



# Potential Conservation of Circadian Clock Proteins in the *phylum* Nematoda as Revealed by Bioinformatic Searches

Andrés Romanowski<sup>1‡</sup>, Matías Javier Garavaglia<sup>2</sup>, María Eugenia Goya<sup>1</sup>, Pablo Daniel Ghiringhelli<sup>2</sup>, Diego Andrés Golombek<sup>1\*</sup>

**1** Laboratorio de Cronobiología, Dto de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Buenos Aires, Argentina, **2** Laboratorio de Ingeniería Genética y Biología Celular y Molecular – Área Virolos de Insectos (LIGBCM – AVI), Instituto de Microbiología Básica y Aplicada (IMBA), Dto de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Buenos Aires, Argentina

## Abstract

Although several circadian rhythms have been described in *C. elegans*, its molecular clock remains elusive. In this work we employed a novel bioinformatic approach, applying probabilistic methodologies, to search for circadian clock proteins of several of the best studied circadian model organisms of different taxa (*Mus musculus*, *Drosophila melanogaster*, *Neurospora crassa*, *Arabidopsis thaliana* and *Synechococcus elongatus*) in the proteomes of *C. elegans* and other members of the *phylum* Nematoda. With this approach we found that the Nematoda contain proteins most related to the core and accessory proteins of the insect and mammalian clocks, which provide new insights into the nematode clock and the evolution of the circadian system.

**Citation:** Romanowski A, Garavaglia MJ, Goya ME, Ghiringhelli PD, Golombek DA (2014) Potential Conservation of Circadian Clock Proteins in the *phylum* Nematoda as Revealed by Bioinformatic Searches. PLoS ONE 9(11): e112871. doi:10.1371/journal.pone.0112871

**Editor:** Urs Albrecht, University of Fribourg, Switzerland

**Received:** August 8, 2014; **Accepted:** October 15, 2014; **Published:** November 14, 2014

**Copyright:** © 2014 Romanowski et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by National Science Agency (ANPCyT), National Research Council (CONICET) and National University of Quilmes (UNQ) grants to DAG and PDG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: dgolombek@gmail.com

‡ Current address: Laboratorio de Genómica Comparativa del Desarrollo Vegetal, Fundación Instituto Leloir, Buenos Aires, Argentina

## Introduction

*C. elegans* is a soil-dwelling nematode subjected to daily changes in environmental variables in its natural habitat. Several daily and circadian rhythms have been described in this organism, including locomotor and swimming activity, tolerance to abiotic and biotic stress, metabolic variables and melatonin synthesis [1–6]. However, little is known about the molecular mechanism that governs its circadian behaviors, in comparison to information related to other model organisms. In particular, early studies in the fruit fly *Drosophila melanogaster* suggested a feedback mechanism involving a transcription-translation oscillator (TTO), based on the observation that the oscillation of the *period* gene at the mRNA and protein level, as well as their interaction, was required to maintain rhythmicity [7]. However, it was not clear whether proteins exerted a direct effect over the mRNA or acted by means of other biochemical signals. This was later complemented with studies in the fungus *Neurospora crassa*, which showed that the FREQUENCY (FRQ) protein was capable of regulating its own transcription, by means of a negative feedback loop [8]. At the present time, the transcriptional-translational feedback loop (TTFL) is considered the fundamental block of every circadian clock and has been identified in every organism studied so far [9], although some exceptions have been reported very recently [10–12].

In mammals, the core clock components are CLOCK, BMAL, PER and CRY. During the day a CLOCK:BMAL heterodimer activates the transcription of the *per* and *cry* genes. Once translated, PER:PER and PER:CRY dimers are formed and then translocate to the nucleus by dusk. These dimers interfere negatively with CLOCK:BMAL and therefore repress the transcription of their own genes. Once the PER:CRY dimer is degraded, CLOCK:BMAL can again induce transcription of *per* and *cry*, thus closing the loop. Posttranscriptional regulation also plays a role fine tuning the circadian molecular machinery. For example, casein kinase 1 epsilon (CKIε) also regulates the clock fulfilling three roles: 1) tagging PER monomers for degradation; 2) promoting PER:PER and PER:CRY translocation to the nucleus; and 3) it is involved in the degradation of this dimers once they have accomplished their repressive roles [13].

In the case of *Drosophila*, the process presents subtle changes regarding the mammalian machinery. The dimer that activates *per* transcription is CLOCK:CYCLE, which also drives TIMELESS (TIM) transcription; TIM:PER dimers accumulate in the cytoplasm, enter the nucleus and repress CLOCK:CYCLE action and, hence *per* and *tim* transcription. The dCRY protein acts as a photoreceptor and, upon light stimulation, blocks the activity of the PER-TIM dimmer by direct interaction with TIM [14].

The homolog to CKIε in flies, DOUBLETIME (DBT), is in charge of tagging PER monomers in the cytoplasm. At dawn,

CRY is activated and induces the degradation of TIM in the PER:TIM dimers. At the same time, DBT induces the degradation of PER [15].

In *Neurospora crassa*, the transcription of the *frq* gene is induced by the WHITE COLLAR (WC) complex, composed of the WC1 and WC2 proteins. The FRQ protein interacts with an RNA helicase, FRH, and this complex then represses the transcription of FRQ. The VIVID photoreceptor acts in a similar way to dCRY: it promotes the degradation of dCLOCK, interacting with the negative element FRQ:FRH and the positive element, the WC complex. Even though the clock proteins of *Neurospora* are not conserved in insects and mammals, they share certain domains, such as the Per-Arnt-Sim (PAS) domain found in PER, CLOCK, WC1-2 and VIVID. Regardless of the actors it is interesting to note that the central TTFL oscillator is very similar in these very different organisms [16]. A similar loop can be found in plants, where TOC1 is repressed by the negative elements LHY y CCA1 [17]. In fact, the same clock architecture is found in the cyanobacteria *Synechococcus aureus*, where the KAI-A protein activates the transcription of the *KaiBC operon*, and the KAI-C product represses it [18]. The cyanobacterial clock can, however, remarkably sustain its rhythmicity in the absence of transcription by means of posttranslational mechanisms, in a loop involving the three KAI proteins [9,19].

The molecular mechanism of the *C. elegans* circadian clock has yet to be elucidated. Previous bioinformatics approaches have identified possible clock homologs. These include the PER homolog protein LIN-42, and the CLOCK homolog protein AHA-1 [20]. Both these proteins have been described to play a key role in the developmental program of *C. elegans* [21–23]. Also, a global genome wide transcriptomics analysis of *Caenorhabditis elegans* N2 nematodes entrained to light:dark (LD) and temperature (T:t, T=high temperature and t=low temperature) has recently shown that several genes are expressed in a circadian manner [24]. The datasets describe that 1817 transcripts are Tt driven; 775 are LD driven; 286 are entrained by Tt and; 406 are entrained by LD. Further analysis of the data showed that the “driven” datasets share 107 transcripts and only 2 genes are entrained by both *Zeitgebers* (*wdr-5.3*, a homolog of mWDR5, a member of a histone methyltransferase complex that in mammals associates with mPER1; and *Y102A5C.6*, a pseudogen). One interpretation could be that *C. elegans* actually has 2 different clocks (differentially entrained by light and temperature). Another interesting and surprising result is that in the conditions used in that report, the mRNA of homologs to clock genes from other organisms does not appear to cycle in *C. elegans*, suggesting that a novel circadian molecular mechanism for nematodes.

There are also some candidates for the entrainment mechanism, including a role for TAX-2, a CNG channel involved in the transduction of temperature and light signals. A mutant strain carrying a mutation in *tax-2* (PR671 strain) showed severely affected transcriptional rhythms. This was studied using a GFP reporter that can be entrained to temperature cycles (*nlp36p::gfp*) and three randomly chosen light-entrainable transcripts. This showed that *tax-2* is necessary to convey light and temperature signals to the clock. It is interesting to note that *lite-1* mutant nematodes (a gene that encodes for the LITE-1 photoreceptor), required for low wavelength light responses, did not show any problems in light synchronization [24].

The bioinformatics efforts performed to search for homologs of the clock genes of other species in *C. elegans* have not been very exhaustive and were performed by searching for protein alignment using the BLASTP algorithm, with mammalian and insect proteins as queries [20,25]. Nowadays, other techniques can be

applied, based in probabilistic methods, which might yield more powerful results than those from BLASTP protein sequence remote similarity search tools. Since the introduction of the BLAST suite in the 1990s, several theoretical advances in the homology search methodology have been made, such as hidden Markov models (HMMs), which allow for searches with probabilistic inference technology that are as fast as BLAST [26–28]. In this way, hidden Markov models can be generated out of protein alignments of similar proteins of known function from different organisms (i.e., the PERIOD proteins of several insects). With this HMM as input a search can be performed in whole proteomes using the *HMMsearch* included in the HMMER suite [28,29]. Also, several complete genomes and proteomes from different species of nematodes of the genus *Caenorhabditis* (including *C. elegans*) and other genus, such as *Ascaris* and *Brugia* [30], are completely annotated. By combining the set of tools of the HMMER suite and the availability of these complete proteomes, we have been able to perform exhaustive searches of core clock and accessory genes similarity between the known model species and the proteomes of 19 nematode species.

Hidden Markov Models (HMM) generated from protein alignments of the clock genes of the aforementioned model organisms allow us to determine similar proteins in nematodes. These HMM models contain the most relevant information (protein domains) that allows these proteins to function as components of the molecular clock. By means of these methods, proteins that fit each generated HMM model can be found in *Caenorhabditis elegans* and other nematodes. In this way, we can determine components shared with the phylum *Nematoda* that could be part of a TTFL biological clock.

This type of approach conveys high predictive value to perform experiments to validate the role of particular genes in the circadian clock of *C. elegans*.

## Results

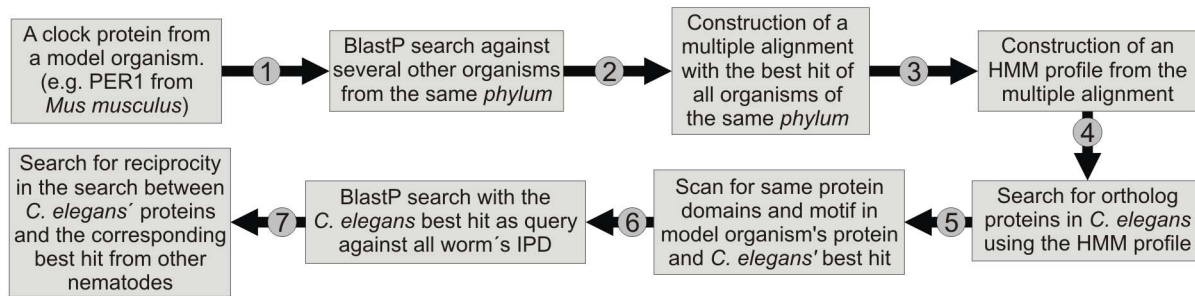
### The proteome of *C. elegans* contains proteins with similitude to clock components from known model organism

Using the workflow depicted in figure 1, we first analyzed the existence of clock components similar to those from the most well-known circadian model organisms of 5 different *phyla*: *Arabidopsis thaliana* (plant), *Synechococcus elongatus* (cyanobacteria), *Neurospora crassa* (fungi), *Drosophila melanogaster* (insect), *Mus musculus* (mammal), in *C. elegans* and then the corresponding ortholog proteins in the rest of the nematodes used in this work.

*C. elegans* is very distantly related to plants as well as to cyanobacteria. As expected, nematodes do not exhibit similarity to any of the core clock proteins of either plants or prokaryotes.

It is interesting to note that the search for similar core clock proteins of fungi revealed, quite unexpectedly, several candidate proteins in the proteome of *C. elegans* that presented a significant similarity to each of the components of the fungal clock, as can be seen in table 1.

FRH, an RNA helicase that interacts with FRQ, had a practically perfect match (E=0) with the *C. elegans* protein ceMTR-4. Both proteins have almost the same length (1006aa and 1026aa, respectively) and the same domains, with the exception of a calcium binding domain (Prosite PS00018, calcium binding domain, *EF-Hand 1*). Nevertheless, in *C. elegans* this protein appears to be involved in the polyadenylation of mRNAs destined to degradation or exosomal trimming and is also a substrate of the RTK-RAS-ERK pathway in vivo [31,32].



**Figure 1. Methodology workflow.** The diagram briefly describes the methods we applied in this work. doi:10.1371/journal.pone.0112871.g001

FWD-1, a protein involved in FRQ degradation, exhibited a high degree of similarity ( $E = 1.2 \times 10^{-80}$ ) with ceLIN-23 (K10B2.1). These proteins are likely to play a similar role in both organisms since they are both proteins with F-Box and WD40 repeats domains. In *C. elegans*, ceLIN-23 is a component of the SCF ubiquitin ligase complex (Skp1, Cullin, F-Box) which is involved in ubiquitin mediated protein degradation. Also, ceLIN-23 is a negative regulator of postembryonic cell division in all tissue types. It also regulates neurite growth in some types of neurons and the abundance of the glutamate receptor ceGLR-1 [33–36].

VIVID (VVD), another negative element of *N. crassa*'s circadian clock, present similarity ( $E = 1.3 \times 10^{-6}$ ) with ceEGL-2 (F16B3.1). VVD is a 186 aa protein which contains a PAS domain, that represses light input and regulates the time setting of the circadian clock. On the other hand, ceEGL-2, is a 956aa potassium channel required for egg laying, muscle activation, defecation, mechanosensation and chemosensation [37–40].

The *N. crassa* clock has two positive elements that comprise the WC complex, WC1 and WC2. WC1 presents similarity ( $E = 1.1 \times 10^{-3}$ ) to a small region of ceEGL-2, probably due to the fact that they both share a PAS domain. WC2, on the other hand, presents similarity to ceHIF-1d ( $E = 3.9 \times 10^{-10}$ ). In *C. elegans*, this protein is involved in oxygen sensing and is required for nematode survival in hypoxic environments (<1% oxygen) [41]. These two proteins probably do not play the same role in both organisms given the fact that WC2 possesses other domains that are absent in ceHIF-1, such as the GATA zinc finger DNA binding.

### *C. elegans* proteome exhibits similarity to insect and mammalian clock proteins

*C. elegans*' proteome includes many proteins with high similarity to those of the core circadian clock of arthropods. The best hits found with the searches performed using the HMM

**Table 1. Protein homologs to core clock proteins of plants, cyanobacteria and fungi.**

Organism	Protein	Clock function	<i>C. elegans</i> homolog	E value
<i>Arabidopsis thaliana</i>	CCA1	central clock	None	None
	COP1	LHY/CCA1 modulator	None	None
	DET1	LHY/CCA1 modulator	None	None
	GI	central clock	None	None
	LHY	central clock	None	None
	LUX	central clock	None	None
	PRR5	central clock	None	None
	PRR7	central clock	None	None
	PRR9	central clock	None	None
	TOC1	central clock	None	None
<i>Synechococcus elongatus</i>	KAI-A	central clock	None	None
	KAI-B	central clock	None	None
	KAI-C	central clock	None	None
<i>Neurospora crassa</i>	FRH	central clock	W08D2.7 (MTR-4)	0
	FRQ	central clock	None	None
	FWD-1	FRQ modulator	K10B2.1 (LIN-23)	1.2e-80
	VVD	central clock	F16B3.1 (EGL-2)	1.3e-06
	WC-1	central clock	F16B3.1 (EGL-2)	1.1e-05
WC-2	central clock	F38A6.3d (HIF-1d)	3.9e-10	

These proteins are the best hits found by probabilistic inference search based on Hidden Markov Models built upon multiple protein alignments of the core clock proteins of plants (*A. thaliana*), prokaryotes (*S. elongatus*) and fungi (*N. crassa*). Common *C. elegans* names are enclosed between parenthesis when available. doi:10.1371/journal.pone.0112871.t001

**Table 2.** Homolog proteins to core clock genes of insects.

Organism	Protein	Clock function	<i>C. elegans</i> homolog	E value
<i>Drosophila melanogaster</i>	PER	central clock	F47F6.1b (LIN-42b)	2.10E-30
	TIM	central clock	Y75B8A.22 (TIM-1)	3.20E-18
	CLK	central clock	C25A1.11a (AHA-1a)	1.70E-21
	CYC	central clock	C25A1.11a (AHA-1a)	4.10E-75
	CRY	central clock	None	None

These proteins are the best hits found by probabilistic inference search based on Hidden Markov Models built upon multiple protein alignments of the core clock proteins of insects (*D. melanogaster*). Common *C. elegans* names are enclosed between parenthesis when available.  
doi:10.1371/journal.pone.0112871.t002

profiles derived from insect protein alignments [15,42,43] are shown in table 2.

The Period (PER) and Timeless (TIM) proteins constitute the principal negative feedback loop of the circadian clock of *Drosophila melanogaster* and other insects. Both proteins have high homology counterparts in *C. elegans*' proteome. ceLIN42b (F47F6.1b) exhibits high similarity to PER ( $E = 2.1 \times 10^{-30}$ ). Both proteins possess PAS domains (PFAM PF00989) and a domain recognized as *PERIOD CIRCADIAN PROTEIN* (PANTHER PTHR11269), although they differ in length: PER has 1218aa and ceLIN-42b is 597aa long. ceTIM-1(Y75B8A.22) presents high similarity to TIM. These proteins are almost the same length (1353aa and 1398aa, respectively) and share the same domains according to analysis by InterProScan. It has been reported, however, that both nematode proteins, ceLIN42 and ceTIM-1, perform heterochronic functions in *C. elegans* [21,44].

The positive elements, Clock (CLK) and Cycle (CYC), show the best hits with the same protein in *C. elegans*: ceAHA-1A (C25A1.11A), with  $E = 1.70 \times 10^{-21}$  and  $E = 4.1 \times 10^{-75}$ , respectively. This is the only protein that has all the domains present in both CLK and CYC. The analysis by InterProScan reveals 2 PAS domains (SMART SM00091), a bHLH domain (PFAM PF00010), nuclear translocation domains (PROSITE PR00785) and one *CIRCADIAN PROTEIN CLOCK/ARNT/BMAL/PAS* domain (PANTHER PTHR23042), among others. The length of the protein is more similar to that of CYC, which is 413aa long, while CLK is 1023aa long (ceAHA-1A is 453aa long). This protein forms heterodimers with other PAS containing proteins, such as ceAHR-1, ceHIF-1 and ceCKY-1, and is also probably capable of forming homodimers, to then stimulate the transcription of its target genes [45,46]. This is also the case in flies, where CLK-CYC dimers activate the transcription of the negative factors and other clock controlled genes [15].

Strikingly, *C. elegans* shows no homologs to the Cryptochrome protein (CRY), involved in blue light photoreception and TIM degradation, nor to CRY2, a potent repressor of CLK-CYC mediated transcription, but insensitive to blue light [43]. It is, however, interesting to note that in recent years a novel family of photoreceptors, unknown to other species, has been described in *C. elegans*. This nematodes exhibit a negative phototactic behavior towards UV and blue light. This response is mediated, at least in part, by the LITE-1 photoreceptor, a G protein coupled receptor with similarity to the gustatory receptors of *Drosophila melanogaster*. This protein activates a guanylate cyclase, producing cGMP, upon activation by low wavelengths of light and, once cGMP levels reach a certain concentration, cGMP-sensitive CNG channels are opened and stimulation of photoreceptor cells occurs. [47].

*C. elegans* also expressed proteins with great similarity to those of the core circadian clock of mammals [48]. The list of the best hits can be found in Table 3.

The positive elements of the mammalian clock are the proteins Clock (CLK) and BMAL. BMAL is similar to dmCYC and there are two variants, BMAL1 and BMAL2. Besides CLK there is also another protein, NPAS2, capable of heterodimerizing with BMAL, and activating the transcription of target genes. The *C. elegans* protein with greatest similarity to CLK, BMAL1, BMAL2 and NPAS2 is ceAHA-1a, with E values of  $1.5 \times 10^{-17}$ ,  $1.8 \times 10^{-65}$ ,  $5.4 \times 10^{-63}$  and  $6.8 \times 10^{-19}$ , respectively. These results are similar to those found in the case of the insect clock, where the protein with best similarity to CLK and CYC was also ceAHA-1a.

The negative elements of the mammalian clock are the proteins Period (PER) and Cryptochrome (CRY). In this case, 3 PER proteins have been described in mammals: PER1, PER2 y PER3. Once again, as was the case with the insect PER protein, the similar to PER2 and PER3 is ceLIN-42B, with an E value of  $2.9 \times 10^{-21}$  and  $5.5 \times 10^{-20}$ , respectively. PER1, on the other hand, was most similar ( $E = 7.5 \times 10^{-17}$ ) to ceLIN-42C. There are no CRY similarities in *C. elegans*.

Even though the similarity scores are high, those found in the case of the insect core clock were higher (compare tables 2 and 3).

### *C. elegans* also conserves many insect and mammalian clock modulators/regulators

Since proteins with high degree of similarity to arthropod core clock proteins were found in *C. elegans*, we broadened the search to include known circadian accessory proteins of *Drosophila melanogaster* [42,49].

The proteins that form accessory loops were the first candidates to search for. These proteins are Vriille, Par Domain Protein 1E and Clockwork Orange.

Vriille (VRI) belongs to the 2nd negative feedback loop of the clock that inhibits CLK transcription. A similar protein, ceATF-2 (K08F2.2), with a score of  $E = 1.7 \times 10^{-17}$  was found. Domain analysis shows that both proteins possess a bZIP domain (PFAM PF07716). It is known that ceATF-2 is a transcription factor that negatively regulate the autophagy genes ceBEC-1 y ceLGG-1 [50], but it is unknown whether it regulates ceAHA-1 (similar to CLK, see table 2) or not.

Par Domain Protein 1e (PDP1e), is involved in the 2nd positive feedback loop of the fly's clock and fulfills an antagonistic role to that of VRI, promoting CLK expression. In this case, ceCES-2 is the most similar protein, with an  $E = 2 \times 10^{-6}$ . Both proteins possess bZIP (PFAM PF07716) domains, although their sizes are very different. While PDP1e is 647aa long, ceCES-2 is only 211 residues long. It is known that this protein regulates apoptosis in

the nematodes and that it is capable of forming dimers with ceATF-2 [51,52].

Clockwork Orange (CWO) is part of a third loop and acts as a transcriptional repressor of CLK target genes, thus modulating the amplitude of *Drosophila*'s circadian clock controlled genes' oscillations. No similarities were found in *C. elegans*.

Clock modulating proteins were then studied: CREB, DoubleTime (DBT), Casein kinase 2 (CK2), Shaggy (SGG), the protein phosphatases 1 and 2 (PP1 y PP2), Supernumerary Limbs (SLMB), Rhodopsin (Rh), Nemo (NMO), Nocte and phospholipase C (NORPA). The similarities and function of these proteins are summarized in table 4.

Proteins related to insect clock output were then studied. These include Takeout (TO), Pigment Dispersing Factor (PDF), Slowpoke (SLO), Narrow Abdomen (NA), Inward Rectifier (IR), Jetlag (Jet), COP9 signalosome (CSN4), Protein kinase A (PKA), Beta Alanil Conjugase Synthase (EBONY), Lark (LRK), Transcription factor FER2 (FER2). The similarities and functions of these proteins are summarized in table 4.

Even though the homology score was higher in the case of the insect core clock proteins, *C. elegans*' proteome was also searched for the mammalian accessory proteins [48].

The positive elements of the second loop are ROR-A and ROR-B. They act positively on BMAL transcription. The protein with higher similarity to ROR-A is ceNHR-23b (C01H6.5b), with  $E = 1.4 \times 10^{-70}$ ; and the similar to ROR-B is ceNHR-23a (C01H6.5a), with  $E = 4.5 \times 10^{-73}$ . In the nematodes this protein acts as a nuclear hormone receptor required for larval molting [53].

REB-ERV alpha (NR1D1 and NR1D2) acts as the negative element of the second loop, inhibiting BMAL transcription. The closest protein similar to NR1D1 and NR1D2 is ceNHR-85a (W05B5.3a), with E values of  $8.1 \times 10^{-47}$  and  $8.7 \times 10^{-55}$ , respectively. This protein is a nuclear hormone receptor involved in egg laying and SDS resistant dauer larva formation [54].

DBP (D-Box binding protein) is part of a second accessory loop, acting through binding to D-Box elements of target genes [55]. The similar to DBP is ceCES-2 (ZK909.4), with  $E = 1.5 \times 10^{-18}$ , which is also the homolog to dmPDP-1e.

We also studied clock-modulating proteins such as CASEIN KINASE 1A, CASEIN KINASE 1D, CASEIN KINASE 1E, DEC1, DEC2, FBXL3, NOCTURNIN (CCRN4L), PPARGC1A. The similarities and functions of these proteins are summarized in table 5.

We then looked for the proteins related to input and output of the clock, including melatonin receptors MEL1A and MEL1B, MELANOPSIN (OPN4), PROKINETICIN (PKC2), VIP and VIP receptor (VPAC2). The similarities and function of these proteins are summarized in table 5. It is interesting to note that previous work from our lab described the existence of melatonin synthesis rhythms and in the ASMT enzyme, which is the rate limiting enzyme of the process [56]. However, the receptors for this humoral signal remain undiscovered. Here we report several melatonin receptor candidates; the one with the highest degree of similarity to the mammalian receptor is ceF41E7.3. This protein is an orphan neuropeptide receptor and could certainly be tested experimentally to verify whether it plays a role in melatonin signaling.

### *C. elegans*' proteome exhibits high similarity to clock accessory proteins of insects and mammals

The previous results show that *C. elegans* expresses the proteins needed to build a clock similar to that of insects with the important exception of a protein similar to dmCRY, the protein responsible for TIM degradation upon blue light stimulation [42]. In the mammalian case, mCRY is part of the core of the clock itself [13].

In the case of clock accessory proteins, *C. elegans* exhibits similarities to form a complete secondary insect-like loop and also most of the insect modulatory and output proteins, with the exception of rhodopsins, dmNOCTE, dmTO y dmEBONY. However, no similarities to dmCWO were found, a protein involved in the third loop of the insect clock. We also found similarities to the proteins of the secondary loops and several of the accessory proteins of the mammalian clock. It is interesting to point out that there were no similarities to VIP, a neuropeptide involved in the synchronization of the mammalian clock neurons, although our methods indicated similarities to its receptor VPAC2. In insects, functions analog to VIP neuropeptide appear to be encoded by PDF. *C. elegans* present two PDF similarities and also a protein similar to a PDF receptor [57,58]. It is tempting to consider PDF as a putative output protein for the *C. elegans* circadian clock, a possibility that remains to be determined experimentally.

The insect and mammalian similarity tables (see tables 2–5) include several accessory proteins with highly elevated and significant E values. This can be explained by the fact that many of those proteins are ionic channels, phosphatases or kinases, all of which are very conserved throughout evolution [59–66] and hence

**Table 3.** Protein homologs to the core clock proteins of mammals.

Organism	Protein	Clock function	<i>C. elegans</i> homolog	E value	
<i>Mus musculus</i>	PER	PER1	central clock	F47F6.1c (LIN-42c)	7.50E-17
		PER2	central clock	F47F6.1b (LIN-42b)	2.90E-21
		PER3	central clock	F47F6.1b (LIN-42b)	5.50E-20
	BMAL	BMAL1	central clock	C25A1.11a (AHA-1a)	1.80E-65
		BMAL2	central clock	C25A1.11a (AHA-1a)	5.40E-63
	CLK	central clock	C25A1.11a (AHA-1a)	1.50E-17	
	NPAS2	central clock	C25A1.11a (AHA-1a)	6.80E-19	
	CRY	CRY1	central clock	None	none
		CRY2	central clock	None	none

These proteins are the best hits found by probabilistic inference search based on Hidden Markov Models built upon multiple protein alignments of the core clock of mammals (*M. musculus*). Common *C. elegans* names are enclosed between parenthesis when available.  
doi:10.1371/journal.pone.0112871.t003

**Table 4.** Homolog proteins to accessory proteins of the insect circadian clock.

Organism	Protein	Clock function	<i>C. elegans</i> homolog	E value
<i>Drosophila melanogaster</i>	VRI	2nd loop	K08F8.2 (ATF-2)	1.70E-17
	PDP1E	2nd loop	ZK909.4 (CES-2)	2.00E-06
	CWO	3rd loop	None	none
	CREB	Modulator	Y41C4A.4d (CRH-1d)	8.10E-43
	DBT	Modulator	C03C10.1 (KIN-19)	5.30E-190
	CK2A	Modulator	B0205.7 (KIN-3)	3.50E-209
	CK2B	Modulator	T01G9.6b (KIN-10)	1.50E-121
	SGG	Modulator	Y18D10A.5 (GSK-3)	3.30E-189
	PP1	PP1b	F29F11.6a (GSP-1a)	1.30E-216
		PP1-13C	F56C9.1 (GSP-2)	1.30E-218
	PP2	PP2a_MTS	Y75B8A.30 (PPH-4.1)	3.90E-177
		PP2a_TWS	F26E4.1 (SUR-6)	4.00E-213
		PP2a_WBT_A	W08G11.4 (PPTR-1)	4.30E-256
		PP2a_WBT_B	C13G3.3c (PPTR-2c)	9.50E-279
	PRMT-5	Modulator	C34E10.5 (PRMT-5)	3.70E-144
	SLMB	Modulator	K10B2.1 (LIN-23)	1.90E-232
	RH1	Modulator	<b>C52B11.3 (DOP-4)</b>	<b>8.60E-34</b>
	RH5	Modulator	<b>C52B11.3 (DOP-4)</b>	<b>4.00E-30</b>
	RH6	Modulator	<b>C52B11.3 (DOP-4)</b>	<b>8.40E-30</b>
	NMO	Modulator	W06F12.1e (LIT-1e)	1.50E-182
	NOCTE	Modulator	None	None
	NORPA	Modulator	B0348.4a (EGL8a)	0
	TO	Output	None	None
	PDF	Output	T07E3.6b (PDF-1)	3.00E-02
	PDFR	Output	C13B9.4a (PDFR-1a)	1.80E-127
	SLO	Output	Y51A2D.19a (SLO-1a)	0
	NA	Output	C11D2.6a (UNC-77a)	0
	IR	Output	M02A10.2a (IRK-2a)	8.50E-159
	JET	Output	C02F5.7a	8.80E-128
	CSN4	Output	Y55F3AM.15 (CSN-4)	7.80E-85
	PKA	Output	ZK909.2h (KIN-1h)	1.60E-217
	EBONY	Output	<b>Y66D12A.14</b>	<b>1.40E-60</b>
	LRK	Output	F18H3.3b (PAB-2b)	1.80E-36
	FER2	Output	B0304.1b (HLH-1b)	1.90E-12
	JAFRAC	Output	F09E5.15a (PRDX-2a)	5.10E-105
	WDS-PB	Output	C14B1.4 (WDR-5.1)	3.50E-165

These proteins are the best hits found by probabilistic inference search based on Hidden Markov Models built upon multiple protein alignments of the accessory proteins of the circadian clock of insects (*D. melanogaster*). Those proteins that did not possess the same domains that the query profile are highlighted (bold). Common *C. elegans* names are enclosed between parenthesis when available.

doi:10.1371/journal.pone.0112871.t004

high E values were expected. On the other hand, proteins that are more clock-specific, as is the case of PRMT-5 (*PROTEIN ARGININE METHYL TRANSFERASE 5*), bear more interest in this type of predictive studies. This protein, which is highly similar ( $E = 3.70E-144$ ) to the *C. elegans* protein C34E10.5, modulates the alternative splicing of clock genes in *D. melanogaster* and *A. thaliana*. Null mutants of this enzyme present severely affected rhythms in both species [67,68]. It would be interesting to study how this protein affects the circadian behavior of *C. elegans*.

### Common clock components in the phylum Nematoda and phylogenetic studies

Once that it was established that the proteome of *C. elegans* contained proteins with similarity to insect and mammalian clock proteins, we searched for homolog proteins in the complete proteomes of other nematodes: *Ascaris suum*, *Brugia malayi*, *Bursaphelenchus xylophilus*, *Caenorhabditis angaria*, *Caenorhabditis brenneri*, *Caenorhabditis briggsae*, *Caenorhabditis japonica*, *Caenorhabditis remanei*, *Caenorhabditis sp7*, *Caenorhabditis sp9*, *Caenorhabditis sp11* and *Pristionchus pacificus*. The results are

summarized in Tables S1 and S2. Thirteen mammalian and thirty two insect *C. elegans*' protein hits were also found in the other nematode proteomes that were analyzed. Interestingly there were 7 proteins common to mammals, insects and nematodes. These common proteins include two core clock components, the positive core clock element Cycle/BMAL and the negative element Period; and four modulators, the circadian clock kinase Dbt/CSNK4; the F-Box modulator Jetlag/FBXL; the transcriptional coactivator Lark/PPARGC1- $\alpha$ ; the peroxiredoxin Jafrac/PRDX-2 and the WDR5 histone modification proteins Wds/WD-Rep. The best conserved protein hits can be seen in table 6 and figure S2.

We also studied the evolutionary relationships between these clock proteins conserved among the three different phyla. Concatemers of these 7 proteins for different species of the three taxa were built from the respective protein alignments and a phylogenetic tree was constructed as described in the materials and methods section. A cytochrome B phylogenetic tree was also built for comparison purposes. As can be seen in figure 2 both trees show three clearly distinct clades, one belonging to each phylum.

There is a longstanding debate about *C. elegans*' relationship with other model organisms [69]. There are two hypotheses, the *coelomata* hypothesis indicates that nematode are closer to insects than to mammals; and the *ecdysozoa* hypothesis states that insects and mammals are closer to each other than either of them to *C. elegans*; indeed, there is evidence supporting both hypotheses. Studies based on 18S RNA sequence state that the nematodes are more closely related to animals that go through the process of

ecdysis. This hypothesis, however, depends on the exclusion of all nematodes with the exception of *Trichinella spiralis*. If the rest of the nematodes are included, then they group in a clade consistent with the *coelomata* hypothesis [70]. RNApol 2 sequence-based trees also support this hypothesis [71]. As shown in figure 2, our results also favor this hypothesis, suggesting that *D. melanogaster* and *M. musculus* are closer to each other, than either of them to *C. elegans*.

The analysis of trees based on a single core clock protein indicates the same kind of grouping (figure 3). One interesting case is that of the TIMELESS homolog. ceTIM-1 is closer to dmTIMEOUT and to mmTIMELESS, that to dmTIMELESS (figure S1). In mammals, this protein is essential for embryonic development and is associated to DNA metabolism. It is also known to associate with peroxiredoxin 2 during cell cycle check points [72]. In *Drosophila*, TIMEOUT is also involved in DNA metabolism, chromosomal cohesion and circadian photoreception [73–75]. This means that *C. elegans*' similarity to TIMELESS is closer to the proteins that perform developmental roles rather than to the protein involved in circadian rhythms; indeed, there is experimental evidence to support ceTIM-1's role in chromosomal cohesion and developmental timing [44,76,77].

#### Circadian *Cis* acting promoter elements analysis

In parallel to the proteins (*trans* acting factors) of the clock, *cis* acting promoter elements modulate the phase of the rhythms found at the mRNA level in other model organisms [78–80]. We searched for the most well-known *cis* elements, including the

**Table 5.** Homolog protein to accessory proteins of the mammalian circadian clock.

Organism	Protein		Clock function	<i>C. elegans</i> homolog	E value
	ROR	ROR-A	2nd loop	C01H6.5b (NHR-23b)	1.40E-70
<i>Mus musculus</i>	REB-ERVa	ROR-B	2nd loop	C01H6.5a (NHR-23a)	4.50E-73
		NR1D1	2nd loop	W05B5.3a (NHR-85a)	8.10E-47
		NR1D2	2nd loop	W05B5.3a (NHR-85a)	8.70E-55
	DBP		3rd loop	ZK909.4 (CES-2)	1.50E-18
	CSNK	CSNK1a	Modulator	C03C10.1 (KIN-19)	4.30E-204
		CSNK1d	Modulator	F46F2.2b (KIN-20b)	1.40E-172
		CSNK1e	Modulator	F46F2.2b (KIN-20b)	5.00E-172
	CCRN4L		Modulator	ZC518.3c (CCR-4c)	3.50E-21
	DEC1		Modulator	<b>Y54G2A.1 (LIN-22)</b>	<b>4.20E-05</b>
	DEC2		Modulator	<b>Y54G2A.1 (LIN-22)</b>	<b>1.50E-05</b>
	FBXL-3		Modulator	C02F5.7a	5.20E-04
	PPARGC1a		Modulator	<b>F18H3.3b (PAB-2b)</b>	<b>6.10E-08</b>
	MEL receptor	MEL1A	Output	F41E7.3 (NPR-6)	1.00E-26
		MEL1B	Output	F41E7.3 (NPR-6)	1.10E-31
	OPN4		Input	<b>T11B7.4e (ALP-1e)</b>	<b>5.20E-61</b>
	PKC2		Output	None	None
	VIP		Output	None	None
	VPAC2		Output	C18B12.2 (SEB-3)	1.60E-37
	PRDX-2		Output	F09E5.15a (PRDX-2a)	1.70E-104
	WD-REP		Output	C14B1.4 (WDR-5.1)	2.70E-159

These proteins are the best hits found by probabilistic inference search based on Hidden Markov Models built upon multiple protein alignments of the accessory proteins of the circadian clock of mammals (*M. musculus*). Those proteins that did not possess the same domains that the query profile are highlighted (bold). Common *C. elegans* names are enclosed between parenthesis when available.

doi:10.1371/journal.pone.0112871.t005

canonic E-Box element, in the promoters of all the genes that codify for proteins similar to insect/mammalian proteins (group 1), insect proteins (group 2) and mammalian proteins (group 3). The analysis revealed several circadian related *cis* acting promoter elements can be found in *C. elegans*.

The morning E-Box element was found to be represented in the promoters of all three groups. This element is bound by bHLH domain containing proteins, such as Cycle/Bmal, and was the first promoter element associated with circadian transcription [55,81]. This element is found in the promoters of the *cycle/bmal*, *lark/pparg1- $\alpha$* , *wds/wd-r-2* and *dbt/csnk4* group 1 gene homologs. Strikingly it is not found in the promoter of the period homolog *lin-42*, although a non-canonic E-Box element is present (data not shown).

The daytime canonic D-Box element was also found in the promoter elements of some of the *C. elegans*' gene homologs (groups 1 and 2). This element is bound by DBP in a repressor-antiphase-to-activator mechanism [55]. It was found in the promoter of the *cycle/bmal* of the group 1 genes homologs.

The nighttime canonic RRE element is bound by RevErbA/ROR binding elements in a repressor-precedes-activator pattern (as is the case for the E-Box element) [55,80]. Interestingly, this element was found in the promoters of the *nhr-85*, the RevErbA homolog, and in the peroxiredoxin-2 gene homolog.

The complete set of results, which includes other circadian expression associated promoter elements (GRE, PPRE, CREB, HSE and CEBP) can be found in Table S3.

## Discussion

The bioinformatics analysis approach used in this *C. elegans* work based on homolog sequence searches from the model organisms *Mus musculus*, *Drosophila melanogaster*, *Neurospora crassa*, *Arabidopsis thaliana* and *Synechococcus elongatus* yielded some novel and unexpected results that differ from simple BlastP search strategies (Table S4). On the one hand, *C. elegans* expresses similar proteins to the ones from the core clock of three of the studied model organisms: *M. musculus*, *D. melanogaster* and *N. crassa*. The high homology to mammal and insect core clock proteins is not surprising if we take into account previous reports and results from our own lab. In fact, most of the studies performed on recent years were focused on the role of these components in the *C. elegans* clock. However, the existence of high similarity to the clock proteins of the fungi *N. crassa* is much more unexpected and has not been reported so far to our knowledge. In particular, the perfect similitude between ncFRH and its possible *C. elegans* counterpart, ceMTR-4, is remarkable. These proteins not only share sequence similitude but also exhibit a similar size (1106aa and 1026aa, respectively) and present almost the same protein domains with just one exception (calcium binding domain, EF-Hand 1). It is also notable that ncFWD-1, which is involved in ncFRQ degradation, exhibited strong sequence similitude to ceLIN-23, which has very similar molecular functions already described in the nematode. Nevertheless, the lack of a protein similar to ncFRQ would hinder the function of a *C. elegans* clock based on the fungal TTFL model. Regarding the rest of the central clock components of *N. crassa*, we must also point out that the WC complex proteins (ncWC1 and ncWC2), that act as positive elements, do not appear to have proteins with strong similitude in *C. elegans*. The same is true in the case of ncVIVID. These results suggest that only part of an ancestral clock is conserved in nematodes and some of the elements could have been substituted by some new, yet unknown, elements, thus closing the regulatory loop. Even though the similitude is not perfect and the

functions described in *C. elegans* for these proteins is different to those expected, they could also be part of the nematode clock by performing additional functions that have not been discovered yet. Even though there are no experimental records to support this hypothesis, the possibility of similitude between *C. elegans*' clock components and those of *N. crassa*'s clock has not been examined. As a first approximation, none of the *C. elegans*' proteins similar to *N. crassa*'s core clock proteins appear to cycle at the mRNA level in the different conditions studied according to the published microarray study [24].

Besides the high similitude to the *N. crassa*'s clock, the results obtained after comparing the proteins that make up the insect and mammalian clock were more conclusive. *C. elegans* exhibits not only the required proteins to build the core clock but also some accessory proteins required for circadian rhythms in both the insect and the mammalian molecular clock. Among the sets of core proteins, ceLIN-42b and ceLIN-42c were of great similarity to PER and ceTIM-1 could be the equivalent of the insect TIM clock protein. Also, ceAHA-1a is the sole protein with similarity to the positive elements of the clock (dmCLK/CYC y mmCLK/BMAL) and includes all of the required domains for the circadian functions of these proteins. It must be pointed out, however, that these homolog proteins exert heterochronic and chromosomal cohesion roles in the nematodes [20,21,44] and their function in the circadian clock of *C. elegans* is not completely clear.

*lin-42* encodes three different isoforms, LIN-42a, LIN-42b and LIN-42c. The b and c isoforms contain PAS domains such as those found in insect and mammalian PER. The experimental evidence suggests that *lin-42* is involved in larval development controlling the multiple changes that occur in different tissues (hypodermis, gonads, sex myoblasts and vulva) [21,22], in molting synchronization [82] and in circadian locomotor activity period determination [3]. It is also known that *lin-42* gene expression cycles throughout development [21,83] and that LIN-42a in particular, whose mRNA peaks during the four larval moltings with a close to 8-h period [82], is required to generate quick, rhythmic and productive molts. In line with this, the forced and constitutive expression of *lin-42a* in wild type nematodes generates anachronistic larval molts and lethargy [82]. The molting defects observed in the *lin-42(ok2385)* (which bears a deletion affecting isoforms a and b) and *lin-42(n1089)* (which bears a deletion affecting isoforms b and c) mutants are completely rescued by *lin-42a* [82]. In addition, the defects in the temporal program of cellular division observed in nematodes carrying the *n1089*, *mg152* and *ve11* mutant alleles can also be fully reversed by the downstream portion of the *lin-42* locus, that codifies for the LIN-42a isoform [22]. However, it must be noted that the three *lin-42* isoforms may interact and influence the pace of larval development. In this sense, the overexpression of LIN-42c causes a dramatically reduction in the pace of development in *lin-42(ok2385)* larvae, suggesting an antagonistic role to that of LIN-42a. Also, the a and b isoforms share SYQ and LT domains, potentially performing redundant roles [82]. Taken together, these results suggest that LIN-42a would be the isoform mainly involved in heterochronic functions, leading to speculate that the PAS-containing b and c isoform may fulfill other functions, one of which could be the regulation of circadian rhythms in the nematodes. In accordance to this hypothesis, adult nematodes that lack PAS containing LIN-42 isoforms exhibit a longer locomotor activity circadian period than wild type [3].

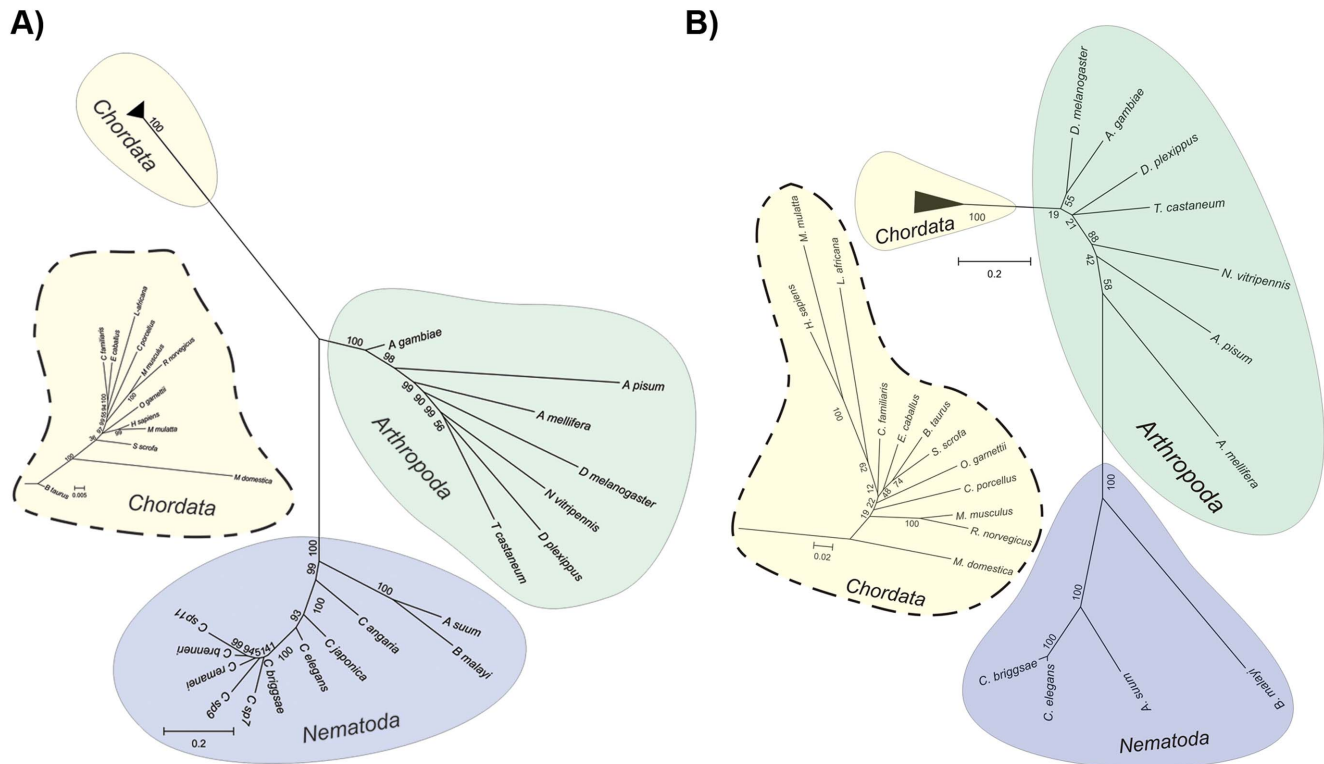
ceTIM-1, another possible negative element found in *C. elegans*, has been implied to be only involved in heterochronic functions. RNAi-mediated studies to knock down ceTIM-1 and ceKIN-20 (homolog to dmDBT and mmCSNK4) indicated that



**Table 6.** Common elements of the insect and mammalian clock also found in nematodes.

Organism	Protein	FBXL	PPARGC1a	PER	WD-Rep	PRDX-2
<i>M. musculus</i>	BMAL1					
<i>D. melanogaster</i>	Cyc	Jetlag	Lark	Period	Wds-PB	Jafrac
<i>C. elegans</i>	C25A1.11a	C02F5.7a	F18H3.3b	F47F6.1c	C14B1.4	F09E5.15a
<i>A. suum</i>	GS_18509	GS_19076	GS_22069	GS_08634	GS_18544	GS_23521
<i>B. malayi</i>	BM21318	BM06465	BM03094	BM06378	BM03793	BM19146
<i>B. xylophilus</i>	BUX.s00397.112	BUX.s00422.570	BUX.s00579.708	BUX.s00397.112	BUX.s00579.553	BUX.s01109.415
<i>C. angaria</i>	CAN07933	CAN15035	CAN13338	CAN12189	CAN22866	CAN05705
<i>C. brenneri</i>	CBN21788	CBN03873	CBN18184	CBN20940	CBN24532	CBN11882
<i>C. briggsae</i>	CBG12219	CBG20206	CBG07431	CBG07211	CBG09206	CBG25150
<i>C. japonica</i>	CJA09659	CJA17198	CJA16600	CJA16851	CJA10837	CJA14459
<i>C. remanei</i>	CRE23963	CRE13355	CRE25417	CRE17943	CRE25440	CRE1768
<i>C. sp7</i>	g19865	g29573	g22800	g31054	g26350	g24613
<i>C. sp9</i>	g12945	g44344	g29591	g23334	g4010	g38213
<i>C. sp11</i>	g9507	g14251	g435	g4740	g7816	g9373
<i>P. pacificus</i>	PPA09278	PPA25500	PPA20677	PPA04587	PPA00309	PPA28346

These proteins are the best hits found by probabilistic inference search based on Hidden Markov Models built upon multiple protein alignments of the circadian clock of mammals (*M. musculus*) and insects (*D. melanogaster*). Those proteins that did not possess the same domains that the query profile are highlighted in red.  
doi:10.1371/journal.pone.0112871.t006



**Figure 2. Phylogenetic analysis.** A) Phylogenetic tree of the conserved clock proteins of insects, mammals and nematodes. B) Phylogenetic tree based on Cytochrome B protein sequences of insects, mammals and nematodes. The insets show a zoomed in view of the Chordata phylum branches.

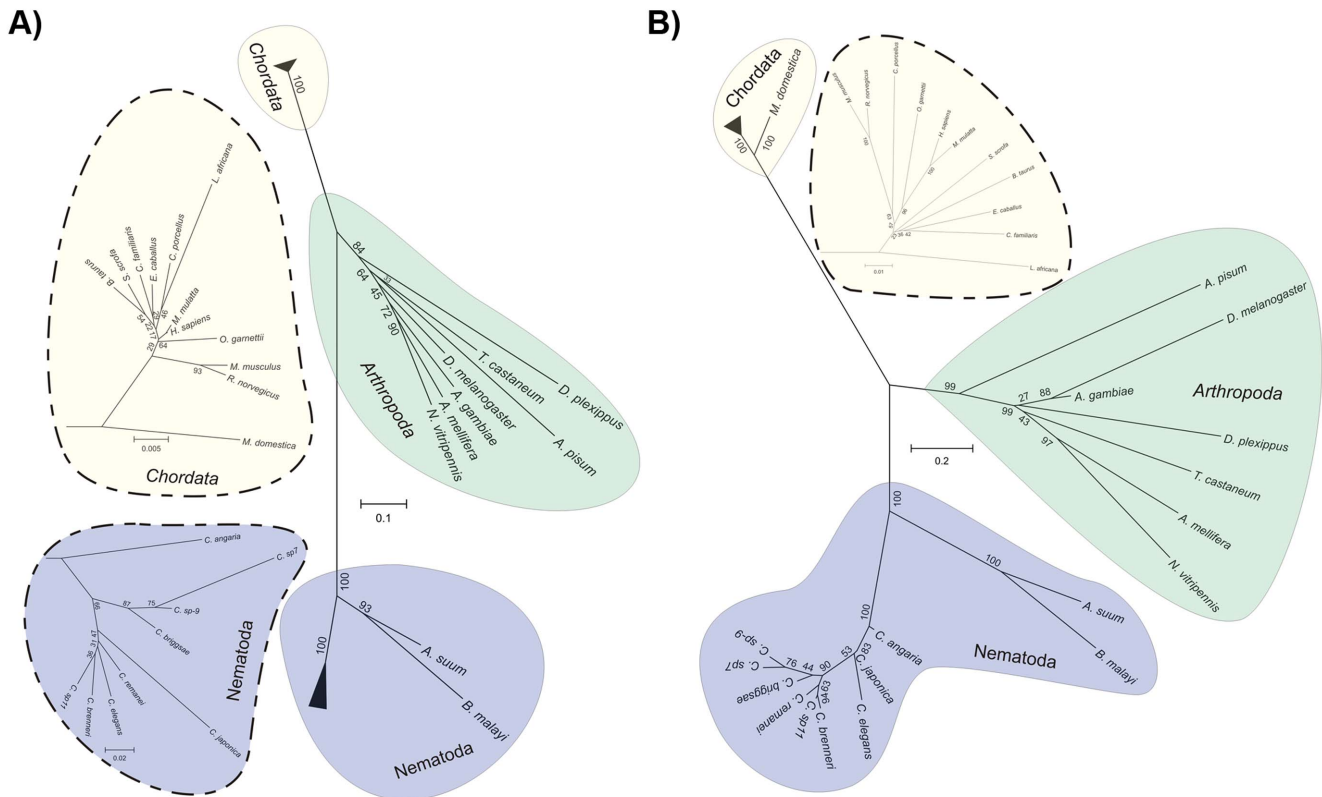
doi:10.1371/journal.pone.0112871.g002

both elements are involved in the seam cell terminal differentiation program during the L4 stage [76]. Also, ceTIM-1 is essential for chromosomal cohesion regulation and is involved in chromosomal location determination of the non-SMZ cohesion subunits [44]. It is interesting to point out that in insects a paralog gene to *timeless*, *timeout*, has been described. *timeless* is a canonic gene of the circadian clock of *D. melanogaster* and other insects, while *timeout* is a multiple-function gene and plays a key role in the maintenance of chromosome integrity, light entrainment of the circadian clock, embryonic development and DNA replication regulation. Studies have shown that in many insects *timeout* circadian expression is light-dependent and also particularly necessary for circadian photoreception in fruit flies [73,84]. In mammals, there is only one *tim* gene, called *timeless*, that is more similar in sequence to insect *timeout* rather than to insect *timeless* (figure S1). This gene is involved in replication termination and cell cycle progression [85–87] and has recently been shown to participate in period determination and circadian rhythmicity [88,89]. It is also known that its expression is rhythmic in the retina [90]. There is also evidence that hymenopteran insects, including ants, bees and wasps, do not have a *timeless* gene and still have a functional circadian clock [84,91]. In *C. elegans*, as is also the case in mammals, there is only one *tim* gene (ceTIM-1) that is closer in sequence to *D. melanogaster timeout* (figure S1). It could be speculated that ceTIM-1 might also fulfill multiple roles in *C. elegans* as in the mentioned examples, even though the function of this protein in the inner workings of the circadian clock has yet to be found.

On the other hand, the absence in *C. elegans* of proteins similar to the Cryptochromes (CRY) of both insects and mammals is striking. Nevertheless, and in accordance to the absence of a CRY

protein, *C. elegans* also completely lacks photolyases, enzymes involved in DNA repair from which cryptochromes evolved [43,92]. As mentioned earlier, the photoreceptor function of CRY in *D. melanogaster* could have been replaced by the convergent evolution of the novel photoreceptor family discovered in *C. elegans*, composed by LITE-1 and GUR-3 [47,93,94]. Even though there is no direct evidence of the role of these proteins in the photic synchronization of the clock, the CNG channel protein TAX-2, which acts downstream of these photoreceptors, is required for the rhythmic expression of diverse transcripts entrained under light: dark 12 h: 12 h cycles [24]. The absence of CRY in the genome of *C. elegans* could mean that the negative transcriptional loop elements conformed by PER/CRY in mammals could have been replaced in the nematode by a PER/TIM dimer, or LIN-42/TIM-1. On the other hand, given the fact that there is no evidence of interaction between these two proteins, the existence of a novel negative element yet to be described cannot be ruled out. In this sense it is worth noting that even though in most insect clocks the PER/TIM dimer fulfills the repressor role, in other insects it is a PER/CRY dimer that performs the same role, even if they also have a TIM protein [43], which infers a great plasticity in the molecular mechanism of the circadian clock.

The most similar protein to the positive elements of the insect and mammalian clocks is ceAHA-1. This protein belongs to the bHLH/PAS protein family, which have the property of binding as an homo or heterodimer to DNA consensus sequences known as an E-Box. There is experimental evidence showing that ceAHA-1 is capable of physical interaction with at least three proteins by means of their PAS domains (ceCKY-1, ceHIF-1 and ceAHR-1) forming heterodimers that translocate to the nucleus and act as



**Figure 3. Core clock proteins phylogenetic analysis.** A) Phylogenetic tree of the conserved proteins mMBAL, dmCYCLE and its homolog ceAHA-1. B) Phylogenetic tree of the conserved PERIOD proteins. A zoomed in view of the branches corresponding to the Chordata (A and B) and Nematoda (A) phylum are shown in the insets. doi:10.1371/journal.pone.0112871.g003

transcription factors. The ceAHA-1/ceAHR-1 dimer binds the TTGCGTG sequence and regulates a subset of genes involved in neuronal development [23,95]; the ceAHA-1/ceCKY-1 dimer regulates gene expression in the pharynx by binding to the TGCCTG sequence [95,96]; finally, the ceAHA-1/ceHIF-1 dimer binds to the sequence CACGTA, to modulate the transcription of genes involved in iron homeostasis [97] and to the TACGTG sequence, to modulate the response to hypoxia [41,95]. There is currently no evidence that links ceAHA-1 to the circadian clock of *C. elegans*; however, given the fact that this protein has bHLH DNA binding domains and PAS protein interaction domains, it would not be surprising if ceAHA-1, acting as a homodimer or as a heterodimer in association with another protein, could play the role of the mammalian CLK/BMAL or insect CLK/CYC heterodimers. Several ceAHA-1-containing dimers regulating the transcription of genes involved in different functions in the nematode have been reported; moreover, the bioinformatics analysis found ceAHA-1a as the most similar protein to the known positive elements. However, there is no experimental evidence that corroborates the existence of ceAHA-1 homodimers in *C. elegans*. In this regard, more experiments will be necessary to study the possible role of ceAHA-1 in the molecular basis of the *C. elegans* circadian clock.

Recently, a global transcriptomic study of nematodes kept under light:dark and temperature (warm:cold) 24 hour cycles did not find significant cycles in the mRNA of these possible core clock genes homologs [24]. It is important to know that this study was performed using microarrays and hence it is not possible to discriminate between the different isoforms of *lin-42* and *aha-1*,

among others. Also, the study was performed analyzing the gene expression levels of whole nematode populations in different timepoints, which could have generated a high level of noise and hence, loss of detection sensibility and the masking of positive results due to the contribution of non-rhythmic tissues. This could have negatively affected the detection of core clock homologs as it is known that in all eukaryotes studied so far, the expression of the core clock genes is rhythmic in certain cells and tissues (and may even change their respective phases in different tissues) and suffer quick dampening in the absence of environmental cues. One such example is the *period* gene of *D. melanogaster*. *period* rhythms were initially described in the heads of adult flies and even though the mRNA in the rest of the fly tissues also shows circadian fluctuations in *per* (with similar phase and amplitude), rhythms from whole body RNA extractions are much noisier. In particular, the *per* rhythm is absent in ovaries, which represents around 70% of the total RNA of female flies. The non-rhythmic *per* mRNA of the ovary represents 26% of the total *per* mRNA and generates low overall *per* rhythms in studies performed with whole female or whole mixed sex flies. Also, *per* rhythms are far more robust in the head under constant conditions than in the rest of the body, and it is a fact that the amplitude of *per* oscillations decreases well over 70% in peripheral clocks after 3 days in constant conditions, whereas they only diminish 30% in head neurons [98,99]. This observation could be valid for some of the genes analyzed in Van der Linden et al [24]. For example, *aha-1* and *atf-2* were scored as cycling under LD conditions (pF24<0.02) but were not robust enough in constant conditions. However, it is also possible that the

**Table 7.** Organisms used in this work.

<b>Nematoda</b>	<b>Tax. ID</b>	<b>Chordata</b>	<b>Tax. ID</b>	<b>Arthropoda</b>	<b>Tax. ID</b>
<i>Caenorhabditis elegans</i>	6239	<i>Mus musculus</i>	10090	<i>Drosophila melanogaster</i>	7227
<i>Ascaris suum</i>	6253	<i>Bos Taurus</i>	9913	<i>Acyrtosiphon pisum</i>	7029
<i>Brugia malayi</i>	6279	<i>Canis lupus familiaris</i>	9615	<i>Anopheles gambiae</i>	7165
<i>Bursaphelenchus xylophilus</i>	6326	<i>Cavia porcellus</i>	10141	<i>Apis mellifera</i>	7460
<i>Caenorhabditis angaria</i>	860376	<i>Equus caballus</i>	9796	<i>Danaus plexippus</i>	13037
<i>Caenorhabditis brenneri</i>	135651	<i>Homo sapiens</i>	9606	<i>Nasonia vitripennis</i>	7425
<i>Caenorhabditis briggsae</i>	6238	<i>Loxodonta africana</i>	9785	<i>Tribolium castaneum</i>	7070
<i>Caenorhabditis japonica</i>	281687	<i>Macaca mulatta</i>	9544		
<i>Caenorhabditis remanei</i>	31234	<i>Monodelphis domestica</i>	13616		
<i>Caenorhabditis sp11.</i>	886184	<i>Ontolemur garnettii</i>	30611		
<i>Caenorhabditis sp7.</i>	870436	<i>Rattus norvegicus</i>	10116		
<i>Caenorhabditis sp9.</i>	870437	<i>Sus scrofa</i>	9823		
<i>Pristionchus pacificus</i>	54126				
<b>Cyanobacteria</b>	<b>Tax. ID</b>	<b>Ascomycota</b>	<b>Tax. ID</b>	<b>Tracheophyta</b>	<b>Tax. ID</b>
<i>Synechococcus elongatus</i>	32046	<i>Neurospora crassa</i>	5141	<i>Arabidopsis thaliana</i>	3702
<i>Acaryochloris marina</i>	155978	<i>Ajellomyces dermatitidis</i>	5039	<i>Brachypodium distachyon</i>	15368
<i>Arthrospira maxima</i>	129910	<i>Arthrotrichy oligospora</i>	13349	<i>Brassica rapa</i>	3711
<i>Arthrospira platensis</i>	118562	<i>Aspergillus terreus</i>	33178	<i>Castanea sativa</i>	21020
<i>Crocospaera watsonii</i>	263511	<i>Beauveria bassiana</i>	176275	<i>Glycine max</i>	3847
<i>Cyanothece sp.</i>	43988	<i>Chaetomium thermophilum</i>	209285	<i>Hordeum vulgare</i>	4513
<i>Leptolyngbya boryana</i>	1184	<i>Claviceps purpurea</i>	5111	<i>Ipomoea nil</i>	35883
<i>Lyngbya aestuarii</i>	118322	<i>Colletotrichum higginsianum</i>	80884	<i>Lemna gibba</i>	4470
<i>Microcoleus chthonoplastes</i>	64178	<i>Cordyceps militaris</i>	73501	<i>Medicago truncatula</i>	3880
<i>Microcystis sp.</i>	1127	<i>Exophiala dermatitidis</i>	5970	<i>Mesembryanthemum crystallinum</i>	3544
<i>Moorea producens</i>	1155739	<i>Fusarium oxysporum</i>	5507	<i>Oryza sativa Japonica</i>	39947
<i>Oscillatoria sp.</i>	1159	<i>Gibberella zeae</i>	5518	<i>Phaseolus vulgaris</i>	3885
<i>Planktothrix rubescens</i>	59512	<i>Glomerella graminicola</i>	31870	<i>Populus nigra</i>	3691
		<i>Leptosphaeria maculans</i>	5022	<i>Ricinus communis</i>	3988
		<i>Metarhizium acridum</i>	92637	<i>Thellungiella halophila</i>	98038
		<i>Myceliophthora thermophila</i>	78579	<i>Triticum aestivum</i>	4565
		<i>Nectria haematococca</i>	140110	<i>Vitis vinifera</i>	29760
		<i>Paracoccidioides brasiliensis</i>	121759	<i>Zea mays</i>	4577
		<i>Podospira anserina</i>	5145		
		<i>Sordaria macrospora</i>	5147		
		<i>Talaromyces stipitatus</i>	28564		
		<i>Thielavia terrestris</i>	35720		
		<i>Trichoderma reesei</i>	51453		
		<i>Verticillium dahliae</i>	27337		

List of organisms from the six different phylum used in this work and their respective taxonomy IDs.  
doi:10.1371/journal.pone.0112871.t007

possible homologs to core clock genes do not have a circadian expression and *C. elegans* may have a novel circadian clock.

Every approach pursued so far to try to discover the molecular basis that dictates the workings of the *C. elegans* clock were based on the widely studied and well understood classic models of *Drosophila melanogaster* and *Mus musculus*. But even if all clocks described to date are based in a transcriptional translational feedback loop (TTFL), with the exception of the aforementioned redox driven oscillators, the proteins that make up the mechanism of these clocks are not conserved in all the studied model organisms. One example is the filamentous fungi *Neurospora*

*crassa* that has served for nearly half a century as a durable model organism for uncovering the basic circadian physiology and molecular biology. In the molecular clock of this organism, the proteins that serve as the core clock of the TTFL, WC, FRQ, FRH and VIVIV, are not conserved in insects or mammals. The molecular clock of *Neurospora crassa* took years of exhaustive research to be discovered. This was done through genetic screens of mutant phenotypes and homology based sequence searches that allowed, for example, the identification of shared PAS domains between WC1/2 and PER/CLOCK.

**Table 8.** List of circadian clock related proteins from the five model organisms used in this work.

Organism	Proteins
<i>Mus musculus</i>	PER1 - PER2 - PER3 - BMAL1 - BMAL2 - CLK - NPAS2 - CRY1 - CRY2 - RORA - RORB - NR1D1 - NR1D2 - DBP CSNK1a - CSNK1d - CSNK1e - CCRN4L - DEC-1 -DEC-2 FBXL-3 -PPARGC1a - MEL1A - MEL1B - OPN4 - PKC2 VIP - VPAC2 - PRDX-2 - WD-REP
<i>Drosophila melanogaster</i>	PER - TIM - CLK - CYC - CRY - VRI - PDP1E - CWO - CREB - DBT - CK2A - CK2B - SGG - PP1b - PP1-13C PP2a_MTS - PP2a_TWS - PP2a_WBT_A - PP2a_WBT_B PRMT-5 - SLMB - RH1 - RH5 - RH6 - NMO - NOCTE - NORPA - TO - PDF - PDFR - SLO - NA - IR - JET - CSN4 - PKA - EBONY - LRK - FER2 - JAFRAC - WDS
<i>Arabidopsis thaliana</i>	CCA1 - COP1 - DET1 - GI - LHY - LUX - PRR5 - PRR7 - PRR9 - TOC1
<i>Neurospora crassa</i>	FRH - FRQ - FWD-1 - VVD - WC-1 - WC-2
<i>Synechococcus elongatus</i>	KAI-A - KAI-B - KAI-C

doi:10.1371/journal.pone.0112871.t008

On the other hand, the lack of robust circadian rhythms in the expression of the principal candidate homolog genes does not mean that the gene products could not be rhythmic through posttranscriptional regulation mechanisms. In this sense, there are cases of circadian rhythms that respond to the PTFL (post-translational feedback loop) model, that are sustained even in the absence of transcription and translation. One such example is seen in the cyanobacteria *Synechococcus elongatus*, where one of the clock components goes through rhythmic cycles of phosphorylation [100]. Another example of a PTFL mechanism that has gained momentum is based on the superoxidation and reduction of peroxirredoxin (PRX), a family of proteins involved in preventing damage from reactive oxygen species [101]. PRXs are highly conserved in a great number of organisms and though this mechanism was first discovered in mammalian red blood cells and in the unicellular green alga *Ostreococcus tauri* [10,11], there is evidence that the rhythm in the oxidative state of these proteins is conserved in all the model organisms from cyanobacteria to plants, fungi, fruit flies and mice. Hence, PRXs have been proposed as universal markers for circadian rhythms [12]. Superoxidation of PRXs is dependent on the redox state of the cell and was shown to be rhythmic even in the complete absence of transcription and of some clock genes necessary to TTFL function. The fact that oxidation rhythms of PRXs have been found in organisms that possess functional TTFLs means that these mechanisms are not mutually exclusive, and to date it is not clear whether this system influences or regulates known TTFL-based clocks. Moreover, it is also unclear whether the oscillation in the oxidative state of the peroxirredoxin is itself part of the PTFL-based clock, or rather an output of another oscillator yet to be found. Recently, an oxidation rhythm of peroxirredoxin-2 (PRDX-2) has been found in *C. elegans* kept under constant darkness and temperature. However, *prdx-2* mRNA is not rhythmic. This oxidation rhythm represents a phylogenetically conserved molecular marker of the circadian clock in *C. elegans* [102] and could be part of a still uncharacterized PTFL-based clock. This does not exclude the possibility of *C. elegans* having a classic TTFL clock and the bioinformatics evidence presented in this work supports the existence of such clock and also of its phylogenetic similarity to the mammalian and insect clock.

Many studies have experimentally shown the existence of circadian rhythms in *C. elegans* at the molecular, physiological and behavioral level. Taken together, the data is indicative of the existence of a biological clock in *C. elegans* that responds to environmental signals (light and temperature cues), but so far the efforts have lacked an integrative analysis to allow for the establishment of the mechanism behind the observations, and the formulation of a model that explains how the circadian clock of

*C. elegans* works. This could be, in part, due to our own human nature and our strictly scientific way of classifying and contrasting new findings with previous theories. We must take careful consideration in *C. elegans* unique characteristics, such as its lack of eyes, soil-dwelling life, low or no contact with sunlight, which make it very different from all the other model organisms studied so far. Indeed, previous studies and the current bioinformatics analysis should be complemented with experimental data from more robust phenotype screenings or the use of high sensitivity reporters that allow for real time and *in vivo* recording of all the candidate genes found.

## Materials and Methods

### Protein databases

In this work we used the genomes and proteomes deposited in WormBase WS230 (table 7). These include *Caenorhabditis elegans* and 8 other members of the *Caenorhabditis* genus, and 2 other species of the genera *Ascaris* and *Brugia*. The proteomes of all these nematodes were extracted from their corresponding WormBase files (<http://www.wormbase.org/pub>). We then created an individual proteome database (IPD) with 11 individual proteomic databases containing the annotated proteome of each nematode. All of them were indexed using the formatdb routine from the NCBI Blast standalone suite v2.2.28.

### Construction of Hidden Markov Models

In order to find remote similarity between model organisms and worm's clock proteins, the proteins involved in the circadian clock of the 5 better characterized chronobiology model organisms, *Drosophila melanogaster*, *Mus musculus*, *Synechococcus elongatus*, *Neurospora crassa* and *Arabidopsis thaliana* (table 8), were considered. Ortholog proteins to each of the core clock proteins of each model organisms were searched using the BlastP service of NCBI (<http://blast.ncbi.nlm.nih.gov/>), using the default cut off E-value. This search was restricted to the following phyla: 12 organisms with completely sequenced genomes belonging to *Chordata*, 7 for *Arthropoda*, 13 for Cyanobacteria, 24 for *Ascomycota*, and 18 for Tracheophyta (table 7). The set of positive hits was then aligned using the Muscle routine of MEGA v5.2 suite with default parameters [59]. Then, specific Hidden Markov Model profiles (HMM profiles) were generated for each alignment using the Hmmbuild routine from the HMMER3 software [28].

### Search of the clock proteins in *Caenorhabditis elegans*

The HMM profiles described above were used as input to search for similar proteins in *Caenorhabditis elegans* using the HMMsearch routine from the same software. This resulted in 5

databases (one for each *Phylum*) containing the best hits for each HMM profile. Each protein of the collection of best hits resulted by the HMMsearch routine in *Caenorhabditis elegans* was further analyzed by InterProScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>). Those proteins that contained the same domains as the query protein from each model organism were kept as *accepted hits* [103]. The default HMMsearch 3.0 inclusion threshold value for the full sequence length was used (0.01).

### Search of the clock orthologous proteins in the others worms

The *accepted hits* from *Caenorhabditis elegans* were then used as queries to perform local homology searches (BlastP software) against the IPDs of the other members of the *phylum Nematoda*. The default standalone blastP cut off E-value was used ( $1E^{-10}$ ). This resulted in a collection of the best hits for each clock protein from all the nematode proteomes. This list was again further refined by comparing the domains found in each protein and that of the accepted hit in *C. elegans* to obtain a list of those proteins conserved in the *phylum*. Finally the best hit for each worm was used as query against *Caenorhabditis elegans* IPD in order to find the reciprocity with an *ad hoc* reciprocal best hit (RBH) routine based in BlastP. The default standalone blastP cut off E-value was used ( $1E^{-10}$ ).

### Analysis of genetic circadian regulatory elements

All *C. elegans*' orthologs that corresponded to accepted clock prototypes were analyzed at the promoter level to search for known circadian regulatory elements with the next IUPAC syntax: E-box (CACGTG), D-box (TTATGYAA); RRE (WAWNTRGGTCA), GRE (ACANNNTGTTCT), PPRE (TGACCY), CREB (TGACGTMA), HSE (NGAANN-GAANNNTTCN), and CBP (TKGNGAAK) [78–80,104]. Genomic sequences 3000 bp upstream of the ATG (translation start site) were downloaded for each putative clock gene with de WormMart tool [105]. Regulatory elements analysis was then performed with the jPREdictor v1.with default parameters [106].

### Phylogenetic analysis

A phylogenetic tree of the *phylum Nematoda* was built using Cytochrome B sequences, as previously described [107]. The sets of orthologous proteins were aligned by the Muscle routine from MEGA v5.2suite. The phylogenetic trees were constructed using the *neighbor-joining* method, the *Poisson* model for amino acid substitutions, a *Pairwise Deletion* for the *Gaps/Missing Data Treatment* and a *Gamma distributed* rate among sites was calculated for each alignment. The percentage of replicate trees where the *taxa* was grouped in the *bootstrap test* (1000 replicates) is shown at the side of each branch. The net distance between *taxa* was determined by the *Poisson* correction model.

## References

1. Simonetta SH, Golombek DA (2007) An automated tracking system for *Caenorhabditis elegans* locomotor behavior and circadian studies application. *J Neurosci Methods* 161: 273–280.
2. Simonetta SH, Romanowski A, Minniti AN, Inestrosa NC, Golombek DA (2008) Circadian stress tolerance in adult *Caenorhabditis elegans*. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 194: 821–828.
3. Simonetta SH, Migliori ML, Romanowski A, Golombek DA (2009) Timing of locomotor activity circadian rhythms in *Caenorhabditis elegans*. *PLoS One* 4: e7571.

## Supporting Information

**Figure S1 Phylogenetic tree of the core clock protein TIMELESS.** The phylogenetic trees were constructed using the *neighbor-joining* method, the *Poisson* model for amino acid substitutions, a *Pairwise Deletion* for the *Gaps/Missing Data Treatment* and a *Gamma distributed* rate among sites was calculated for each alignment. The percentage of replicate trees where the *taxa* was grouped in the *bootstrap test* (1000 replicates) is shown at the side of each branch. The net distance between *taxa* was determined by the *Poisson* correction model. (TIF)

**Figure S2 Similar proteins are found among mammals, insects and nematodes.** The figure shows the seven *C. elegans*' proteins that are conserved among the clocks of mammals and insects, depicted in a: A) mammalian like clock model; and, B) insect (*Drosophila*) like clock model. (TIF)

**Table S1 Conserved mammalian *C. elegans*' hits in all nematode species.** This table shows the 12 mammalian protein hits that are conserved in all other nematode species. *C. elegans*' best hit for each mammalian protein is used as a query to search each proteome. The database, protein hit, percentage of identity, alignment length, E value and bit score are detailed in the table. (XLSX)

**Table S2 Conserved insect *C. elegans*' hits in all nematode species.** This table shows the 32 insect protein hits that are conserved in all other nematode species. *C. elegans*' best hit for each insect protein is used as a query to search each proteome. The database, protein hit, percentage of identity, alignment length, E value and bit score are detailed in the table. (XLSX)

**Table S3 Promoter elements analysis.** This table shows the occurrence of putative circadian promoter elements in the genes of the conserved mammal/insect components, similar to insect components and similar to mammal components. (XLSX)

**Table S4 Comparison between BlastP and HMM search strategies.** The table highlights the proteins that have been identified by the HMM search which were not found by the classical blastP approach. (XLSX)

**File S1 HMM profiles used in this work and HMMsearch results.** (ZIP)

## Author Contributions

Conceived and designed the experiments: AR MJG MEG PDG DAG. Performed the experiments: AR MJG MEG. Analyzed the data: AR MJG MEG DAG. Contributed reagents/materials/analysis tools: MJG PDG. Wrote the paper: AR MJG MEG DG PDG.

7. Hardin PE, Hall JC, Rosbash M (1990) Feedback of the *Drosophila* period gene product on circadian cycling of its messenger RNA levels. *Nature* 343: 536–540.
8. Aronson BD, Johnson KA, Loros JJ, Dunlap JC (1994) Negative feedback defining a circadian clock: autoregulation of the clock gene frequency. *Science* 263: 1578–1584.
9. Brown SA, Kowalska E, Dallmann R (2012) (Re)inventing the circadian feedback loop. *Dev Cell* 22: 477–487.
10. O'Neill JS, Reddy AB (2011) Circadian clocks in human red blood cells. *Nature* 469: 498–503.
11. O'Neill JS, van Ooijen G, Dixon LE, Troein C, Corellou F, et al. (2011) Circadian rhythms persist without transcription in a eukaryote. *Nature* 469: 554–558.
12. Edgar RS, Green EW, Zhao Y, van Ooijen G, Olmedo M, et al. (2012) Peroxiredoxins are conserved markers of circadian rhythms. *Nature* 485: 459–464.
13. Ko CH, Takahashi JS (2006) Molecular components of the mammalian circadian clock. *Hum Mol Genet* 15 Spec No 2: R271–277.
14. Ceriani MF, Darlington TK, Staknis D, Mas P, Petti AA, et al. (1999) Light-dependent sequestration of TIMELESS by CRYPTOCHROME. *Science* 285: 553–556.
15. Claridge-Chang A, Wijnen H, Naef F, Boothroyd C, Rajewsky N, et al. (2001) Circadian regulation of gene expression systems in the *Drosophila* head. *Neuron* 32: 657–671.
16. Baker CL, Loros JJ, Dunlap JC (2012) The circadian clock of *Neurospora crassa*. *FEMS Microbiol Rev* 36: 95–110.
17. McWatters HG, Devlin PF (2011) Timing in plants—a rhythmic arrangement. *FEBS Lett* 585: 1474–1484.
18. Ishiura M, Kutsuna S, Aoki S, Iwasaki H, Andersson CR, et al. (1998) Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria. *Science* 281: 1519–1523.
19. Kageyama H, Nishiwaki T, Nakajima M, Iwasaki H, Oyama T, et al. (2006) Cyanobacterial circadian pacemaker: Kai protein complex dynamics in the KaiC phosphorylation cycle in vitro. *Mol Cell* 23: 161–171.
20. Hasegawa K, Saigusa T, Tamai Y (2005) *Caenorhabditis elegans* opens up new insights into circadian clock mechanisms. *Chronobiol Int* 22: 1–19.
21. Jeon M, Gardner HF, Miller EA, Deshler J, Rougvi AE (1999) Similarity of the *C. elegans* developmental timing protein LIN-42 to circadian rhythm proteins. *Science* 286: 1141–1146.
22. Tennessen JM, Gardner HF, Volk ML, Rougvi AE (2006) Novel heterochronic functions of the *Caenorhabditis elegans* period-related protein LIN-42. *Dev Biol* 289: 30–43.
23. Qin H, Powell-Coffman JA (2004) The *Caenorhabditis elegans* aryl hydrocarbon receptor, AHR-1, regulates neuronal development. *Dev Biol* 270: 64–75.
24. van der Linden AM, Beverly M, Kadener S, Rodriguez J, Wasserman S, et al. (2010) Genome-wide analysis of light- and temperature-entrained circadian transcripts in *Caenorhabditis elegans*. *PLoS Biol* 8: e1000503.
25. Temmerman L, Meelkop E, Janssen T, Bogaerts A, Lindemans M, et al. (2011) *C. elegans* homologs of insect clock proteins: a tale of many stories. *Ann N Y Acad Sci* 1220: 137–148.
26. Eddy SR (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol* 7: e1002195.
27. Johnson LS, Eddy SR, Portugaly E (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* 11: 431.
28. Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39: W29–37.
29. Ebersberger I, Strauss S, von Haeseler A (2009) HaMSTR: profile hidden markov model based search for orthologs in ESTs. *BMC Evol Biol* 9: 157.
30. WormBase (2012) Species Available Through WormBase.
31. Jia H, Wang X, Liu F, Guenther UP, Srinivasan S, et al. (2011) The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. *Cell* 145: 890–901.
32. Arur S, Ohmachi M, Nayak S, Hayes M, Miranda A, et al. (2009) Multiple ERK substrates execute single biological processes in *Caenorhabditis elegans* germ-line development. *Proc Natl Acad Sci U S A* 106: 4776–4781.
33. Kipreos ET, Gohel SP, Hedgecock EM (2000) The *C. elegans* F-box/WD-repeat protein LIN-23 functions to limit cell division during development. *Development* 127: 5071–5082.
34. Fay DS, Keenan S, Han M (2002) *fzr-1* and *lin-35/Rb* function redundantly to control cell proliferation in *C. elegans* as revealed by a nonbiased synthetic screen. *Genes Dev* 16: 503–517.
35. Mehta N, Loria PM, Hobert O (2004) A genetic screen for neurite outgrowth mutants in *Caenorhabditis elegans* reveals a new function for the F-box ubiquitin ligase component LIN-23. *Genetics* 166: 1253–1267.
36. Dreier L, Burbea M, Kaplan JM (2005) LIN-23-mediated degradation of beta-catenin regulates the abundance of GLR-1 glutamate receptors in the ventral nerve cord of *C. elegans*. *Neuron* 46: 51–64.
37. Trent C, Tsuing N, Horvitz HR (1983) Egg-laying defective mutants of the nematode *Caenorhabditis elegans*. *Genetics* 104: 619–647.
38. Weinshenker D, Wei A, Salkoff L, Thomas JH (1999) Block of an ether-a-go-go-like K(+) channel by imipramine rescues *egl-2* excitation defects in *Caenorhabditis elegans*. *J Neurosci* 19: 9831–9840.
39. Wes PD, Bargmann CI (2001) *C. elegans* odour discrimination requires asymmetric diversity in olfactory neurons. *Nature* 410: 698–701.
40. LeBoeuf B, Guo X, Garcia LR (2011) The effects of transient starvation persist through direct interactions between CaMKII and ether-a-go-go K+ channels in *C. elegans* males. *Neuroscience* 175: 1–17.
41. Jiang H, Guo R, Powell-Coffman JA (2001) The *Caenorhabditis elegans* *hif-1* gene encodes a bHLH-PAS protein that is required for adaptation to hypoxia. *Proc Natl Acad Sci U S A* 98: 7916–7921.
42. Tomioka K, Matsumoto A (2010) A comparative view of insect circadian clock systems. *Cell Mol Life Sci* 67: 1397–1406.
43. Yuan Q, Metterville D, Briscoe AD, Reppert SM (2007) Insect cryptochromes: gene duplication and loss define diverse ways to construct insect circadian clocks. *Mol Biol Evol* 24: 948–955.
44. Chan RC, Chan A, Jeon M, Wu TF, Pasqualone D, et al. (2003) Chromosome cohesion is regulated by a clock gene paralogue TIM-1. *Nature* 423: 1002–1009.
45. Powell-Coffman JA, Bradfield CA, Wood WB (1998) *Caenorhabditis elegans* orthologs of the aryl hydrocarbon receptor and its heterodimerization partner the aryl hydrocarbon receptor nuclear translocator. *Proc Natl Acad Sci U S A* 95: 2844–2849.
46. Ooe N, Saito K, Oeda K, Nakatuka I, Kaneko H (2007) Characterization of *Drosophila* and *Caenorhabditis elegans* NXF-like-factors, putative homologs of mammalian NXF. *Gene* 400: 122–130.
47. Liu J, Ward A, Gao J, Dong Y, Nishio N, et al. (2010) *C. elegans* phototransduction requires a G protein-dependent cGMP pathway and a taste receptor homolog. *Nat Neurosci* 13: 715–722.
48. Takahashi JS, Lowrey PL (2011) Genetics of Circadian Rhythms in Mammalian Model Organisms. In: Brody S, editor. *The Genetics of Circadian Systems*.
49. Sandrelli F, Costa R, Kyriacou CP, Rosato E (2008) Comparative analysis of circadian clock genes in insects. *Insect Mol Biol* 17: 447–463.
50. Erdelyi P, Borsos E, Takacs-Vellai K, Kovacs T, Kovacs AL, et al. (2011) Shared developmental roles and transcriptional control of autophagy and apoptosis in *Caenorhabditis elegans*. *J Cell Sci* 124: 1510–1518.
51. Hatzold J, Conradt B (2008) Control of apoptosis by asymmetric cell division. *PLoS Biol* 6: e84.
52. Wang X, Jia H, Chamberlin HM (2006) The bZip proteins CES-2 and ATF-2 alter the timing of transcription for a cell-specific target gene in *C. elegans*. *Dev Biol* 289: 456–465.
53. Kostrouchova M, Krause M, Kostrouch Z, Rall JE (2001) Nuclear hormone receptor CHR3 is a critical regulator of all four larval molts of the nematode *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* 98: 7360–7365.
54. Gissendanner CR, Crossgrove K, Kraus KA, Maina CV, Sluder AE (2004) Expression and function of conserved nuclear receptor genes in *Caenorhabditis elegans*. *Dev Biol* 266: 399–416.
55. Ueda HR, Hayashi S, Chen W, Sano M, Machida M, et al. (2005) System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nat Genet* 37: 187–192.
56. Migliori ML, Romanowski A, Simonetta SH, Valdez D, Guido M, et al. (2011) Daily variation in melatonin synthesis and arylalkylamine N-acetyltransferase activity in the nematode *Caenorhabditis elegans*. *J Pineal Res*.
57. Janssen T, Husson SJ, Meelkop E, Temmerman L, Lindemans M, et al. (2009) Discovery and characterization of a conserved pigment dispersing factor-like neuropeptide pathway in *Caenorhabditis elegans*. *J Neurochem* 111: 228–241.
58. Meelkop E, Temmerman L, Janssen T, Suetens N, Beets I, et al. (2012) PDF receptor signaling in *Caenorhabditis elegans* modulates locomotion and egg-laying. *Mol Cell Endocrinol* 361: 232–240.
59. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
60. Kelly RJ, Vincent DE, Friedberg I (2010) IPRStats: visualization of the functional potential of an InterProScan run. *BMC Bioinformatics* 11 Suppl 12: S13.
61. Leclerc GM, Boockfor FR (2005) Pulses of prolactin promoter activity depend on a noncanonical E-box that can bind the circadian proteins CLOCK and BMAL1. *Endocrinology* 146: 2782–2790.
62. Darlington TK, Lyons LC, Hardin PE, Kay SA (2000) The period E-box is sufficient to drive circadian oscillation of transcription in vivo. *J Biol Rhythms* 15: 462–471.
63. Hanks SK, Hunter T (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 9: 576–596.
64. Andersen JN, Mortensen OH, Peters GH, Drake PG, Iversen LF, et al. (2001) Structural and evolutionary relationships among protein tyrosine phosphatase domains. *Mol Cell Biol* 21: 7117–7136.
65. Miskei M, Adam C, Kovacs L, Karanyi Z, Dombradi V (2011) Molecular evolution of phosphoprotein phosphatases in *Drosophila*. *PLoS One* 6: e22218.
66. Cohen PT, Brevin ND, Hughes V, Mann DJ (1990) Protein serine/threonine phosphatases; an expanding family. *FEBS Lett* 268: 355–359.
67. Petrillo E, Sanchez SE, Kornbliht AR, Yanovsky MJ (2011) Alternative splicing adds a new loop to the circadian clock. *Commun Integr Biol* 4: 284–286.

68. Sanchez SE, Petrillo E, Beckwith EJ, Zhang X, Rugnone ML, et al. (2010) A methyl transferase links the circadian clock to the regulation of alternative splicing. *Nature* 468: 112–116.
69. Fitch DHA (2005) Introduction to nematode evolution and ecology. In: Community TCeR, editor. *WormBook*.
70. Aguinaldo AM, Turbeville JM, Linford LS, Rivera MC, Garey JR, et al. (1997) Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* 387: 489–493.
71. Sidow A, Thomas WK (1994) A molecular evolutionary framework for eukaryotic model organisms. *Curr Biol* 4: 596–603.
72. Gotter AL, Suppa C, Emanuel BS (2007) Mammalian TIMELESS and Tipin are evolutionarily conserved replication fork-associated factors. *J Mol Biol* 366: 36–52.
73. Benna C, Bonaccorsi S, Wulbeck C, Helfrich-Forster C, Gatti M, et al. (2010) *Drosophila* timeless2 is required for chromosome stability and circadian photoreception. *Curr Biol* 20: 346–352.
74. Gotter AL (2006) A Timeless debate: resolving TIM's noncircadian roles with possible clock function. *Neuroreport* 17: 1229–1233.
75. McFarlane RJ, Mian S, Dalgaard JZ (2010) The many facets of the Tim-Tipin protein families' roles in chromosome biology. *Cell Cycle* 9: 700–705.
76. Banerjee D, Kwok A, Lin SY, Slack FJ (2005) Developmental timing in *C. elegans* is regulated by kin-20 and tim-1, homologs of core circadian clock genes. *Dev Cell* 8: 287–295.
77. Golden A, Cohen-Fix O (2003) Getting (chromosomes) loaded—a new role for timeless. *Dev Cell* 5: 7–9.
78. Kumaki Y, Ukai-Tadenuma M, Uno KD, Nishio J, Masumoto KH, et al. (2008) Analysis and synthesis of high-amplitude Cis-elements in the mammalian circadian clock. *Proc Natl Acad Sci U S A* 105: 14946–14951.
79. Munoz E, Baler R (2003) The circadian E-box: when perfect is not good enough. *Chronobiol Int* 20: 371–388.
80. Ukai-Tadenuma M, Kasukawa T, Ueda HR (2008) Proof-by-synthesis of the transcriptional logic of mammalian circadian clocks. *Nat Cell Biol* 10: 1154–1163.
81. Hao H, Allen DL, Hardin PE (1997) A circadian enhancer mediates PER-dependent mRNA cycling in *Drosophila melanogaster*. *Mol Cell Biol* 17: 3687–3693.
82. Monsalve GC, Van Buskirk C, Frand AR (2011) LIN-42/PERIOD controls cyclical and developmental progression of *C. elegans* molts. *Curr Biol* 21: 2033–2045.
83. Hendriks GJ, Gaidatzis D, Aeschmann F, Grosshans H (2014) Extensive Oscillatory Gene Expression during *C. elegans* Larval Development. *Mol Cell* 53: 380–392.
84. Gu HF, Xiao JH, Niu LM, Wang B, Ma GC, et al. (2014) Adaptive evolution of the circadian gene timeout in insects. *Sci Rep* 4: 4212.
85. Unsal-Kacmaz K, Chastain PD, Qu PP, Minoo P, Cordeiro-Stone M, et al. (2007) The human Tim/Tipin complex coordinates an Intra-S checkpoint response to UV that slows replication fork displacement. *Mol Cell Biol* 27: 3131–3142.
86. Yang X, Wood PA, Hrushesky WJ (2010) Mammalian TIMELESS is required for ATM-dependent CHK2 activation and G2/M checkpoint control. *J Biol Chem* 285: 3030–3034.
87. Unsal-Kacmaz K, Mullen TE, Kaufmann WK, Sancar A (2005) Coupling of human circadian and cell cycles by the timeless protein. *Mol Cell Biol* 25: 3109–3116.
88. Engelen E, Janssens RC, Yagita K, Smits VA, van der Horst GT, et al. (2013) Mammalian TIMELESS is involved in period determination and DNA damage-dependent phase advancing of the circadian clock. *PLoS One* 8: e56623.
89. Barnes JW, Tischkau SA, Barnes JA, Mitchell JW, Burgoon PW, et al. (2003) Requirement of mammalian Timeless for circadian rhythmicity. *Science* 302: 439–442.
90. Takumi T, Nagamine Y, Miyake S, Matsubara C, Taguchi K, et al. (1999) A mammalian ortholog of *Drosophila* timeless, highly expressed in SCN and retina, forms a complex with mPER1. *Genes Cells* 4: 67–75.
91. Shemesh Y, Cohen M, Bloch G (2007) Natural plasticity in circadian rhythms is mediated by reorganization in the molecular clockwork in honeybees. *FASEB J* 21: 2304–2311.
92. Sancar A (2000) Cryptochrome: the second photoactive pigment in the eye and its role in circadian photoreception. *Annu Rev Biochem* 69: 31–67.
93. Edwards SL, Charlie NK, Milfort MC, Brown BS, Gravlin CN, et al. (2008) A novel molecular solution for ultraviolet light detection in *Caenorhabditis elegans*. *PLoS Biol* 6: e198.
94. Ward A, Liu J, Feng Z, Xu XZ (2008) Light-sensitive neurons and channels mediate phototaxis in *C. elegans*. *Nat Neurosci* 11: 916–922.
95. Crews ST (2003) PAS proteins: regulators and sensors of development and physiology. Boston: Kluwer Academic Publishers. xiii, 263 p. p.
96. Qin H, Zhai Z, Powell-Coffman JA (2006) The *Caenorhabditis elegans* AHR-1 transcription complex controls expression of soluble guanylate cyclase genes in the URX neurons and regulates aggregation behavior. *Dev Biol* 298: 606–615.
97. Romney SJ, Newman BS, Thacker C, Leibold EA (2011) HIF-1 regulates iron homeostasis in *Caenorhabditis elegans* by activation and inhibition of genes involved in iron uptake and storage. *PLoS Genet* 7: e1002394.
98. Hardin PE (1994) Analysis of period mRNA cycling in *Drosophila* head and body tissues indicates that body oscillators behave differently from head oscillators. *Mol Cell Biol* 14: 7211–7218.
99. Plautz JD, Kaneko M, Hall JC, Kay SA (1997) Independent photoreceptive circadian clocks throughout *Drosophila*. *Science* 278: 1632–1635.
100. Dong G, Kim YI, Golden SS (2010) Simplicity and complexity in the cyanobacterial circadian clock mechanism. *Curr Opin Genet Dev* 20: 619–625.
101. Rhee SG, Woo HA, Kil IS, Bac SH (2012) Peroxiredoxin functions as a peroxidase and a regulator and sensor of local peroxides. *J Biol Chem* 287: 4403–4410.
102. Olmedo M, O'Neill JS, Edgar RS, Valekunja UK, Reddy AB, et al. (2012) Circadian regulation of olfaction and an evolutionarily conserved, nontranscriptional marker in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* 109: 20479–20484.
103. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, et al. (2005) InterProScan: protein domains identifier. *Nucleic Acids Res* 33: W116–120.
104. Yan J, Wang H, Liu Y, Shao C (2008) Analysis of gene regulatory networks in the mammalian circadian rhythm. *PLoS Comput Biol* 4: e1000193.
105. Schwarz EM, Antoshechkin I, Bastiani C, Bieri T, Blasiar D, et al. (2006) WormBase: better software, richer content. *Nucleic Acids Res* 34: D475–478.
106. Fiedler T, Rehmsmeier M (2006) jPREdictor: a versatile tool for the prediction of cis-regulatory elements. *Nucleic Acids Res* 34: W546–550.
107. Sudhaus W, Fitch D (2001) Comparative studies on the phylogeny and systematics of the rhabditidae (nematoda). *J Nematol* 33: 1–70.