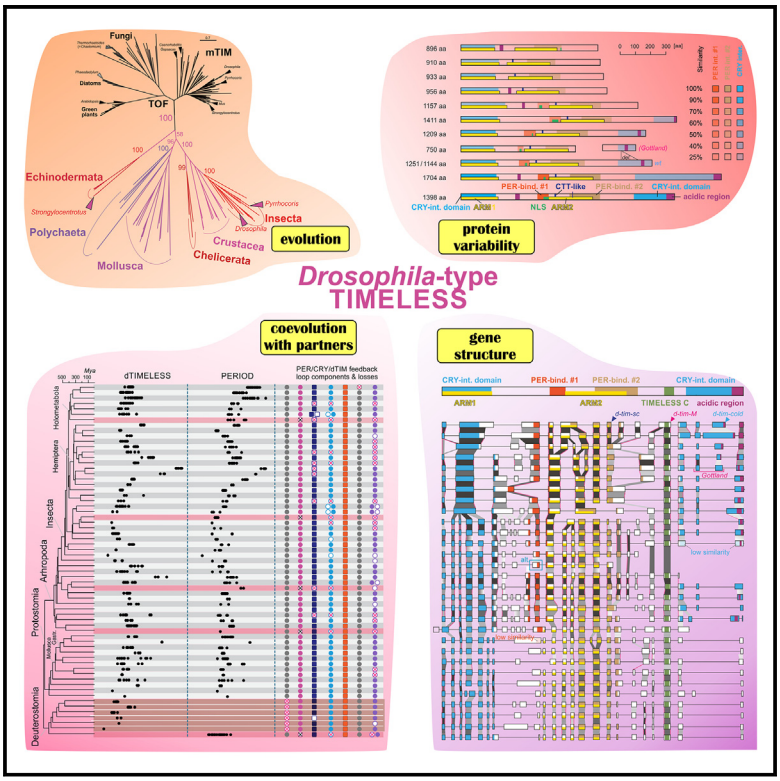
iScience

Article

Coevolution of *Drosophila*-type timeless with partner clock proteins

Graphical abstract



Highlights

- Reconstruction of the origin of dTIM from mTIM/TOF
- Evolution of dTIM domains was reconstructed
- Losses of dTIM were compared to losses of partner proteins (i.e., PER, dCRY, and JET)
- Four lineage-specific dTIM duplications were found, the most extreme in *Daphnia*

Authors

Enrico Bullo, Ping Chen, Ivan Fiala, Vlastimil Smýkal, David Doležel

Correspondence

david.dolezel@entu.cas.cz

In brief

Genetics; Molecular biology; Neuroscience; Evolutionary biology

Bullo et al., 2025, iScience *28*, 112338
May 16, 2025
https://doi.org/10.1016/j.isci.2025.112338

CellPress

iScience

CellPress
OPEN ACCESS

Article

Coevolution of *Drosophila*-type timeless with partner clock proteins

Enrico Bullo,[1,2,3] Ping Chen,[1,2,3] Ivan Fiala,[1] Vlastimil Smýkal,[1] and David Doležel[1,4,*]
[1]Biology Centre of the Czech Academy of Sciences, 37005 České Budějovice, Czech Republic
[2]Faculty of Science, University of South Bohemia, 37005 České Budějovice, Czech Republic
[3]These authors contributed equally
[4]Lead contact
*Correspondence: david.dolezel@entu.cas.cz
https://doi.org/10.1016/j.isci.2025.112338

SUMMARY

Drosophila-type timeless (dTIM) is a key clock protein in fruit flies, regulating rhythmicity and light-mediated entrainment. However, functional experiments indicate that its contribution to the clock differs in various insects. Therefore, we conducted a comprehensive phylogenetic analysis of dTIM across animals and dated its origin, gene duplications, and losses. We identified variable and conserved protein domains and pinpointed animal lineages that underwent the biggest changes in dTIM. While dTIM modifications are only mildly affected by changes in the PER protein, even the complete loss of PER in echinoderms had no impact on dTIM. However, changes in dTIM always co-occur with the loss of CRYPTOCHROMES or JETLAG. This is exemplified by the remarkably accelerated evolution of dTIM in phylloxera and aphids. Finally, alternative *d-tim* splicing, characteristic of *Drosophila melanogaster* temperature-dependent function, is conserved to some extent in Diptera, albeit with unique alterations. Altogether, this study pinpoints major changes that shaped dTIM evolution.

INTRODUCTION

The majority of life forms are exposed to periodic alteration of day and night. Anticipating environmental changes such as dawn, dusk, or midday heat provides a significant advantage to various organisms. Circadian clocks, the molecular devices 'ticking' with an approximately 24-h period, have evolved independently in cyanobacteria, plants, fungi, and animals.[1,2] Despite their independent origins, some molecular mechanisms appear to be conserved.[3] The animal circadian clock, best characterized in the fruit fly *Drosophila melanogaster* and mouse *Mus musculus*, operates through several interconnected transcription-translation feedback loops (TTFLs).[2,4] These clock components are broadly conserved across bilaterian animals, with some traceable back to cnidarians.[5,6] Key transcription factors, such as CLOCK and CYCLE (with the transactivation domain of CYCLE uniquely lost in *Drosophila* and other cyclorrhaphan flies) or CLOCK and BMAL1 (in most other species), are integral to the circadian clock.[7,8] These transcription factors belong to the basic-helix-loop-helix (bHLH) Period-Arnt-Singleminded (PAS) protein family and act as critical activators of the clock.[9–12]

CLK and CYC/BMAL1 drive the expression of both positive and negative regulators in the circadian clock. In *Drosophila*, the key negative regulators are the PERIOD (PER) and *Drosophila*-type TIMELESS (dTIM) proteins. In the original descriptions and much of the literature, this protein/gene is often referred to simply as TIM/tim. To avoid confusion with the mammalian type of the protein/gene (mTIM/m-tim), we strictly use the "d" or "m" prefix throughout this text.

In contrast to *Drosophila*, the mammalian clock relies on PER proteins (three paralogs in mice) and mammalian-type CRYPTOCHROMEs (mCRYs; two paralogs in mice), sometimes referred to as CRY2-type.[13,14] CRYs are phylogenetically related to photolyases, although mCRYs have lost their ability to detect light.[15] The *Drosophila*-type CRY (dCRY, also known as CRY1-type) serves as a key photoreceptor in brain neurons responsible for clock entrainment.[16–18] However, dCRY may also function as a repressor in the peripheral clock of *Drosophila*.[19] This latter role facilitated its identification in a screen for mutants affecting periodic luciferase oscillations.[20]

Over time, various methods and approaches have elucidated the molecular mechanisms of the clock, including protein-protein interactions. Classical mutagenesis and genetic mapping were instrumental in identifying the first circadian clock components, while techniques like yeast two-hybrid assays, co-immunoprecipitation, and co-transfection experiments (comparing wild-type and mutant versions) revealed interactions among these components. For example, several key domains have been identified within the dTIM protein: two regions essential for binding PER (PER-binding #1 and PER-binding #2), a cytoplasmic localization domain (CLD), a nuclear localization signal (NLS), and a region where Importin binds[21–23] (Figure 1C). The dynamics of PER-dTIM interaction have been studied in great detail using fluorescence resonance energy transfer (FRET) in *Drosophila* Schneider 2 (S2) cells. While the PER-dTIM interaction is essential for the

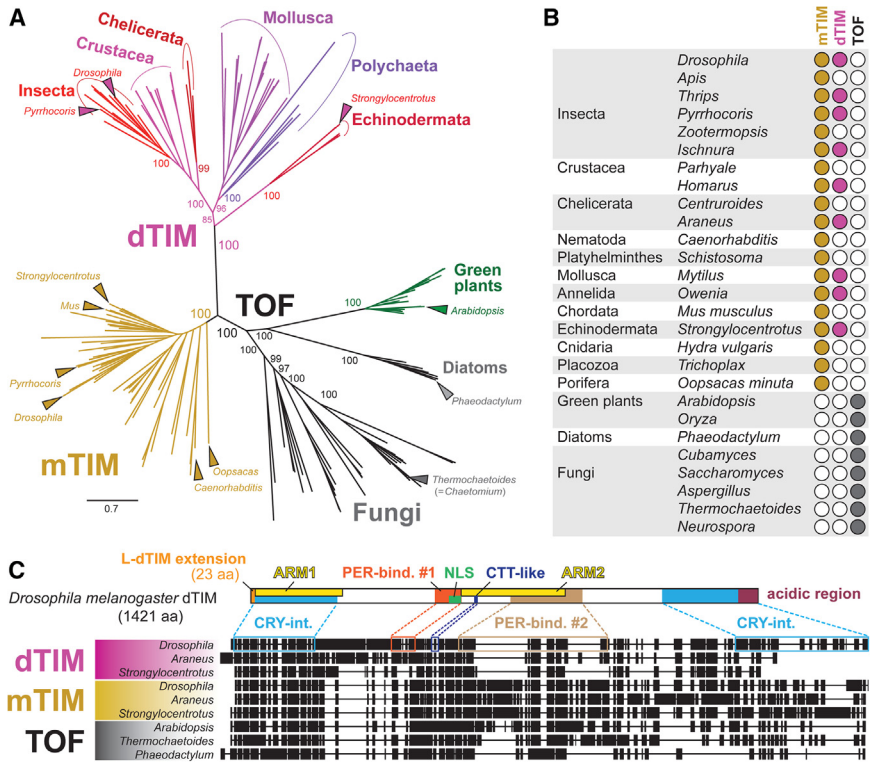


Figure 1. Phylogeny of dTIM (*Drosophila*-type timeless), mTIM (mammalian-type timeless), and TOF (Topoisomerase 1-associated Factor)

(A) A tree inferred from protein sequence alignment using RAxML maximum likelihood GAMMA-based model. Major groups are highlighted, and arrowheads indicate the position of particular model species. Bootstrap values (inferred from 500 replicates) are shown in %. For the detailed tree including all bootstrap values see Figure S1.

(B) Distribution and mutual exclusivity of mTIM and TOF suggest that these proteins are orthologs specific to animals (mTIM) and plant/fungi/diatoms (TOF), respectively. dTIM is found only in Bilateria, including two basal deuterostome lineages (Hemichordata, Echinodermata) and the majority of Protostomia. Therefore, dTIM has been independently lost in chordates, platyhelminths, nematodes, scorpions, termites, hymenopterans, and amphipod Crustacea (e.g., *Parhyale*). For evidence supporting gene loss, see Table S1.

(C) A schematic model of *Drosophila* dTIM illustrates its major functional domains. Protein alignment of dTIM, mTIM, and TOF highlights expansions and deletions, including a dTIM-specific deletion within PER-interaction domain #2 (light brown), which is conserved from the sea urchin (*Strongylocentrotus*) to *Drosophila*.

nuclear translocation of both proteins, once inside the nucleus either the topology of the proteins or their interaction changes.[24,25] Additionally, nuclear export signals are critical for the proper functioning of both dTIM and PER.[26,27] The stability of dTIM is further regulated by phosphorylation, with shaggy (GSK) and CK2 playing major roles in this process.[28]

The interaction between PER and dTIM is crucial not only for maintaining circadian rhythms under constant conditions but also as part of the entrainment (clock synchronization) mechanism. Since PER is stabilized by dTIM, depletion of dTIM in the morning results in a phase advance, whereas evening depletion triggers a phase delay. Light-mediated degradation of dTIM requires interaction between dTIM, dCRY, and JETLAG.[29] Flies with mutated or absent dCRY[30,31] or mutated JETLAG[32,33] remain rhythmic under constant light—conditions that cause arrhythmicity in wild-type flies. JETLAG, an F box protein with leucine-rich repeats, targets dTIM for light-induced degradation. Meanwhile, dCRY polyubiquitination and subsequent degradation are triggered by Ramshackle (BRWD3).[34] Beyond its role in light input, dCRY also contributes to the robustness of rhythmicity under constant conditions at low temperatures in *Drosophila*.[31] In the monarch butterfly, *Danaus plexippus*, dCRY depletion reduces eclosion rhythmicity and impairs the robustness of periodic clock gene expression in the brain.[35] Similarly, in the silkworm *Bombyx mori*, a dCRY knockout strain exhibited arrhythmic eclosion, and mutant females were unable to properly measure photoperiod.[36]

The role of dTIM in seasonality is supported by studies on a unique mutant in the drosophilid fly *Chymomyza costata*. This species inhabits temperate regions and must synchronize its development with seasonal changes. Long photoperiods promote direct larval development during early instars, whereas short photoperiods induce a developmental arrest known as diapause.[37] Photoperiodic time measurement requires a functional *d-tim* gene, as indicated by mapping a non-photoperiodic diapause mutation to the *timeless* gene. This mutation involves an approximately 3-kbp deletion in the promoter region, removing *cis*-regulatory sequences essential for transcription.[38,39] There has been an ongoing debate regarding whether the reproductive arrest triggered by low temperature and short photoperiods in *Drosophila melanogaster* constitutes diapause or quiescence. The definition of diapause as an anticipatory mechanism to prepare for harsh conditions[40] suggests the latter. Importantly, *d-tim* alleles significantly influence the incidence of diapause/quiescence,[41] and the genetic interaction between *d-tim* and protein tyrosine phosphatase *eyes absent* (*eya*) is involved in regulating this state.[42] Surprisingly, mutations in dTIM's partner, PER, do not appear to affect reproductive arrest in *D. melanogaster*,[43] which contrasts with the synergistic role of these two proteins in the circadian clock. In *Bombyx mori*, dTIM, PER, CLK, and BMAL1 are all critical components of the photoperiodic clock, although the precise mechanism by which they function remains unclear.[44]

For some proteins, such as dCRY, mCRY, CLK, and BMAL1, their structures have been determined using crystallography.[11,45–47] In the case of PER, only a portion of the protein has been successfully crystallized.[48] However, for other proteins, including dTIM, the 3D structure has remained elusive.

While related proteins such as Topoisomerase 1-associated Factor (TOF1) from the fungus *Thermochaetoides* have been partially resolved through crystallography,[49] the full structure of dTIM was only recently resolved as part of the dCRY-dTIM dimer using cryogenic electron microscopy (Cryo-EM).[23] The interaction between dCRY and dTIM involves two distinct regions of dTIM: the amino-terminal Armadillo 1 (ARM1) repeats and a C-terminal CRY-binding domain. Upon light illumination, large-scale rearrangements in dCRY are coupled with conformational changes in its flavin cofactor. This process results in the release of the autoinhibitory C-terminal tail (CTT) from the dCRY pocket, which is then replaced by the N terminus of dTIM. This terminal motif is highly conserved across dTIM species, and its addition of only 23 amino acids significantly reduces light sensitivity in flies.[41,50–53]

Interestingly, the paralogous mTIM has been shown to play a role in the circadian clock, albeit to a much lesser extent than dTIM. It influences light sensitivity in *Drosophila*, rhythmicity in crickets and linden bugs, and more recently has been implicated in adjusting activity phases in humans and mice.[54–57] However, the functions of dTIM have generally diverged from those of mTIM and TOF1. TOF1 directly binds to double-stranded DNA as a component of the large replisome complex,[58] while mTIM stabilizes replication forks, pairs sister chromatids, and regulates the S phase of the cell cycle.[59] Moreover, despite the overall structural similarities among dTIM, mTIM, and TOF1, specific details appear to be unique to each protein type.[23]

Despite the relatively well-characterized role of dTIM in the *Drosophila* clock, its function in other organisms appears to be only partially conserved. In vertebrates and several arthropod lineages, dTIM and dCRY have been completely lost,[55,60] with mCRY serving as the partner of PER in these species.[15] Furthermore, vertebrates have undergone multiple rounds of genome duplication, resulting in the multiplication of many clock components, including PERs and mCRYs. In non-*Drosophila* insects, various combinations of mCRY, dTIM, JET, and FBXL3 have been identified.[55,61,62] Functional experiments suggest that the roles of these components in the pacemaker system of some species differ significantly from those in *Drosophila*. The non-essentiality of certain components for clock rhythmicity, as demonstrated for dTIM in the linden bug *Pyrrhocoris apterus*, cricket, firebrat, and cockroaches, suggests potential transitions from one clock setup to another.[55,63–65]

The growing availability of transcriptomic and genomic data from all major animal lineages has enabled systematic analyses of the evolution of specific clock proteins. In this study, we focused on a detailed comparison of dTIM, aiming to identify its origin and map its occurrences and losses across the animal phylogeny. We analyzed conserved and variable regions of the protein, evaluating them in the context of *d-tim* gene evolution, including its exon/intron structure. Additionally, we examined the coevolution of dTIM with its partner proteins—PER, dCRY, mCRY, and JETLAG—and, by mapping these changes onto the animal phylogeny, highlighted potential key interactions that have shaped the evolution of the animal circadian clock. Finally, we investigated alternative splicing of *d-tim* in Diptera, proposing conserved molecular mechanisms that may regulate dTIM in this insect order.

RESULTS

dTIM is an ortholog of animal mTIM and TOF1 from plants, fungi, and diatoms

We systematically searched for dTIM, its related mammalian-type TIM (mTIM), and TOF1 proteins across diverse organisms and reconstructed the phylogeny of the identified sequences (Figures 1A and S1). Consistent with previous reports, dTIM was identified in most Protostomia and two basal deuterostome groups, Echinodermata and Hemichordata. For the latter group, however, dTIM was only found in the acorn worm *Ptychodera flava*, while it was absent in another hemichordate, *Saccoglossus kowalevskii*. Notably, we did not retrieve dTIM from any Cnidaria, Placozoa, or Porifera, suggesting that dTIM likely originated at the dawn of Bilateria.

The paralogous mTIM, a protein essential for development and participating in circadian clock function in mammals and insects,[54,66,67] is found in all animals, including "prebilaterian" groups such as Cnidaria, Placozoa, and Porifera. Furthermore, the occurrence of mTIM is mutually exclusive with that of TOF1, a protein found in plants, diatoms, and fungi (where it is known as Swi1 in *Schizosaccharomyces pombe*). This suggests that mTIM and TOF1 are one-to-one orthologs (Figure 1B). Protein alignment reveals conserved dTIM-specific deletions in the region containing PER-binding domain #2 (PER-bind. #2), which clearly distinguishes dTIM from mTIM (Figure 1C). The similarity among mTIMs and TOF1 proteins is highest within the first ~600 amino acids, whereas the C-terminal region is notably variable.

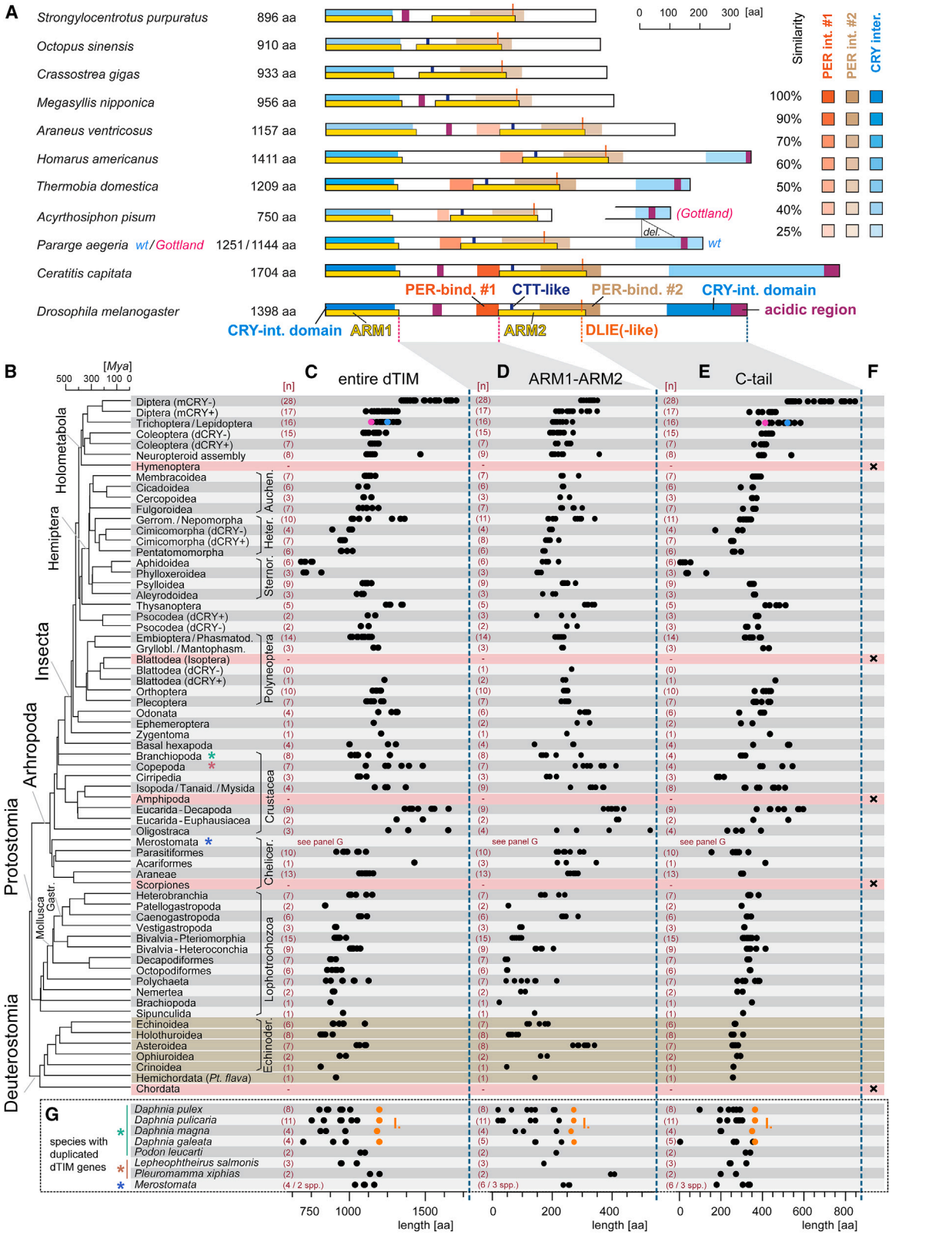
Independent losses of dTIM

In addition to the well-supported and previously described losses of dTIM in Chordata, Hymenoptera, and termites (with the possible exception of early-branching termite species such as *Porotermes*),[55,60] we did not identify this gene in Amphipoda (Crustacea), Scorpiones (Chelicerata), Platyhelminthes, or Nematoda (Figures 1B and 2F; see Figure S1; Table S1 for details). The *Caenorhabditis elegans* TIM homolog TIM-1, previously analyzed in the context of chronobiology,[68] belongs to the mTIM clade (Figure 1A).

Variability in dTIM

The available dataset allowed us to explore the sequence variability of dTIM proteins. The length of dTIM ranged from 683 to 757 amino acids (aa) in aphids and 874–920 aa in cephalopod mollusks, reaching up to 1666–1704 aa in the Tephritidae family of Diptera (including *Ceratitis* and *Bactrocera* genera) (Figure 2C; Table S2). Since *Drosophila* dTIM is by far the best-characterized member of this group—with extensive functional assays,[21,22] genetic mutant studies,[26,69–72] and structural analyses[23]—we used it as a reference for domain and region terminology in our interspecies comparisons.

The most conserved regions across species are the two armadillo domains (Figure 2A). The first domain, ARM1, includes a region that interacts with the importin 1 protein, where several mutations impacting the free-running period have been reported.[22,28,73] ARM1 is nearly identical to the first CRY-interaction domain, whereas the second armadillo domain (ARM2) partially overlaps with PER-binding site #2 (PER-bind. #2), which



(legend on next page)

was functionally identified using *Drosophila* S2 cell cultures.[21] ARM2 also contains a short CTT-like sequence (DTSHFFWLVT in *Drosophila melanogaster*). This motif is located far from the dCRY-dTIM interface and is therefore unlikely to influence the binding of these two proteins.[23] The motif is found in nearly all analyzed dTIMs (Figures S2A, S2B, and S3).

In contrast to PER-bind. #2, PER-binding site #1 (PER-bind. #1), located between the two armadillo domains, exhibits greater sequence variability and can be detected in insects, Crustacea, and, to some extent, chelicerates. The absence of PER-bind. #1 correlates with a shorter central part of the protein in certain Mollusca and basal Deuterostomia (Echinodermata and Hemichordata). In these groups, only the acidic region of the N-terminal part of PER-bind. #1 is partially conserved (Figures 2, S2, S4, and S5). Additional variability in the central region (between the armadillo domains) is particularly pronounced in some Crustacea, an ancient and highly diverse group from which Hexapoda (including insects) evolved.

The most variable part of dTIM proteins is located in the C-terminal region. In *D. melanogaster*, this region contains a CRY-interaction domain and a C-terminal acidic region responsible for cytoplasmic localization, also known as the cytoplasmic localization domain (CLD).[21,23,74] This region has notably expanded in a subset of Diptera, including *Drosophila* and *Ceratitis*. In contrast, the shortest sequences are found in aphids and phylloxera, which differ dramatically from their sister group, Psylloidea (Figure 2E). Remarkably, the geographical allele Gottland of the Speckled Wood butterfly (*Pararge aegeria*) is characterized by a substantial deletion[75] that encompasses the CRY-interaction domain (Figures 2A and 2E), suggesting that this allelic variant may result in altered light sensitivity. The CRY-interaction domain is recognizable even in the most ancestral insect, *Thermobia*, and is detected in some, but not all, Crustacea (Figures S2 and S6). Interestingly, variability among paralogs within individual species—most prominently in the *Daphnia* genus, which has up to 8–11 paralogs—centers on the inter-armadillo region and the C-terminal tail. These regions are also the most variable in interspecific comparisons (Figures 3 and S6).

Independent duplications of dTIM

In several crustacean and chelicerate species, multiple *d-tim* paralogs were identified. While two paralogs were most commonly observed, *Daphnia* species exhibited varying numbers of paralogs, reaching as many as 11 in *Daphnia pulicaria*. Furthermore, the lengths of *Daphnia* dTIM paralogs varied within species, showing differences in inter-armadillo distance and the C-tail (Figure 2G). We performed a detailed phylogenetic analysis to investigate the origin of these duplications and assess whether *d-tim* gene duplication could contribute to a more general pattern in dTIM evolution. As shown in Figure 3 and well supported by bootstrap analysis, *d-tim* duplication can be mapped to only five lineages. In horseshoe crabs (*Limulus polyphemus*, *Carcinoscorpius rotundicauda*, and *Tachypleus tridentatus*; Figure S7; Table S3), two *d-tim* paralogs are consistent with the genome duplication events in this lineage, which also includes circadian transcription factor BMAL1 and other bHLH-PAS proteins.[12] In copepods (Crustacea), three closely related *Lepeophtheirus salmonis* paralogs (Figure S8) resulted from duplications independent of duplication resulting in two paralogs found in *Pleuromamma* and *Metridia*. In branchiopods, the duplication that produced two *d-tim* genes in *Podon* occurred independently of the series of duplications specific to *Daphnia* species. Four primary dTIM clusters (I–IV) are present across all species in the *Daphnia* genus. The longest proteins are found in cluster I (Figure 2G), which branches at the base of the *Daphnia*-specific tree expansion. Proteins from cluster I also contain a complete set of functional domains, the longest inter-armadillo region, and the longest C-tail (Figures 2G and 3). Furthermore, the *Daphnia pulex* dTIM paralog A, belonging to cluster I, is cyclically expressed.[76] Clusters II, III, and IV include additional species-specific duplications, resulting in eight paralogs in *D. pulex* and eleven in *D. pulicaria*.

Substitution rate in dTIM and coevolution with partner proteins

To provide a clear graphical representation of protein sequence variability, we plotted the substitution rate per amino acid position per million years for each analyzed dTIM protein. This method is minimally influenced by in-frame deletions or variations at phylogenetically variable positions, as evidenced by the minimal differences in substitution rates between mCRY+ and mCRY- Diptera (Figure 4B), despite these groups showing significant variation in the length of their C-terminus (Figure 2E). However, this approach effectively highlights taxa with alterations in phylogenetically conserved positions, such as aphids and phylloxera. A notable difference in substitution rates is observed among *Daphnia* paralogs. The lowest values are

Figure 2. Differences in dTIM proteins

(A) Selected examples of dTIM with highlighted functional domains, annotated based on sequence similarity, illustrate the major varieties of dTIM. Similarities in PER-binding and CRY-interaction domains are represented as shades of specific colors. Note the two alleles in *Pararge aegeria* differ in the CRY-interaction domain. CTT-like is a 10-aa motif similar to the C-terminal tail (CTT) recognized in dCRY of *D. melanogaster*. For additional annotated protein models and detailed similarity values, see Figures S2A, S2B, and S5.

(B) The tree illustrates the current view of phylogeny among groups of Bilateria for which dTIM was analyzed. Brick-red horizontal background highlights lineages where dTIM has most likely been lost. Variability in dTIM protein lengths is visualized across different regions: (C) the entire dTIM length, (D) the central portion between the ARM1 and ARM2 domains, and (E) the terminal region between the DLIE motif and the C-terminus. Each dot corresponds to a protein measurement (in amino acids) from a species within a specific taxonomic group. Values for *P. aegeria* (Lepidoptera) alleles are highlighted in magenta and turquoise, matching the colors in (A). For exact measurements, see Table S2.

(F) The loss of dTIM is depicted as "x."

(G) Species with multiple dTIM paralogs are shown, with each plotted on a separate line. The entire length of dTIM in Merostomata (horseshoe crabs) is plotted for four proteins in two species (4/2 spp.), whereas six proteins from three species are grouped together (6/3 spp.) for the ARM1-ARM2 and C-tail lengths. Orange dots correspond to *Daphnia* paralogs from Cluster I, proposed as the most complete dTIM variants in this genus. The number of values in each group and category is indicated by a magenta 'n'.

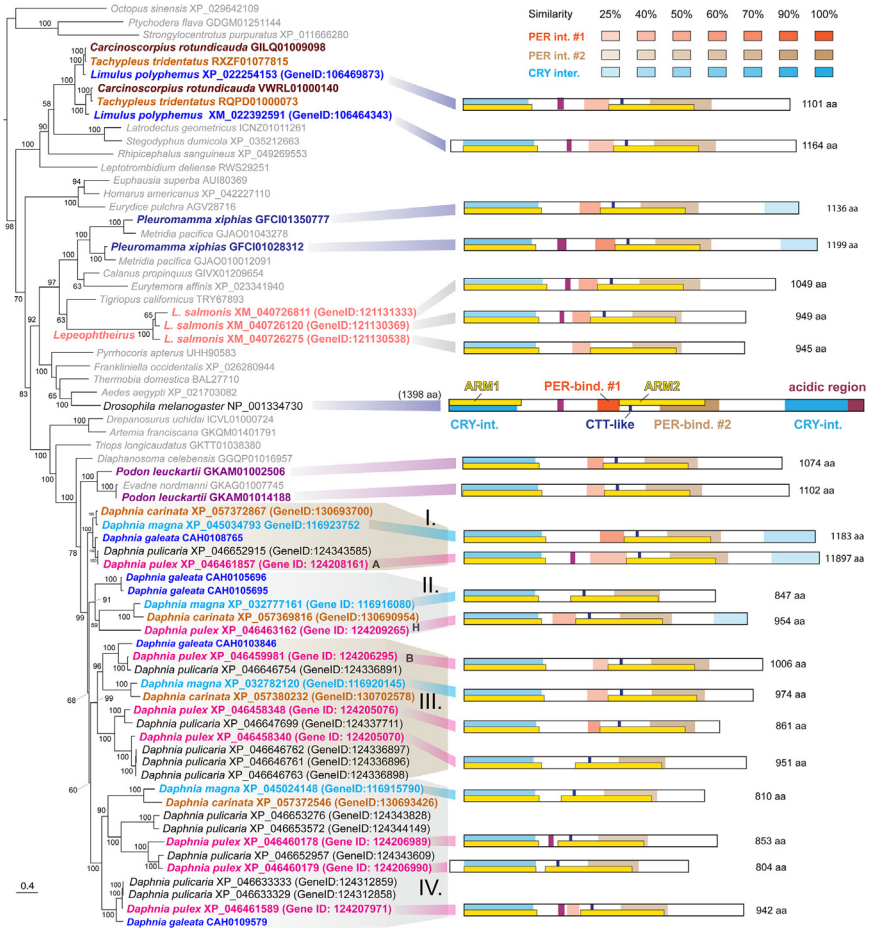


Figure 3. Independent dTIM gene duplications were detected in horseshoe crabs (*Limulus*, *Tachypleus*, and *Carcinoscorpius*), copepods (*Pleuromamma*, *Metridia*, and *Lepeophtheirus*), and branchiopods (*Podon* and *Daphnia*)

The phylogenetic tree inferred using RAxML, with bootstrap values (500 replicates, shown in %), indicates that these duplications are lineage specific. In the *Daphnia* genus, four primary dTIM clusters (I.-IV.) are present across all species. Additional species-specific duplications resulted in 8 paralogs in *D. pulex* and 11 in *D. pulicaria*. Protein models show conserved similarities in PER-binding and CRY-interaction domains. Labels A, B, and H for *D. pulex* paralogs follow the terminology of Bernatovicz et al.[76]

detected for paralogs from cluster I across all four analyzed *Daphnia* species, further supporting the hypothesis that this paralog is the closest to the ancestral dTIM (Figure 4E).

To further explore the potential causes of dTIM variability, we examined dTIM-interacting proteins. PER, a key partner in the *Drosophila* clock, is present in nearly all Bilateria, with the unique exceptions of two basal deuterostomian groups: Hemichordata and Echinodermata. Notably, even the loss of PER did not affect the substitution rate of dTIM. Conversely, the highest substitution rates observed in dTIM from aphids and phylloxera were mirrored by increased variability in PER within these species, suggesting coevolution of the two proteins (Figures 4C and 4D). The dTIM duplication in *Daphnia* did not affect the substitution rate in PER, which reached values comparable to those observed in branchiopods with only one dTIM paralog. Duplications of the *per* gene are rare. In addition to the well-established *per* duplication in vertebrates, *per* duplications have been identified in three lineages: horseshoe crabs (*Merostomata*), *Podon leucartii*, and *Pleuromamma xiphias* (Figure 4E; Table S4).

The loss of *d-tim* is consistently accompanied by the loss of *d-cry*, as observed in Hymenoptera, Isoptera, Amphipoda, Scorpiones, and Chordata (Figure 4D). However, several instances exist where *d-cry* has been independently lost while *d-tim* remains present. These cases include certain Coleoptera (most Polyphaga, including the red flour beetle *Tribolium*), a subset of Blattodea (including *Periplaneta americana*[77]), and a subset of Heteroptera. Notably, within Heteroptera, *d-cry* has been lost at least twice independently, as evidenced by its complete absence in the order Pentatomomorpha and parts of Cimicomorpha (family Cimicidae). This conclusion is further supported by synteny analysis. For example, *Lethocerus indicus* (Nepomorpha), representing an ancestral lineage of Heteroptera, retains the *d-cry* gene. Comparative analysis of the *d-cry* locus in *Lethocerus* and *Riptortus pedestris* suggests that *d-cry* loss in the latter species resulted from a combination of several chromosomal inversions. In contrast, different chromosomal rearrangements led to the loss of *d-cry* in the genus *Cimex*. Meanwhile, the sister lineage of Cimicomorpha, Miridae, still retains the *d-cry* gene (Figures 5A, 5B, and S9).

Our analysis reveals that the loss of *d-cry* is consistently accompanied by the loss of *jetlag* in several taxa, including Hymenoptera, a subset of Coleoptera, Phthiraptera (a subset of Psocodea), Amphipoda, Parasitiformes (a group within Acari that includes ticks), and a subset of Heteroptera (Figure 4D). Remarkably, the two independent losses of *d-cry* in Heteroptera (Pentatomomorpha and Miridae) are also associated with the loss of the *jetlag* gene. Gene synteny analysis suggests that independent chromosomal rearrangements underlie these losses of *jetlag* in Pentatomomorpha and Cimicidae (Figures 5C and S9). Despite the loss of *d-cry*, the organization of dTIM domains (Figure S2A) appears to be very similar among *Lygus* (where *d-cry* is still present), *Cimex* (where *d-cry* has been lost), and *Pyrrhocoris* (where *d-cry* has been independently lost). Notably, dTIM in the early-branching species *Lethocerus indicus* differs significantly from these species as it retains an elongated CRY-interaction domain at its C-terminus.

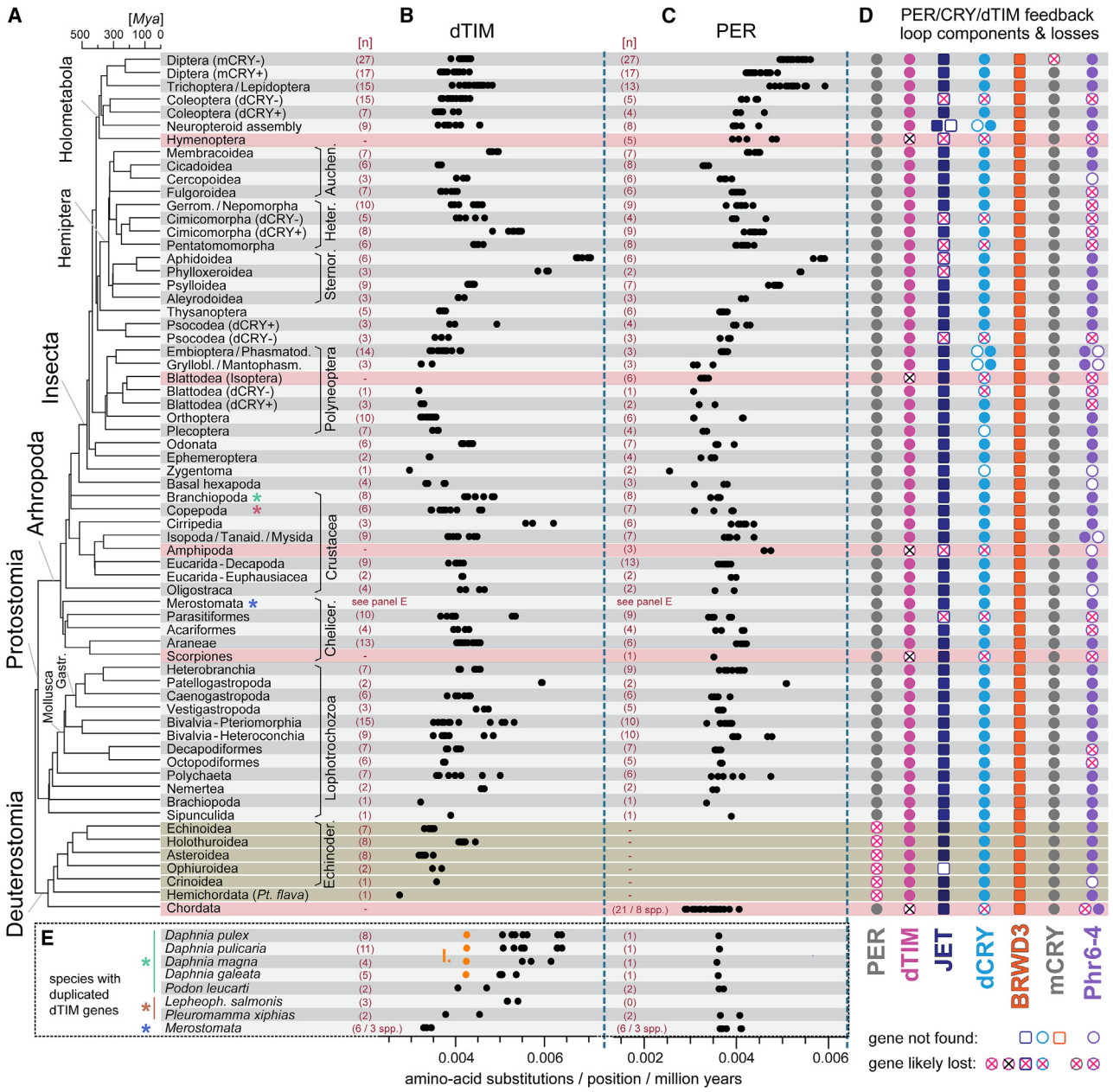


Figure 4. Coevolution of dTIM with partner proteins

(A) The tree illustrates the current view of phylogeny among the analyzed groups of Bilateria.

(B and C) Numerical depiction of protein changes, plotted as substitutions per position per million years of evolution. Each dot corresponds to a protein from a species within a particular higher taxon: (B) dTIM, and (C) PER, with the number of values in each group and category indicated by a magenta n.

(D) Presence (filled circle), absence (empty circle), and loss (x) of the key components of the PER/mCRY/dTIM negative feedback loop in specific lineages. Loss of dTIM is highlighted by a brick-red horizontal background, while loss of PER (in Hemichordata and Echinodermata) is shown with a gray-brown background. For accession numbers and exact values, see and Tables S2, S4, and S5; for evidence supporting gene loss, see Table S1.

(E) Species with multiple dTIM paralogs are shown, each plotted on a separate line. Orange dots represent *Daphnia* paralogs from Cluster I, proposed as the most complete and ancestral dTIM variants in this genus.

The losses of *d-cry* and *jetlag* do not align perfectly across all bilaterian species. For instance, *jetlag* (annotated as F box and LRR protein 15 in chordates) is retained in Chordata and Blattodea (Figure 4D). Termites (Isoptera), a subset of Blattodea that have progressively lost *d-cry* and even *d-tim*, still possess *jetlag*. Interestingly, a unique and substantial modification of *d-tim* in Aphidomorpha (Phyloxeroidea and Aphidoidea) is not associated with the loss of *d-cry*. The *d-cry* gene is present in all Aphidomorpha species, and dCRY protein was detected in the brains of the pea aphid *Acyrthosiphon pisum* via

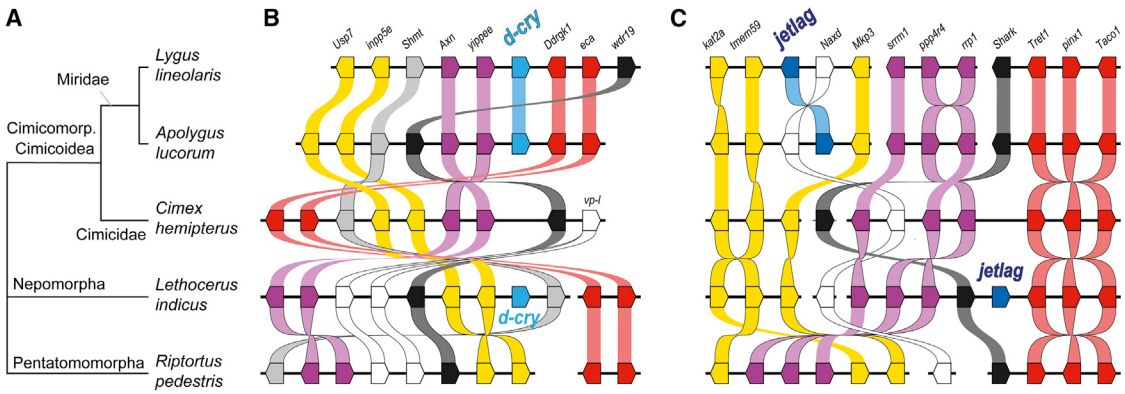


Figure 5. Gene synteny illustrates the independent loss of *Drosophila*-type *cryptochrome* (*d-cry*) and *jetlag* (*jet*) genes in *Cimex hemipterus* and *Riptortus pedestris*

(A) The phylogenetic relationship among five analyzed (pan)heteropteran species is presented alongside (B) *d-cry* and (C) *jet* gene syntenies. The genomic positions of *d-cry* and *jet* in *L. lineolaris*/*A. lucorum* (Miridae) and *L. indicus* (Nepomorpha) differ, as indicated by distinct sets of neighboring protein-coding genes. Although synteny genes can be mapped to *C. hemipterus* (Cimicidae) and *R. pedestris* (Pentatomorpha), their composition and orientation vary, supporting the independent losses of *d-cry* and *jet*. Horizontal black lines represent genomic contigs/scaffolds (not to scale). Arrow-like boxes indicate protein-coding genes and their orientation. Orthologous genes are color-coded and interconnected. For detailed synteny depictions, see Figure S9; Tables S6 and S7.

immunohistochemistry.[78] This suggests that the rapid evolution of dTIM in Aphidomorpha is lineage-specific, as their sister groups, Psylloidea and Aleyroidea, possess dTIM with a generally conserved domain structure (Figure S2A). Notably, the loss of *jetlag* corresponds precisely to the accelerated accumulation of changes in dTIM within Aphidomorpha. Moreover, the PER protein also shows an elevated substitution rate in this lineage (Figure 4).

The C-terminal tail (CTT) is a key feature of dCRY in *Drosophila*, where it binds alongside an FAD cofactor in the protein.[79] The CTT is highly conserved across all dCRYs, although in some species, such as hemichordates and aphids, a substantial extension is present (Figure S3).

Changes in gene structure underlie dTIM evolution

To gain further insight into dTIM changes, we compared the *d-tim* gene structure across selected representative species. Even visual inspection revealed a complex remodeling of *d-tim* exons in most Crustacea, with only the early-branching *Darwinula* showing a structure similar to both the horseshoe crab *Limulus* (Chelicerata) and the firebrat *Lepisma* (Insecta). Similarly, exon fusion in *Orchesella* (Collembola) appears to be lineage-specific, which is why these species are presented in a separate box (Figure 6). Importantly, a generally conserved exon pattern is observed in the majority of the species analyzed, with several exons traceable from echinoderms to insects. Consistent with the variability in dTIM protein, the most conserved regions encompass the exons encoding the ARM1 and ARM2 domains. In the case of ARM1, a Holometabola-specific fusion of four exons is observed, while a similar, though less pronounced, exon fusion can be identified in ARM2.

The dTIM protein region between the ARM1 and ARM2 domains exhibits significant variability. Notably, some level of similarity is detected in insects, particularly in the PER-binding #1 region; however, beyond insects, this portion of dTIM shows high divergence. Similarly, the corresponding gene structure is not conserved, as exons encoding this region vary in length and are only minimally conserved within insects and show no detectable similarity in Deuterostomia or ancestral Protostomia, including lineages such as Mollusca, Nemertea, Polychaeta, and Chelicerata. Within insects, the variability of the inter-armadillo region is most pronounced in *A. pisum* and *D. vitifoliae*, where similarity falls below the plotted 31% threshold. This highlights the remarkably high degree of modification of dTIM in both Phylloxeroidea and Aphidoidea (Figures 6 and S4). Furthermore, early work identified alternative splicing of the *d-tim* exon encoding the first part of the PER-binding #1 region in two aphid species[80]; this splicing is distinct from several alternative splicing isoforms detected in *D. melanogaster* (Figure 6, top).

The C-tail of dTIM, located downstream of the ARM2 domain, contains a few conserved exons. In *D. melanogaster*, a Drosophilidae-specific fusion to the upstream exon has been detected (Figure 7). This region encodes the TIMELESS C domain (Pfam: PF05029, Figure S10), where similarity is observed even between dTIM and mTIM.[81] When mutated, various degrees of temperature-dependent changes in the free-running period have been described.[26,72] Interestingly, a major truncation in Aphidomorpha removes the C-tail and even the otherwise highly conserved TIMELESS C domain (Figure 6).

The high variability in the protein tails downstream of the TIMELESS C domain is evident at the gene structure level, where only two exons are detected in Deuterostomia (*Strongylocentrotus* and *Acanthaster*), Mollusca (*Crassostrea*), Annelida (*Owenia*), and Chelicerata (*Limulus*). In insects, tail expansion correlates with similarities to the predicted CRY-interaction domain in *D. melanogaster.* This domain has been independently lost in *P. apterus* and Aphidomorpha. The latter is a sister group to *Bemisia tabaci*, a species in which only low similarity (19%) to the CRY-interaction domain is detected (Figure 6). A complete C-tail is found in *Lethocerus indicus*, a sister taxon

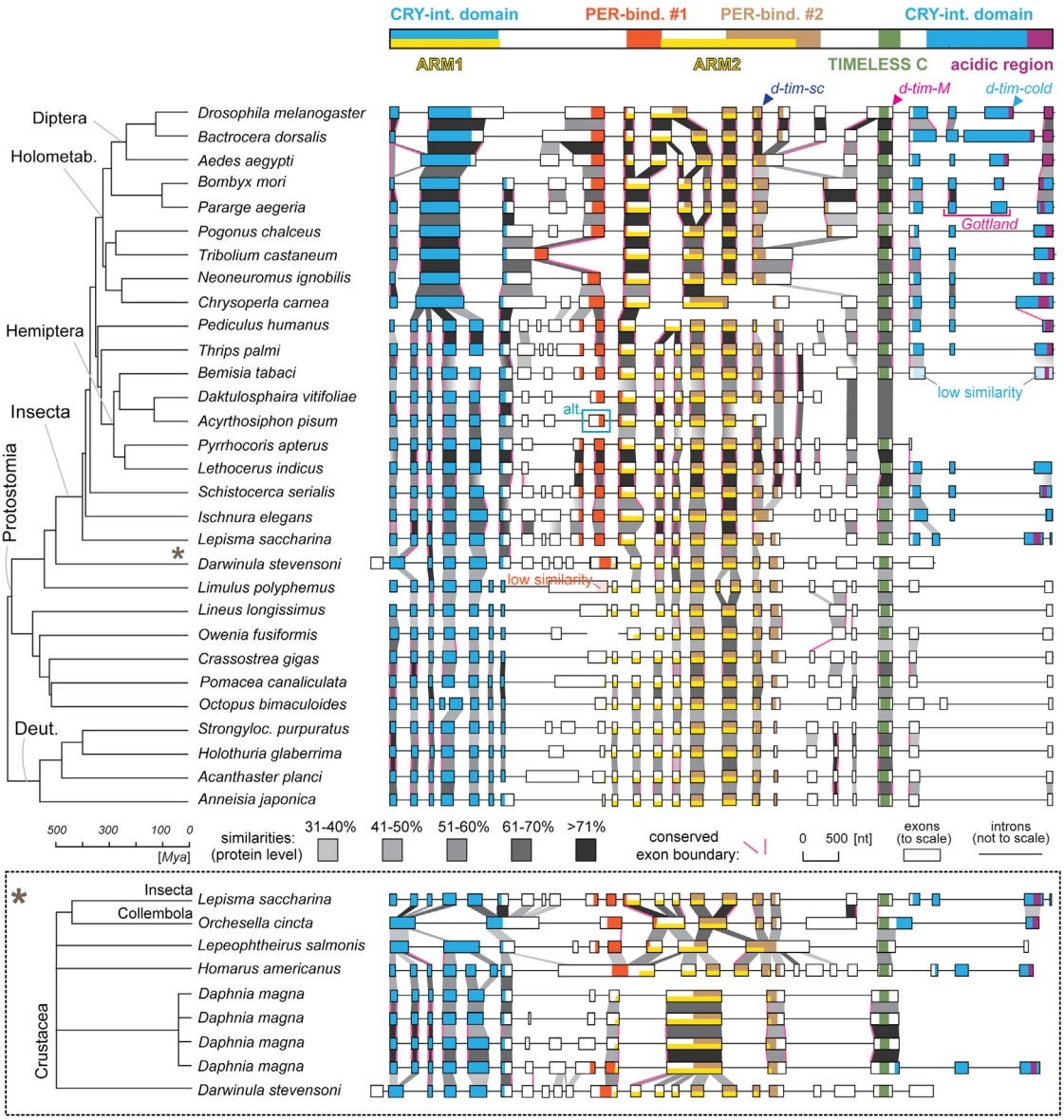


Figure 6. Gene structures and protein domains of dTIMs in Bilateria

Introns, exons, and homologous regions illustrate the conserved and variable parts of the *d-tim* gene. Vertical blocks in shades of gray represent sequence similarities at the protein level, while vertical magenta lines indicate conserved exon boundaries. Protein domains and functional regions are annotated based on their similarity to *Drosophila* dTIM (shown at the top). For clarity, the complexity specific to Crustacea is presented in a separate box (*), where *Lepisma* and *Darwinula* appear in both analyses. The highest variability is observed in the PER-binding region #1 and the CRY-interaction domain located at the C-terminus. Within Insecta, the CRY-interaction domain has been lost in *Acyrthosiphon pisum* and *Daktulosphaira vitifoliae*, and independently in *Pyrrhocoris apterus*. In *Pararge aegeria*, a unique geographical allele (Gottland) results from a deletion spanning two exons (highlighted with a magenta horizontal line). Thresholds for plotting specific regions were as follows: similarities greater than 40% for ARM2, and PER-binding regions (except for *Limulus*, where similarity is 38%), and greater than 20% for the CRY-interaction domain (with the lowest similarity observed in *Bemisia*, at 19%). Intron retention in three *Drosophila* splicing isoforms is marked with arrowheads, and alternative exon removal in *A. pisum*[80] is indicated by a turquoise rectangle with a note "alt".

to *P. apterus*, where truncation removes the entire CRY-interaction domain downstream of the TIMELESS C domain. In Lepidoptera, gene analysis of the speckled wood butterfly (*P. aegeria*) reveals that the loss of two entire exons in the Gottland allele results in a substantial in-frame deletion in the CRY-interaction domain. In contrast, the expansion of dTIM in Diptera lacking mCRY (mCRY- Diptera) is due to exon extensions, as illustrated in *Bactrocera dorsalis* and *Musca domestica* (Figures 6 and 7).

N-terminal extensions are rare in dTIM proteins

Gene model analysis suggests that *Darwinula d-tim* may encode a protein extended by 61 amino acids at the N-terminus. Similar N-terminal extensions were found in

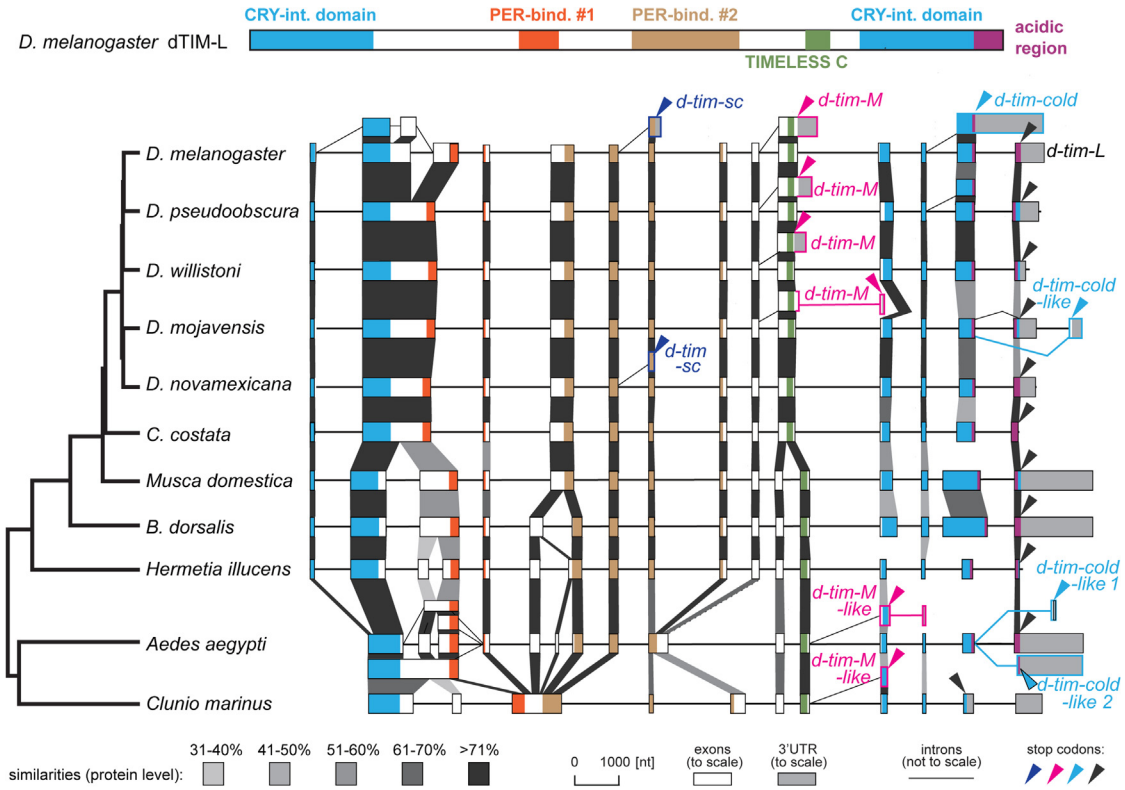


Figure 7. *d-tim* gene structures in Diptera with highlighted alternative splicing isoforms

The similarities and dTIM domains are depicted identically as in Figure 6. Four major splicing isoforms exist in *D. melanogaster*: The longest isoform is the canonical *d-tim-L*. Another isoform, *d-tim-cold* (turquoise), is produced by the retention of the last intron at low temperatures, resulting in the absence of the last exon. At high temperatures, retention of another intron generates the *d-tim-M* isoform (magenta). Additionally, the *d-tim-sc* (blue) isoform is abundant at low temperatures. Identical or comparable isoforms have been observed in some other dipteran species. For example, modified exon splicing produces *tim-cold*-like isoforms in *D. mojavensis* and *Aedes*. In *Aedes* and *Clunio*, alternative splicing occurs one exon downstream of the position typical for *d-tim-M* isoform. Therefore, we refer to this variant as *d-tim-M*-like isoforms. Despite this difference, the effect on *d-tim* mRNA destabilization may be similar to the impact of *d-tim-M* in *D. melanogaster*.

Macrobrachium (Crustacea) and some dTIM paralogs of *Daphnia pulex* and *Limulus* (Figures S2B and S6). Importantly, even a shorter 23 amino acid extension in *D. melanogaster* has a significant impact on its interaction with dCRY, resulting in lower sensitivity to light, including altered behavioral rhythmicity under constant dim light.[23,51–53] In *D. melanogaster*, the N-terminal extension is encoded by an in-frame alternative start codon in the *ls-tim* allele.[50] Since this allele is believed to be of recent origin, specific to *D. melanogaster*,[41] it was surprising that BLAST-P identified several L-dTIM protein sequences in GenBank for various *Drosophila* species (*D. ananassae* EDV31195; *Drosophila kikkawai* XP_041631545). However, closer inspection revealed that these protein sequences are clear annotation errors, as the available genomes only encode short (s-dTIM) proteins.

Conserved features of the alternative splicing pattern in Diptera

Alternative splicing is a key regulatory mechanism of *d-tim* in *D. melanogaster* at both low and high temperatures.[82–86] Splicing isoforms include *d-tim-sc, d-tim-M*, and *d-tim-cold*. The full-length *d-tim* transcript has been designated as *d-tim-L*,[85] which should not be confused with an allele encoding proteins extended at the N-terminus by 23 aa. This extended protein has been labeled as L-TIM[41] and referred to in this article as L-dTIM.

The combination of genomic sequences and available transcriptome shotgun assemblies (TSAs) allowed us to investigate whether a similar splicing pattern exists in other Diptera (Figure 7). The *d-tim-sc* (*tim-short and cold*), resulting from an alternative polyadenylation site, was identified only in *D. novamexicana* and confirmed in *D. melanogaster*. A second low-temperature-dependent splicing regulation of *d-tim* in *D. melanogaster* is the retention of the last intron (located between the penultimate and final exon), which results in a premature stop codon that removes the acidic region at the C-terminus. This portion of the dTIM protein is phosphorylated by Casein Kinase 2, which in turn inhibits nuclear export and impacts CLOCK transcriptional activity.[74] Importantly, three out of four phosphorylation sites are conserved in the full-length protein (dTIM-L) in the analyzed Diptera, as well as in other insects such as the cricket *Gryllus bimaculatus* and the most ancestral insect species, the firebrat *Thermobia domestica*. Notably, alternative splicing severely impacts this region

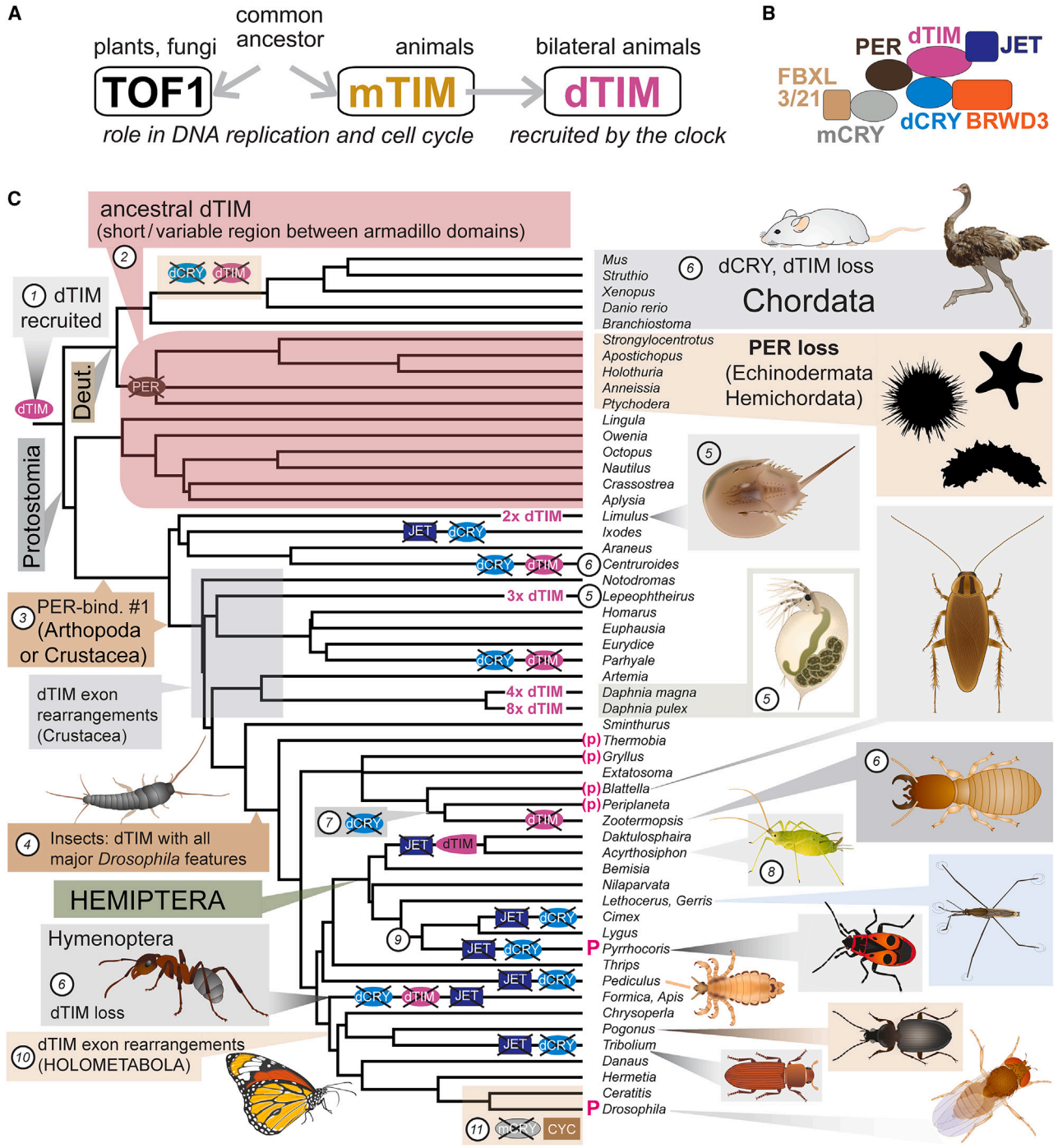


Figure 8. Summary of major changes in dTIM evolution

(A) dTIM originated from mTIM by gene duplication in the common ancestor of Bilateria.

(B) Schematic depiction of key components in the negative PER/mCRY/dTIM feedback loop in circadian clocks.

(C) Evolutionary changes in dTIM and its interacting clock components were mapped onto the phylogeny of Bilateria: (1) dTIM was recruited in the common ancestor of Protostomia and Deuterostomia and (2) dTIM is retained in ancestral lineages of both Protostomia and Deuterostomia. (3) The PER-binding domain #1 evolved either in Arthropoda or Crustacea. (4) All features characteristic for *Drosophila* dTIM appeared in ancestral insect dTIM, and these key features are retained in dTIM of most (but not all) insects. (5) Lineage-specific duplications resulted in multiple dTIM paralogs in *Limulus* (and other horseshoe crabs), *Lepeophtheirus*, and *Daphnia*. (6) dTIM has been lost in several lineages, including Chordata, scorpions (*Centruroides*), Amphipoda (*Parhyale*), Hymenoptera, and Isoptera (*Zootermopsis*). These losses consistently coincide with the absence of dCRY. (7) In *Blattodea*, the loss of dCRY precedes the loss of dTIM in termites. (8) Changes in the dTIM sequence in aphids and *Phylloxera* correlate with the loss of JET. Independent loss of dCRY and JET occurred in two Heteroptera lineages

(legend continued on next page)

(Figure S11). Although the last intron is generally conserved across several species, splicing was minimally affected in *Musca*.[87] However, we identified two distinct alternative splicing patterns in the mosquito *Aedes aegypti* and the fly *Drosophila mojavensis* that either remove or modify the C-terminal part of dTIM. In the first pattern, an alternative splice site within the exon removes a portion of the protein. In the second, the penultimate canonical exon is fused with an alternative last exon located at the very end of the gene, resulting in an altered amino acid sequence.

The last major dTIM isoform, *medium* (*d-tim-M*), is expressed at high temperatures and is characterized by the retention of an intron in the center of the mRNA. The resulting transcript does not appear to be translated; instead, this mechanism reduces the pool of dTIM-coding mRNAs via nonsense-mediated decay (NMD).[85] A comparable splicing pattern was identified in three *Drosophila* species, with a slight modification observed in *D. mojavensis*. Interestingly, in *Aedes* and *Clunio*, alternative splicing is detected one exon downstream of the canonical *d-tim-M* isoform (Figure 7). Although no experimental data are available, and it is unknown how temperature affects this splicing pattern, the similarity with the *d-tim-M* isoform suggests that a potentially comparable regulatory mechanism might also be present in *Clunio* and *Aedes*.

DISCUSSION

Although the major role of circadian clocks—measuring 24 h—is conserved among organisms, the molecular machinery achieving this task appears to be variable among Bilateria, a group relying on homologous clock genes. This study focused on the variability and evolution of one of these proteins, dTIM, by exploiting available genomic and transcriptomic data (see Figure 8 for a summary). While dTIM originated from mTIM in early Bilateria, its functional participation in the circadian clock is confirmed in only several insect species representing the orders Diptera, Heteroptera, Orthoptera, and Zygentoma. However, its role varies even among these groups. In *Drosophila*, dTIM is an essential clock component, whereas in other species, such as the linden bug *P. apterus*, the cricket *G. bimaculatus*, and the firebrat *T. domestica* (which belongs to the most ancestral insect group, Zygentoma), it serves as a mere modulator of clock pace. A key feature of dTIM in *Drosophila* is its interaction with the PER protein, which was mapped in the seminal study using the *Drosophila Schneider 2* cell line.[21] Out of the two identified PER-binding sites, one is conserved in dTIMs across all studied lineages. Since a substantial part of this site overlaps with the ARM2 domain, it is difficult to determine which of these functions is responsible for the high conservation. The other site, PER-bind #1, is well conserved in Insecta and some Crustacea (although the latter also contain species with variable and divergent dTIM sequences), and it is still identified in Chelicerata, albeit with lower similarity.

In Annelida and Mollusca, the protein region corresponding to PER-bind #1 is substantially modified, inviting us to speculate whether dTIM even interacts with PER in these organisms. In Hemichordata and Echinodermata, two lineages that lost PER completely, the level of modifications in the region between ARM1 and ARM2 is comparable to that detected in Mollusca and Annelida. Furthermore, low substitution rates per amino acid position are observed in dTIM of Mollusca, Annelida, Hemichordata, and Echinodermata. Thus, even the absence of PER had minimal impact on dTIM evolution in early branching Deuterostomia.

Not much is known about the clocks of these classes of organisms, but two independent studies on *Crassostrea gigas* (Mollusca) and *Platynereis dumerilii* (Annelida) report the presence of the full core set of circadian oscillator gene orthologs from both *Drosophila* and *mouse*.[88,89] The first study proposes a putative molecular oscillator model for *C. gigas* with intermediate features between the *Drosophila* and *mouse* clocks, in which dTIM interacts with PER and CRY proteins to inhibit the CLOCK/BMAL complex. Our findings suggest that this model may not be entirely accurate or that PER-binding domain #2 and CRY-binding domains are sufficient for TIM and PER to interact in these organisms. However, it is important to note that these are marine animals, exposed to diverse environmental conditions and possessing multiple types of rhythms within the same species (circadian, circatidal, circalunar, etc.). As a result, elucidating specific molecular interactions in such a complex system is challenging. Two major theories have emerged: either different rhythms rely on similar molecular components, or independent clocks with distinct molecular mechanisms coexist within a single organism.[90] In case of crustaceans, as illustrated by *Parhyale*, circadian clock genes also orchestrate the circatidal rhythm.[91]

A second role of dTIM, light-mediated entrainment, is documented in great detail in *D. melanogaster*, where the light-triggered dTIM-dCRY interaction—elegantly described in the yeast system, later in *Drosophila* S2 cells and flies, and most recently by Cryo-EM[23,92,93]—results in ubiquitination and subsequent degradation of both proteins, a process in which JET and BRWD3 play a key role.[29,34] Except for BRWD3, a gene essential during development,[94] multiple well-supported cases of gene loss are reported for dTIM, dCRY, and JET. Usually, but not always, several components are lost simultaneously. For example, the absence of dTIM is always accompanied by the loss of dCRY (at least 5 independent cases).

However, the absence of dCRY does not always coincide with the loss of dTIM. A simultaneous absence of dCRY and JET is reported in 7 independent cases. These examples of gene losses suggest that once one of the components is missing, the system is compromised, and additional components are often quickly lost. Particularly intriguing is the independent loss of both genes in *Pentatomomorpha* and *Cimicidae*. Notably, the absence of dCRY and JET aligns with the relatively small impact of light as

(9), suggesting a mutation in their common ancestor that impacted dCRY and JET function in the circadian clock. (10) A major fusion of *d-tim* exons is observed in Holometabolous insects. (11) The loss of mCRY and simultaneous modification of BMAL1 into CYC dates to Cyclorrhapha (e.g., *Drosophila*). Species with circadian clock phenotypes assessed from *d-tim* genetic mutants are marked with a magenta 'P'. Phenotypes derived from *d-tim* RNAi silencing experiments are indicated with '(p)'.

the entraining cue in the linden bug *P. apterus*.[95] It will be interesting to determine the light entrainment capacity in Miridae, a lineage where dCRY and JET are still present. In general, RNA interference seems to work reliably in true bugs, making it possible to explore the specific roles of these proteins. The loss of dCRY and JET is typical for a large portion of Coleoptera, including the red flour beetle *Tribolium castaneum*. Although this species lives in environments with minimal light, its behavior is rhythmic both under constant darkness and constant light, albeit not very robust.[96]

In the case of aphids and phylloxera, JET has been lost, while dCRY is still present. It is unclear whether the loss of JET triggered the accumulation of changes in dTIM, further permitting the loss of several exons at the 3′ end of the gene, or if some changes are influenced by coevolution between dTIM and PER proteins. Nevertheless, the major truncation of the C-terminal region suggests that the interaction between dTIM and dCRY is likely affected. Supporting this hypothesis, immunohistochemical localization of dCRY in the pea aphid (*A. pisum*) revealed that the protein remains stable upon light exposure.[78] In two aphid species, alternative splicing of *d-tim* transcripts was identified.[80] However, this splicing pattern differs from the alternative splicing observed in *Drosophila* (Figure 6). Interestingly, the affected exon encodes part of the PER-bind #1 site, which could strongly impact the interaction properties of the resulting dTIM protein with PER. It will be interesting to see whether this alternative splicing is affected by temperature. A recent study indicates that, at the genetic level, the major divergence in the aphid clock primarily involves loss of JET, while other circadian clock genes are present.[97]

Particularly interesting is the gradual loss of clock components in Blattodea. All Blattodea appear to have lost the *6-4 photolyase* (*phr6-4*) gene, an *m-cry* paralog that is not involved in the circadian clock. While one major lineage, Blaberoidea (represented by *Blattella*), still possesses *d-cry*, this gene has been lost in *Periplaneta* (representing Blattoidea). Termites (*Isoptera*), which from a phylogenetic perspective are a subset of Blattoidea,[98] have also lost *d-tim*. However, at least some early-branching species, such as *Porotermes*, may still possess this gene.

There are several lineages for which transcriptomic data suggest possible gene loss; however, well-assembled genomes, which are necessary to confirm such claims, are not available. For example, dCRY has not been identified in any representatives of Plecoptera or Raphidioptera. It is also important to note that some groups of organisms are insufficiently represented in GenBank, which hampers meaningful reconstruction of changes in the clock setup. For instance, Amphipoda, a group of Crustacea represented by *Parhyale hawaiensis*, lacks both dCRY and dTIM.[91] To illustrate this gap in the available data, note that *Eurydice pulchra*, the closest species with dCRY and dTIM, diverged from amphipods more than 350 million years ago—around the same time when the ancestors of humans and *Xenopus* separated (Figure 8 and references used to refine the tree).

What are the potential consequences and selective forces driving the observed variability in the negative loop components of the circadian clock? Changes and losses affecting dTIM, dCRY, and JET are expected to influence the sensitivity of clocks to light, although parallel light input pathways involving opsins are also utilized.[18] Reduced sensitivity to light might confer advantages under long photoperiods; however, the entrainment capacities of the clock may also depend on the neuroanatomy of the system, as demonstrated in flies from northern latitudes.[99] Therefore, conducting functional analyses in specific organisms is essential to uncover the exact importance of particular genes.

It is important to note that the free-running period is a valuable marker of circadian clock properties. However, the actual selection pressure may act on adjusting the daily preference for locomotor activity. Clock genes also play a role in photoperiodic time measurement. Therefore, in cases such as geographic variability in the free-running period, it can be challenging to disentangle the effects of selection pressure on daily activity profiles from potential selection pressures acting on seasonal timing mechanisms. Both dCRY+ and dCRY- insects exhibit robust photoperiodic timers.[38,44,100–102] The interplay between daily anticipatory mechanisms and seasonal responses introduces additional selection pressures on circadian clock genes. Moreover, it is crucial to highlight that some circadian clock genes are also involved in regulating rhythms beyond the 24-h range, such as the ∼13-h tidal clock and circalunar (monthly) clocks.[89,91,103,104]

Limitations of the study

A cornerstone of studies like this is access to high-quality, well-assembled genomes with gene annotations supported by comprehensive transcriptomic data. Alternatively, transcriptomes that capture a diverse range of tissues and conditions serve as invaluable tools. However, predicting the complete structure of genes in such cases is challenging, if not impossible. Consequently, this study is inherently shaped by the limitations of data availability and biased toward well-sequenced groups of organisms. For example, insects—at least a significant subset of them—provide reliable resources to investigate connections among the studied clock proteins. In contrast, several phylogenetically important groups, including numerous crustacean lineages, polychaeta, brachiopods, sipunculids, and acorn worms, remain poorly represented in genomic and transcriptomic databases.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, David Dolezel at david.dolezel@entu.cas.cz.

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank Martina Hajdušková (www.biographix.cz) for the color schemes used in Figure 8 and phylopic (https://www.phylopic.org/images) for schemes of starfish, sea urchin, and sea cucumber. We thank the current members of our laboratory for their supportive and nice working atmosphere. The work was supported by Czech Science Foundation (GACR, 22-10088S) to D.D.

AUTHOR CONTRIBUTIONS

D.D. designed and guided the study, E.B. performed the protein comparisons, I.F. performed the phylogenetic analyses, P.C. reconstructed gene models under the supervision of V.S. who also performed synteny analyses. D.D. wrote the paper and finalized the figures with input from all coauthors.

DECLARATION OF INTERESTS

The authors declare no competing interest.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Gene identification and datasets
  - Phylogenetic analyses
  - Gene loss
  - Prediction of protein domains
  - The length of the entire protein, ARM1-ARM2 region, and C-tail length
  - Substitutions per amino acid per million years
  - Gene models
  - Exon homology, similarity, and dTIM domain localization in gene models
  - Gene syntenies
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.112338.

Received: October 8, 2024
Revised: February 4, 2025
Accepted: March 28, 2025
Published: April 2, 2025

REFERENCES

1. Bhadra, U., Thakkar, N., Das, P., and Pal Bhadra, M. (2017). Evolution of circadian rhythms: from bacteria to human. Sleep Med. *35*, 49–61. https://doi.org/10.1016/j.sleep.2017.04.008.
2. Dunlap, J.C. (1999). Molecular bases for circadian clocks. Cell *96*, 271–290. https://doi.org/10.1016/s0092-8674(00)80566-8.
3. Edgar, R.S., Green, E.W., Zhao, Y., van Ooijen, G., Olmedo, M., Qin, X., Xu, Y., Pan, M., Valekunja, U.K., Feeney, K.A., et al. (2012). Peroxiredoxins are conserved markers of circadian rhythms. Nature *485*, 459–464. https://doi.org/10.1038/nature11088.
4. Mendoza-Viveros, L., Bouchard-Cannon, P., Hegazi, S., Cheng, A.H., Pastore, S., and Cheng, H.Y.M. (2017). Molecular modulators of the circadian clock: lessons from flies and mice. Cell. Mol. Life Sci. *74*, 1035–1059. https://doi.org/10.1007/s00018-016-2378-8.
5. Kwiatkowski, E.R., and Emery, P. (2024). Cnidarians are CLOCKing in. Elife *13*, e98512. https://doi.org/10.7554/eLife.98512.
6. Aguillon, R., Rinsky, M., Simon-Blecher, N., Doniger, T., Appelbaum, L., and Levy, O. (2024). CLOCK evolved in cnidaria to synchronize internal rhythms with diel environmental cues. Elife *12*, RP89499. https://doi.org/10.7554/eLife.89499.
7. Thakkar, N., Giesecke, A., Bazalova, O., Martinek, J., Smykal, V., Stanewsky, R., and Dolezel, D. (2022). Evolution of casein kinase 1 and functional analysis of new doubletime mutants in Drosophila. Front. Physiol. *13*, 1062632. https://doi.org/10.3389/fphys.2022.1062632.
8. Tomioka, K., and Matsumoto, A. (2015). Circadian molecular clockworks in non-model insects. Curr. Opin. Insect Sci. *7*, 58–64. https://doi.org/10.1016/j.cois.2014.12.006.
9. Allada, R., White, N.E., So, W.V., Hall, J.C., and Rosbash, M. (1998). A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless. Cell *93*, 791–804. https://doi.org/10.1016/S0092-8674(00)81440-3.
10. Rutila, J.E., Suri, V., Le, M., So, W.V., Rosbash, M., and Hall, J.C. (1998). CYCLE is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of Drosophila period and timeless. Cell *93*, 805–814. https://doi.org/10.1016/S0092-8674(00)81441-5.
11. Michael, A.K., Stoos, L., Crosby, P., Eggers, N., Nie, X.Y., Makasheva, K., Minnich, M., Healy, K.L., Weiss, J., Kempf, G., et al. (2023). Cooperation between bHLH transcription factors and histones for DNA access. Nature *619*, 385–393. https://doi.org/10.1038/s41586-023-06282-3.
12. Tumova, S., Dolezel, D., and M, J. (2024). Conserved and Unique Roles of bHLH-PAS Transcription Factors in Insects – From Clock to Hormone Reception. J. Mol. Biol. *436*, 1–25. https://doi.org/10.1016/j.jmb.2023.168332.
13. Kume, K., Zylka, M.J., Sriram, S., Shearman, L.P., Weaver, D.R., Jin, X., Maywood, E.S., Hastings, M.H., and Reppert, S.M. (1999). mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. Cell *98*, 193–205. https://doi.org/10.1016/S0092-8674(00)81014-4.
14. Putker, M., Wong, D.C.S., Seinkmane, E., Rzechorzek, N.M., Zeng, A., Hoyle, N.P., Chesham, J.E., Edwards, M.D., Feeney, K.A., Fischer, R., et al. (2021). CRYPTOCHROMES confer robustness, not rhythmicity, to circadian timekeeping. EMBO J. *40*, e106745. https://doi.org/10.15252/embj.2020106745.
15. Yuan, Q., Metterville, D., Briscoe, A.D., and Reppert, S.M. (2007). Insect cryptochromes: Gene duplication and loss define diverse ways to construct insect circadian clocks. Mol. Biol. Evol. *24*, 948–955. https://doi.org/10.1093/molbev/msm011.
16. Emery, P., So, W.V., Kaneko, M., Hall, J.C., and Rosbash, M. (1998). CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity. Cell *95*, 669–679. https://doi.org/10.1016/S0092-8674(00)81637-2.
17. Emery, P., Stanewsky, R., Helfrich-Förster, C., Emery-Le, M., Hall, J.C., and Rosbash, M. (2000). Drosophila CRY is a deep brain circadian photoreceptor. Neuron *26*, 493–504. https://doi.org/10.1016/S0896-6273(00)81181-2.
18. Helfrich-Forster, C., Winter, C., Hofbauer, A., Hall, J.C., and Stanewsky, R. (2001). The circadian clock of fruit flies is blind after elimination of all known photoreceptors. Neuron *30*, 249–261. https://doi.org/10.1016/S0896-6273(01)00277-X.
19. Collins, B., Mazzoni, E.O., Stanewsky, R., and Blau, J. (2006). Drosophila CRYPTOCHROME is a circadian transcriptional repressor. Curr. Biol. *16*, 441–449. https://doi.org/10.1016/j.cub.2006.01.034.
20. Stanewsky, R., Kaneko, M., Emery, P., Beretta, B., Wager-Smith, K., Kay, S.A., Rosbash, M., and Hall, J.C. (1998). The cry(b) mutation identifies cryptochrome as a circadian photoreceptor in Drosophila. Cell *95*, 681–692. https://doi.org/10.1016/S0092-8674(00)81638-4.

21. Saez, L., and Young, M.W. (1996). Regulation of nuclear entry of the Drosophila clock proteins period and timeless. Neuron *17*, 911–920. https://doi.org/10.1016/S0896-6273(00)80222-6.

22. Jang, A.R., Moravcevic, K., Saez, L., Young, M.W., and Sehgal, A. (2015). Drosophila TIM Binds Importin alpha1, and Acts as an Adapter to Transport PER to the Nucleus. PLoS Genet. *11*, e1004974. https://doi.org/10.1371/journal.pgen.1004974.

23. Lin, C., Feng, S., DeOliveira, C.C., and Crane, B.R. (2023). Cryptochrome-Timeless structure reveals circadian clock timing mechanisms. Nature *617*, 194–199. https://doi.org/10.1038/s41586-023-06009-4.

24. Meyer, P., Saez, L., and Young, M.W. (2006). PER-TIM interactions in living Drosophila cells: An interval timer for the circadian clock. Science *311*, 226–229. https://doi.org/10.1126/science.1118126.

25. Saez, L., Derasmo, M., Meyer, P., Stieglitz, J., and Young, M.W. (2011). A Key Temporal Delay in the Circadian Cycle of Drosophila Is Mediated by a Nuclear Localization Signal in the Timeless Protein. Genetics *188*, 591–600. https://doi.org/10.1534/genetics.111.127225.

26. Singh, S., Giesecke, A., Damulewicz, M., Fexova, S., Mazzotta, G.M., Stanewsky, R., and Dolezel, D. (2019). New Drosophila Circadian Clock Mutants Affecting Temperature Compensation Induced by Targeted Mutagenesis of Timeless. Front. Physiol. *10*, 1442. https://doi.org/10.3389/fphys.2019.01442.

27. Giesecke, A., Johnstone, P.S., Lamaze, A., Landskron, J., Atay, E., Chen, K.F., Wolf, E., Top, D., and Stanewsky, R. (2023). A novel period mutation implicating nuclear export in temperature compensation of the Drosophila circadian clock. Curr. Biol. *33*, 336–350.e5. https://doi.org/10.1016/j.cub.2022.12.011.

28. Top, D., Harms, E., Syed, S., Adams, E.L., and Saez, L. (2016). GSK-3 and CK2 Kinases Converge on Timeless to Regulate the Master Clock. Cell Rep. *16*, 357–367. https://doi.org/10.1016/j.celrep.2016.06.005.

29. Peschel, N., Chen, K.F., Szabo, G., and Stanewsky, R. (2009). Light-Dependent Interactions between the Drosophila Circadian Clock Factors Cryptochrome, Jetlag, and Timeless. Curr. Biol. *19*, 241–247. https://doi.org/10.1016/j.cub.2008.12.042.

30. Emery, P., Stanewsky, R., Hall, J.C., and Rosbash, M. (2000). Drosophila cryptochromes - A unique circadian-rhythm photoreceptor. Nature *404*, 456–457. https://doi.org/10.1038/35006558.

31. Dolezelova, E., Dolezel, D., and Hall, J.C. (2007). Rhythm defects caused by newly engineered null mutations in Drosophila's cryptochrome gene. Genetics *177*, 329–345. https://doi.org/10.1534/genetics.107.076513.

32. Peschel, N., Veleri, S., and Stanewsky, R. (2006). Veela defines a molecular link between Cryptochrome and Timeless in the light-input pathway to Drosophila's circadian clock. Proc. Natl. Acad. Sci. USA *103*, 17313–17318. https://doi.org/10.1073/pnas.0606675103.

33. Koh, K., Zheng, X., and Sehgal, A. (2006). JETLAG resets the Drosophila circadian clock by promoting light-induced degradation of TIMELESS. Science *312*, 1809–1812. https://doi.org/10.1126/science.1124951.

34. Ozturk, N., VanVickle-Chavez, S.J., Akileswaran, L., Van Gelder, R.N., and Sancar, A. (2013). Ramshackle (Brwd3) promotes light-induced ubiquitylation of Drosophila Cryptochrome by DDB1-CUL4-ROC1 E3 ligase complex. Proc. Natl. Acad. Sci. USA *110*, 4980–4985. https://doi.org/10.1073/pnas.1303234110.

35. Iiams, S.E., Wan, G., Zhang, J., Lugena, A.B., Zhang, Y., Hayden, A.N., and Merlin, C. (2024). Loss of functional cryptochrome 1 reduces robustness of 24-hour behavioral rhythms in monarch butterflies. iScience *27*, 108980. https://doi.org/10.1016/j.isci.2024.108980.

36. Tobita, H., and Kiuchi, T. (2024). Knockout of cryptochrome 1 disrupts circadian rhythm and photoperiodic diapause induction in the silkworm, Bombyx mori. Insect Biochem. Mol. Biol. *172*, 104153. https://doi.org/10.1016/j.ibmb.2024.104153.

37. Poupardin, R., Schöttner, K., Korbelová, J., Provazník, J., Doležel, D., Pavlinic, D., Beneš, V., and Koštál, V. (2015). Early transcriptional events linked to induction of diapause revealed by RNAseq in larvae of drosophilid fly, Chymomyza costata. BMC Genom. *16*, 720. https://doi.org/10.1186/s12864-015-1907-4.

38. Stehlik, J., Zavodska, R., Shimada, K., Sauman, I., and Kostal, V. (2008). Photoperiodic induction of diapause requires regulated transcription of timeless in the larval brain of Chymomyza costata. J. Biol. Rhythm. *23*, 129–139. https://doi.org/10.1177/0748730407313364.

39. Kobelkova, A., Bajgar, A., and Dolezel, D. (2010). Functional Molecular Analysis of a Circadian Clock Gene timeless Promoter from the Drosophilid Fly Chymomyza costata. J. Biol. Rhythm. *25*, 399–409. https://doi.org/10.1177/0748730410385283.

40. Kostal, V. (2006). Eco-physiological phases of insect diapause. J. Insect Physiol. *52*, 113–127. https://doi.org/10.1016/j.jinsphys.2005.09.008.

41. Tauber, E., Zordan, M., Sandrelli, F., Pegoraro, M., Osterwalder, N., Breda, C., Daga, A., Selmin, A., Monger, K., Benna, C., et al. (2007). Natural selection favors a newly derived timeless allele in Drosophila melanogaster. Science *316*, 1895–1898. https://doi.org/10.1126/science.1138412.

42. Abrieux, A., Xue, Y., Cai, Y., Lewald, K.M., Nguyen, H.N., Zhang, Y., and Chiu, J.C. (2020). EYES ABSENT and TIMELESS integrate photoperiodic and temperature cues to regulate seasonal physiology in Drosophila. Proc. Natl. Acad. Sci. USA *117*, 15293–15304. https://doi.org/10.1073/pnas.2004262117.

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

44. Tobita, H., and Kiuchi, T. (2022). Knockouts of positive and negative elements of the circadian clock disrupt photoperiodic diapause induction in the silkworm, Bombyx mori. Insect Biochem. Mol. Biol. *149*, 103842. https://doi.org/10.1016/j.ibmb.2022.103842.

45. Zoltowski, B.D., Vaidya, A.T., Top, D., Widom, J., Young, M.W., and Crane, B.R. (2011). Structure of full-length Drosophila cryptochrome. Nature *480*, 396–399. https://doi.org/10.1038/Nature10618.

46. Levy, C., Zoltowski, B.D., Jones, A.R., Vaidya, A.T., Top, D., Widom, J., Young, M.W., Scrutton, N.S., Crane, B.R., and Leys, D. (2013). Updated structure of Drosophila cryptochrome. Nature *495*, E3–E4. https://doi.org/10.1038/nature11995.

47. Czarna, A., Berndt, A., Singh, H.R., Grudziecki, A., Ladurner, A.G., Timinszky, G., Kramer, A., and Wolf, E. (2013). Structures of Drosophila cryptochrome and mouse cryptochrome1 provide insight into circadian function. Cell *153*, 1394–1405. https://doi.org/10.1016/j.cell.2013.05.011.

48. Yildiz, O., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S., Urbanke, C., Sassone-Corsi, P., and Wolf, E. (2005). Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. Mol. Cell *17*, 69–82. https://doi.org/10.1016/j.molcel.2004.11.022.

49. Grabarczyk, D. B. (2020). Crystal structure and interactions of the Tof1-Csm3 (Timeless-Tipin) fork protection complex. Nucleic Acids Res. *48*, 6996–7004. https://doi.org/10.1093/nar/gkaa456.

50. Rosato, E., Trevisan, A., Sandrelli, F., Zordan, M., Kyriacou, C.P., and Costa, R. (1997). Conceptual translation of timeless reveals alternative initiating methionines in Drosophila. Nucleic Acids Res. *25*, 455–458. https://doi.org/10.1093/nar/25.3.455.

51. Sandrelli, F., Tauber, E., Pegoraro, M., Mazzotta, G., Cisotto, P., Landskron, J., Stanewsky, R., Piccin, A., Rosato, E., Zordan, M., et al. (2007). A molecular basis for natural selection at the timeless locus in Drosophila melanogaster. Science *316*, 1898–1900. https://doi.org/10.1126/science.1138426.

52. Lamaze, A., Chen, C., Leleux, S., Xu, M., George, R., and Stanewsky, R. (2022). A natural timeless polymorphism allowing circadian clock synchronization in "white nights". Nat. Commun. *13*, 1724. https://doi.org/10.1038/s41467-022-29293-6.

53. Deppisch, P., Prutscher, J.M., Pegoraro, M., Tauber, E., Wegener, C., and Helfrich-Förster, C. (2022). Adaptation of Drosophila melanogaster to Long Photoperiods of High-Latitude Summers Is Facilitated by the Is-Timeless Allele. J. Biol. Rhythms *37*, 185–201. https://doi.org/10.1177/07487304221082448.

54. Benna, C., Bonaccorsi, S., Wülbeck, C., Helfrich-Förster, C., Gatti, M., Kyriacou, C.P., Costa, R., and Sandrelli, F. (2010). Drosophila timeless2 is required for chromosome stability and circadian photoreception. Curr. Biol. *20*, 346–352. https://doi.org/10.1016/j.cub.2009.12.048.

55. Kotwica-Rolinska, J., Chodakova, L., Smykal, V., Damulewicz, M., Provaznik, J., Wu, B.C., Hejnikova, M., Chvalova, D., and Dolezel, D. (2022). Loss of Timeless Underlies an Evolutionary Transition within the Circadian Clock. Mol. Biol. Evol. *39*, msab346. https://doi.org/10.1093/molbev/msab346.

56. Nose, M., Tokuoka, A., Bando, T., and Tomioka, K. (2018). timeless2 plays an important role in reproduction and circadian rhythms in the cricket Gryllus bimaculatus. J. Insect Physiol. *105*, 9–17. https://doi.org/10.1016/j.jinsphys.2017.12.007.

57. Kurien, P., Hsu, P.K., Leon, J., Wu, D., McMahon, T., Shi, G., Xu, Y., Lipzen, A., Pennacchio, L.A., Jones, C.R., et al. (2019). TIMELESS mutation alters phase responsiveness and causes advanced sleep phase. Proc. Natl. Acad. Sci. USA *116*, 12045–12053. https://doi.org/10.1073/pnas.1819110116.

58. Baretic, D., Jenkyn-Bedford, M., Aria, V., Cannone, G., Skehel, M., and Yeeles, J.T.P. (2020). Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork. Mol. Cell *78*, 926–940.e13. https://doi.org/10.1016/j.molcel.2020.04.012.

59. Unsal-Kacmaz, K., Mullen, T.E., Kaufmann, W.K., and Sancar, A. (2005). Coupling of human circadian and cell cycles by the timeless protein. Mol. Cell Biol. *25*, 3109–3116. https://doi.org/10.1128/MCB.25.8.3109-3116.2005.

60. Rubin, E.B., Shemesh, Y., Cohen, M., Elgavish, S., Robertson, H.M., and 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 Res. *16*, 1352–1365. https://doi.org/10.1101/gr.5094806.

61. Dolezel, D. (2023). Molecular Mechanism of the Circadian Clock. In Insect Chronobiology, 1st edn., H. Numata and K. Tomioka, eds. (Springer), pp. 49–84. https://doi.org/10.1007/978-981-99-0726-7_4.

62. Deppisch, P., Kirsch, V., Helfrich-Forster, C., and Senthilan, P.R. (2023). Contribution of cryptochromes and photolyases for insect life under sunlight. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *209*, 373–389. https://doi.org/10.1007/s00359-022-01607-5.

63. Kamae, Y., and Tomioka, K. (2012). Timeless is an essential component of the circadian clock in a primitive insect, the firebrat Thermobia domestica. J. Biol. Rhythms *27*, 126–134. https://doi.org/10.1177/0748730411435997.

64. Danbara, Y., Sakamoto, T., Uryu, O., and Tomioka, K. (2010). RNA interference of timeless gene does not disrupt circadian locomotor rhythms in the cricket Gryllus bimaculatus. J. Insect Physiol. *56*, 1738–1745. https://doi.org/10.1016/j.jinsphys.2010.07.002.

65. Werckenthin, A., Huber, J., Arnold, T., Koziarek, S., Plath, M.J.A., Plath, J.A., Stursberg, O., Herzel, H., and Stengl, M. (2020). Neither per, nor tim1, nor cry2 alone are essential components of the molecular circadian clockwork in the Madeira cockroach. PLoS One *15*, e0235930. https://doi.org/10.1371/journal.pone.0235930.

66. Gotter, A.L., Manganaro, T., Weaver, D.R., Kolakowski, L.F., Possidente, B., Sriram, S., MacLaughlin, D.T., and Reppert, S.M. (2000). A time-less function for mouse Timeless. Nat. Neurosci. *3*, 755–756.

67. Barnes, J.W., Tischkau, S.A., Barnes, J.A., Mitchell, J.W., Burgoon, P.W., Hickok, J.R., and Gillette, M.U. (2003). Requirement of mammalian Timeless for circadian rhythmicity. Science *302*, 439–442. https://doi.org/10.1126/science.1086593.

68. Hasegawa, K., Saigusa, T., and Tamai, Y. (2005). Caenorhabditis elegans opens up new insights into circadian clock mechanisms. Chronobiol. Int. *22*, 1–19. https://doi.org/10.1081/cbi-200038149.

69. Sehgal, A., Rothenfluhhilfiker, A., Hunterensor, M., Chen, Y., Myers, M.P., and Young, M.W. (1995). Rhythmic Expression of Timeless - a Basis for Promoting Circadian Cycles in Period Gene Autoregulation. Science *270*, 808–810. https://doi.org/10.1126/science.270.5237.808.

70. Rothenfluh, A., Abodeely, M., Price, J.L., and Young, M.W. (2000). Isolation and analysis of six timeless alleles that cause short- or long-period circadian rhythms in Drosophila. Genetics *156*, 665–675. https://doi.org/10.1093/genetics/156.2.665.

71. Rothenfluh, A., Young, M.W., and Saez, L. (2000). A TIMELESS-independent function for PERIOD proteins in the Drosophila clock. Neuron *26*, 505–514. https://doi.org/10.1016/S0896-6273(00)81182-4.

72. Matsumoto, A., Tomioka, K., Chiba, Y., and Tanimura, T. (1999). timrit Lengthens circadian period in a temperature-dependent manner through suppression of PERIOD protein cycling and nuclear localization. Mol. Cell Biol. *19*, 4343–4354. https://doi.org/10.1128/MCB.19.6.4343.

73. Thakkar, N., Hejzlarova, A., Brabec, V., and Dolezel, D. (2023). Germline Editing of Drosophila Using CRISPR-Cas9-based Cytosine and Adenine Base Editors. CRISPR J. *6*, 557–569. https://doi.org/10.1089/crispr.2023.0026.

74. Cai, Y.D., Xue, Y., Truong, C.C., Del Carmen-Li, J., Ochoa, C., Vanselow, J.T., Murphy, K.A., Li, Y.H., Liu, X., Kunimoto, B.L., et al. (2021). CK2 Inhibits TIMELESS Nuclear Export and Modulates CLOCK Transcriptional Activity to Regulate Circadian Rhythms. Curr. Biol. *31*, 502–514.e7. https://doi.org/10.1016/j.cub.2020.10.061.

75. Lindestad, O., Nylin, S., Wheat, C.W., and Gotthard, K. (2022). Local adaptation of life cycles in a butterfly is associated with variation in several circadian clock genes. Mol. Ecol. *31*, 1461–1475. https://doi.org/10.1111/mec.16331.

76. Bernatowicz, P.P., Kotwica-Rolinska, J., Joachimiak, E., Sikora, A., Polanska, M.A., Pijanowska, J., and Bębas, P. (2016). Temporal Expression of the Clock Genes in the Water Flea Daphnia pulex (Crustacea: Cladocera). J. Exp. Zool. A Ecol. Genet. Physiol. *325*, 233–254. https://doi.org/10.1002/jez.2015.

77. Bazalova, O., Kvicalova, M., Valkova, T., Slaby, P., Bartos, P., Netusil, R., Tomanova, K., Braeunig, P., Lee, H.J., Sauman, I., et al. (2016). Cryptochrome 2 mediates directional magnetoreception in cockroaches. Proc. Natl. Acad. Sci. USA *113*, 1660–1665. https://doi.org/10.1073/pnas.1518622113.

78. Colizzi, F.S., Beer, K., Cuti, P., Deppisch, P., Martínez Torres, D., Yoshii, T., and Helfrich-Förster, C. (2021). Antibodies Against the Clock Proteins Period and Cryptochrome Reveal the Neuronal Organization of the Circadian Clock in the Pea Aphid. Front. Physiol. *12*, 705048. https://doi.org/10.3389/fphys.2021.705048.

79. Vaidya, A. T., Top, D., Manahan, C.C., Tokuda, J.M., Zhang, S., Pollack, L., Young, M.W., and Crane, B.R. (2013). Flavin reduction activates Drosophila cryptochrome. Proc. Natl. Acad. Sci. USA *110*, 20455–20460. https://doi.org/10.1073/pnas.1313336110.

80. Barbera, M., Collantes-Alegre, J.M., and Martinez-Torres, D. (2017). Characterisation, analysis of expression and localisation of circadian clock genes from the perspective of photoperiodism in the aphid Acyrthosiphon pisum. Insect Biochem. Mol. Biol. *83*, 54–67. https://doi.org/10.1016/j.ibmb.2017.02.006.

81. Koike, N., Hida, A., Numano, R., Hirose, M., Sakaki, Y., and Tei, H. (1998). Identification of the mammalian homologues of the Drosophila timeless gene, Timeless1. FEBS Lett. *441*, 427–431. https://doi.org/10.1016/s0014-5793(98)01597-x.

82. Boothroyd, C.E., Wijnen, H., Naef, F., Saez, L., and Young, M.W. (2007). Integration of light and temperature in the regulation of circadian gene expression in Drosophila. PLoS Genet. *3*, e54. https://doi.org/10.1371/journal.pgen.0030054.

83. Montelli, S., Mazzotta, G., Vanin, S., Caccin, L., Corrà, S., De Pittà, C., Boothroyd, C., Green, E.W., Kyriacou, C.P., and Costa, R. (2015). period and timeless mRNA Splicing Profiles under Natural Conditions in Drosophila melanogaster. J. Biol. Rhythms *30*, 217–227. https://doi.org/10.1177/0748730415583575.

84. Shakhmantsir, I., Nayak, S., Grant, G.R., and Sehgal, A. (2018). Spliceosome factors target timeless (tim) mRNA to control clock protein accumulation and circadian behavior in Drosophila. Elife *7*, e39821. https://doi.org/10.7554/eLife.39821.

85. Martin Anduaga, A., Evantal, N., Patop, I.L., Bartok, O., Weiss, R., and Kadener, S. (2019). Thermosensitive alternative splicing senses and mediates temperature adaptation in Drosophila. Elife *8*, e44642. https://doi.org/10.7554/eLife.44642.

86. Foley, L.E., Ling, J., Joshi, R., Evantal, N., Kadener, S., and Emery, P. (2019). Drosophila PSI controls circadian period and the phase of circadian behavior under temperature cycle via tim splicing. Elife *8*, e50063. https://doi.org/10.7554/eLife.50063.

87. Bazalova, O., and Dolezel, D. (2017). Daily Activity of the Housefly, Musca domestica, Is Influenced by Temperature Independent of 3' UTR period Gene Splicing. G3 (Bethesda) *7*, 2637–2649. https://doi.org/10.1534/g3.117.042374.

88. Perrigault, M., and Tran, D. (2017). Identification of the Molecular Clockwork of the Oyster Crassostrea gigas. PLoS One *12*, e0169790. https://doi.org/10.1371/journal.pone.0169790.

89. Zantke, J., Ishikawa-Fujiwara, T., Arboleda, E., Lohs, C., Schipany, K., Hallay, N., Straw, A.D., Todo, T., and Tessmar-Raible, K. (2013). Circadian and circalunar clock interactions in a marine annelid. Cell Rep. *5*, 99–113. https://doi.org/10.1016/j.celrep.2013.08.031.

90. Palmer, J.D. (1997). Dueling hypotheses: circatidal versus circalunidian battle basics. Chronobiol. Int. *14*, 337–346. https://doi.org/10.3109/07420529709001455.

91. Kwiatkowski, E.R., Schnytzer, Y., Rosenthal, J.J.C., and Emery, P. (2023). Behavioral circatidal rhythms require Bmal1 in Parhyale hawaiensis. Curr. Biol. *33*, 1867–1882.e5. https://doi.org/10.1016/j.cub.2023.03.015.

92. Ceriani, M.F., Darlington, T.K., Staknis, D., Más, P., Petti, A.A., Weitz, C.J., and Kay, S.A. (1999). Light-dependent sequestration of TIMELESS by CRYPTOCHROME. Science *285*, 553–556. https://doi.org/10.1126/science.285.5427.553.

93. Busza, A., Emery-Le, M., Rosbash, M., and Emery, P. (2004). Roles of the two Drosophila CRYPTOCHROME structural domains in circadian photoreception. Science *304*, 1503–1506. https://doi.org/10.1126/science.1096973.

94. D'Costa, A., Reifegerste, R., Sierra, S., and Moses, K. (2006). The Drosophila ramshackle gene encodes a chromatin-associated protein required for cell morphology in the developing eye. Mech. Dev. *123*, 591–604. https://doi.org/10.1016/j.mod.2006.06.007.

95. Kaniewska, M.M., Vaněčková, H., Doležel, D., and Kotwica-Rolinska, J. (2020). Light and Temperature Synchronizes Locomotor Activity in the Linden Bug, Pyrrhocoris apterus. Front. Physiol. *11*, 242. https://doi.org/10.3389/fphys.2020.00242.

96. R, R., Pruser, T., Schulz, N.K.E., Mayer, P.M.F., Ogueta, M., Stanewsky, R., and Kurtz, J. (2024). Deciphering a Beetle Clock: Individual and Sex-Dependent Variation in Daily Activity Patterns. J. Biol. Rhythm. *39*, 484–501. https://doi.org/10.1177/07487304241263619.

97. Smýkal, V., Tobita, H., and Dolezel, D. (2025). Evolution of circadian clock and light-input pathway genes in Hemiptera. Insect Biochem. Mol. Biol. *180*, 104298. https://doi.org/10.1016/j.ibmb.2025.104298.

98. Evangelista, D.A., Wipfler, B., Bethoux, O., Donath, A., Fujita, M., Kohli, M.K., Legendre, F., Liu, S., Machida, R., Misof, B., et al. (2019). An integrative phylogenomic approach illuminates the evolutionary history of cockroaches and termites (Blattodea). Proc. Biol. Sci. *286*, 20182076. https://doi.org/10.1098/rspb.2018.2076.

99. Bertolini, E., Schubert, F.K., Zanini, D., Sehadová, H., Helfrich-Förster, C., and Menegazzi, P. (2019). Life at High Latitudes Does Not Require Circadian Behavioral Rhythmicity under Constant Darkness. Curr. Biol. *29*, 3928–3936.e3. https://doi.org/10.1016/j.cub.2019.09.032.

100. Ikeno, T., Tanaka, S.I., Numata, H., and Goto, S.G. (2010). Photoperiodic diapause under the control of circadian clock genes in an insect. BMC Biol. *8*, 116. https://doi.org/10.1186/1741-7007-8-116.

101. Urbanova, V., Bazalova, O., Vaneckova, H., and Dolezel, D. (2016). Photoperiod regulates growth of male accessory glands through juvenile hormone signaling in the linden bug, Pyrrhocoris apterus. Insect Biochem. Mol. Biol. *70*, 184–190. https://doi.org/10.1016/j.ibmb.2016.01.003.

102. Kotwica-Rolinska, J., Damulewicz, M., Chodakova, L., Kristofova, L., and Dolezel, D. (2022). Pigment Dispersing Factor Is a Circadian Clock Output and Regulates Photoperiodic Response in the Linden Bug, Pyrrhocoris apterus. Front. Physiol. *13*, 884909. https://doi.org/10.3389/fphys.2022.884909.

103. Takekata, H., Numata, H., Shiga, S., and Goto, S.G. (2014). Silencing the circadian clock gene Clock using RNAi reveals dissociation of the circatidal clock from the circadian clock in the mangrove cricket. J. Insect Physiol. *68*, 16–22. https://doi.org/10.1016/j.jinsphys.2014.06.012.

104. Zhang, L., Green, E.W., Webster, S.G., Hastings, M.H., Wilcockson, D.C., and Kyriacou, C.P. (2023). Correction: The circadian clock gene bmal1 is necessary for co-ordinated circatidal rhythms in the marine isopod Eurydice pulchra (Leach). PLoS Genet. *19*, e1011047. https://doi.org/10.1371/journal.pgen.1011047.

105. Stöver, B.C., and Müller, K.F. (2010). TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinf. *11*, 7. https://doi.org/10.1186/1471-2105-11-7.

106. Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G., and Hedges, S.B. (2022). TimeTree 5: An Expanded Resource for Species Divergence Times. Mol. Biol. Evol. *39*, msac174. https://doi.org/10.1093/molbev/msac174.

107. Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. *26*, 1641–1650. https://doi.org/10.1093/molbev/msp077.

108. Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., et al. (2014). Phylogenomics resolves the timing and pattern of insect evolution. Science *346*, 763–767. https://doi.org/10.1126/science.1257570.

109. von Reumont, B.M., Jenner, R.A., Wills, M.A., Dell'ampio, E., Pass, G., Ebersberger, I., Meyer, B., Koenemann, S., Iliffe, T.M., Stamatakis, A., et al. (2012). Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. Mol. Biol. Evol. *29*, 1031–1045. https://doi.org/10.1093/molbev/msr270.

110. Thomas, G.W.C., Dohmen, E., Hughes, D.S.T., Murali, S.C., Poelchau, M., Glastad, K., Anstead, C.A., Ayoub, N.A., Batterham, P., Bellair, M., et al. (2020). Gene content evolution in the arthropods. Genome Biol. *21*, 15. https://doi.org/10.1186/s13059-019-1925-7.

111. Bernot, J.P., Owen, C.L., Wolfe, J.M., Meland, K., Olesen, J., and Crandall, K.A. (2023). Major Revisions in Pancrustacean Phylogeny and Evidence of Sensitivity to Taxon Sampling. Mol. Biol. Evol. *40*, msad175. https://doi.org/10.1093/molbev/msad175.

112. Wipfler, B., Letsch, H., Frandsen, P.B., Kapli, P., Mayer, C., Bartel, D., Buckley, T.R., Donath, A., Edgerly-Rooks, J.S., Fujita, M., et al. (2019). Evolutionary history of Polyneoptera and its implications for our understanding of early winged insects. Proc. Natl. Acad. Sci. USA *116*, 3024–3029. https://doi.org/10.1073/pnas.1817794116.

113. Johnson, K.P., Dietrich, C.H., Friedrich, F., Beutel, R.G., Wipfler, B., Peters, R.S., Allen, J.M., Petersen, M., Donath, A., Walden, K.K.O., et al. (2018). Phylogenomics and the evolution of hemipteroid insects. Proc. Natl. Acad. Sci. USA *115*, 12775–12780. https://doi.org/10.1073/pnas.1815820115.

114. McKenna, D.D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D.J., Donath, A., Escalona, H.E., Friedrich, F., Letsch, H., et al. (2019). The evolution and genomic basis of beetle diversity. Proc. Natl. Acad. Sci. USA *116*, 24729–24737. https://doi.org/10.1073/pnas.1909655116.

115. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

116. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

117. Smykal, V., Pivarci, M., Provaznik, J., Bazalova, O., Jedlicka, P., Luksan, O., Horak, A., Vaneckova, H., Benes, V., Fiala, I., et al. (2020). Complex Evolution of Insect Insulin Receptors and Homologous Decoy Receptors, and Functional Significance of Their Multiplicity. Mol. Biol. Evol. *37*, 1775–1789. https://doi.org/10.1093/molbev/msaa048.

118. Li, H. (2021). New strategies to improve minimap2 alignment accuracy. Bioinformatics *37*, 4572–4574. https://doi.org/10.1093/bioinformatics/btab705.

119. Wulbeck, C., Szabo, G., Shafer, O.T., Helfrich-Forster, C., and Stanewsky, R. (2005). The novel Drosophila tim(blind) mutation affects behavioral rhythms but not periodic eclosion. Genetics *169*, 751–766. https://doi.org/10.1534/genetics.104.036244.

120. Smykal, V., and Dolezel, D. (2023). Evolution of proteins involved in the final steps of juvenile hormone synthesis. J. Insect Physiol. *145*, 104487. https://doi.org/10.1016/j.jinsphys.2023.104487.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Sequence files	https://doi.org/10.5061/dryad.44j0zpcq0	DRYAD https://datadryad.org/stash
Software and algorithms		
NLS prediction: psort II	https://psort.hgc.jp/	
NLS prediction: NLStradamus	http://www.moseslab.csb.utoronto.ca/NLStradamus/	
MEME Suite	https://meme-suite.org/meme/	
Geneious Prime 21.0.3	https://www.geneious.com/ https://assets.geneious.com/documentation/geneious/release_notes.html#v2021.0	Biomatters, Auckland, New Zealand
TreeGraph 2	http://treegraph.bioinfweb.info/	Stöver and Müller[105]
Prism 7	https://www.graphpad.com	GraphPad Software, La Jolla, CA, USA
CorelDRAW X6	https://www.coreldraw.com/	CorelDRAW
Adobe Illustrator (version 6)	https://www.adobe.com	Adobe
Other		
TIMETREE 5	https://timetree.org/	Kumar et al.[106]
NCBI	https://www.ncbi.nlm.nih.gov/	

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

No new samples were collected for this study.

METHOD DETAILS

Gene identification and datasets

A systematic search for clock components was conducted, building on previous studies exploring the evolution, duplication, and loss of circadian clock components.[7,55] To identify circadian clock proteins and genes encoding dTIM, mTIM, TOF1, PERIOD, dCRY, mCRY, 6-4 Photolyase, JETLAG, and BRWD3 in Bilateria/Metazoa, GenBank (NCBI) protein and genomic databases, as well as transcriptome shotgun assemblies (TSA), were utilized. BLASTP and tBLASTn algorithms were applied with taxon-restricted searches targeting specific lineages at the levels of orders, suborders, infraorders, and, in some cases, families. In certain instances, annotated genomes or whole-genome shotgun contigs (wgs) were also examined. Protein sequences of circadian clock genes from *Drosophila melanogaster*, *Danaus plexippus*, and *Pyrrhocoris apterus* served as initial queries. Reciprocal and lineage-focused searches incorporated queries representing identified proteins from related taxa. To detect duplicate hits (common in TSAs) or closely related proteins (e.g., bHLH-PAS proteins instead of PERIOD), the E-INS-i algorithm in MAFFT was used for alignment, followed by FastTree analysis,[107] both conducted in Geneious Prime 21.0.3 (Biomatters, New Zealand).

Phylogenetic analyses

To identify specific types of TIM, CRY, FBXL, or PAS proteins (e.g., distinguishing PER from bHLH PAS proteins), sequences were aligned using the MAFFT algorithm in Geneious Prime 21.0.3 (Biomatters, New Zealand). Representative datasets containing target proteins and related types were included. Ambiguously aligned regions were trimmed, and phylogenetic analyses were conducted using RAxML with a maximum likelihood GAMMA-based model in Geneious Prime 21.0.3.

The Metazoan phylogenies presented in Figures 2, 4, 5, 6, 7, 8, and S9 were retrieved using TIMETREE 5[106] and cross-referenced with recent molecular phylogenomic studies. These included works focused on insects,[108] chelicerates, and crustaceans.[109–111] For insects, phylogenomic studies specific to Polyneoptera,[112] the hemipteroid assembly,[113] and Coleoptera[114] were used to refine the corresponding sections of the phylogeny.

Gene loss

While it is impossible to definitively prove the absence of a gene, in some cases, gene loss is the most plausible explanation. Recent advances in phylogenomics, the availability of extensive TSA data, and an increasing number of sequenced genomes have enabled a

systematic exploration of circadian clock genes across major Bilateria groups (Protostomia and Deuterostomia). Our analysis focuses on lineage-specific gene losses that are strongly supported by data from multiple species, whole-genome assemblies, and deep transcriptome sequencing. Evidence for gene loss is summarized for specific genes and animal groups in Table S1.

Finally, the CTT-like motif, located on the ARM2 domain, and CTT motif located on dCRY, were annotated on *D. melanogaster* dTIM and dCRY sequences following the boundaries defined by Lin et al.[23] After MAFFT alignment with the other dTIM sequences in the dataset, it was considered conserved when the degree of similarity exceeded 50%.

Prediction of protein domains

dTIM functional and binding domains were originally annotated based on the sequence of *Drosophila melanogaster* dTIM, following the boundaries defined by Cryo-EM[23] or identified in cell-based experiments.[21] Accordingly, the dCRY-interaction domains, ARMADILLO repeats (ARM1 and ARM2), and PER-binding sites 1 (PER-bind #1) and 2 (PER-bind #2) were annotated in the *D. melanogaster* dTIM protein isoform P (NP_001334730).

This annotated sequence was then aligned to each sequence in the dataset using MAFFT[115] (Algorithm: E-INS-I, scoring matrix: BLOSUM80) within the Geneious Prime 2024 software. The obtained similarities were plotted as intensities corresponding to numerical values in Figures 2 and 3. The PER-bind #1 domain was annotated when the total length of the region was at least 35 amino acids and the degree of similarity exceeded 40%. The CRY-interaction domain was highlighted when similarities exceeded 20%.

In the gene models (Figures 6 and 7), only two shades were used: regions corresponding to ARM1, ARM2, and PER-binding sites were highlighted when similarity exceeded 40%. For *Limulus*, a value of 38% was depicted as "low similarity" using a paler shade of red. The CRY-interaction domain was highlighted when similarity exceeded 20%. For *Bemisia*, "low similarity" was represented by a paler shade of blue, highlighting values of 19%. In the supplementary figures depicting protein models (Figures S2, S4, and S6), specific numerical values were presented next to domain annotations when a single shade was used for each domain.

Nuclear localization signal (NLS) domains were predicted for each protein sequence using PSORT II and NLStradamus prediction software. Acidic domains were annotated based on the following criteria: a sequence located in the corresponding protein region that is at least half the size of the reference sequence, contains at least 20% acidic residues (D, E), and includes less than 10% basic residues (K, H, R). To further investigate the acidic domains, their sequence motifs were scanned. Conserved acidic motifs, presented in Figure S5, were identified using Gapped Local Alignment of Motifs (GLAM2) within the MEME Suite. This analysis focused on two distinct regions of 35 insect dTIM proteins: the sequences between ARM1 and ARM2 and the region covering the C-terminal tails.

The length of the entire protein, ARM1-ARM2 region, and C-tail length

The protein sequence was considered (likely) complete when the N-terminal part included a complete ARM1 domain starting with a methionine and the ARM2 domain was present. If a TSA (transcriptome shotgun assembly) sequence was used, the stop codon indicated the predicted C-terminal end. In the case of multiple paralogs in *Daphnia*, only protein sequences longer than 600 amino acids (aa) and containing both ARM domains were further analyzed.

To determine the length of the variable region (in amino acids) in the central part of the protein, all protein sequences in the dataset were aligned using MAFFT,[115] v7.490, Algorithm: E-INS-I, scoring matrix: BLOSUM80). Conserved motifs corresponding to the *Drosophila melanogaster* ARM1 and ARM2 domains were identified and the number of aa separating ARM1 and ARM2 calculated. Additionally, conserved motifs corresponding to *Drosophila* YKDQ (located in ARM1) and LLLR (located in PER-bind #2 and ARM2) were identified in each dTIM protein as a parallel measurement.

To measure the length of the C-terminal tail, a conserved motif corresponding to *Drosophila* DLIE (located at the C-terminal end of ARM2) was identified in each dTIM. The number of amino acids between the DLIE-like motif and the C-terminus was then calculated and plotted (Figure 2E). For motif positions, see Figure S2. These values were plotted as dots representing each sequence distance in PRISM 7 for all proteins in the dataset, with exact values provided in Table S2. If a species lacked one or more conserved motifs (due to partial sequences or deletions), the length of the corresponding region was not calculated and was annotated as "n.d." (non-determined) in Table S2.

Substitutions per amino acid per million years

To calculate substitution rates per amino acid position for dTIM and PER proteins, we used the following approach. Protein sequences (dTIM or PER) were aligned using the E-INS-i algorithm in MAFFT.[115] The complete alignments were used to infer phylogenetic trees with RAxML[116] under the PROTGAMMAJTT model, ensuring the topology matched the evolutionary relationships of the organisms. Constraint trees were created in TreeGraph 2.[105] For Crustacea, where phylogenies are still debated, we enforced monophyly with insects but did not specify internal branching. Similarly, Polyneoptera were constrained as monophyletic without defining their internal topologies.

The resulting unrooted trees were swapped to position Protostomia and Deuterostomia as sister groups. Branch lengths were extracted and summed from the Protostomia/Deuterostomia split to terminal species. These values were then divided by 700 million years (the estimated divergence time of Protostomia and Deuterostomia) to compute substitution rates per amino acid position per million years. The calculated rates are presented in Tables S2 and S4. These values were plotted in Prism 7 (GraphPad Software, La Jolla, CA, USA), with each dot on the plot representing a single protein.

Gene models

In brief, similarities in gene structure were assessed as previously[117] when the gene models were either directly downloaded from GenBank species' Whole-genome shotgun contigs (wgs) or Representative genomes (RefSeq genomes). Alternatively, gene models were manually reconstructed using genomic and transcriptomic data from GenBank, following two distinct scenarios: (1) Long genomic contigs without annotated *timeless* genes. The wgs contigs were manually annotated by mapping Transcriptome Shotgun Assembly (TSA) and/or non-redundant nucleotide (nr/nt) *d-tim* sequences by the Minimap2 mapper[118] within Geneious Prime 2024 software (Biomatters, Auckland, New Zealand). *Lethocerus indicus tim* gene model was reconstructed by mapping TSA sequences from three related species: *Trichocorixa calva, Gelastocoris ocwulatus, Buenoa margaritacea*, to an unannotated genomic scaffold. Similarly, *Neoneuromus ignobilis tim* gene was annotated using TSA from *Protohermes xanthodes*. In *Pogonus chalceus*, two partial non-overlapping TSA sequences cover the majority of the coding sequence (CDS). The missing region spanning parts of exons 7 and 8 was reconstructed by mapping TSAs from related beetle species *Amphizoa insolens, Tribolium castaneum, and Sinaspidytes wrasei*. (2) Fragmented genomic contigs without annotated *timeless* genes. The *tim* exons were annotated and contigs scaffolded using *tim* CDS or TSA sequences. In the case of *Lepisma saccharina*, *Thermobia domestica d-tim* TSAs were mapped to the seven *L. saccharina* unannotated contigs. See Table S8 for the accession number of genomic and TSA sequences used for gene reconstruction and Table S9 for accession numbers used to annotate the CDS in gene models.

Exon homology, similarity, and dTIM domain localization in gene models

The *tim* CDSs of representative species were translated and aligned using MAFFT[115] within Geneious Prime 2024 software to determine conserved domains and conserved/homologous exons and exon boundaries. Two protein alignments were used: ''species-to-species pair alignments'' and ''all representative proteins alignment''. The longest *d-tim* isoforms of 29 representative species were analyzed in Figure 6. Available *d-tim* isoforms were included in Figure 7, where 28 TIM sequences from 11 dipteran species were aligned.

Two exon boundaries were determined as conserved through the species-to-species dTIM/dTIM pair alignment if a minimum of three out of five amino acids (aa) located at the exon/intron boundary were identical. The similarity of amino acids was analyzed within: One pair of homologous exons if at least one side of the exon border could be aligned, or several exons homologous to one 'fused' exon if both sides of the exon border could be aligned, according to species-to-species pair exon alignment using MAFFT (% Similarity, auto algorithm, scoring matrix: BLOSUM90 with threshold 1). A semi-quantitative scale with five similarity ranges (30-40; 41-50, 51-60, 61-70; 71-100 %) was used to visualize exon/exon similarity as quadrilaterals of five different levels of greyscale in Figures 6 and 7.

CRYPTOCHROME (CRY)-interaction domains and PERIOD (PER)-binding sites were mapped and annotated in dTIM protein representative sequences utilizing domains defined in *Drosophila melanogaster* dTIM.[23,119] The figure was drawn using Adobe Illustrator (version 6) software.

Gene syntenies

In brief, gene syntenies were performed as in a recent study exploring gene duplications.[120] We compared gene synteny of genes neighboring *Drosophila*-type *cryptochrome* (*d-cry*) in heteropteran species *Apolygus lucorum*, *Lygus lineolaris* and *Lethocerus indicus* with species lacking *d-cry*, *Cimex hemipterus* and *Riptortus pedestris*. First, candidate *d-cry*-containing contigs were identified in *A. lucorum*, *L. lineolaris* and *L. indicus* genomes by using BLAST (Basic Local Alignment Search Tool) with *L. hesperus* dCRY-encoding Transcribed Sequence Assembly (TSA, acc # GDHC01010508.1) sequence as a query. Four syntenic protein-coding genes were selected in *A. lucorum*: *Axin* (*Axn*, GeneID: GE061_006998), *yippee-like* (*yippee*, GE061_006997), *DDRGK domain-containing protein 1* (*Ddrgk1*, GE061_006990), and *eclair* (*eca*, GE061_006989). Next, *L. indicus* 'd-cry' contig was pre-annotated using the plugin Augustus (version 0.1.1) within Geneious Prime (2024.0.5) with *Rhodnius prolixus* as a reference species. Then, four other syntenic genes neighboring *L. indicus d-cry* were selected: *venom protease-like* (*vp-l*), *WD repeat-containing protein 19* (*wdr19*), *ubiquitin carboxyl-terminal hydrolase 7* (*Usp7*), and *inositol polyphosphate 5-phosphatase E* (*inpp5e*).

Protein sequences of all eight syntenic genes and *L. hesperus* dCRY were used as queries to search in TSA databases to identify orthologs of the syntenic genes. *Lygus hesperus* and *Lygus lineolaris* TSA databases were searched for *L. lineolaris* and *A. lucorum*, *Cimex lectularius* for *Cimex hemipterus*, *Lethocerus indicus* and *Belostoma flumineum* for *L. indicus*, and *Riptortus pedestris* for *R. pedestris*. Gene-specific TSA sequences were then used as BLAST queries to identify syntenic genes' genomic contigs and to map syntenic genes to the contigs. TSAs were also mapped to the *A. lucorum* annotated genome to verify syntenic genes' identity and correct/build gene models. The accession numbers of genomic contigs and representative TSAs are in Table S6.

The same approach applied to *d-cry* was used for *jetlag* (*jet*) gene synteny, including the same species set used for TSA and genomic search and gene mapping, respectively. *Lygus hesperus jet* TSA (acc # GBRD01002318) was used to identify and map *jet* gene-containing contigs in *A. lucorum*, *L. lineolaris* and *L. indicus* genomes. TSAs of three protein-coding genes upstream of *jet*: *histone acetyltransferase kat2a* (*kat2a*, GE061_007944), *transmembrane protein 59-like* (*tmem59*, GE061_007947), *ATP-dependent (S)-NAD(P)H-hydrate dehydratase* (*Naxd*, GE061_007948) and one *jet* downstream gene: *Dual specificity protein phosphatase* (*Mkp3*, GE061_007950) were selected as syntenic genes from *A. lucorum* genomic contig. Similarly to *L. indicus d-cry*, four upstream *jet* syntenic genes: *serine/arginine repetitive matrix protein 1* (*srrm1*), *ribosomal RNA processing protein 1 homolog* (*rrp1*), *serine/threonine-protein phosphatase 4 regulatory subunit 4* (*ppp4r4*) and *tyrosine-protein kinase Shark* (*Shark*), and three *jet* downstream

genes: *facilitated trehalose transporter Tret1* (*Tret1*), *PIN2/TERF1-interacting telomerase inhibitor 1* (*pinx1*) and *Translational activator of cytochrome c oxidase 1* (*Taco1*), were selected as *jet* syntenic genes from *L. indicus*. The accession numbers of *jet* genomic contigs and representative TSAs are in Table S7. The final gene syntenies were drawn in CorelDRAW X6 (16.4.0.1280).

QUANTIFICATION AND STATISTICAL ANALYSIS

The length measurements were performed in Geneious Prime (2024.0.5) software as described above. The robustness of phylogenetic trees was evaluated as described above. No statistical analyses were performed.