers rely on two separate chronometric mechanisms. *Proceedings of the Royal Society B: Biological Sciences* **266**:749–752 DOI 10.1098/rspb.1999.0700.
- Ugolini A, Somigli S, Pasquali V, Renzi P. 2007.** Locomotor activity rhythm and sun compass orientation in the sandhopper *Talitrus saltator* are related. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology* **193**:1259–1263 DOI 10.1007/s00359-007-0277-z.
- Ugolini A, Tiribilli B, Boddi V. 2002.** The sun compass of the sandhopper *Talitrus saltator*: the speed of the chronometric mechanism depends on the hours of light. *Journal of Experimental Biology* **205**:3225–3230.
- Wagner GP, Kin K, Lynch VJ. 2012.** Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences* **131**:281–285 DOI 10.1007/s12064-012-0162-3.
- Wilcockson DC, Zhang L, Hastings MH, Kyriacou CP, Webster SG. 2011.** A novel form of pigment-dispersing hormone in the central nervous system of the intertidal marine isopod, *Eurydice pulchra* (leach). *Journal of Comparative Neurology* **519**:562–575 DOI 10.1002/cne.22533.
- Xie W, Lei YY, Fu W, Yang ZX, Zhu X, Guo ZJ, Wu QJ, Wang SL, Xu BY, Zhou XG, Zhang YJ. 2012.** Tissue specific transcriptome profiling of *Plutella Xylostella* third instar larval midgut. *International Journal of Biological Sciences* **8**:1142–1155 DOI 10.7150/ijbs.4588.
- Yang JS, Dai ZM, Yang F, Yang WJ. 2006.** Molecular cloning of Clock cDNA from the prawn, *Macrobrachium rosenbergii*. *Brain Research* **1067**:13–24 DOI 10.1016/j.brainres.2005.10.003.
- Zantke J, Ishikawa-Fujiwara T, Arboleda E, Lohs C, Schipany K, Hallay N, Straw AD, Todo T, Tessmar-Raible K. 2013.** Circadian and circalunar clock interactions in a marine annelid. *Cell Reports* **5**:99–113 DOI 10.1016/j.celrep.2013.08.031.
- Zhang L, Hastings MH, Green EW, Tauber E, Sladek M, Webster SG, Kyriacou CP, Wilcockson DC. 2013.** Dissociation of circadian and circatidal timekeeping in the marine crustacean *Eurydice pulchra*. *Current Biology* **23**:1863–1873 DOI 10.1016/j.cub.2013.08.038.
- Zhu H, Yuan Q, Briscoe AD, Froy O, Casselman A, Reppert SM. 2005.** The two CRYs of the butterfly. *Current Biology* **15**:R953–R954 DOI 10.1016/j.cub.2005.11.030.