ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Revealing the profound influence of diapause on gene expression: Insights from the annual transcriptome of the copepod *Calanus finmarchicus*

Laura Payton[1,2,3] | Kim S. Last[4] | Jordan Grigor[4] | Céline Noirot[5] | Lukas Hüppe[1,2,6,7] | David V. P. Conway[8] | Mona Dannemeyer[2] | David Wilcockson[9] | Bettina Meyer[1,2,7]

[1]Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

[2]Section Polar Biological Oceanography, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

[3]CNRS, Univ. Bordeaux, Bordeaux INP, EPOC, UMR 5805, Arcachon, F-33120, France

[4]Scottish Association for Marine Science, Oban, Argyll, UK

[5]Plateforme Bio-Informatique GenoToul, MIAT, INRAE, UR875 Mathématiques et Informatique Appliquées Toulouse, Castanet-Tolosan, France

[6]Neurobiology and Genetics, Theodor-Boveri Institute, Biocentre, University of Würzburg, Würzburg, Germany

[7]Helmholtz Institute for Functional Marine Biodiversity (HIFMB) at the Carl von Ossietzky University Oldenburg, Oldenburg, Germany

[8]Marine Biological Association of the UK, the Laboratory, Plymouth, UK

[9]Department of Life Sciences, Aberystwyth University, Aberystwyth, UK

Correspondence
Laura Payton, CNRS, Univ. Bordeaux, Bordeaux INP, EPOC, UMR 5805, Arcachon, F-33120 France.
Email: laura.payton@u-bordeaux.fr

Abstract

Annual rhythms are observed in living organisms with numerous ecological implications. In the zooplanktonic copepod *Calanus finmarchicus*, such rhythms are crucial regarding its phenology, body lipid accumulation, and global carbon storage. Climate change drives annual biological rhythms out of phase with the prevailing environmental conditions with yet unknown but potentially catastrophic consequences. However, the molecular dynamics underlying phenology are still poorly described. In a rhythmic analysis of *C. finmarchicus* annual gene expression, results reveal that more than 90% of the transcriptome shows significant annual rhythms, with abrupt and dramatic upheaval between the active and diapause life cycle states. This work explores the implication of the circadian clock in the annual timing, which may control epigenetic mechanisms to profoundly modulate gene expression in response to calendar time. Results also suggest an increased light sensitivity during diapause that would ensure the photoperiodic entrainment of the endogenous annual clock.

KEYWORDS
annual rhythms, circadian clock, copepods, diapause, transcriptome, zooplankton

1 | INTRODUCTION

The annual rotation of the Earth around the Sun is the cause of periodic, predictable changes in risks and opportunities for living organisms. Thus, numerous species have evolved annual rhythms (Denlinger, 2022a; Helm & Lincoln, 2017; Ikegami & Yoshimura, 2017) conceptualized mechanistically by the endogenous circannual clock, synchronized by reliable external cues. This clock enables organisms to anticipate seasonal events and initiate biological processes accordingly. While external cues such as food availability, predator presence, or temperature can fine-tune the timing of annual rhythms, change in day length (photoperiod) is the most reliable and thus the most important external factor synchronizing the endogenous circannual timing mechanism (Baumgartner & Tarrant, 2017; Denlinger, 2022a; Helm & Lincoln, 2017; Ikegami & Yoshimura, 2017). Perhaps the most profound annual biological rhythms in organisms include those which involve long periods of dormancy, hibernation, or diapause (i.e. in microalgae, mammals, and arthropods, respectively). Within arthropods, diapause not only has been studied extensively in insects but has also been well described in crustaceans and particularly marine copepods, which are among the most numerous metazoan animals on earth (Baumgartner & Tarrant, 2017; Denlinger, 2022b). Despite their ecological importance, copepods are non-model invertebrates for which genomic resources are still limited. However, in recent years, the sequencing of de novo transcriptomes has illuminated the way for molecular investigations of copepod physiology and copepod responses to environmental change (Baumgartner & Tarrant, 2017; Lenz et al., 2014; Lenz, Lieberman, et al., 2021; Lizano et al., 2022; Tarrant et al., 2019).

Copepods of the genus *Calanus*, of which three congeners in the North Atlantic and Arctic (*C. hyperboreus*, *C. glacialis*, and *C. finmarchicus*) play a key role in marine ecosystems. Whilst *C. hyperboreus* and *C. glacialis* dominate the Arctic, it is *C. finmarchicus* which dominates the zooplankton biomass in the North Atlantic basins of the Labrador Sea, the Norwegian Sea and adjacent areas North (Helaouët & Beaugrand, 2007). *C. finmarchicus* are primary consumers, rich in lipids, and in turn sustain higher trophic levels, including fish, seabirds, and whales, conferring them a central position in the North Atlantic and Arctic food webs. Their life cycle is well described (Baumgartner & Tarrant, 2017) and characterized by an active phase synchronized with the annual peak of food availability and a diapause state between the active phases (Baumgartner & Tarrant, 2017). During the few months of the active state, they perform diel vertical migrations (DVM), rising to the surface at night where they feed on phytoplankton and accumulate body lipids up to 70% of body volume (Helaouët & Beaugrand, 2007). After this period of DVM, *C. finmarchicus* enters diapause characterized by a subsequent vertical migration typically to mesopelagic and bathypelagic depths (although some will diapause in shallow fjord or shelf environments) for several months (Bandara et al., 2021), where temperatures are generally cooler than in the epipelagic with food and predators less abundant or absent. Due to the transport and sequestration of a significant amount of carbon into deep aquatic layers, this seasonal diapause migration is significant in global biogeochemical cycling

(Baumgartner & Tarrant, 2017). Emergence from diapause occurs synchronously in the following year. After the return to the epipelagic, *C. finmachicus* moults into the adult male/female stage, mates, and produces offspring production whose peak abundance coincides with the spring phytoplankton bloom, before the whole life cycle begins again (Baumgartner & Tarrant, 2017; Helaouët & Beaugrand, 2007).

Anticipating the optimal time for entering and exiting diapause is of central importance for the success of the copepod's life cycle and may not solely depend on direct responses to environmental parameters. Indeed, waiting for direct cues of the adverse season (with no food availability) to initiate diapause could prove deadly, because of the substantial energy reserves required for diapause emergence several months later. It is also highly adaptive to anticipate the spring algae bloom to emerge from diapause and reproduce to ensure the synchronization of offspring and peak phytoplankton bloom, which ensures the survival of the subsequent generation. Indeed, waiting for the algal bloom to directly cue emergence from diapause would not leave sufficient time to reach the juvenile stages that are physiologically ready to accumulate energy. Finally, a synchronous emergence of the population from diapause, that is, emergence from diapause over a comparably short time span within a given population, is crucial to optimize mating and reproduction. Therefore, in a highly seasonal food environment, we postulate that the evolution of an endogenous circannual clock that controls the timing of biological processes, such as regulating neuroendocrine pathways that control diapausing or non-diapausing development, is adaptive. However, environmental changes induced by climate change might disrupt the synchrony between the annual rhythms of organisms and their environment, impacting copepods fitness and resulting in population shifts which can lead to changes in biogeochemical cycles as well as in functional biodiversity, due to the central role of copepods in the food web (Baumgartner & Tarrant, 2017).

The putative circannual clock mechanism is assumed to drive the rhythmic expression of downstream genes whose rhythmic translation and function ultimately underlie phenology. While transcriptomic rhythms are increasingly studied at the daily and tidal scales, including in copepods (Payton et al., 2021), studies at the annual scale are still rare (Dopico et al., 2015; Nagano et al., 2019). Here, we performed a rigorous and demanding sampling campaign to investigate the annual transcriptome of *C. finmarchicus* in order to reveal the dynamics of gene expression and to gain insights into the mechanisms underlying diapause control and phenology.

2 | MATERIALS AND METHODS

2.1 | Sampling and site description

Sampling was conducted in Bonawe deep, the deepest point (≈150m) of Loch Etive, a fjordic sea loch on the west coast of Scotland, UK (56.45° N, 5.18° W) (Figure 1a). All sampling complied with the international ethical standards. The detailed sampling procedure is available in Payton, Noirot, et al. (2022). Briefly, net sampling was conducted with the R/V Seol Mara (Scottish Association for Marine
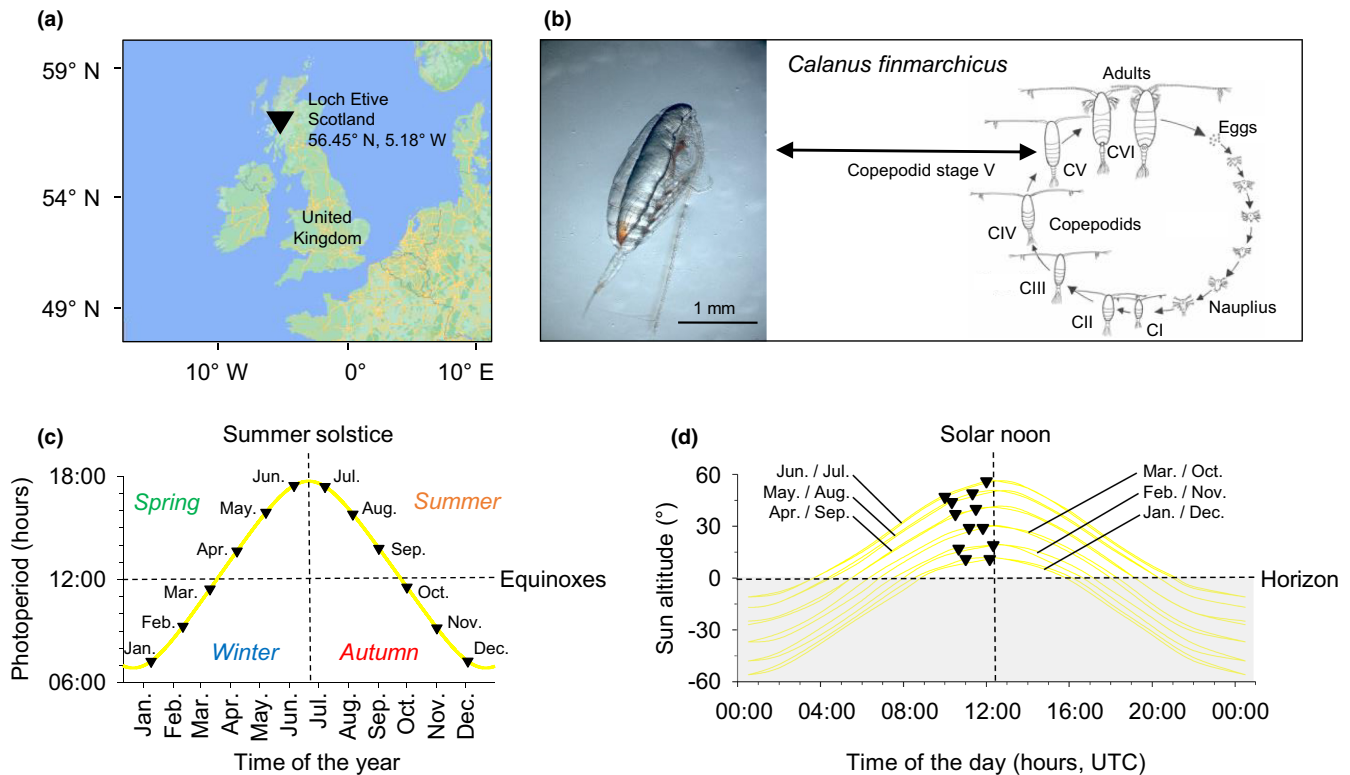
FIGURE 1 Sampling description. (a) Sampling site: Loch Etive, Scotland, UK (56.45° N, 5.18° W). (b) Life cycle of the study species, the zooplanktonic copepod *Calanus finmarchicus* (drawings from NERC ZIMNES project). Photo of a *C. finmarchicus* copepodid stage V (CV), the stage of interest, with its lipid sac. (c) The yellow line shows the annual photoperiod, and the 12 sampling times are indicated by dark triangles. Sampling was performed between the 1st and the 12th of each month of the year 2019, covering the entire annual photoperiod cycle (12 sampling times, indicated by dark triangles). (d) The yellow lines show the daily sun altitude cycle occurring on the day of each sampling time. Each sample (indicated by dark triangle) was performed between 10:00 am and 12:15 pm (UTC), while solar noon is at ~12:30 pm (UTC).

Science, SAMS) with a WP2-net equipped with an opening/closing mechanism hauled vertically through the water column from 140 m to 50 m depth. Monthly sampling was carried out between the 1st and the 12th of each month of 2019. At the sampling site, the annual minimal daily light duration was ≈6 h 50 min (at winter solstice), and the maximum was ≈17 h 45 min (at summer solstice). The 12 monthly samples covered the four seasons (Figure 1c). January, February, and March samples covered the winter season (short days <12 h that progressively got longer). April, May, and June samples covered the spring season (long days >12 h that also progressively got longer). July, August, and September samples covered the summer season (long days >12 h that progressively got shorter). Finally, October, November, and December samples covered the autumn season (short days <12 h that progressively get shorter). Besides shorter daylight duration, shorter photoperiods are associated with lower sun altitude above the horizon at solar noon and lower sun altitude below the horizon at solar midnight (Figure 1d). Each sample was collected between 10:00 am and 12:15 pm (UTC), while solar noon is at ≈12:30 pm (UTC). Finally, each sample was collected on a waxing moon (between New Moon and First Quarter Moon) and within 2 h 46 min around high tide. Monthly conductivity, temperature, and depth (CTD) measurements were taken (Last & Dumont, 2021), and very small changes in

temperature (11.4 ± 0.3°C) and salinity (27.8 ± 0.1 psu) were observed between months at the sampling depth.

2.2 | Copepods populations storage and characterization

Net samples were fixed as bulk samples in RNA*later* stabilization solution (Ambion) within 5 min after retrieval. They were incubated at 2–4°C for 12 h before being transferred to −80°C for measurement of RNA content and transcriptional profiling. Duplicate net samples were collected at each time point with identical methodology and fixed in 4% buffered formalin solution for species identification, copepodid stage composition, and measurement of mean prosome length of CV copepodids. In the laboratory, formaldehyde was filtered and rinsed, and the samples were transferred to a round-bottomed flask and made up to 300 mL with tap water. After careful sample mixing, sub-samples were taken with a 5 mL stempel pipette and transferred to a Bogorov tray. The first 200 *Calanus* were identified to copepodid stage (CI-CV and CVI female and male). The first 100 CV copepodids were picked out from the Bogorov tray to a separate petri dish for prosome measurements. Further sub-samples

were taken if numbers were not sufficient. Copepods were measured under a stereo microscope using a calibrated eyepiece micrometre. Each copepod was orientated similarly for measurement, head to the left, on their right-hand side. In all months, *C. finmarchicus* was the only Calanus species found (no *C. helgolandicus*), supporting the assumption of extremely low risk of congener contamination (Payton, Noirot, et al., 2022). *C. glacialis* does not occur in this region.

2.3 | RNA extraction

RNA extraction was performed on the CV copepodid stage, the main overwintering stage that represents the bulk of the population over the year. Briefly, three replicates of 25 CV *C. finmarchicus* were sorted for each time point, and RNA were extracted using a Phenol/Chloroform-based single-step extraction in combination with a spin column based on solid-phase extraction (Direct-zol™ RNA MiniPrep Kit; Zymo Research, USA). The pool of 25 individuals allowed to get the sufficient amount of RNA required for RNA sequencing, to increase the representativeness of the population (75 copepods per time point, 900 copepods in total) and to decrease the effect of individual variability. The monthly relative total RNA content per CV copepodid individual was estimated based on pooled values and normalized by the mean minimal annual value.

2.4 | RNA sequencing

The detailed procedure for generation of the transcriptomic dataset is available in Payton, Noirot, et al. (2022), and datasets are available in NCBI Sequence Read Archive (Payton et al., 2022a) and Figshare (Payton et al., 2022b). RNA sequencing was performed at the GeT-PlaGe core facility, INRAE Toulouse, on normalized amounts of messenger RNA. The 36 RNA sequencing libraries (three replicates per month) were prepared according to Illumina's protocols using the Illumina TruSeq Stranded mRNA sample prep kit to sequence mRNA on a NovaSeq S4 lane (Illumina) using a paired-end read length of 2×150 pb with the Illumina NovaSeq Reagent Kits. The RNA sequencing libraries' read qualities were evaluated by checking the number of expected sequences, the GC percentage, the presence of adaptor, the overexpressed sequences, and putative contamination using FastQC (Payton, Noirot, et al., 2022). The quantification matrix was generated by alignment to the reference transcriptome (Lenz et al., 2014) using Salmon (Payton, Noirot, et al., 2022). The reference transcriptome included only unique components, that is, one or many individual assembled sequences ("contigs") that may differ in size which clustered (Lenz et al., 2014). In situations in which a given numbered unique components contained multiple contigs, the longest contig was selected for the reference (Lenz et al., 2014). An outlier component (a remaining 18S ribosomal RNA gene) in one of the libraries was removed, and only components with more than 1 CPM (Count Per Million) in at least one sample were kept using edgeR (Payton, Noirot, et al., 2022), resulting in a matrix of 35,357

unique components and an average of 99 ± 3 million mapped reads (mean $\pm$ SE) per sample. An annotation matrix was generated by combining two existing annotation works (Bioproject PRJNA628886; Payton et al., 2020 and Bioproject PRJNA236528; Lenz et al., 2014). Finally, a down-sampling normalization was applied to homogenize the sequencing depth to 66.7 million reads per sample for all samples, while retaining 35,357 unique components per sample. All further analyses in this study were made on that resulting dataset. In this manuscript, the use of "component" or "transcript" refers to "unique component". In our dataset, 64.0% of components have been assigned an annotation (Payton, Noirot, et al., 2022), and 44.5% have been attributed for a Gene Ontology (GO) term.

2.5 | PCA analysis

Principal component analysis (PCA) was performed using "fviz_pca()" provided by ggplot2 on RStudio (R Core Team, 2021). It was performed on pseudo-counts from the normalized dataset (down-sampling normalization, 35,357 components per sample).

2.6 | Rhythmic analysis

Rhythmic analysis of the 35,357 components' expression over the year was performed in RStudio (R Core Team, 2021), using RAIN package. RAIN was specifically designed to detect rhythms in biological datasets independent of waveform by using a non-parametric approach (Thaben & Westermark, 2014). The 35,357 components have been tested using the three samples per time point as replicates and the "independent" mode. As the time interval between two time points is 1 month, a period of "12" was tested as an annual period length. Different waveforms were tested, and the one yielding the most significant result was selected for each transcript. The false discovery rate of the p-values was corrected using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995). Heatmap was plotted using PAST software (https://past.en.lo4d.com/windows). Relative levels of expression (pseudo-counts normalized from 0 to 1) were calculated for each transcript by normalizing monthly levels of expression as (monthly level – minimum monthly level)/(maximum monthly level – minimum monthly level). Amplitude fold change of rhythmic components is calculated as (maximum monthly level – minimum monthly level)/minimum monthly level. In Figures 5–7, temporal profiles are normalized to the minimum, meaning (monthly level–minimum monthly level), where value 1 refers to the minimal annual value. Results of rhythmic analysis are available in Table S1a,b. Specific results presented in Figures 5–7 are also available in Tables S2d and S3.

2.7 | Gene ontology

Gene Ontology (GO) analyses of biological processes were performed in RStudio (R Core Team, 2021), using the pGO package

(Alexa et al., 2006). Enrichment analyses were performed using Fisher's exact test, "weight" method (Alexa et al., 2006), and "ratio" parameter, thanks to custom scripts (https://forgemia.inra.fr/bios4 biol/bioinfo-utils/-/blob/master/bin/GOEnrichment.R). Results of Gene Ontology analysis are available in Table S1c.

2.8 | Real-time quantitative PCR

RNA sequencing results were verified through real-time quantitative qPCR analysis. The same samples used for RNA sequencing were normalized to 1.3 μg total RNA equivalent and reverse transcribed to cDNA using Thermo Scientific RevertAid First Strand cDNA synthesis kit (Thermo Scientific). The expression level of five circadian core clock genes (*Clock*, *Cycle*, *Period*, *Timeless*, and *Cryptochrome 2*) relative to the geometric mean of five reference genes (*Elongation factor 1-*, *RNA polymerase*, *16 s rRNA*, *Ribosomal protein L 32*, and *Beta actin*) were measured using the PowerUp™ SYBR® Green Master Mix (ThermoFisher, USA) on a ViiA™ 7 system (Applied Biosystems, USA), on 10 ng cDNA per reaction in a 20 μL final volume. Primer sequences are available in Table S2a. They were designed from Christie et al. (2013) and Lenz et al. (2014), and already used in previous studies (Häfker et al., 2017, 2018; Payton et al., 2021). Raw data are available in Table S2b. Final monthly mean data ± standard errors of both real-time quantitative PCR and RNA sequencing analyses are available in Table S2c. Real-time quantitative PCR time series were tested for rhythmicity following the same procedure used for RNA sequencing, and results are available in Table S2d.

3 | RESULTS AND DISCUSSION

3.1 | Sampling campaign to investigate the annual transcriptome of *Calanus finmarchicus*

The sampling of *C. finmarchicus* was performed each month over 1 year spanning the annual photoperiodic cycle (Figure 1c), at the deepest point (≈150 m) of Loch Etive on the Scottish west coast (UK; 56.45° N, 5.18° W) (Figure 1a). We focused on the *Calanus* stage CV, the fifth copepodid pre-adult stage (Figure 1b), which experiences diapause and has the longest duration of all the ontogenetic stages, which range from few hours to a few weeks in duration (Baumgartner & Tarrant, 2017). This pre-adult diapause is the most common and ecologically important type of postembryonic diapause in marine copepods, characteristic of several species, and specifically the Calanidae and Eucalanidae families, which dominate copepod biomass in polar, subpolar, and many temperate environments (Baumgartner & Tarrant, 2017). Previous studies have shown that multiple environmental factors synchronize biological rhythms in *C. finmarchicus* and other marine species (Häfker et al., 2017; Last et al., 2016; Payton et al., 2021; Tessmar-Raible et al., 2011). Therefore, organisms were sampled at comparable phases of daily, synodic month, and tidal cycles, that is, around noon (Figure 1d),

between New Moon and First Quarter Moon, and consistently near the time of high tide (Payton, Noirot, et al., 2022).

3.2 | Characterization of the annual life cycle

In order to characterize the annual life cycle of the sampled copepods, we identified the monthly *Calanus* copepodid stages composition of the sampled population (Figure 2a), the monthly mean prosome length of CV individuals (Figure 2b), and the relative total RNA content per CV (Figure 2c). In the first third of the sampling period, the development from CV to adults started in February. It progressed in March, as indicated by the appearance of adults in the community (3.0% in February, 34.5% in March) (Figure 2a). April was the month with the lowest relative abundance of CV within the population (16.5%), while 71.5% of the *C. finmarchicus* community were adults. The mean prosome length of CV in March (intermediate between February and
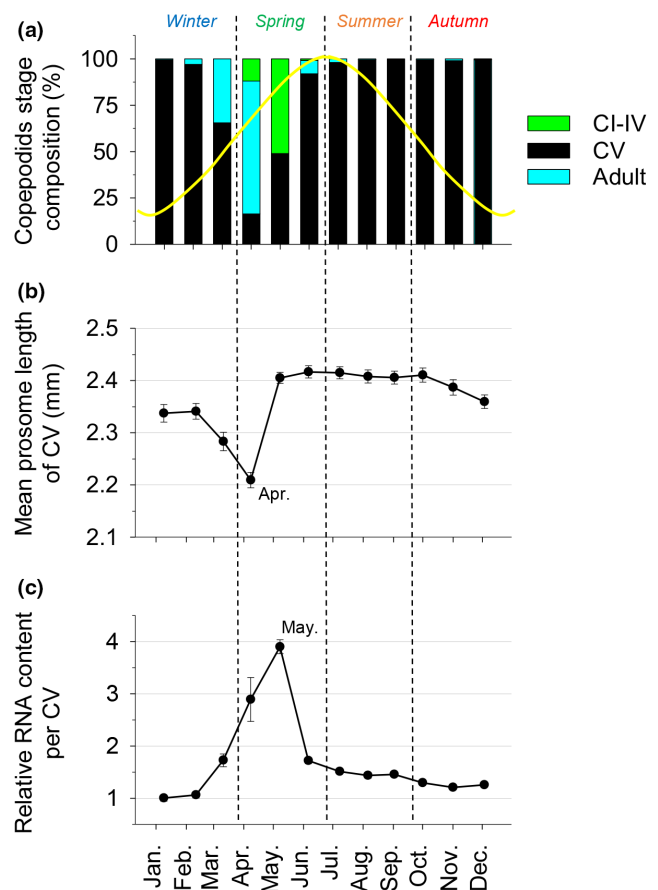

FIGURE 2 Characterization of the annual life cycle. (a) Monthly copepodids stage composition at the sampling site, expressed as the percentage of copepodid I to IV (green), copepodid V (CV, black), and adults (CVI female and male, blue) (*n* = 200 individuals per month). The yellow line indicates the photoperiod, and the dotted lines indicate the spring equinox, summer solstice, and autumn equinox respectively. (b) Monthly mean prosome length (mm) of CV (± SE, *n* = 100 copepodids CV per month). (c) Monthly relative RNA content per CV (± SE, *n* = 3 pools of 25 CV individuals per month), where the value 1 refers to the minimal annual value.

April) suggests the emergence of a new generation of CV (i.e. small animals) that dominates in April (Figure 2b). Thus, the CI-CIV stages most likely appeared before the March sampling but missed in this study due to likely short-lived peak that was missed with relatively infrequent sampling. The CV proportion increased again in May representing about half (49%) of the population, while the rest represented the smaller stages (CI-CIV) with no adults present at this time (*C. finmarchicus* is semelparous). In addition, CV prosome length increased by ~8% from April to May, indicating the growth of the new CV generation. From June, CV dominated the community (92% to 100% of the copepodids), with very little variation in CV prosome length. A small proportion of adult *C. finmacrhcus* were collected in June (7%), July (2%), August (0.5%), and October to January (1.0%) suggesting that some of the population may undergo a different life strategy and produce more than one generation per year (Baumgartner & Tarrant, 2017), where a second generation has been observed to descend to diapause in October (Häfker et al., 2018). The monthly relative RNA content per CV reflects transcriptional activity (Figure 2c). A clear annual pattern appeared, with an increase from March to a maximum in May, which coincides with the time of the phytoplankton peak in Loch Etive (Häfker et al., 2018; Last & Dumont, 2021), with a relative RNA content 3.9 times higher than in January. Increased RNA yields were previously observed in *C. finmarchicus* and may be

associated with favourable food conditions or an increase in metabolic activity (Häfker et al., 2018; Wagner et al., 1998). The increased RNA content is associated with the new generation of CV that appears in March (mixed with the initial CV population) and dominated the CV community in April and May. In May, the RNA content was highest, coinciding with the highest prosome length of the new CV generation (Figure 1a–c). Subsequently, the RNA contents decreased rapidly, suggesting the switch from the active to the diapause state. This assumption is supported by previous studies at Loch Etive (Clark et al., 2013; Häfker et al., 2018; Mauchline, 1987), showing the major relocation of *C. finmarchicus* biomass from the upper water column to deep layers in June (Häfker et al., 2018).

3.3 | Insight into the annual transcriptome

To reveal the dynamics of gene expression underlying the observed annual life cycle (Figure 2), we generated an extensive transcriptomic dataset of the monthly expression of 35,357 unique components in the whole body of *Calanus* stage CV over one full year (Payton, Noirot, et al., 2022). Principal component analysis (PCA) was used to visualize and investigate the variance of the annual gene expression dataset (Figure 3). The first principal component (PC1)
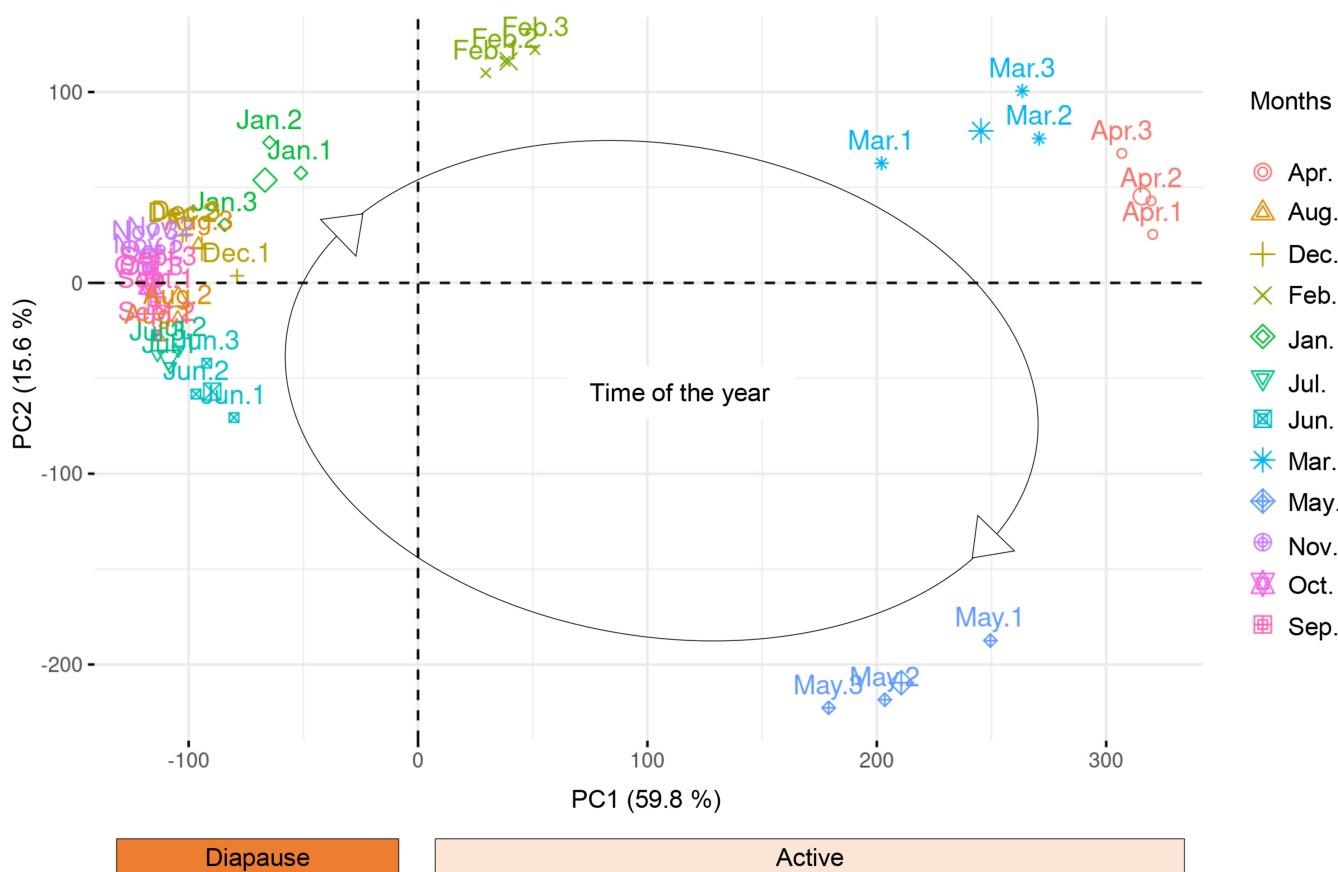

FIGURE 3 Principal component analysis (PCA) of the annual transcriptome. PCA made on the 35,357 unique components' pseudo-counts of the 36 samples, i.e. 3 replicates of 25 *Calanus* stage CV per month, from January to December 2019. The black circle is manually added to indicate the time of the year. The two orange rectangles indicate the putative diapause (dark orange) and active (pale orange) states.

explains 59.8% of the dataset's variance, while the second (PC2) captures 15.6% of the variance, together explaining 75.4% of the dataset's variance. The three replicates per month are well clustered. Strikingly, the display of the 12 groups of monthly replicates shows a clear temporal organization along PC1 and PC2, which follows the time of the year depicted in Figure 3. June to January clustered together on the left part of PC1 axis, distributed in chronological order along the vertical PC2 axis, from bottom to top. In contrast, February to May are clustered on the right part of PC1 axis. In view of previous results (Figure 2), we assume that samples on the left part of PC1 axis cluster 'diapause' transcriptomic profiles, while samples on the right part indicate 'active' transcriptomic profiles. This analysis indicates that copepods showed an 'active' transcriptomic profile in February, when the post-diapause maturation to adult stage was initiated, with no associated increase in RNA content (Figure 2). Interestingly, a high variance is noted between 'active' transcriptomic profiles, compared to the low variance between 'diapause' transcriptomic profiles. For instance, while March and April are relatively clustered, February and May are clearly dissociated. This may illustrate the specificity of biological processes, that is, a post-diapause state of the previous CV generation in February and a pre-diapause state of the new CV generation in May. Whilst the 'diapause' transcriptomic profiles are closely clustered, they still show a clear separation between months. In particular, June (below PC1 axis) and January (above PC1 axis) profiles dissociate from the other months, corresponding to the first and last month of 'diapause', respectively. These 2 months correspond to the diapause initiation and termination phases described in insects (Denlinger, 2022b). In summary, our approach highlights that in *C. finmarchus*, there is an annual modulation of gene expression with dramatic dissociation between 'active' and 'diapause' transcriptomic profiles and distinct transition phases into and out of diapause. Although previous studies have described transcriptomic investigations in copepods (Clark et al., 2013; Freese et al., 2016; Häfker et al., 2018; Lenz, Lieberman, et al., 2021; Lenz, Roncalli, et al., 2021; Roncalli et al., 2018, 2020, 2021; Semmouri et al., 2020; Tarrant et al., 2014, 2016, 2019), this is the first time, to our knowledge, that a comprehensive annual transcriptomic profile has been revealed.

3.4 | Annual rhythms in the copepod transcriptome

To further characterize the annual transcriptomic dynamics, a robust chronobiological statistical analysis of all the transcriptome components (i.e. 35,357) using RAIN algorithm (Thaben & Westermark, 2014) was performed. Our analysis detected 32,070 components (90.7%) with a significant 12-month rhythm of expression (adjusted *p*-value = .05; Figure 4a,b, Table S1a,b). Among significantly rhythmic components, 29,950 (93.4%) show amplitudes of oscillations (fold change) above 1, which is twofold (Figure 4c). Annual transcriptomic dynamics revealed in Figure 4a corroborate and explain the clear distinction between monthly transcriptomic profiles revealed with the PCA analysis (Figure 3). Again, two clear

patterns of peak expression are observed in the heatmap: components with peak of expression from June to January (P-Jun., P-Jul., P-Aug., P-Sep., P-Oct., P-Nov., P-Dec., and P-Jan. groups), assigned to the 'diapause' profiles and components with peak of expression from February to May (P-Feb., P-Mar., P-Apr., and P-May. groups) assigned to the 'active' profiles. Overall, components with a peak of expression during one of the 'diapause' months are also relatively highly expressed during the other months of diapause. On the contrary, components with a peak of expression during one of the 'active' months are often only weakly expressed during months of diapause, and vice versa (Figure 4a). It is noteworthy that well-marked monthly-specific transcriptomic patterns are observed during 'active' months, while during 'diapause' months these are more homogeneous (Figure 4a). This corroborates the higher variance between 'active' profiles compared to 'diapause' profiles observed in the PCA analysis (Figure 3). Interestingly, the February pattern is clearly associated to an 'active' profile but shows a subtle persistence of 'diapause' profile. Indeed, genes that have a peak of expression in February (P-Feb. group) are also expressed, even weakly, during diapause months. These results support the hypothesis of a specific transition status of CV in February, corresponding to an 'active' post-diapause state of the initial community with no associated increase of RNA content (Figures 2 and 3). Nevertheless, many genes peaking in February (P-Feb.) are also expressed in March and April. However, they are very weakly expressed in May, the last month to accumulate energy resources before diapause, when CV from the new generation have reached their higher prosome length and show the highest RNA content (Figure 2). Indeed, results reveal that in May, a very specific group of components is expressed that at other times of the year show very weak levels of expression, comprising other months of 'active' state (P-May. group). This particular pattern corroborates the specific position of May samples in the PCA analysis (Figure 3). On the whole, the rhythmic analysis highlights that while the 'active' state duration represents only one-third of the year, 49.7%, of the rhythmic components (i.e. 15,942) have a peak of expression during these 4 months, underlining the transcriptomic diversity during the 'active' state (Figure 4d). To conclude, these results reveal widespread annual transcriptional rhythms in a diapausing species, highlighting a sharp switch between diapause and active states, but also a strong specific modulation between active months.

3.5 | Annual dynamics of biological processes

To explore the annual dynamics of biological processes, we generated a matrix that combines the rhythmic analysis results of all the components (35,357) with two existing transcriptomic annotation works in *C. finmarchicus* (Payton, Noirot, et al., 2022; Payton et al., 2022b) (Table S1b). We performed GO enrichment analysis on each group of components according to their expression peak month. Table 1 shows the four most significant enriched GO terms per group, focusing on biological processes, while the full lists of significantly enriched GO terms are available in Table S1c. In February, the 'active'
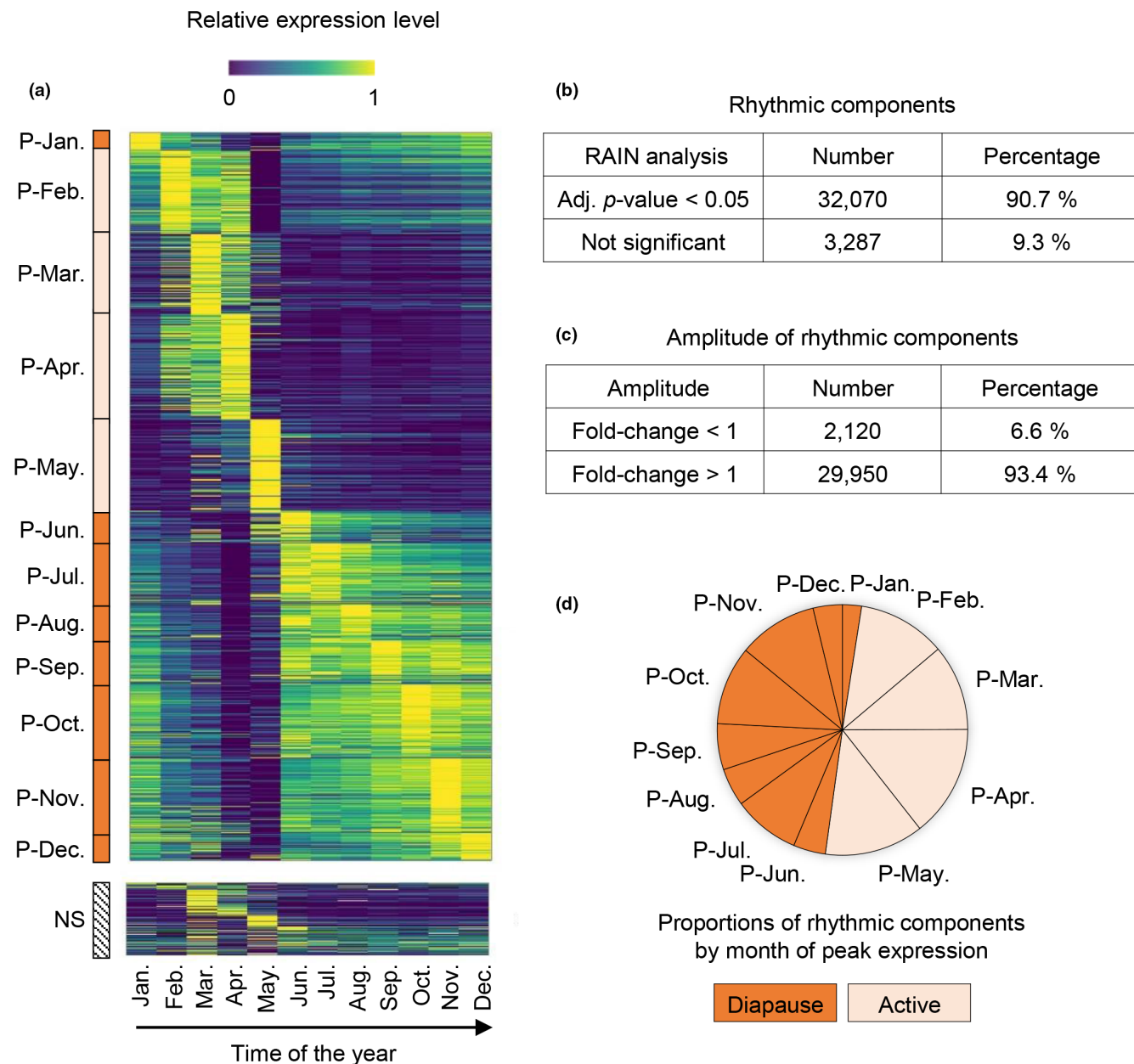
FIGURE 4 Rhythmic analysis of *Calanus finmarchicus* annual transcriptome. (a) Heatmap of the annual transcriptome (35,357 unique components) organized according to rhythmic analysis' results. Each line shows the relative expression level (pseudo-counts normalized from 0 to 1) of one unique component from January to December 2020 (from left to right). "P-Month" notes associated with orange rectangles indicate the monthly peak of expression during the year of the corresponding significantly rhythmic group of components. Dark and pale orange indicate that the corresponding group of components have a peak of expression during the putative diapause or active state, respectively. "NS" indicates the group of components with no significant annual rhythm. (b) Table with summarized rhythmic analysis results. See Table S1a,b for detailed results. (c) Amplitude of significantly rhythmic components, expressed as the number of components with a fold-change higher or lower than 1 between the minimal and the maximal annual expression values. A fold-change of 1 meaning twofold. (d) Proportions of rhythmic components by month of peak expression, and diapause (dark orange) or active (pale orange) states.

post-diapause month, enrichments are found for processes related to *nucleic acid metabolic process*, including *gene expression*, *mRNA processing*, *snRNA processing*; concomitant with processes such as *chromosome organization*, *chromatin remodelling*, *histone modification*, or *methylation*, which evokes epigenetic activity, discussed below. While epigenetic processes are mechanisms capable of switching gene activity (Stevenson, 2018), it is noteworthy that these results

are observed in February, when the switch from 'diapause' to 'active' transcriptomic profiles occurred (Figures 3 and 4). The enrichment of *spermatogenesis* is also observable in February (Table S1c), when gonad development is active and confirmed by previous microscopic and transcriptomic studies (Lenz, Roncalli, et al., 2021; Niehoff & Hirche, 1996; Tarrant et al., 2014, 2016). In both March and April, *chitin metabolic* processes are enriched. These are related to the

TABLE 1 List of the four most significantly enriched GO terms (Biological Processes) by month of peak expression.

Month peak of expression	GO (biological process)	GO definition	Enrichment p-value
P-Jan.	GO:0006412	Translation (protein synthesis)	1.6E−05
	GO:0006099	Tricarboxylic acid cycle (TCA cycle)	3.5E−03
	GO:0006457	Protein folding	3.5E−03
	GO:0006207	'De novo' pyrimidine nucleobase biosynthetic process	6.2E−03
P-Feb.	GO:0090304	Nucleic acid metabolic process	1.7E−45
	GO:0009059	Macromolecule biosynthetic process	5.4E−40
	GO:0006974	Cellular response to DNA damage stimulus	4.3E−32
	GO:0010467	Gene expression	7.2E−30
P-Mar.	GO:0006030	Chitin metabolic process	1.3E−17
	GO:0030163	Protein catabolic process	1.6E−04
	GO:0045454	Cell redox homeostasis	9.2E−04
	GO:0050482	Arachidonic acid secretion	1.1E−03
P-Apr.	GO:0015074	DNA integration	1.4E−07
	GO:0015986	ATP synthesis coupled proton transport	9.6E−07
	GO:0006260	DNA replication	4.7E−06
	GO:0006030	Chitin metabolic process	5.8E−06
P-May.	GO:0008152	Metabolic process	2.0E−23
	GO:0005975	Carbohydrate metabolic process	2.4E−21
	GO:0055085	Transmembrane transport	1.2E−11
	GO:0009166	Nucleotide catabolic process	9.8E−07
P-Jun.	GO:0005975	Carbohydrate metabolic process	1.1E−04
	GO:0034314	Arp2/3 Complex-mediated Actin nucleation	3.4E−03
	GO:0018095	Protein polyglutamylation	4.2E−03
	GO:0009306	Protein secretion	4.3E−03
P-Jul.	GO:0007154	Cell communication	2.1E−10
	GO:0006508	Proteolysis	1.3E−05
	GO:0070588	Calcium ion transmembrane transport	9.4E−05
	GO:0016266	O-glycan processing	1.5E−03
P-Aug.	GO:0006508	Proteolysis	8.5E−14
	GO:0016539	Intein-mediated protein splicing	9.7E−03
	GO:0006869	Lipid transport	1.3E−02
	GO:0006749	Glutathione metabolic process	2.4E−02
P-Sep.	GO:0007186	G protein-coupled receptor signalling pathway	9.2E−12
	GO:0030001	Metal ion transport	4.0E−05
	GO:0006811	Ion transport	4.2E−05
	GO:0033227	dsRNA transport	5.6E−05
P-Oct.	GO:0007186	G protein-coupled receptor signalling pathway	3.9E−06
	GO:0006355	Regulation of transcription, DNA-templated	5.0E−04
	GO:0042559	Pteridine-containing compound biosynthetic process	2.4E−03
	GO:0051260	Protein homooligomerization	4.5E−03
P-Nov.	GO:2000112	Regulation of cellular macromolecule biosynthetic process	8.4E−08
	GO:0006813	Potassium ion transport	7.8E−07
	GO:0009190	Cyclic nucleotide biosynthetic process	3.1E−06
	GO:0051260	Protein homooligomerization	4.9E−06
P-Dec.	GO:0006355	Regulation of transcription, DNA-templated	1.6E−03
	GO:0017157	Regulation of exocytosis	1.7E−03
	GO:0030178	Negative regulation of Wnt signalling pathway	1.7E−03
	GO:0007269	Neurotransmitter secretion	2.0E−03

Note: "P-Jan." means the group of components with a significant annual rhythm, with an expression peak in January. The full list is available in Table S1c. Dark and pale orange indicate the diapause or active state, respectively.
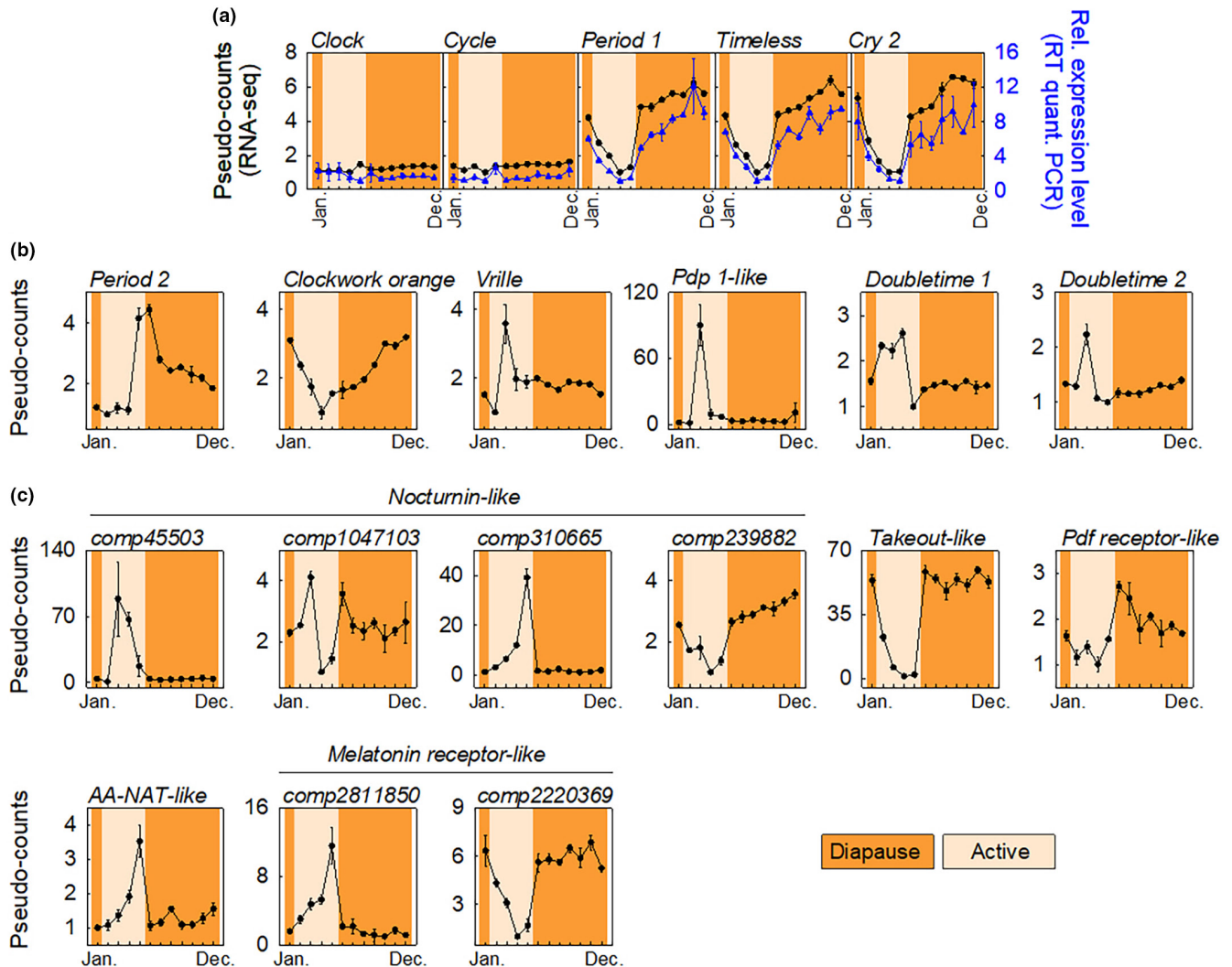
FIGURE 5 Annual expression profiles of circadian clock and circadian clock-related components. (a) Annual expression profiles of five core circadian clock genes investigated by both RNA sequencing (in black) and real-time quantitative PCR analysis (in blue). Detailed results are available in Table S2. (b) Annual expression profiles of additional circadian clock components investigated by RNA sequencing. (c) Annual expression profiles of circadian clock-related components investigated by RNA sequencing. Detailed results and transcript ID of Figure 1b,c are available in Table S3a. Except for Figure 1a, all components are significant for annual rhythm. All temporal profiles are normalized to the minimum, where the value 1 refers to the minimal annual value. The orange rectangles indicate the putative diapause (dark orange) and active (pale orange) states.

remodelling of cuticle (Tarrant et al., 2014), as would be expected for frequent moulting during ontogenic development, and their peak expression coincides with the observed growth in prosome length of *Calanus* stage CV during March and April (Figure 2). In April, *lipid storage* is enriched, as expected for the preparation of diapause (Baumgartner & Tarrant, 2017). During lipid metabolism for diapause, low-energy algal carbohydrates and proteins are converted into high-energy lipids, stored both as triacyclglycerols in oil droplets throughout the body and as wax esters in the oil sac that can occupy a large part of the body cavity (see photo Figure 1b). Enrichment analysis shows the high metabolic, feeding, and energy accumulation activities in May, coincident with the peak of phytoplankton (Häfker et al., 2018), as previously observed to a smaller extent (targeted transcriptomic approach) in Loch Etive (Häfker et al., 2018).

Metabolic, *carbohydrate metabolic*, *lipid metabolic*, *organic acid catabolic*, *acetyl-CoA biosynthetic*, and *cGMP biosynthetic* are among the enriched biological processes during the pre-diapause month. As previously discussed, the fall in RNA contents suggested a drastic decrease in metabolic activity after May (Figure 2). Nevertheless, *carbohydrate metabolic* processes are still enriched in June, such as *proteolysis* from June to August, which would suggest the persistence of feeding activity at least during diapause initiation, as previously observed in insects (Denlinger, 2022b). Although further investigations on specific genes are now necessary to support this suggestion, the persistence of feeding activity would corroborate the relatively high behavioural activity level observed in organisms from Loch Etive during the early diapause (July and August), compared with organisms experiencing diapause in deeper waters, with
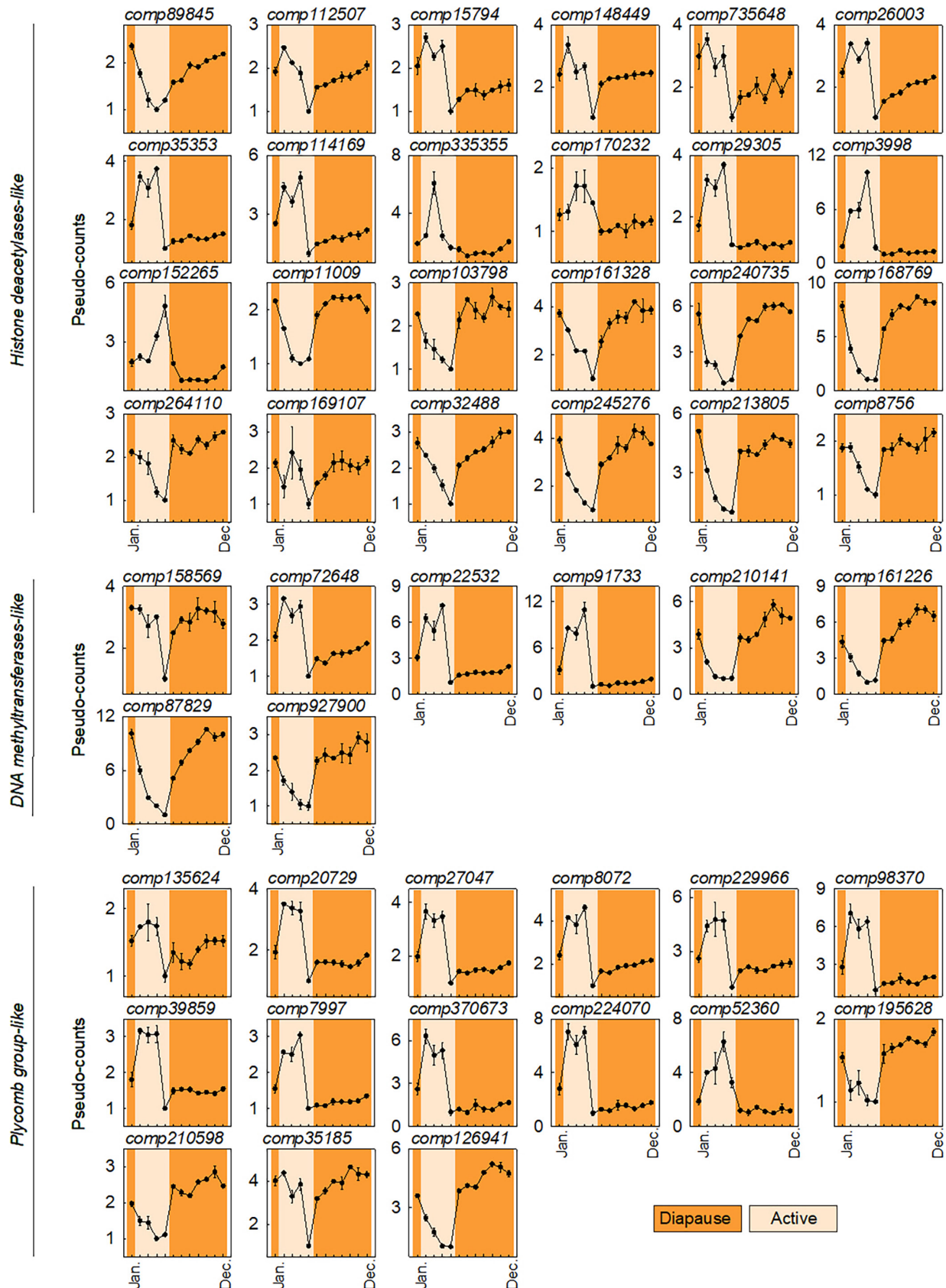
FIGURE 6 Annual expression profiles of epigenetic-related components. All temporal profiles are significant for annual rhythm and normalized to the minimum, where the value 1 refers to the minimal annual value. Components are ordered according to their monthly peak of expression. The orange rectangles indicate the putative diapause (dark orange) and active (pale orange) states. Detailed results and transcript ID are available in Table S3b.
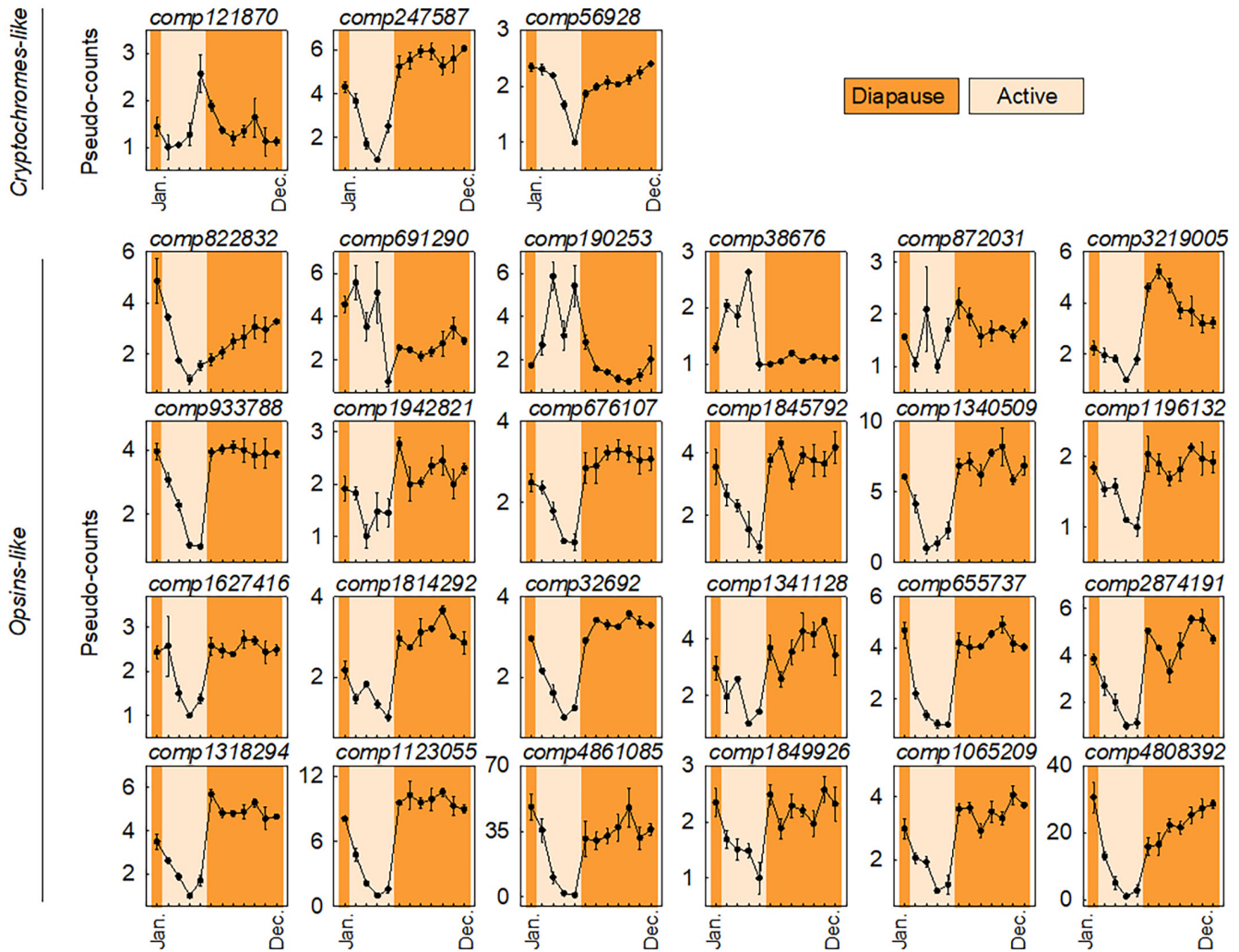
FIGURE 7 Annual expression profiles of light reception-related components. All temporal profiles are significant for annual rhythm and normalized to the minimum, where the value 1 refers to the minimal annual value. Components are ordered according to their monthly peak of expression. The orange rectangles indicate the putative diapause (dark orange) and active (pale orange) states. Detailed results and transcript ID are available in Table S3c.

less food availability (Coguiec et al., 2023; Grigor et al., 2022). The *G protein-coupled receptor* (*GPCR*) *pathway* is the most significantly enriched process in both September and October, and it was also enriched in August. GPCRs are involved in the perception of extracellular signals relayed to heterotrimeric G proteins and transduced intracellularly to appropriate downstream effectors (Liu et al., 2021). Interestingly, they have previously been described as receptor for a diapause hormone controlling the arrest development in the silkworm *Bombyx* (Homma et al., 2006). Among GPCRs are also opsins, involved in light reception, which corroborates with the enrichment of *phototransduction* process observed in October. Signalling processes appear to have a significant role during diapause, consistent with the diversity of networks that collectively generate the diapause phenotype in insects (Denlinger, 2022c). Among them, *Notch*, *JAK–STAT*, *TOR*, and *Wnt signalling pathways* have already been associated with molecular pathways that regulate diapause in insects (Denlinger, 2022c). Finally, it is notable that many processes related to ion transport are enriched during diapause months (*iron ion*

transport and *inorganic anion transport* in June; *calcium ion transmembrane transport* in July; *metal ion transport*, *ion transport*, and *cellular calcium homeostasis* in September; *potassium ion transport* in October and November; *calcium ion* and *ion transport* in November) and in March (*inorganic anion transmembrane transport*, *regulation of pH*), the month with a putative mix between post-diapausing organisms and CV copepodids from the new generation. These processes may be involved in buoyancy regulation during and after diapause, for which ions exchange is commonly used in marine taxa (Baumgartner & Tarrant, 2017; Pond, 2012).

3.6 | Mechanisms programming annual rhythms

Despite the apparent fitness and ecological importance of annual rhythms, the endogenous circannual clock mechanism underlying phenology is still poorly understood, probably since studying biological processes over such a long time period is logistically

challenging (Baumgartner & Tarrant, 2017; Denlinger, 2022a; Helm & Lincoln, 2017; Ikegami & Yoshimura, 2017). Nevertheless, the necessity of circadian clock genes for photoperiodic responses is well described in many groups (Denlinger, 2022a; Helm & Lincoln, 2017; Schwarzenberger et al., 2020). In particular, the circadian system may be involved in the photoperiodic time measurement (Denlinger, 2022a; Helm & Lincoln, 2017; Schwarzenberger et al., 2020). Circadian clock genes are initially described to be involved in the generation of daily rhythms. In *C. finmarchicus* (Christie et al., 2013), this mechanism has been suggested to be similar to the ancestral clock model known from the monarch butterfly *Danaus plexippus* (Denlinger, 2022a; Häfker et al., 2017; Zhu et al., 2005). Here, we investigated the annual oscillations of five core circadian clock genes and compared our RNA sequencing results with real-time quantitative PCR results (Figure 5a, Table S2). Both techniques led to comparable gene expression profiles, attesting to the reliability of the dataset. Although RAIN algorithm detects significant rhythms in *Clock* and *Cycle* annual profiles, the positive transcriptional regulators of the circadian clock, no clear annual patterns appear and very low amplitudes of oscillation are notable for these two genes (Table S2c,d). Contrary to this, very clear annual oscillations are observed in the three negative regulators of the circadian clock: *Period 1*, *Timeless*, and *Cry 2*, that are known to inhibit (as a protein heterotrimer) the action of the positive regulator CLOCK-CYCLE (acting as a protein heterodimer) in the monarch butterfly *Danaus plexippus* (Zhu et al., 2005). In *C. finmarchicus*, the overall annual pattern is common to the three components. It is characterized by an abrupt increase of expression in June (the first month of 'diapause'), maintained throughout 'diapause' (until January), before progressively decreasing in expression during the 'active' months reaching a minimum in April and May. Temporal profiles of *Period 1*, *Timeless*, and *Cry 2* corroborate previous results obtained from *C. finmarchicus* in Loch Etive (Häfker et al., 2018). An increased expression of *Period* during diapause compared to non-diapause organisms has also been observed in flies *Drosophila montana* and *D. virilis* (Kankare et al., 2016; Salminen et al., 2015). Interestingly, the higher expression of the negative regulators of the circadian clock (*Period*, *Timeless*, *Cry 2*) during diapause would not be associated with higher circadian clock oscillations at the daily scale at this time of the year. Indeed, daily oscillations of circadian clock genes in *C. finmarchicus* are maximal in May in Loch Etive, and hypothesized to be inactivated during diapause (Häfker et al., 2018).

In order to increase the breadth of investigations beyond the main core circadian clock genes, additional clock and clock-associated gene profiles were investigated revealing further evidence of an annual timing machinery (Figure 5b, Table S3a for detailed results and transcript ID). The additional clock and clock-associated genes have previously been characterized in silico in *C. finmarchicus* (Christie et al., 2013) except *Par Domain Protein 1* (*Pdp 1*). Noteworthy is that we identify a second *Period* component (*Period 2*) (Christie et al., 2013; Häfker et al., 2017) with an annual profile which differs from that of *Period 1*, suggesting dissociation between the two genes, at least during annual timing.

Clockwork orange is presumed to act as an additional repressor of the CLOCK-CYCLE protein complex in an independent feedback loop (Helfrich-Förster, 2017), and shows a comparable annual pattern with *Period 1*, *Timeless*, and *Cry 2*, despite a gradual increase during the 'diapause' phase rather than rapid increases of expression between May and June. *Vrille* and *Pdp 1-like* constitute a putative additional loop of the circadian clock in *C. finmarchicus* based on model species (Denlinger, 2022a; Häfker et al., 2017; Zhu et al., 2005). Both show a peak of expression in March, during the arrival of the next copepodids CV generation. Notably, the amplitude of oscillation in *Pdp 1-like* is approximately 90 times higher in March than in other months, suggesting that this gene may be an important component of the circannual clock. Although not significantly rhythmic, *Eyes absent-like* (*Eya-like*), a circadian clock-related gene involved in diapause regulation in insects (Denlinger, 2022a; Hidalgo et al., 2023), shows a similar expression profile with a peak of expression in March (Table S3a). Similar profiles are also observed in *Doubletime 2*, while *Doubletime 1*'s peak expression is from February to April. DOUBLETIMES are kinases involved in posttranscriptional modifications of circadian clock components, especially important for setting the period and amplitude of the circadian clock (Helfrich-Förster, 2017).

We also identified putative circadian clock-controlled components previously identified in the annual rhythms in other species (Figure 5c, Table S3a for detailed results and transcript ID). We reveal four *Nocturnin-like* and one *Takeout-like* component showing significant annual rhythms, with particularly high amplitude profiles for three (*Nocturnin-like comp45503* and *comp310665*, and *Takeout-like*). These genes are found downstream of the clock in the mosquito *Culex pipiens* and are suspected to be involved in diapause regulation (Chang & Meuti, 2020; Denlinger, 2022a). In *C. pipiens*, *Nocturnin* possibly plays a role in increasing fatty acid synthesis (Chang & Meuti, 2020; Denlinger, 2022a). *Takeout* is known to play a role in diapause-related functions such as neuroendocrine signalling pathway (Chang & Meuti, 2020; Denlinger, 2022a). In the same mosquito species, *Pigment dispersing factor* (*Pdf*) (or *pigment dispersing hormone* (*Pdh*) in crustaceans) is a circadian-related gene for which the knock-down results in diapause induction disruptions (Denlinger, 2022a; Meuti et al., 2015). In *Drosophila*, it has been proposed that PDF protein together with EYA protein are responsible for seasonal timing (Hidalgo et al., 2023). PDF signalling would negatively modulate EYA levels to regulate seasonal physiology (Hidalgo et al., 2023). Here, we did not find annotation for *Pdf*, but we identified one *Pdf receptor-like*, with higher expression during diapause and with maximal expression during diapause initiation. Consequently, our results suggest a higher abundance of PDF during 'diapause', and thus an antagonist regulation of PDF and EYA (with *Eya-like's* peak of expression in March, Table S3a) in *Calanus*, such as has been observed in *Drosophila* (Hidalgo et al., 2023). However, while EYA and PDF promote, respectively, 'winter' and 'summer' physiology in *Drosophila* (Hidalgo et al., 2023), the seasonal patterns would be the opposites in *Calanus*. Finally, melatonin has been reported in several of

invertebrates and vertebrates, as well as plants and bacteria, and shown to be involved in temporal signalling for daily, lunar, and annual rhythms (Denlinger, 2022a; Ikegami & Yoshimura, 2017; Reiter, 1993), including in zooplankton species (Schwarzenberger et al., 2023; Tosches et al., 2014). It is synthesized from serotonin by the enzymes arylalkylamine-N-acetyltransferase (AA-NAT) and hydroxyindole-O-methyltransferase (HIOMT), two outputs of the circadian clock. No *HIOMT-like* components were annotated in our data, but one *AA-NAT-like*, and two *Melatonin receptor-like* components were identified. Interestingly, *AA-NAT-like* shows a progressive increase of expression during the 'active' state and reaches a maximum in May, before a steep decrease in June and a lower level of expression during diapause (except a second small peak observed in October). While a comparable pattern is observed (with no peak in October) in the *Melatonin receptor-like comp2811850*, a very different pattern is observed in the *Melatonin receptor-like comp2220369* that shows maximum expression levels during 'diapause'.

How can a circannual clock control such drastic annual transcriptomic changes, as highlighted in this study (Figures 3 and 4)? Activation of gene expression requires access to promotor areas, which depend on DNA conformation. The latter is regulated by epigenetic mechanisms that control chromatin structure and genome landscape through methylation and acetylation. Epigenetic processes are usually investigated in relation to environmental conditions inducing permanent and inheritable genetic modifications. However, the role of epigenetics in circadian biology has indicated that epigenetic modifications exhibit daily reversible variations (Ripperger & Merrow, 2011; Sahar & Sassone-Corsi, 2013; Zhu & Belden, 2020). Indeed, histone acetylation (generated by the protein CLOCK; Doi et al., 2006) and deacetylation, as well as methylation processes, have all been shown to be part of the clock-controlled regulatory processes that generate circadian oscillations in gene expression (Ripperger & Merrow, 2011; Sahar & Sassone-Corsi, 2013; Zhu & Belden, 2020). More recently, epigenetic mechanisms were found to be involved in diapause regulation and even proposed to be integral components for timing endogenous circannual rhythms (Denlinger, 2022c; Reynolds, 2017; Stevenson, 2018; Stevenson & Lincoln, 2017; Wood et al., 2020). In our study, no histone acetylases were annotated (except *Clock*), but we did identify 24 histone *deacetylases-like*, all showing significant annual rhythms, with various profiles and peaks of expression from January to December (Figure 6, Table S3b for detailed results and transcript ID). We also identified eight *DNA methyltransferase-like*, all with significant annual rhythmic expression profiles. Four of these were annotated to be *Enhancer of Zest-like*, coding for histone methyltransferase enzymes (EZH), which have a direct role in maintaining circadian rhythms (Etchegaray et al., 2006). EZH are part of the Polycomb group (PcG) proteins, which elicit chromatin-mediated transcriptional repression by both the deacetylation and methylation of histone, involved in the core clockwork mechanism of mammals (Etchegaray et al., 2006). Fifteen *PcG proteins-like* have been identified here in *C. finmarchicus* with significant annual

rhythms. In summary, all identified genes have been found to show annual rhythms. These findings may suggest that epigenetic regulation could be involved in the architecture of the circannual system of *C. finmarchicus*.

Annual photoperiod changes are the most reliable external cue (Zeitgeber) to ensure the synchronous emergence from diapause necessary for mating (Denlinger, 2022a; Helm & Lincoln, 2017). However, diapausing copepods rest in deep layers, 400–2000 m in the open ocean (Baumgartner & Tarrant, 2017). Even though diapause depth in Loch Etive is far shallower (145 m), light intensity is strongly attenuated due to the input of humic compounds from rivers (Häfker et al., 2018; Wood et al., 1973) with light penetration to depth severely restricted. In this context, how can the circannual clock be synchronized by photoperiod during diapause? Here, we found an enrichment of the *phototransduction* process during diapause in October (Table S1c), implying increased light sensitivity. Widening our search, we looked for annual expression of phototransduction genes in copepods (Porter et al., 2017) (Figure 7, Table S3c for detailed results and transcript ID). We identified three *Cryptochrome-like*, all showing a significant annual rhythm with three different temporal profiles. *Comp121870* shows a peak in May, such as one of the *Nocturnin-like* (comp310665, Figure 5) as well as *AA-NAT-like* and one of the *Melatonin receptors* (comp2811850, Figure 5), while the two others have a peak of expression during 'diapause'. We also identified 27 *Opsin-like*, including *rhodopsins-like*, *pteropsin-like*, *peropsin-like*, *green-sensitive opsin-like*, *pinopsin-like*, *short-wavelength sensitive* (*sws*) *1 opsin-like*, and only three were not significant for an annual rhythm (Table S3c). Among the 24 *Opsin-like* with significant annual dynamics, 21 show a peak during diapause, that is, 87.5%. These results clearly suggest that an increase of light sensitivity during diapause in *C. finmarchicus* is likely and that would compensate for extremely low light intensities at diapause depths, which would ensure the sustainability of the annual clock entrainment by photoperiodic signals. These results corroborate the extremely low sensitivity thresholds that adequately provide the necessary photoperiodic signals in diapausing insects, such as the ones of the oriental fruit fly moth within the core of an apple, pupae of the silk moth in its dense cocoon, or parasitoids developing within the body of their host (Denlinger, 2022a).

3.7 | Annual rhythms in a changing environment

Copepods have diverse life-history strategies with population adaptations finely tuned over evolutionary time scales to adapt them to their local environments. While diapause termination is synchronized in a given copepod population, as also observed in insects, plasticity in timing may occur in diapause initiation in a given population (Baumgartner & Tarrant, 2017; Denlinger, 2022b). By analogy, in *Drosophila*, plasticity in annual rhythms is possible due to different *Timeless* isoforms suggested to confer an adaptive advantage at the population level in mitigating annual environmental

variability (Sandrelli et al., 2007). However, global climate change rapidly alters the annual dynamics of biotic (e.g. timing of phytoplankton blooms, seasonal migrations, etc.) and abiotic (e.g. temperature, precipitation, sea ice extent, etc.) environmental cycles. Consequently, mismatches may occur between organisms' annual rhythms and their environment, impairing the fitness of populations (Kharouba & Wolkovich, 2020). For example, an earlier algae bloom would impair the survival of the upcoming generation (Søreide et al., 2010). Indeed, our study revealed profound month-specific gene expression. Specifically, gene expression in May, when the peak of phytoplankton occurs, illustrated a very specific physiological preparation optimizing food intake and energy accumulation at this time. We revealed that in *C. finmarchus* there is a profound reorganization of gene expression between the diapause and the active state with overt endogenous physiological adaption. Understanding how the synchronization arises and what short-term adaptability of endogenous mechanisms exist in circannual timing is now particularly relevant at a time of rapid climate change and where environmental cycles are changing their amplitude and phase relationship to the annual calendar (Walker II et al., 2019).

AUTHOR CONTRIBUTIONS

L.P. designed the study, participated in the samples collection, extracted RNA, performed data analysis and wrote the manuscript; K.S.L. designed the study, collected samples, participated in the data analysis and the writing of the manuscript and supervised the study; J.G. collected samples and reviewed the manuscript; C.N. performed reads quality assessment, reads alignment on transcriptome and reviewed the manuscript; L.H. participated in the samples collection, sorted the copepods and reviewed the manuscript; D.V.P.C. identified copepods species and stages, measured the copepods length, and reviewed the manuscript; M.D. sorted the copepods, performed RT-qPCR analysis, and reviewed the manuscript; D.W. participated in the data analysis and reviewed the manuscript; B.M. designed the study, participated in the data analysis and the writing of the manuscript and supervised the study.

CONFLICT OF INTEREST STATEMENT

We have no conflict of interest.

DATA AVAILABILITY STATEMENT

All data generated or analysed during this study are included in this published article and its Supporting information files. The datasets analysed during the current study are available in the NCBI Sequence Read Archive repository, as BioProject PRJNA734260 (Payton et al., 2022a) and in Figshare collection 5450397 (Payton et al., 2022b), and described in Payton, Noirot, et al. (2022)

BENEFIT-SHARING

Benefits Generated: Benefits from this research accrue from the sharing of our data and results on public databases as described above.

ORCID

Laura Payton https://orcid.org/0000-0001-5090-0929

REFERENCES

Alexa, A., Rahnenführer, J., & Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, *22*, 1600–1607.

Bandara, K., Varpe, Ø., Wijewardene, L., Tverberg, V., & Eiane, K. (2021). Two hundred years of zooplankton vertical migration research. *Biological Reviews of the Cambridge Philosophical Society*, *96*, 1547–1589. https://doi.org/10.1111/brv.12715

Baumgartner, M. F., & Tarrant, A. M. (2017). The physiology and ecology of diapause in marine copepods. *Annual Review of Marine Science*, *9*, 387–411.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, *57*, 289–300.

Chang, V., & Meuti, M. E. (2020). Circadian transcription factors differentially regulate features of the adult overwintering diapause in the northern house mosquito, *Culex pipiens. Insect Biochemical Molecular Biology*, *121*, 103365.

Christie, A. E., Fontanilla, T. M., Nesbit, K. T., & Lenz, P. H. (2013). Prediction of the protein components of a putative *Calanus finmarchicus* (Crustacea, Copepoda) circadian signaling system using a de novo assembled transcriptome. *Comparative Biochemistry and Physiology. Part D, Genomics & Proteomics*, *8*, 165–193.

Clark, K. A. J., Brierley, A. S., Pond, D. W., & Smith, V. J. (2013). Changes in seasonal expression patterns of ecdysone receptor, retinoid X receptor and an A-type allatostatin in the copepod, *Calanus finmarchicus*, in a sea loch environment: An investigation of possible mediators of diapause. *General and Comparative Endocrinology*, *189*, 66–73.

Coguiec, E., Last, K. S., Cohen, J. H., Hobbs, L., Choquet, M., Ershova, E., Berge, J., & Daase, M. (2023). Photoperiodism and overwintering in boreal and sub-Arctic *Calanus finmarchicus* populations. *Marine Ecology Progress Series*, *712*, 49–65.

Denlinger, D. L. (2022a). Interpreting seasonal cues to program diapause entry. In D. L. Denlinger (Ed.), *Insect diapause* (pp. 57–120). Cambridge University Press.

Denlinger, D. L. (2022b). Confronting the challenges of a seasonal environment. In D. L. Denlinger (Ed.), *Insect diapause* (pp. 1–18). Cambridge University Press.

Denlinger, D. L. (2022c). Molecular signaling pathways that regulate diapause. In D. L. Denlinger (Ed.), *Insect diapause* (pp. 240–292). Cambridge University Press.

Doi, M., Hirayama, J., & Sassone-Corsi, P. (2006). Circadian regulator CLOCK is a histone acetyltransferase. *Cell*, *125*, 497–508.

Dopico, X. C., Evangelou, M., Ferreira, R. C., Guo, H., Pekalski, M. L., Smyth, D. J., Cooper, N., Burren, O. S., Fulford, A. J., Hennig,

B. J., Prentice, A. M., Ziegler, A. G., Bonifacio, E., Wallace, C., & Todd, J. A. (2015). Widespread seasonal gene expression reveals annual differences in human immunity and physiology. *Nature Communications*, 6, 7000.

Etchegaray, J.-P., Yang, X., DeBruyne, J. P., Peters, A. H. F. M., Weaver, D. R., Jenuwein, T., & Reppert, S. M. (2006). The polycomb group protein EZH2 is required for mammalian circadian clock function. *The Journal of Biological Chemistry*, 281, 21209–21215.

Freese, D., Søreide, J. E., & Niehoff, B. (2016). A year-round study on digestive enzymes in the Arctic copepod *Calanus glacialis*: Implications for its capability to adjust to changing environmental conditions. *Polar Biology*, 39, 2241–2252.

Grigor, J. J., Freer, J. J., Tarling, G. A., Cohen, J. H., & Last, K. S. (2022). Swimming activity as an indicator of seasonal diapause in the copepod *Calanus finmarchicus*. *Frontiers in Marine Science*, 9, 9528.

Häfker, N. S., Meyer, B., Last, K. S., Pond, D. W., Hüppe, L., & Teschke, M. (2017). Circadian clock involvement in zooplankton diel vertical migration. *Current Biology*, 27, 2194–2201.e3.

Häfker, N. S., Teschke, M., Last, K. S., Pond, D. W., Hüppe, L., & Meyer, B. (2018). *Calanus finmarchicus* seasonal cycle and diapause in relation to gene expression, physiology, and endogenous clocks. *Limnology and Oceanography*, 63, 2815–2838.

Helaouët, P., & Beaugrand, G. (2007). Macroecology of *Calanus finmarchicus* and *C. Helgolandicus* in the North Atlantic Ocean and adjacent seas. *Marine Ecology Progress Series*, 345, 147–165.

Helfrich-Förster, C. (2017). The drosophila clock system. In V. Kumar (Ed.), *Biological timekeeping: Clocks, rhythms and behaviour* (pp. 133–176). Springer India.

Helm, B., & Lincoln, G. A. (2017). Circannual rhythms anticipate the Earth's annual periodicity. In V. Kumar (Ed.), *Biological timekeeping: Clocks, rhythms and behaviour* (pp. 545–569). Springer India.

Hidalgo, S., Anguiano, M., Tabuloc, C. A., & Chiu, J. C. (2023). Seasonal cues act through the circadian clock and pigment-dispersing factor to control EYES ABSENT and downstream physiological changes. *Current Biology*, 20, 1006–1008.

Homma, T., Watanabe, K., Tsurumaru, S., Kataoka, H., Imai, K., Kamba, M., Niimi, T., Yamashita, O., & Yaginuma, T. (2006). G protein-coupled receptor for diapause hormone, an inducer of Bombyx embryonic diapause. *Biochemical and Biophysical Research Communications*, 344, 386–393.

Ikegami, K., & Yoshimura, T. (2017). Molecular mechanism regulating seasonality. In V. Kumar (Ed.), *Biological timekeeping: Clocks, rhythms and behaviour* (pp. 589–605). Springer India.

Kankare, M., Parker, D. J., Merisalo, M., Salminen, T. S., & Hoikkala, A. (2016). Transcriptional differences between Diapausing and non-Diapausing *D. Montana* females reared under the same photoperiod and temperature. *PLoS One*, 11, e0161852.

Kharouba, H. M., & Wolkovich, E. M. (2020). Disconnects between ecological theory and data in phenological mismatch research. *Nature Climate Change*, 10, 406–415.

Last, K. S., & Dumont, E. (2021). CTD time-series data collection from Loch Etive, Scotland from September 2018 to October 2020.

Last, K. S., Hobbs, L., Berge, J., Brierley, A. S., & Cottier, F. (2016). Moonlight drives ocean-scale mass vertical migration of zooplankton during the Arctic winter. *Current Biology*, 26, 244–251.

Lenz, P. H., Lieberman, B., Cieslak, M. C., Roncalli, V., & Hartline, D. K. (2021). Transcriptomics and metatranscriptomics in zooplankton: Wave of the future? *Journal of Plankton Research*, 43, 3–9.

Lenz, P. H., Roncalli, V., Cieslak, M. C., Tarrant, A. M., Castelfranco, A. M., & Hartline, D. K. (2021). Diapause vs. reproductive programs: Transcriptional phenotypes in a keystone copepod. *Communications Biology*, 4, 1–13.

Lenz, P. H., Roncalli, V., Hassett, R. P., Wu, L. S., Cieslak, M. C., Hartline, D. K., & Christie, A. E. (2014). De novo assembly of a transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) – The dominant zooplankter of the North Atlantic Ocean. *PLoS One*, 9, e88589.

Liu, N., Wang, Y., Li, T., & Feng, X. (2021). G-protein coupled receptors (GPCRs): Signaling pathways, characterization, and functions in insect physiology and toxicology. *International Journal of Molecular Sciences*, 22, 5260.

Lizano, A. M., Smolina, I., Choquet, M., Kopp, M., & Hoarau, G. (2022). Insights into the species evolution of *Calanus copepods* in the northern seas revealed by de novo transcriptome sequencing. *Ecology and Evolution*, 12, e8606.

Mauchline, J. (1987). Zooplankton in the Oban area and especially in upper loch Etive. *Scott Marine Biological Association*, 144, 1–91.

Meuti, M. E., Stone, M., Ikeno, T., & Denlinger, D. L. (2015). Functional circadian clock genes are essential for the overwintering diapause of the northern house mosquito, *Culex pipiens*. *Journal of Experimental Biology*, 218, 412–422.

Nagano, A. J., Kawagoe, T., Sugisaka, J., Honjo, M. N., Iwayama, K., & Kudoh, H. (2019). Annual transcriptome dynamics in natural environments reveals plant seasonal adaptation. *Nature Plants*, 5, 74–83.

Niehoff, B., & Hirche, H.-J. (1996). Oogenesis and gonad maturation in the copepod *Calanus finmarchicus* and the prediction of egg production from preserved samples. *Polar Biology*, 16, 601–612.

Payton, L., Hüppe, L., Noirot, C., Hoede, C., Last, K. S., Wilcockson, D., Ershova, E., Valière, S., & Meyer, B. (2021). Widely rhythmic transcriptome in *Calanus finmarchicus* during the high Arctic summer solstice period. *iScience*, 24, 101927.

Payton, L., Noirot, C., Hoede, C., Hüppe, L., Last, K., Wilcockson, D., Ershova, E. A., Valière, S., & Meyer, B. (2020). Daily transcriptomes of the copepod *Calanus finmarchicus* during the summer solstice at high Arctic latitudes. *Scientific Data*, 7, 415.

Payton, L., Noirot, C., Last, K. S., Grigor, J., Hüppe, L., Conway, D. V. P., Dannemeyer, M., Suin, A., & Meyer, B. (2022). Annual transcriptome of a key zooplankton species, the copepod *Calanus finmarchicus*. *Ecology and Evolution*, 12, e8605.

Payton, L., Payton, L., Noirot, C., Last, K. S., Grigor, J., Hüppe, L., Conway, D. V. P., Dannemeyer, M., Suin, A., & Meyer, B. (2022a). BioProject PRJNA734260. Deep sequencing of the annual transcriptome of a key zooplanktonic species: The copepod *Calanus finmarchicus*. *NCBI Sequence Read Archive*. http://www.ncbi.nlm.nih.gov/bioproject/734260

Payton, L., Payton, L., Noirot, C., Last, K. S., Grigor, J., Hüppe, L., Conway, D. V. P., Dannemeyer, M., Suin, A., & Meyer, B. (2022b). Deep sequencing of the annual transcriptome of a key zooplanktonic species: The copepod Calanus finmarchicus. *Figshare*. https://doi.org/10.6084/m9.figshare.c.5450397

Pond, D. W. (2012). The physical properties of lipids and their role in controlling the distribution of zooplankton in the oceans. *Journal of Plankton Research*, 34, 443–453.

Porter, M. L., Steck, M., Roncalli, V., & Lenz, P. H. (2017). Molecular characterization of copepod photoreception. *The Biological Bulletin*, 233, 96–110.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reiter, R. J. (1993). The melatonin rhythm: Both a clock and a calendar. *Experientia*, 49, 654–664.

Reynolds, J. A. (2017). Chapter five - epigenetic influences on diapause. In H. Verlinden (Ed.), *Advances in insect physiology* (Vol. 53, pp. 115–144). Academic Press.

Ripperger, J. A., & Merrow, M. (2011). Perfect timing: Epigenetic regulation of the circadian clock. *FEBS Letters*, 585, 1406–1411.

Roncalli, V., Cieslak, M. C., Castelfranco, A. M., Hopcroft, R. R., Hartline, D. K., & Lenz, P. H. (2021). Post-diapause transcriptomic restarts: Insight from a high-latitude copepod. *BMC Genomics*, 22, 409.

Roncalli, V., Cieslak, M. C., Hopcroft, R. R., & Lenz, P. H. (2020). Capital breeding in a diapausing copepod: A transcriptomics analysis. *Frontiers in Marine Science*, 7, 56.

Roncalli, V., Sommer, S. A., Cieslak, M. C., Clarke, C., Hopcroft, R. R., & Lenz, P. H. (2018). Physiological characterization of the emergence from diapause: A transcriptomics approach. *Scientific Reports*, 8, 1–15.

Sahar, S., & Sassone-Corsi, P. (2013). The epigenetic language of circadian clocks. In A. Kramer & M. Merrow (Eds.), *Circadian Clocks* (pp. 29–44). springer.

Salminen, T. S., Vesala, L., Laiho, A., Merisalo, M., Hoikkala, A., & Kankare, M. (2015). Seasonal gene expression kinetics between diapause phases in Drosophila virilis group species and overwintering differences between diapausing and non-diapausing females. *Scientific Reports*, 5, 11197.

Sandrelli, F., Tauber, E., Pegoraro, M., Mazzotta, G., Cisotto, P., Landskron, J., Stanewsky, R., Piccin, A., Rosato, E., Zordan, M., Costa, R., & Kyriacou, C. P. (2007). A molecular basis for natural selection at the timeless locus in *Drosophila melanogaster. Science*, 316, 1898–1900.

Schwarzenberger, A., Bartolin, P., & Wacker, A. (2023). Photoperiodic adaptation of aanat and clock gene expression in seasonal populations of *Daphnia pulex. Chronobiology International*, 20, 1–9.

Schwarzenberger, A., Chen, L., & Weiss, L. C. (2020). The expression of circadian clock genes in Daphnia magna diapause. *Scientific Reports*, 10, 19928.

Semmouri, I., de Schamphelaere, K. A. C., van Nieuwerburgh, F., Deforce, D., Janssen, C. R., & Asselman, J. (2020). Spatio-temporal patterns in the gene expression of the calanoid copepod *Temora longicornis* in the Belgian part of the North Sea. *Marine Environmental Research*, 160, 105037.

Søreide, J. E., Leu, E., Berge, J., Graeve, M., & Falk-Petersen, S. (2010). Timing of blooms, algal food quality and *Calanus glacialis* reproduction and growth in a changing Arctic. *Global Change Biology*, 16, 3154–3163.

Stevenson, T. J. (2018). Epigenetic regulation of biological rhythms: An evolutionary ancient molecular timer. *Trends in Genetics*, 34, 90–100.

Stevenson, T. J., & Lincoln, G. A. (2017). Epigenetic mechanisms regulating circannual rhythms. In V. Kumar (Ed.), *Biological timekeeping: Clocks, rhythms and behaviour* (pp. 607–623). Springer India.

Tarrant, A. M., Baumgartner, M. F., Hansen, B. H., Altin, D., Nordtug, T., & Olsen, A. J. (2014). Transcriptional profiling of reproductive development, lipid storage and molting throughout the last juvenile stage of the marine copepod *Calanus finmarchicus. Frontiers in Zoology*, 11, 91.

Tarrant, A. M., Baumgartner, M. F., Lysiak, N. S. J., Altin, D., Størseth, T. R., & Hansen, B. H. (2016). Transcriptional profiling of metabolic transitions during development and diapause preparation in the copepod *Calanus finmarchicus. Integrative and Comparative Biology*, 56, 1157–1169.

Tarrant, A. M., Nilsson, B., & Hansen, B. W. (2019). Molecular physiology of copepods - from biomarkers to transcriptomes and back again. *Comparative Biochemistry and Physiology. Part D, Genomics & Proteomics*, 30, 230–247.

Tessmar-Raible, K., Raible, F., & Arboleda, E. (2011). Another place, another timer: Marine species and the rhythms of life. *BioEssays*, 33, 165–172.

Thaben, P. F., & Westermark, P. O. (2014). Detecting rhythms in time series with RAIN. *Journal of Biological Rhythms*, 29, 391–400.

Tosches, M. A., Bucher, D., Vopalensky, P., & Arendt, D. (2014). Melatonin signaling controls circadian swimming behavior in marine zooplankton. *Cell*, 159, 46–57.

Wagner, M., Durbin, E., & Buckley, L. (1998). RNA:DNA ratios as indicators of nutritional condition in the copepod *Calanus finmarchicus. Marine Ecology Progress Series*, 162, 173–181.

Walker, W. H., II, Meléndez-Fernández, O. H., Nelson, R. J., & Reiter, R. J. (2019). Global climate change and invariable photoperiods: A mismatch that jeopardizes animal fitness. *Ecology and Evolution*, 9, 10044–10054.

Wood, B. J. B., Tett, P. B., & Edwards, A. (1973). An introduction to the phytoplankton, primary production and relevant hydrography of loch Etive. *Journal of Ecology*, 61, 569–585.

Wood, S. H., Hindle, M. M., Mizoro, Y., Cheng, Y., Saer, B. R. C., Miedzinska, K., Christian, H. C., Begley, N., McNeilly, J., McNeilly, A. S., Meddle, S. L., Burt, D. W., & Loudon, A. S. I. (2020). Circadian clock mechanism driving mammalian photoperiodism. *Nature Communications*, 11, 4291.

Zhu, H., Yuan, Q., Froy, O., Casselman, A., & Reppert, S. M. (2005). The two CRYs of the butterfly. *Current Biology*, 15, R953–R954.

Zhu, Q., & Belden, W. J. (2020). Molecular regulation of circadian chromatin. *Journal of Molecular Biology*, 432, 3466–3482.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.