



**HAL**  
open science

## Annual transcriptome of a key zooplankton species, the copepod *Calanus finmarchicus*

Laura Payton, Céline Noirot, Kim S Last, Jordan Grigor, Lukas Hüppe, David V P Conway, Mona Dannemeyer, Amandine Suin, Bettina Meyer

### ► To cite this version:

Laura Payton, Céline Noirot, Kim S Last, Jordan Grigor, Lukas Hüppe, et al.. Annual transcriptome of a key zooplankton species, the copepod *Calanus finmarchicus*. *Ecology and Evolution*, 2022, 12 (2), pp.e8605. 10.1002/ece3.8605 . hal-04096227

**HAL Id: hal-04096227**

**<https://hal.inrae.fr/hal-04096227>**

Submitted on 12 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## RESEARCH ARTICLE

# Annual transcriptome of a key zooplankton species, the copepod *Calanus finmarchicus*

Laura Payton<sup>1,2</sup>  | Céline Noirot<sup>3</sup>  | Kim S. Last<sup>4</sup>  | Jordan Grigor<sup>4</sup>  |  
Lukas Hüppe<sup>1,2,5,6</sup>  | David V. P. Conway<sup>7</sup> | Mona Dannemeyer<sup>2</sup> | Amandine Suin<sup>8</sup> |  
Bettina Meyer<sup>1,2,6</sup> 

<sup>1</sup>Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

<sup>2</sup>Section Polar Biological Oceanography, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

<sup>3</sup>Plateforme bio-informatique GenoToul, MIAT, INRAE, UR875 Mathématiques et Informatique Appliquées Toulouse, Castanet-Tolosan, France

<sup>4</sup>Scottish Association for Marine Science, Oban, UK

<sup>5</sup>Neurobiology and Genetics, Theodor-Boveri Institute, Biocentre, University of Würzburg, Würzburg, Germany

<sup>6</sup>Helmholtz Institute for Functional Marine Biodiversity (HIFMB), University of Oldenburg, Oldenburg, Germany

<sup>7</sup>Marine Biological Association of the UK, Plymouth, UK

<sup>8</sup>Plateforme Génomique, INRAE US 1426 GeT-PlaGe, Centre INRAE de Toulouse Occitanie, Castanet-Tolosan, France

## Correspondence

Laura Payton and Bettina Meyer, Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University of Oldenburg, Oldenburg 26111, Germany.  
Emails: laura.payton@uni-oldenburg.de; bettina.meyer@awi.de

## Funding information

Bundesministerium für Bildung und Forschung, Grant/Award Number: 03F0803A; Natural Environment Research Council, Grant/Award Number: NE/R012733/1

## Abstract

The copepod *Calanus finmarchicus* (Crustacea, Copepoda) is a key zooplanktonic species with a crucial position in the North Atlantic food web and significant contributor to ocean carbon flux. Like many other high latitude animals, it has evolved a programmed arrested development called diapause to cope with long periods of limited food supply, while growth and reproduction are timed to take advantage of seasonal peaks in primary production. However, anthropogenic warming is inducing changes in the expected timing of phytoplankton blooms, suggesting phenological mismatches with negative consequences for the N. Atlantic ecosystem. While diapause mechanisms are mainly studied in terrestrial arthropods, specifically on laboratory model species, such as the fruit fly *Drosophila*, the molecular investigations of annual rhythms in wild marine species remain fragmentary. Here we performed a rigorous year-long monthly sampling campaign of *C. finmarchicus* in a Scottish Loch (UK; 56.45°N, 5.18°W) to generate an annual transcriptome. The mRNA of 36 samples (monthly triplicate of 25 individuals) have been deeply sequenced with an average depth of  $137 \pm 4$  million reads (mean  $\pm$  SE) per sample, aligned to the reference transcriptome, and filtered. We detail the quality assessment of the datasets and provide a high-quality resource for the investigation of wild annual transcriptomic rhythms (35,357 components) in a key diapausing zooplanktonic species.

## KEYWORDS

annual rhythms, copepods, diapause, transcriptome, wildlife, zooplankton

## TAXONOMY CLASSIFICATION

Ecophysiology

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

The calanoid copepod *Calanus finmarchicus* (Crustacea, Copepoda) is a key ecological species in the North Atlantic that often dominates zooplankton biomass between ~40° and 80°N (Helaouët & Beaugrand, 2007). Primarily herbivorous, this species has a central position in the food web as it converts sugars from micro-algae (phytoplankton) into energy-rich lipids that sustain higher consumers including marine fish larvae, seabirds, and whales (Beaugrand et al., 2003; Berge et al., 2012; Prokopchuk & Sentyabov, 2006). Its high abundance and primary consumer position also makes it an important contributor to ocean carbon flux (Archibald et al., 2019). Despite this, copepods are among important non-model invertebrates for which genomic resources are still limited, principally due to their large genomes (Bron et al., 2011; Choquet et al., 2019). However, the de novo transcriptome of *C. finmarchicus* (Lenz et al., 2014) represents the most useful resource to date for molecular investigations in this ecologically important species.

Seasonal adaptations have evolved in zooplankton to deal with annual biotic and abiotic cycles, such as long periods of limited food resources (Baumgartner & Tarrant, 2017; Daase et al., 2021; Lenz & Roncalli, 2019). Copepods within the Calanidae and Eucalanidae families synchronize reproduction and growth with periods of high food availability, during which they accumulate large energetic reserves in sunlit surface waters (Baumgartner & Tarrant, 2017). This follows a migration of copepods into deeper water where they enter a prolonged phase of inactivity referred to as diapause, characterized by a non-feeding state with reduced metabolism lasting 8–10 months. The energy reserves accumulated during the spring production phase sustain the animals during periods of low food supply and are crucial to fuel moulting to adult stages and reproduction in the following year. In *C. finmarchicus*, the seasonal cycle begins prior to, and during, the spring bloom, when the juvenile pre-adult copepodid stage CV emerge from diapause, molt into adults, mate and spawn eggs (Baumgartner & Tarrant, 2017; Niehoff et al., 2002). The resulting offspring then develop through several nauplii and juvenile copepodid stages. The energy resource provided by the spring bloom is optimized by large lipid accumulations, specifically by the last copepodid stage CV. At this stage, CVs either directly migrate into deeper water where they enter diapause or molt into adults to produce another generation that enters diapause at a later time (Häfker et al., 2018; Johnson et al., 2008; Tarrant et al., 2014, 2016). The number of generations and timing of diapause vary according to populations and geographic locations (Häfker et al., 2018; Melle et al., 2014).

While the seasonal timing of the copepod life cycle is crucial for its survival and reproduction, broad geographic temperature increases caused by global climate change alter the seasonal dynamics of biotic and abiotic environmental cycles (Helm et al., 2013; Kharouba & Wolkovich, 2020). In particular, shifts in seasonal timing of phytoplankton blooms have been observed (Friedland et al., 2018). Concurrently, climate change induces the shift of *C. finmarchicus* to higher latitudes (Murphy et al., 2016; Reygondeau & Beaugrand,

2011; Saikkonen et al., 2012) with new residences predicted in high boreal and Arctic seas (Tarling et al., 2021). Consequently, climate change may lead to phenological mismatches, where the timing of critical events between interacting species that have been finely tuned over evolutionary time scales becomes desynchronized (Helm et al., 2013; Kharouba & Wolkovich, 2020). Considering the key position of lipid-rich copepods in the food web (Archibald et al., 2019; Beaugrand et al., 2003; Berge et al., 2012; Prokopchuk & Sentyabov, 2006), such mismatches may ultimately have consequences on energy flow pathways in ecosystems and carbon sequestration (Friedland et al., 2018). While our current understanding of the life cycle and ecology of copepods has emerged mostly from population dynamics and physiological studies (Baumgartner & Tarrant, 2017; Lenz & Roncalli, 2019), seasonal investigations on the transcriptional level are still rare and fragmentary (Clark et al., 2013; 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), as too are the processes controlling the timing of the annual life cycle (Baumgartner & Tarrant, 2017; Lenz & Roncalli, 2019).

Studying the seasonal life cycle of *C. finmarchicus* in its natural habitat, the open ocean, is extremely challenging. Populations are migrating vertically via diel vertical migrations (DVM) and horizontally via advection in ocean currents, and different copepod species frequently co-occur. However, an isolated population of *C. finmarchicus* can be found on the west coast of Scotland in Loch Etive. In the upper basin, including Bonawe deep (~150 m), *C. finmarchicus* dominates the mesoplankton community (Häfker et al., 2018; Hill, 2009; Mauchline, 1987), as less than 4% of CV and CVI stage individuals have been identified as *Calanus helgolandicus* (Hill, 2009). Indeed, the Loch Etive *Calanus* species composition was analyzed in 2016 genetically, and no *C. helgolandicus* was detected (Choquet et al., 2017). It has been suggested that *C. helgolandicus* is intolerant to low salinities, such as found in Loch Etive (Helaouët & Beaugrand, 2007; Hill, 2009). Primary production in this loch is comparatively low (Wood et al., 1973), and terrestrial organic matter run-off probably provides a supplementary food source for *C. finmarchicus*. Häfker et al. (2018) have described a peak of phytoplankton production in May, followed by a sharp decrease until August, and very low Chl *a* concentration from November to March. Previous studies agreed on one single synchronized cohort of *C. finmarchicus* each year (Clark et al., 2012, 2013; Hill, 2009; Mauchline, 1987), which would start to descend in preparation for diapause (100–150 m depth) in July, while Häfker et al. (2018) argued for a possible second generation descending to diapause depth in October. During their active phase, *C. finmarchicus* perform DVM, with a daytime depth ~70 m (Häfker et al., 2018). Finally, the water column at Bonawe deep is highly stratified. Salinity and temperature in the top 50 m are variable seasonally due to the fluctuations in terrestrial freshwater run-off and tidal influence. Below 50 m, the water column is mixed with temperature and salinity extremely stable over the year (Häfker et al., 2018; Hill, 2009).

In this study, we took advantage of Loch Etive to perform a complete annual transcriptome of the copepod *C. finmarchicus*. The

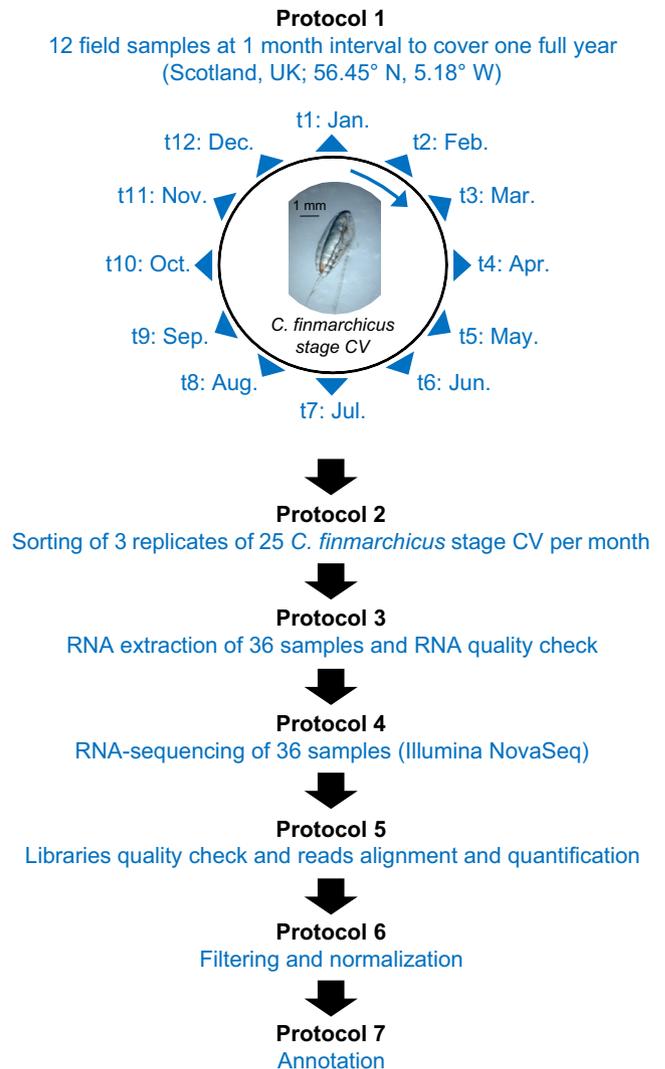
sampling campaign was conducted at Bonawe deep (UK; 56.45°N, 5.18°W) between 50 and 140 m depth, under extremely stable oceanographic conditions, but where both active and diapausing copepods are found. As copepods use multiple environmental cycles to synchronize their biological rhythms, from gene expression to behavior, the annual sampling was made at comparable phases of daily and lunar cycles (tidal and monthly cycles) (Häfker et al., 2017; Ibáñez-Tejero et al., 2018; Last et al., 2016; Payton et al., 2021; Schmitt et al., 2011). The genetic analysis focused on the copepodite pre-adult stage CV, which experiences all seasons and has a key role both in the active surface feeding and deep-water diapause states. The 36 mRNA samples (3 replicates of 25 individuals per month), analyzed with a high sequencing depth, yielded a total of 4.9 billion reads, with an average of  $137 \pm 4$  million reads (mean  $\pm$  SE) per sample. In addition, we provide a quality assessment and the alignment to the reference transcriptome (Lenz et al., 2014), allowing the monthly quantification of 35,357 unique “components” in the *C. finmarchicus* transcriptome and the first investigation of annual transcriptomic rhythms in a key zooplanktonic species. Finally, existing annotations have been combined for further transcriptomic data exploration.

## 2 | MATERIALS AND METHODS

### 2.1 | Sampling design

Sampling design and analysis strategy are presented in Figure 1, Table 1, and Table S1.

Sampling was conducted over a year, with monthly samples from January to December 2019 (between the 1st and the 12th of each month), resulting in 12 time points. Sampling covered the winter (January–March), spring (April–June), summer (July–September), and autumn (October–December) seasons. Sampling was performed between 10:00 am and 12:15 pm, while local solar local noon is at ~12:30 pm, in order to be at a comparable phase of the daily rhythm (all times noted in UTC) (detailed in Table S2). Furthermore, sampling was performed on waxing moons (between New Moon and First Quarter Moon) and within 2:46 h of high tide, to provide comparable phases of lunar cycles (Table S2). However, while all samples were taken around high tide, some were during the flood tide and others during the ebb tide (Table S2). Copepods were collected from the RV *Calanus* (Scottish Association for Marine Science, SAMS) using a WP2-net equipped with an opening/closing mechanism (200  $\mu$ m mesh size/50  $\mu$ m cod end) hauled vertically through the water column at 0.5 ms from 140 m to 50 m depth. Samples were immediately (within 5 min after retrieval) fixed as bulk samples in RNAlater stabilization solution (Ambion). A 12 h period of incubation at 2–4°C was allowed to soak the samples thoroughly with the stabilization solution before they were transferred to –80°C for further transport and storage. Duplicate net sampling, with identical methodology, was carried out at the same time for species identification samples. These were fixed in 4% buffered formalin solution.



**FIGURE 1** Overview of the experimental workflow used to generate the transcriptomic data output. Sample details are available in Tables S1 and S2

### 2.2 | Site description

Sampling was conducted in Bonawe deep, the deepest point (~150 m) of Loch Etive, a sea loch on the western coast of Scotland, UK (56.45°N, 5.18°W). Two sills with a depth of ~7 m and ~13 m at Connel and Taynult villages respectively limit the exchange with the open ocean (Häfker et al., 2018). The deeper waters of the sampling station typically cycle from normoxic to hypoxic and due to the sills, overturning of these water masses happens irregularly every ~18 months when spring tides in spring/autumn co-occur with low precipitation (Edwards & Edelsten, 1977). Monthly CTD (conductivity, temperature, and depth) measurements were performed and archived by the British Oceanographic Data Centre (BODC) (Last & Dumont, 2021). As previously shown (Häfker et al., 2018), large seasonal changes in temperature and salinity occur above 50 m which is highly stratified, while only small seasonal changes occur in deeper waters (mixed), where net sampling was conducted. The average

TABLE 1 Summary of sampling and sequencing strategy. Details are available in Tables S1 and S2

Number of time points	Sampling frequency	Total duration of sampling	Number of replicates per time point	Total number of samples	Sequencing strategy	Reads	Platform	Volume
12	1 month	1 year	3	36	RNA-seq	Paired-end 2 × 150 pb	Illumina NovaSeq	1 lane S4

temperature and salinity from 140 m to 50 m depth over the year were  $11.4 \pm 0.3^\circ\text{C}$  and  $27.8 \pm 0.1$  psu (mean  $\pm$  SE) respectively (detailed in Table S2).

### 2.3 | Species identification

Formalin solution was removed from the monthly community samples by draining them through a 100 micron nylon gauze and then rinsing the sample with tap water. Animals were transferred to a round-bottomed flask and diluted to 300 ml with tap water. After careful sample mixing, sub-samples were taken with a 5 ml stem-pipette and transferred to a Bogorov tray. The first 200 *Calanus* were confirmed morphologically to corroborate species.

### 2.4 | Copepod sorting

Molecular analysis was performed on the CV copepodite stage of *C. finmarchicus*, the main overwintering stage that represents the bulk of the population over the year. Copepods were sorted at  $2^\circ\text{C}$  under a stereo microscope for stage (CV) using morphological characteristics. For each time point, 3 replicates of 25 *C. finmarchicus* CV were sorted ("Jan.1," "Jan.2," "Jan.3"; to "Dec.1," "Dec.2," "Dec.3") (Figure 1). The choice to pool 25 individuals was made to: 1) get the sufficient amount of RNA required for RNA sequencing and 2) increase the number of individuals thereby increasing the representativeness of the population (75 copepods per time point, 900 copepods in total) and to decrease the effect of individual variability.

### 2.5 | RNA extraction

Each replicate was transferred to a 2 ml Precellys<sup>®</sup> homogenization tube (Bertin Instruments, France), containing a mix of 1.4 mm and 2.8 mm ceramic beads and homogenized in 600  $\mu\text{l}$  of TRIzol<sup>®</sup> reagent (ThermoFisher Scientific, USA) with a Precellys<sup>®</sup> 24 Tissue Homogenizer (Bertin Instruments, France), using  $2 \times 15$  s of homogenization at 1400  $\times g$  with a 10 s. break between. For RNA extraction, a Phenol/Chloroform-based single-step extraction in combination with a spin column based on solid phase extraction (Direct-zol<sup>™</sup> RNA MiniPrep Kit, Zymo Research, USA) was used. Genomic DNA was removed by DNase I digestion on column as part of the RNA extraction kit and total RNA was eluted in ultra-pure water. Samples were sent to GeT-PlaGe core facility in dried-ice for RNA sequencing.

### 2.6 | RNA sequencing

RNA sequencing was performed at the GeT-PlaGe core facility, INRAE Toulouse. The 36 RNA sequencing libraries were prepared according to Illumina's protocols using the Illumina TruSeq Stranded mRNA sample prep kit to analyze mRNA. Briefly, mRNAs were selected using

poly-T beads. Then, RNAs were fragmented to generate double-stranded cDNA, and adaptors were ligated to be sequenced. Eleven cycles of PCR were applied to amplify libraries. Library quality was assessed using a fragment analyzer (Advanced Analytical Technologies, Inc.), and libraries were quantified by qPCR using the Kapa Library Quantification Kit (Roche). RNA sequencing experiments were performed on a NovaSeq S4 lane (Illumina) using a paired-end read length of  $2 \times 150$  pb with the Illumina NovaSeq Reagent Kits. A total of 4,928,221,676 reads were sequenced. Raw RNA sequencing data are available as BioProject PRJNA734260 (Payton et al., 2022a).

## 2.7 | Reads alignment and quantification

Thirty-six RNA sequencing libraries ("Jan.1" to "Dec.3") were obtained (Figure 1, Table 1, and Table S1). The number of reads per library was between 93 million and 191 million with an average of  $137 \pm 4$  million (mean  $\pm$  SE) reads (Table 2, Table S3). The RNA sequencing libraries' read qualities were evaluated by checking the number of expected sequence, the GC percentage, the presence of adaptor, and the overexpressed sequences using FastQC (Andrews, 2010). Contamination was checked by aligning reads against *E. coli*, Yeast, and PhiX genomes. No further contamination investigations have been necessary considering the satisfactory reads alignment to the de novo reference transcriptome described thereafter and in the Section 3.

The *C. finmarchicus* de novo transcriptome (Lenz et al., 2014), based on different life stages and deposited to Bioproject PRJNA236528, was used as the reference transcriptome. It is composed of 206,012 contigs comprising 96,090 unique components ("comps") and presents good results of quality assessment, with a nearly complete BUSCO set (Lenz et al., 2014; Tarrant et al., 2019). To generate the quantification matrix, reads alignment to the reference transcriptome (Lenz et al., 2014) and quantification was performed with Salmon (Patro et al., 2017). An average of  $102 \pm 3$  million reads (mean  $\pm$  SE) were mapped to the 96,090 unique comps (Lenz et al., 2014) (Table 2, Table S3).

## 2.8 | Filtering

The quantification matrix revealed one outlier component (comp92\_c0\_seq1; corresponding to the unique transcript GAXK01170082.1). The outlier comp was searched in the NCBI nucleotide collection (the collection consists of GenBank, EMBL, DDBJ, PDB, and RefSeq sequences) using basic local alignment search tool (BLAST) and was found to have 99.98% identity with a *C. finmarchicus* 18S ribosomal RNA gene (accession number MF993124.1). This result illustrates

that despite the selection of mRNA using poly-T beads, a small fraction of ribosomal RNA (rRNA) can remain and be variable between libraries (Abernathy & Overturf, 2016). 26.2% and 28.8% of mapped reads were aligned to this component in the samples "Mar.1" and "Jul.3," respectively, against  $0.4 \pm 0.1\%$  (mean  $\pm$  SE) in all other samples, comprising the associated replicates "Mar.2," "Mar.3," "Jul.1," and "Jul.2" (Table S3). For these reasons, and because of the clear increase in replicate correlations after filtering this outlier (see in Section 3), this component was removed. Additionally, the matrix was filtered with edgeR (Robinson et al., 2010) and only "comps" with more than 1 CPM (Count Per Million) in at least one sample were kept, providing a matrix of 35,357 unique "comps" and an average of  $99 \pm 3$  mapped reads (mean  $\pm$  SE) per sample (Table 2, Table S3). The quantification matrix for the 35,357 components after filtering is available in Figshare (Payton et al., 2022b).

## 2.9 | Annotation

An annotation matrix was generated by combining two existing annotation works: Bioproject PRJNA628886 (Payton et al., 2020) and Bioproject PRJNA236528 (Lenz et al., 2014) (Table 3). Bioproject PRJNA628886 (Payton et al., 2020) consisted of the selection of the best annotation hit of *C. finmarchicus* reference transcriptome (Lenz et al., 2014) against NR, TrEMBL, and Swissprot databases, using DIAMOND (Buchfink et al., 2015) (<https://doi.org/10.6084/m9.figshare.c.5127704>). Bioproject PRJNA236528 (Lenz et al., 2014) consisted of an annotation of the *C. finmarchicus* reference transcriptome (Lenz et al., 2014) against the non-redundant (nr) protein database (<https://doi.org/10.5061/dryad.11978>). Details of the number of annotated components per databases are provided in Table 3. Joining the two annotation works led to 22,643 components over 35,357 with at least one annotation (Table 3). GO (Gene Ontology) annotation was done using Bioproject PRJNA628886 (Payton et al., 2020), in which InterProScan (Jones et al., 2014) provided annotations from many protein signature databases (<https://doi.org/10.6084/m9.figshare.c.5127704>). 15,726 components over 35,357 were assigned to a GO annotation (Table 3). The combined annotation matrix of the 35,357 components is available in Figshare (Payton et al., 2022b).

## 2.10 | Normalization

Non-uniformities in sequencing depths or library sizes (the total number of mapped reads) between samples mean that the observed counts are not directly comparable between samples. A down-sampling normalization was then proposed, constituting

**TABLE 2** Mapping results on the reference transcriptome (96,090 comps) before and after filtering. Details are available in Table S3

	Average number of aligned reads (million, mean $\pm$ SE)	Number of comps
Mapping results before filtering	$102 \pm 3$	96,090
Mapping results after filtering	$99 \pm 3$	35,357

**TABLE 3** Number (and percentage) of annotated comps in the working transcriptome (35,357 comps), using annotation works from Bioproject PRJNA628886 (Payton et al., 2020) and Bioproject PRJNA236528 (Lenz et al., 2014)

Dataset	Annotation source	Database	Annotated comps	Comps with at least one annotation
Working transcriptome (35,357 comps)	Bioproject PRJNA628886	NR	9779 (27.7%)	22,643 (64.0%)
		Trembl	7868 (22.3%)	
		SwissProt	3803 (10.8%)	
	Bioproject PRJNA236528	Non-redundant database	21,938 (62.0%)	
	Bioproject PRJNA628886	InterProScan (GO)	15,726 (44.5%)	15,726 (44.5%)

down-sampling the mapped reads to the lowest number among the 36 samples (after filtering, Table S3), that is, to 66.7 million reads per sample for all samples, in order to adjust for differences in sequencing depth among samples. This normalization, previously shown suitable for rhythmic studies (Hughes et al., 2017; Koike et al., 2012; Li et al., 2015; Payton et al., 2021), is appropriate considering the high sequencing depth of the sample with the smaller library size (66.7 million reads after filtering, Table S3) (Li et al., 2015). This was performed with package R metaseqR (<https://rdrr.io/bioc/metaseqR/man/downsample.counts.html>). The quantification matrix of the 35,357 components after filtering and down-sampling normalization is available in Figshare (Payton et al., 2022b).

## 2.11 | Code availability

Parameters to software tools involved are described in the following paragraph.

FastQC: version 0,11,2, --nogroup --casava.

Salmon : vesrion 1.4.0

- indexation : salmon index -t c-finmarchicus\_96k\_ref.fasta -i transcripts\_index -k 31

- quantification : salmon quant -i transcripts\_index -l IU -1 read1.fastq.gz -2 read2.fastq.gz --

validateMappings -o output -p 8

EdgeR: version 3.26.5. Function cpm()

metaseqR: version 1.26.0. Function `downsample.counts` with `option`

seed=42<https://rdrr.io/bioc/metaseqR/man/downsample.counts.html>

## 3 | RESULTS

### 3.1 | Species identification results

In all months no *C. helgolandicus*, only *C. finmarchicus*, were found between 140 and 50 m depth, where molecular sampling was performed. This result supports the assumption of extremely low risk of congener contamination.

### 3.2 | Extraction and RNA integrity

RNA extraction procedures were performed with randomization of samples to ensure reliable and unbiased data production. RNA purity was assessed by OD measurements with a NanoDrop 8000 spectrophotometer (ThermoFisher Scientific, USA), and all 260/280 and 260/230 OD ratio was superior to 1.95. RNA integrity was evaluated with a Fragment Analyzer (Advanced Analytical Technologies, Inc., Iowa, USA; RNA Kit (15nt) Standard Sensitivity, Agilent). Due to a non-conventional 28S/18S ribosomal ratio in this species, sample quality was evaluated on the electropherogram (Asai et al., 2015; DeLeo et al., 2018). No degradation in the inter region was observed. Total RNA samples were stored at  $-80^{\circ}\text{C}$ .

### 3.3 | Raw reads assessment and quantification overview

All samples passed the FastQC (Andrews, 2010) "base quality control." No relevant contamination hit was found after the alignment

against *E. coli*, Yeast, and PhiX. Mapping of the sequenced reads against the 96,090 unique "comps" of the reference transcriptome (Lenz et al., 2014) yielded an overall alignment of  $102 \pm 3$  million reads (mean  $\pm$  SE), that is,  $74.5 \pm 0.3\%$  (mean  $\pm$  SE) of total reads (Table 2, Table S3). This result is in accordance with the previously observed result on the de novo transcriptome of *C. finmarchicus* study (Lenz et al., 2014), where the alignment rate against unique comps was of 75%.

### 3.4 | Effect of filtering and normalization on counts distribution

Samples counts distribution before and after removing the outlier component (comp92\_c0\_seq1; corresponding to the unique transcript GAXK01170082.1) clearly highlighted the overrepresentation of the latter in samples "Mar.1" and "Jul.3" (Figure 2a vs. Figure 2b). Moreover, filtering components expressed below 1 CPM in all samples clearly increased counts distribution (Figure 2b vs. Figure 2c). Finally, the down-sampling normalization aligned the distribution of counts across samples, without observable impacts on the proportion of very low and very high expressed components (Figure 2c vs. Figure 2d) implying that this is a satisfactory approach.

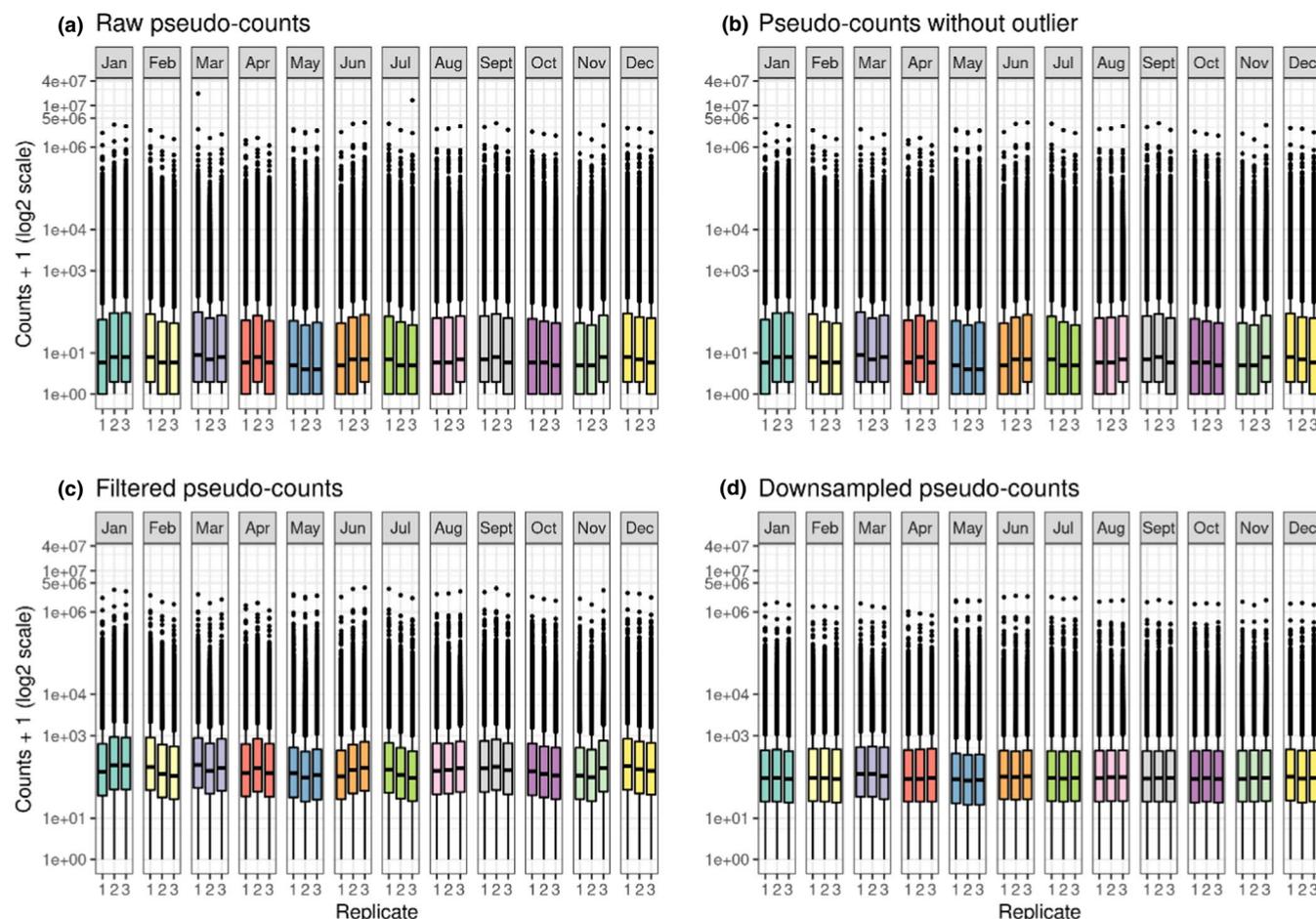
### 3.5 | Correlation between replicates

Filtering the outlier component led to an increase in replicate correlations for March and July samples (Figure 3, Table S4). While

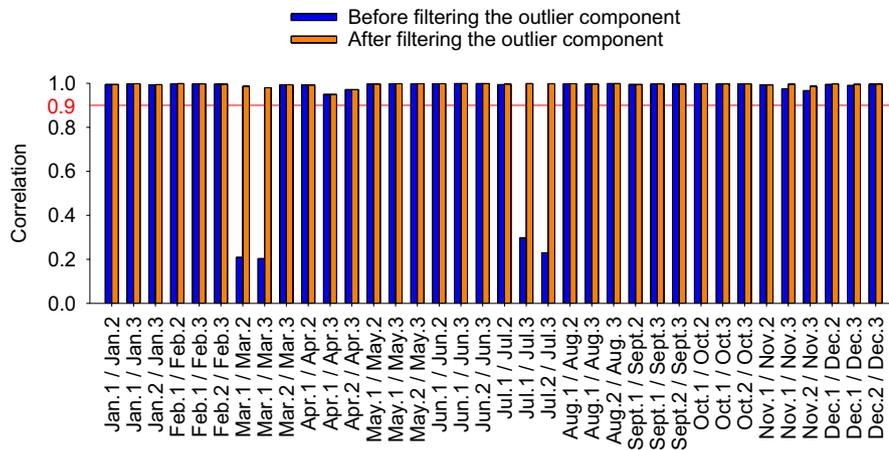
correlations between "Mar. 1/Mar. 2," "Mar. 1/Mar. 3" and "Jul. 1/Jul. 3," "Jul. 2/Jul. 3" were of 0.21, 0.20, 0.30, and 0.23 before filtering; they increased to 0.99, 0.98, 1.00, and 1.00 respectively after filtering the outlier component, validating our methodology. Note that the minimal replicate correlation after filtering is 0.95 ("Apr. 1/Apr. 3"), revealing the good comparability of replicates in this study (Figure 3, Table S4).

### 3.6 | Proportion of components expressed above the threshold of 1 CPM

Of the 96,090 components of the reference transcriptome (Lenz et al., 2014), 36.8% (35,357) were expressed above the threshold of 1 CPM (Table 2). This result corroborates previously observed results on the *C. finmarchicus* transcriptome (Lenz et al., 2014; Payton et al., 2020). Thus, a large proportion of the whole contigs (63.2%) exhibited an extremely low level of expression, representing only  $0.39 \pm 0.01\%$  (mean  $\pm$  SE) of total aligned reads per sample (Table S3).



**FIGURE 2** Counts distribution. Counts (expressed as pseudo-counts, i.e. (counts + 1) at the  $\log_2$  scale) distribution: (a) before filtering the outlier component (96,090 "comps"), (b) after filtering the outlier component (96,089 "comps"), (c) after filtering components with less than 1 CPM in all samples (35,357 "comps"), and (d) after down-sampling normalization (35,357 "comps")



**FIGURE 3** Correlation between replicates before and after filtering the outlier component (comp92\_c0\_seq1; GAXK01170082.1). Data are available in Table S4

## 4 | CONCLUSION

This year-round transcriptome dataset is of general interest for the understanding of annual oscillations of gene expression, specifically in organisms which perform diapause. Our study presents a significant contribution in providing the first in situ annual transcriptomic resource of a key zooplanktonic species, the copepod *C. finmarchicus*. This is significant since seasonal transcriptomic investigations in wild marine organisms are rare (Clark et al., 2013; 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). The proposed dataset is of interest and of high quality due to the unique and rigorous temporal resolution of sampling spanning the entire year and thereby facilitating rhythmic analysis (Thaben & Westermark, 2014). Furthermore, there is a high depth of the RNA sequencing with resulting high-quality sequencing data. Indeed, the average sequencing depth of  $137 \pm 4$  million reads (mean  $\pm$ SE) per sample optimizes the investigation of gene expression in a species with a large genome (Bron et al., 2011; Choquet et al., 2019). We provide the raw reads, the quantification matrix after filtering, both before and after the normalization, and finally the combined annotation of the 35,357 components. Ultimately, this dataset will, among others, allow new insights into the understanding of mechanisms of animal diapause (Baumgartner & Tarrant, 2017; Hand et al., 2016), (circ)annual rhythms (Goto, 2013; Helm & Lincoln, 2017; Helm et al., 2013), and the consequences of climate-driven environmental changes (Kharouba & Wolkovich, 2020).

## ACKNOWLEDGMENTS

This work was supported by the CHASE project, part of the Changing Arctic Ocean program, jointly funded by the UKRI Natural Environment Research Council (NERC, project number: NE/R012733/1) and the German Federal Ministry of Education and Research (BMBF, project number: 03F0803A). We thank Chris Beveridge (SAMS Aquarium Manager) and Norman Smith and crew (R/V Seol Mara) for lab support and animal collection, Estelle Dumont (SAMS) for oceanographic data processing and archiving, and Elise Maigne (INRAE) for help with figures. Open Access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST

We have no conflict of interest.

## AUTHOR CONTRIBUTIONS

**Laura Payton:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Project administration (equal); Supervision (lead); Visualization (equal); Writing – original draft (lead); Writing – review & editing (equal). **Céline Noirot:** Data curation (lead); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – review & editing (equal). **Kim S. Last:** Conceptualization (equal); Data curation (equal); Funding acquisition (lead); Investigation (lead); Methodology (equal); Project administration (equal); Resources (equal); Supervision (lead); Writing – review & editing (equal). **Jordan Grigor:** Conceptualization (equal); Investigation (equal); Methodology (equal); Writing – review & editing (equal). **Lukas Hüppe:** Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – review & editing (equal). **David V. P. Conway:** Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Writing – review & editing (equal). **Mona Dannemeyer:** Formal analysis (equal); Methodology (equal); Writing – review & editing (equal). **Amandine Suin:** Investigation (equal); Methodology (equal); Writing – review & editing (equal). **Bettina Meyer:** Conceptualization (equal); Funding acquisition (lead); Methodology (equal); Project administration (equal); Supervision (lead); Writing – review & editing (equal).

## OPEN RESEARCH BADGES



This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at <http://www.ncbi.nlm.nih.gov/bioproject/734260>; <https://doi.org/10.6084/m9.figshare.c.5450397.v1>.

## DATA AVAILABILITY STATEMENT

Raw reads were gathered in the BioProject PRJNA734260 (Payton et al., 2022a; <http://www.ncbi.nlm.nih.gov/bioproject/734260>)

which includes all BioSamples used for the study (detailed in Table S1). We also provide the following in Figshare (Payton et al., 2022b; <https://doi.org/10.6084/m9.figshare.c.5450397.v1>): the quantification matrix for the 35,357 components after filtering; the quantification matrix of the 35,357 components after filtering and down-sampling normalization; the combined annotation of the 35,357 components. The publication complies with relevant national laws. We have no competing interests. Benefits generated 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>

Céline Noirot  <https://orcid.org/0000-0002-9861-1496>

Kim S. Last  <https://orcid.org/0000-0001-9402-2347>

Jordan Grigor  <https://orcid.org/0000-0001-8227-0773>

Lukas Hüppe  <https://orcid.org/0000-0002-7793-9046>

Bettina Meyer  <https://orcid.org/0000-0001-6804-9896>

## REFERENCES

- Abernathy, J., & Overturf, K. (2016). Comparison of ribosomal RNA removal methods for transcriptome sequencing workflows in teleost fish. *Animal Biotechnology*, 27(1), 60–65. <https://doi.org/10.1080/10495398.2015.1086365>
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute.
- Archibald, K. M., Siegel, D. A., & Doney, S. C. (2019). Modeling the impact of zooplankton diel vertical migration on the carbon export flux of the biological pump. *Global Biogeochemical Cycles*, 33(2), 181–199. <https://doi.org/10.1029/2018GB005983>
- Asai, S., Ianora, A., Lauritano, C., Lindeque, P. K., & Carotenuto, Y. (2015). High-quality RNA extraction from copepods for Next Generation Sequencing: A comparative study. *Marine Genomics*, 24(Pt 1), 115–118. <https://doi.org/10.1016/j.margen.2014.12.004>
- Baumgartner, M. F., & Tarrant, A. M. (2017). The physiology and ecology of diapause in marine copepods. *Annual Review of Marine Science*, 9(1), 387–411. <https://doi.org/10.1146/annurev-marine-010816-060505>
- Beaugrand, G., Brander, K. M., Lindley, J. A., Souissi, S., & Reid, P. C. (2003). Plankton effect on cod recruitment in the North Sea. *Nature*, 426(6967), 661–664. <https://doi.org/10.1038/nature02164>
- Berge, J., Gabrielsen, T. M., Moline, M., & Renaud, P. E. (2012). Evolution of the Arctic *Calanus* complex: An Arctic marine avocado? *Journal of Plankton Research*, 34(3), 191–195. <https://doi.org/10.1093/plankt/fbr103>
- Bron, J. E., Frisch, D., Goetze, E., Johnson, S. C., Lee, C. E., & Wyngaard, G. A. (2011). Observing copepods through a genomic lens. *Frontiers in Zoology*, 8(1), 22. <https://doi.org/10.1186/1742-9994-8-22>
- Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59–60. <https://doi.org/10.1038/nmeth.3176>
- Choquet, M., Hatlebakk, M., Dhanasiri, A. K. S., Kosobokova, K., Smolina, I., Søreide, J. E., Svensen, C., Melle, W., Kwaśniewski, S., Eiane, K., Daase, M., Tverberg, V., Skreslet, S., Bucklin, A., & Hoarau, G. (2017). Genetics redraws pelagic biogeography of *Calanus*. *Biology Letters*, 13(12), 20170588. <https://doi.org/10.1098/rsbl.2017.0588>
- Choquet, M., Smolina, I., Dhanasiri, A. K. S., Blanco-Bercial, L., Kopp, M., Jueterbock, A., Sundaram, A. Y. M., & Hoarau, G. (2019). Towards population genomics in non-model species with large genomes: A case study of the marine zooplankton *Calanus finmarchicus*. *Royal Society Open Science*, 6(2), 180608. <https://doi.org/10.1098/rsos.180608>
- Clark, K. A. J., Brierley, A. S., & Pond, D. W. (2012). Composition of wax esters is linked to diapause behavior of *Calanus finmarchicus* in a sea loch environment. *Limnology and Oceanography*, 57(1), 65–75. <https://doi.org/10.4319/lo.2012.57.1.0065>
- 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. <https://doi.org/10.1016/j.ygcen.2013.04.002>
- Daase, M., Berge, J., Søreide, J. E., & Falk-Petersen, S. (2021). Ecology of Arctic pelagic communities. In D. N. Thomas (Ed.), *Arctic ecology* (pp. 219–259). John Wiley & Sons Ltd.
- DeLeo, D. M., Pérez-Moreno, J. L., Vázquez-Miranda, H., & Bracken-Grisson, H. D. (2018). RNA profile diversity across arthropoda: Guidelines, methodological artifacts, and expected outcomes. *Biology Methods and Protocols*, 3(1), bpy012. <https://doi.org/10.1093/biomethods/bpy012>
- Edwards, A., & Edelman, D. J. (1977). Deep water renewal of Loch Etive: A three basin Scottish fjord. *Estuarine and Coastal Marine Science*, 5(5), 575–595. [https://doi.org/10.1016/0302-3524\(77\)90085-8](https://doi.org/10.1016/0302-3524(77)90085-8)
- Friedland, K. D., Mouw, C. B., Asch, R. G., Ferreira, A. S. A., Henson, S., Hyde, K. J. W., Morse, R. E., Thomas, A. C., & Brady, D. C. (2018). Phenology and time series trends of the dominant seasonal phytoplankton bloom across global scales. *Global Ecology and Biogeography*, 27(5), 551–569. <https://doi.org/10.1111/geb.12717>
- Goto, S. G. (2013). Roles of circadian clock genes in insect photoperiodism. *Entomological Science*, 16(1), 1–16. <https://doi.org/10.1111/ens.12000>
- Häfer, 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: CB*, 27(14), 2194–2201.e3. <https://doi.org/10.1016/j.cub.2017.06.025>
- Häfer, 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(6), 2815–2838. <https://doi.org/10.1002/lno.11011>
- Hand, S. C., Denlinger, D. L., Podrabsky, J. E., & Roy, R. (2016). Mechanisms of animal diapause: Recent developments from nematodes, crustaceans, insects, and fish. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 310(11), R1193–R1211. <https://doi.org/10.1152/ajpregu.00250.2015>
- 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. <https://doi.org/10.3354/meps06775>
- 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.
- Helm, B., Ben-Shlomo, R., Sheriff, M. J., Hut, R. A., Foster, R., Barnes, B. M., & Dominoni, D. (2013). Annual rhythms that underlie phenology: Biological time-keeping meets environmental change. *Proceedings of the Royal Society B: Biological Sciences*, 280(1765), 20130016. <https://doi.org/10.1098/rspb.2013.0016>
- Hill, K. A. J. (2009). *Changes in gene expression, lipid class and fatty acid composition associated with diapause in the marine copepod Calanus finmarchicus from Loch Etive, Scotland*. Thesis, PhD Doctor of Philosophy, University of St Andrews.
- Hughes, M. E., Abruzzi, K. C., Allada, R., Anafi, R., Arpat, A. B., Asher, G., Baldi, P., de Bekker, C., Bell-Pedersen, D., Blau, J., Brown, S., Ceriani, M. F., Chen, Z., Chiu, J. C., Cox, J., Crowell, A. M., DeBruyne, J. P., Dijk, D.-J., DiTacchio, L., ... Hogenesch, J. B. (2017). Guidelines for

- genome-scale analysis of biological rhythms. *Journal of Biological Rhythms*, 32(5), 380–393. <https://doi.org/10.1177/0748730417728663>
- Ibáñez-Tejero, L., Ladah, L. B., Sánchez-Velasco, L., Barton, E. D., & Filonov, A. (2018). Vertical distribution of zooplankton biomass during internal tidal forcing under mesoscale conditions of upwelling and relaxation. *Continental Shelf Research*, 171, 1–11. <https://doi.org/10.1016/j.csr.2018.10.003>
- Johnson, C. L., Leising, A. W., Runge, J. A., Head, E. J. H., Pepin, P., Plourde, S., & Durbin, E. G. (2008). Characteristics of *Calanus finmarchicus* dormancy patterns in the Northwest Atlantic. *ICES Journal of Marine Science*, 65(3), 339–350. <https://doi.org/10.1093/icesjms/fsm171>
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, 30(9), 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>
- Kharouba, H. M., & Wolkovich, E. M. (2020). Disconnects between ecological theory and data in phenological mismatch research. *Nature Climate Change*, 10(5), 406–415. <https://doi.org/10.1038/s41558-020-0752-x>
- Koike, N., Yoo, S.-H., Huang, H.-C., Kumar, V., Lee, C., Kim, T.-K., & Takahashi, J. S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science*, 338(6105), 349–354. <https://doi.org/10.1126/science.1226339>
- Last, K. S., & Dumont, E. (2021). CTD time-series data collection from Loch Etive, Scotland from September 2018 to October 2020. NERC EDS British Oceanographic Data Centre NOC. [https://www.bodc.ac.uk/data/published\\_data\\_library/catalogue/10.5285/c7b3dfbf-4a88-6e29-e053-6c86abc01b4f/](https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/c7b3dfbf-4a88-6e29-e053-6c86abc01b4f/)
- 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(2), 244–251. <https://doi.org/10.1016/j.cub.2015.11.038>
- 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(1), 3–9. <https://doi.org/10.1093/plankt/fbaa058>
- Lenz, P. H., & Roncalli, V. (2019). Diapause within the context of life-history strategies in calanid copepods (Calanoida: Crustacea). *The Biological Bulletin*, 237(2), 170–179. <https://doi.org/10.1086/705160>
- 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), 1–13. <https://doi.org/10.1038/s42003-021-01946-0>
- 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 zooplankton of the North Atlantic Ocean. *PLoS One*, 9(2), e88589. <https://doi.org/10.1371/journal.pone.0088589>
- Li, J., Grant, G. R., Hogenesch, J. B., & Hughes, M. E. (2015). Considerations for RNA-seq analysis of circadian rhythms. *Methods in Enzymology*, 551, 349–367. <https://doi.org/10.1016/bs.mie.2014.10.020>
- Mauchline, J. (1987). Zooplankton in the Oban area and especially in upper Loch Etive. *Scottish Marine Biological Association Internal Reports*, 144, 1–91.
- Melle, W., Runge, J., Head, E., Plourde, S., Castellani, C., Licandro, P., Pierson, J., Jonasdottir, S., Johnson, C., Broms, C., Debes, H., Falkenhaus, T., Gaard, E., Gislason, A., Heath, M., Niehoff, B., Nielsen, T. G., Pepin, P., Stenevik, E. K., & Chust, G. (2014). The North Atlantic Ocean as habitat for *Calanus finmarchicus*: Environmental factors and life history traits. *Progress in Oceanography*, 129, 244–284. <https://doi.org/10.1016/j.pocean.2014.04.026>
- Murphy, E. J., Cavanagh, R. D., Drinkwater, K. F., Grant, S. M., Heymans, J. J., Hofmann, E. E., Hunt, G. L., & Johnston, N. M. (2016). Understanding the structure and functioning of polar pelagic ecosystems to predict the impacts of change. *Proceedings of the Royal Society B: Biological Sciences*, 283(1844), 20161646. <https://doi.org/10.1098/rspb.2016.1646>
- Niehoff, B., Madsen, S., Hansen, B., & Nielsen, T. (2002). Reproductive cycles of three dominant *Calanus* species in Disko Bay, West Greenland. *Marine Biology*, 140(3), 567–576. <https://doi.org/10.1007/s00227-001-0731-3>
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4), 417–419. <https://doi.org/10.1038/nmeth.4197>
- 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(1), 101927. <https://doi.org/10.1016/j.isci.2020.101927>
- 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(1), 415. <https://doi.org/10.1038/s41597-020-00751-4>
- 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., 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.v1>
- Prokopchuk, I., & Sentyabov, E. (2006). Diets of herring, mackerel, and blue whiting in the Norwegian Sea in relation to *Calanus finmarchicus* distribution and temperature conditions. *ICES Journal of Marine Science*, 63(1), 117–127. <https://doi.org/10.1016/j.icesjms.2005.08.005>
- Reygondeau, G., & Beaugrand, G. (2011). Future climate-driven shifts in distribution of *Calanus finmarchicus*. *Global Change Biology*, 17(2), 756–766. <https://doi.org/10.1111/j.1365-2486.2010.02310.x>
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- 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(1), 409. <https://doi.org/10.1186/s12864-021-07557-7>
- 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. <https://doi.org/10.3389/fmars.2020.00056>
- 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), 1–15. <https://doi.org/10.1038/s41598-018-30873-0>
- Saikkonen, K., Taulavuori, K., Hyvönen, T., Gundel, P. E., Hamilton, C. E., Vänninen, I., Nissinen, A., & Helander, M. (2012). Climate change-driven species' range shifts filtered by photoperiodism. *Nature Climate Change*, 2(4), 239–242. <https://doi.org/10.1038/nclimate1430>
- Schmitt, F. G., Devreker, D., Dur, G., & Souissi, S. (2011). Direct evidence of tidally oriented behavior of the copepod *Eurytemora affinis* in

- the Seine estuary. *Ecological Research*, 26(4), 773–780. <https://doi.org/10.1007/s11284-011-0841-4>
- Semmouri, I., De Schampelaere, 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.
- Tarling, G. A., Freer, J. J., Banas, N. S., Belcher, A., Blackwell, M., Castellani, C., Cook, K. B., Cottier, F. R., Daase, M., Johnson, M. L., Last, K. S., Lindeque, P. K., Mayor, D. J., Mitchell, E., Parry, H. E., Speirs, D. C., Stowasser, G., & Wootton, M. (2022). Can a key boreal *Calanus* copepod species now complete its life-cycle in the Arctic? Evidence and implications for Arctic food-webs. *Ambio*, 51(2), 333–344. <https://doi.org/10.1007/s13280-021-01667-y>
- 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(1), 91. <https://doi.org/10.1186/s12983-014-0091-8>
- 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(6), 1157–1169. <https://doi.org/10.1093/icb/icw060>
- 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 and Proteomics*, 30, 230–247. <https://doi.org/10.1016/j.cbd.2019.03.005>
- Thaben, P. F., & Westermark, P. O. (2014). Detecting rhythms in time series with RAIN. *Journal of Biological Rhythms*, 29(6), 391–400. <https://doi.org/10.1177/0748730414553029>
- 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(2), 569–585. <https://doi.org/10.2307/2259045>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** 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. <https://doi.org/10.1002/ece3.8605>