Analysis of the transcriptional pathways associated with the induction of quiescent embryonic arrest in the calanoid copepod Acartia tonsa

Acebal, Miguel Cifuentes; Hansen, Benni Winding; Jørgensen, Tue Sparholt; Dalgaard, Louise Torp

Published in:
Developmental Biology

DOI:
10.1016/j.ydbio.2023.09.004

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Acebal, M. C., Hansen, B. W., Jørgensen, T. S., & Dalgaard, L. T. (2023). Analysis of the transcriptional pathways associated with the induction of quiescent embryonic arrest in the calanoid copepod *Acartia tonsa*. *Developmental Biology*, *504*, 38-48. https://doi.org/10.1016/j.ydbio.2023.09.004

Developmental Biology 504 (2023) 38–48

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Analysis of the transcriptional pathways associated with the induction of quiescent embryonic arrest in the calanoid copepod *Acartia tonsa*

Miguel Cifuentes Acebal [a], Benni Winding Hansen [a,**], Tue Sparholt Jørgensen [a,b,1], Louise Torp Dalgaard [a,*]

[a] Department of Science and Environment, Roskilde University, Universitetsvej 1, DK-4000, Roskilde, Denmark
[b] Department of Environmental Science - Environmental Microbiology and Biotechnology, Aarhus University, Frederiksborgvej 399, DK-4000, Roskilde, Denmark

ARTICLE INFO

Keywords:
Development
Embryonic dormancy
Arrest
Diapause

ABSTRACT

The copepod species *Acartia tonsa* (Dana)(Crustacea) have the unique ability to induce quiescent embryonic dormancy if adverse environmental conditions occur; a characteristic shared by 41 other species belonging to the superfamily Centropagoida in the Calanoida order. However, the transcriptional changes characterizing this process are not known. Here, we compare the transcriptome of embryos in arrested quiescence with the normal development to identify pathways and differentially regulated transcripts involved in quiescent embryogenesis. Quiescence was induced by incubating eggs at 4 °C with anoxia for 26 h(hr), while eggs undergoing normal immediate development were incubated at 16.9 °C in normoxia for 7 h (where gastrulation occurs) or 14 h (where organogenesis occurs) before collecting for RNA extraction and analysis by RNA-sequencing. Results indicate that the expression profile of the quiescent embryo is not as different from the normal embryonic gastrulation as initially expected: None of the mapped transcripts is uniquely expressed in quiescence. Moreover, in quiescence a large proportion of the annotated transcripts display expression values halfway in-between the normal, immediate developmental stages of gastrulation and organogenesis. In depth comparison between the organogenesis stage and quiescent samples, reveal a high degree of divergence, confirming that a developmental arrest has been induced through quiescence. Specifically: Stress response transcripts are prominent in the quiescent phase with a transcript like the mammalian autophagy gene *Sequestosome-1/p62* (SQSTM) being upregulated. The present analysis provides a better understanding of the molecular mechanisms characterizing the quiescent embryonic state of *A. tonsa*.

1. Introduction

Copepods, which include approximately 14,000 species, are among the most abundant metazoans in aquatic environments and play a key role in maintaining trophic webs and transfer of carbon and energy in the ecosystem (Gee, 1989; Walter and Boxshall, 2022; Wassmànn, 1997). The pelagic species *Acartia tonsa*, which belongs to the superfamily Centropagoidae, order Calanoida, is one of the model species for copepod studies (Leandro et al., 2006; Lopes et al., 2021; Saiz and Kiørboe, 1995). Moreover, *A. tonsa* has recently attracted attention for being one of the very few copepods whose genome has been sequenced (Jørgensen et al., 2019a). Additionally, recent studies have focused on its developmental biology as well as on the stress response pathways of the adults through the use of transcriptomics (Jørgensen et al., 2019a; Nilsson et al., 2018). Nevertheless, there is still a large knowledge gap regarding the molecular and developmental biology of this species. This is especially remarkable when investigating the capability of inducing embryonic dormancy, a unique trait of 42 species in the Calanoida order (Holm et al., 2018). Even though some developmental studies have been carried out on Calanoid copepods, these were focused on the post-embryonic development and less is known about early copepod development, in part due to the technical constraints that the egg chorion constitutes (Hansen, 2019; Jørgensen et al., 2019a; Nilsson and Hansen, 2018; Zirbel et al., 2007). With the growing accessibility of new and powerful sequencing techniques, further studies of the developmental

* Corresponding author.
** Corresponding author.
E-mail addresses: bhansen@ruc.dk (B.W. Hansen), ltd@ruc.dk (L.T. Dalgaard).
[1] Present affiliation: The Novo Nordisk Foundation Center for Biosustainability (DTU Biosustain) at the Technical University of Denmark Building 220, Kemitorvet, DK-2800 Kgs. Lyngby, Denmark.

https://doi.org/10.1016/j.ydbio.2023.09.004
Received 19 December 2022; Received in revised form 3 September 2023; Accepted 12 September 2023
Available online 20 September 2023

Abbreviations

DE	Differential expression
DET	Differentially expressed transcripts

changes occurring during the embryogenesis of *A. tonsa* become feasible.

Under normal environmental conditions for copepods (i.e., 16.9 °C and normoxia), copepod eggs will follow the traditional developmental program of animals, also known as subitaneous embryogenesis. Opposed to subitaneous development, under adverse environmental conditions, some copepods early embryos can induce embryonic arrest resulting in embryonic dormancy. A remarkable feature of the so far described Calanoid copepods with this trait is their ability to induce embryonic arrest at more than one time point in development. Egg dormancy is a well-known process present in different groups of animals, including arthropods. In fact, embryonic dormancy ranges from insects to crustaceans and has been proposed to have an active role in their ecological success (Diniz et al., 2017; Hairston and Kearns, 2002; Hand et al., 2016; Hansen, 2019; Jørgensen et al., 2019a; Lees, 1956; Marcus, 1996). However, as opposed to egg dormancy, embryonic dormancy is restricted to few animal species. Embryonic dormancy in copepods would serve to increase population resilience via the burial of some of the eggs in the sediment, thus creating a gene reserve that survives during adverse environmental conditions. Therefore, the capability of embryonic dormancy allows the animals to recover their original pelagic population size faster once the right conditions are restored (Hairston and Kearns, 2002; Hansen, 2019). A better understanding of copepod embryonic dormancy could greatly benefit the scientific community not only in crustacean research, where the ecological implications are clear, but also in other arthropods, where the molecular process could be similar, such as in mosquitoes where egg dormancy has been reported (Diniz et al., 2017). Therefore, the findings reported here may not only be relevant to the species of interest but also for a broader public interested in developmental biology.

The types of copepod embryonic dormancy that exist and their differences were recently discussed by Nilsson et al. (2018) and Hansen, BW (Hansen, 2019). In brief, it is widely accepted that two categories of resting eggs are produced in copepods: Diapause and quiescent. Diapause eggs have an arrest induced by the female before laying them. In a recent study, a transcriptomic analysis of diapause in calanoid copepods was carried out to determine the gene expression and pathways present through their developmental program (Roncalli et al., 2020). Meanwhile, quiescent eggs are those in which the development is interrupted not by a genetic or hormonal response signaled by the female, but resulting from an environmental factor that is sensed by the embryo itself while undergoing development, such as cold temperatures, anoxia, or hypoxia (Hand and Podrabsky, 2000; Hansen, 2019; Jørgensen et al., 2019b). The response to low oxygen concentrations and the increase in stress response proteins (*Heat Shock protein 70* and *ferritin*) during quiescence have been documented in recent years (Jørgensen et al., 2019b; Nilsson et al., 2014). Nevertheless, a transcriptome-wide analysis of the quiescent developmental program in copepods has not been performed to date.

In the present study, a *de novo* assembled transcriptome (GSE210554) from the calanoid copepod *A. tonsa*, which contains a good representation of transcripts derived from subitaneous as well as quiescent embryonic stages, was used to study the transcriptional profile of embryos in the quiescent stage. Thanks to previous studies, it is known that quiescence can occur either at gastrulation or at later stages in embryogenesis, if the environmental cues set in after gastrulation (Nilsson and Hansen, 2018). Moreover, quiescent embryos will progress slower into the consecutive developmental stages taking almost four times more time than the subitaneous eggs to reach gastrulation (7 h (hr) for subitaneous and 20+ hr for quiescent). For the species *A. tonsa,* the entire subitaneous program takes about 48 h to complete, while the quiescent embryos remain viable for months if the environmental conditions persist (Hansen, 2019; Jørgensen et al., 2019b).

Therefore, a direct comparison between quiescent gastrulation and its subitaneous counterpart, as well as the embryologic stages immediately after, are relevant to obtain a better understanding of quiescence. Hence, the current contribution presents a direct transcriptome comparison between quiescent and subitaneous embryos in gastrulation. For this we used the most complete transcriptome of *A. tonsa* to date (GSE210554) (Acebal et al., 2023) focusing on a critical moment in the establishment and success of embryonic arrest. Furthermore, we present several candidate genes potentially involved in quiescence, such as *sqstm1,* a homologue to the mammalian sequestosome/p62 and *tret1*, a facilitated trehalose transporter, in the quiescent stage of *A. tonsa*. SQSTM1 is originally described as a gene involved in selective autophagy, NF-kB activation, and response to oxidative stress damage, which mediates life-span extension in *Drosophila melanogaster* (Aparicio et al., 2019). Nevertheless, this gene can be found in most mammalians and even insects, indicating a conserved role in evolution. The results open an exciting opportunity for understanding how autophagy mediates quiescent development in copepods and possibly other arthropods.

2. Materials and methods

2.1. Copepod culture

Copepods were cultured, and the eggs later collected, as described previously (Acebal et al., 2023). In brief, *Acartia tonsa (DFH-ATI),* obtained in 1981 in Øresund (N 56"/E12"; Denmark) (Støttrup et al., 1986), continuously cultured since, were used (Jørgensen et al., 2019a, 2019b; Nilsson and Hansen, 2018). A copepod culture was set up with salinity 32; 16.9 °C; no light and was fed daily *ad libitum* with a monoculture of the microalgae *Rhodomonas salina* (>800 μg C L^{-1} as in Berggreen et al., 1988). The seawater used in the experiments was UV light treated and filtered through a 0.2 μm mesh size filter.

The outline of the experimental approach is given in Fig. 1. In brief, we compared the subitaneous developmental stages of gastrulation (occurring at 7 h after spawning) and organogenesis (at 14 h after spawning) with embryos kept cold and anoxic for 26 h, which is a treatment known to induce quiescent development.

The collection of samples was achieved after cleaning the 60 L copepod tank bottom. With this approach, it was possible to know the exact time range since the first eggs were spawned. Eggs were obtained by collecting them from the tank bottom every hour. The eggs were then isolated by filtering using a 54 μm mesh and allowed to continue development until the target timepoint was reached (Table 1): for subitaneous eggs, 7 or 14 h; for quiescent eggs 26 h, after which they were collected. Subitaneous eggs were allowed to develop normally in the same environment as the main culture (salinity 32, 16.9 °C, darkness). Quiescent samples (Q) were immediately stored in cold (4 °C)/anoxic conditions during the whole incubation period (26 h), following the procedure described by Nilsson and Hansen (2018) that describe in detail the distribution of embryonic stages over time in both subitaneous and quiescent development for the same *A. tonsa* strain as used in the present study. Moreover, these environmental conditions have been noted to be required for the induction of embryonic arrest in the species (Jørgensen et al., 2019b). A total of fifteen samples were prepared with each timepoint having five technical replicates (Table 1).

2.2. RNA purification and RNA sequencing

To prepare the samples and technical replicates, a previously described procedure was used (Acebal et al., 2023): Five batches of approximately 1000 eggs per timepoint were counted and transferred to separate 1.5 mL Eppendorf tubes to produce the technical replicates.

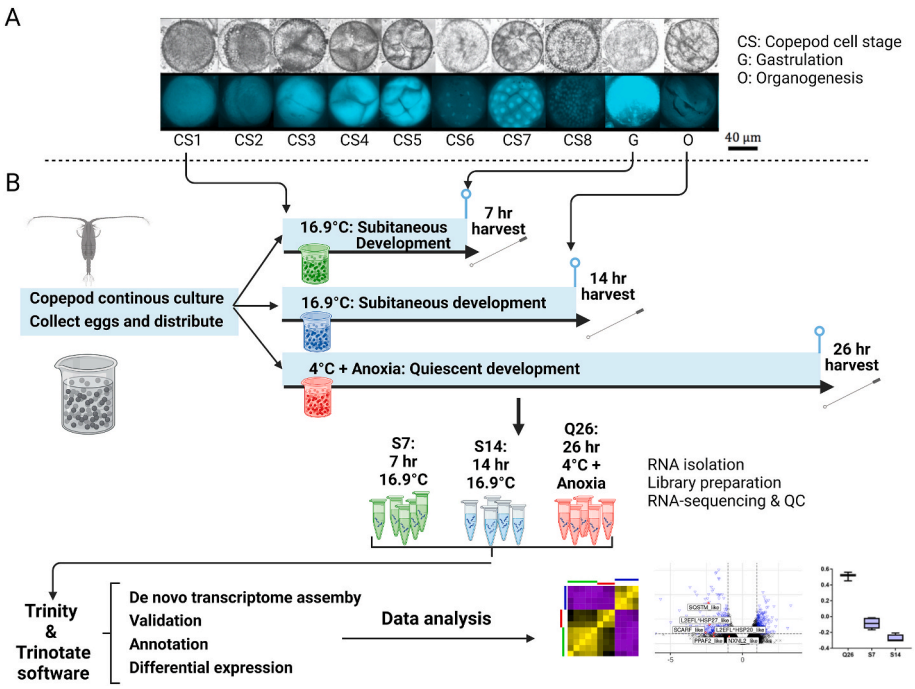

Fig. 1. A) Normal subitaneous development stages of *A. tonsa* from spawning (copepod cell stage 1, CS1) over gastrulation (G) to organogenesis (O). Top are phase-contrast images and below are dapi-stained fluorescent images displaying cell nuclei. With permission Nilsson and Hansen (2018) (CC BY 4.0). B) Workflow for *A. tonsa* sample collection and bioinformatics analysis. This figure illustrates the process of collecting egg samples from *A. tonsa* at various time points (7 h, 14 h, 26 h), represented by the colors green, blue, and red, respectively, followed by RNA-extraction and RNA-sequencing. The subsequent bioinformatics analysis, conducted using the Trinity-Trinotate pipeline, is also depicted in the diagram.

Table 1
Sample and RNA sequencing characteristics.

Sample Name	Time point in development	Development type	Developmental stage	RNA conc. (ng/µl)	% Duplicate reads	% GC	Raw reads (10^6)	Strandedness (% F)
7A	6–7 h	Subitaneous	Gastrulation	39.8	68.6%	45%	27.5	64.6%
7B	6–7 h	Subitaneous	Gastrulation	31.7	71.1%	45%	25.2	62.5%
7C	6–7 h	Subitaneous	Gastrulation	27.1	65.4%	46%	23.9	64.3%
7D	6–7 h	Subitaneous	Gastrulation	34.5	64.0%	44%	23.8	66.0%
7E	6–7 h	Subitaneous	Gastrulation	14.7	66.8%	46%	20.0	66.9%
14A	13–14 h	Subitaneous	Organogenesis	34.3	67.1%	45%	26.7	64.0%
14B	13–14 h	Subitaneous	Organogenesis	38.4	67.8%	45%	30.1	69.0%
14C	13–14 h	Subitaneous	Organogenesis	18.7	63.5%	44%	23.4	67.5%
14D*	13–14 h	Subitaneous	Organogenesis	57.0	82.7%	45%	22.8	60.0%
14E	13–14 h	Subitaneous	Organogenesis	51.0	58.5%	45%	19.3	72.3%
Q26A (ColdA)	25–26 h	Quiescent	Gastrulation	30.9	63.9%	45%	18.2	74.8%
Q26B* (ColdB)	25–26 h	Quiescent	Gastrulation	41.6	60.8%	44%	7.0	70.8%
Q26C (ColdC)	25–26 h	Quiescent	Gastrulation	38.5	61.5%	45%	22.7	70.0%
Q26C (ColdD)	25–26 h	Quiescent	Gastrulation	20.3	61.3%	45%	23.2	68.9%
Q26E* (ColdE)	25–26 h	Quiescent	Gastrulation	43.0	91.7%	45%	26.6	35.8%

From here, they were pelleted by centrifugation (10.000 rpm, 30s), and the supernatant was removed. Finally, aiming to preserve the RNA contained within the eggs to be degraded through RNAse activity, 50 µL of RNA Later (ThermoFisher Scientific, Massachusetts, USA) were supplied to each tube.

RNA was extracted using RNeasy mini kit (Qiagen GmbH, Hilden, Germany) following manufacturers' protocol and RNA yield measured in a Qubit fluorometer (ThermoFisher Scientific, Massachusetts, USA). Library synthesis (TruSeq stranded mRNA kit (Illumina, California, USA)) were made using 10 µL of RNA when the RNA concentration was ≥20 ng/µL and 20 µL when the RNA concentration was <20 ng/µL for the 15 libraries (Table 1). The samples were sequenced on an Illumina NextSeq500 (Illumina, California, USA) using a 75-cycle kit, giving a 1 × 75 single-read length. Additional subitaneous samples were produced and are discussed in greater detail elsewhere (Acebal et al., 2023).

2.2.1. Data analysis and statistics

The transcriptome assembly (GSE210554) (Acebal et al., 2023), was used for the present analysis. In short, we used the software Trinity and Trinotate together in a commonly used pipeline for producing *de novo* transcriptome assemblies and annotating gene names to transcripts based on data from RNA-sequencing experiments (Bryant et al., 2017; Grabherr et al., 2011) (Fig. 1). We used default settings; except for Trinity where a forward strand specific library was specified. The

Trinotate annotation was achieved with the use of Blastx and Blastp hits against the Uniprot database, Pfam protein domains identified with HMMER and TransDecoder was used to predict coding regions (Bateman et al., 2021; Haas et al., 2013; Mistry et al., 2021). The Trinotate annotation also retrieved the Gene ontology (GO) terms of the annotated transcripts from the Uniprot database.

Quality control (QC) of the libraries was done using FastQC, and QC of the assembly was performed using the *trinitystats. pl* script to calculate N50 and assembly length. We estimated completeness of the assembly using Benchmarking Universal Single-Copy Orthologs (BUSCO), which was done against Arthropoda odb10 database to obtain the BUSCO score. BUSCO is a well-established tool that allows the study of the transcriptome assembly completeness by using highly conserved evolutionary sequences as query (Andrews, 2010; Grabherr et al., 2011; Simao et al., 2015). We removed libraries that did not meet the quality criteria established by Acebal et al. (2023) (number of reads $<10^7$, or the duplicate count >80%) (Table 1).

The QC of the data revealed that three samples (S14 (sample 14D) and Q26 (samples cold B and cold E) (Table 1) failed to pass the established criteria and they did not cluster together with their replicate groups. Moreover, the QC assessment revealed unexpected values for the strandedness with numbers ranging from 60% to 73%, and one library (cold E) with 35%. These values although showing a signal for strandedness are not expected for a forward stranded library, nor for an unstranded one. Other metrics analyzed remained consistent throughout the samples, with the GC percentage being close to 45% and 65% for the duplication of reads. Regarding the quality of the assembly, it has been thoroughly discussed previously (Acebal et al., 2023). Overall, the metrics were consistent with other assemblies, with a similar GC content (38.34%) and N50 value (1,233) (Table 2). Moreover, this is the longest assembly for *A. tonsa* with more than 180, 000, 000 assembled bases. Additionally, the transcriptome yielded a BUSCO score of 95.8% complete BUSCOs, which is larger than for previous assemblies for this species (85% HAGX01 & 91.7% GFWY00000000.1). These results validate the data for further analysis as the single 1 × 75 bp sequencing is offset by a greater sequence variety. It is possible that the higher BUSCO score includes genes unique to embryogenesis that are underrepresented in previous assemblies.

For the statistical analysis of the transcriptome, we focused on the comparative analysis of subitaneous embryonic development at 7 h (S7, Gastrulation) and 14 h (S14, Organogenesis) versus gastrulation in quiescent embryogenesis (Q26, 26 h). The selection of these timepoints for the analysis and the establishment of developmental stages was based upon previously published data obtained through microscopy and DAPI staining (Nilsson and Hansen, 2018). The timepoints were studied using a conditional maximum likelihood pairwise analysis to identify the transcripts that characterize each of the development types and stages studied (S7, S14 and Q26). This analysis was performed using the Bioconductor package edgeR through the Trinity Perl script for differential expression (DE) analysis (Grabherr et al., 2011; Robinson et al., 2010), Additionally, the Bejnamini-Hochberg procedure was used to calculate the false discovery rate (FDR) (Benjamini and Hochberg, 1995).

For a transcript to be considered DE, a 0.01-alpha level of significance was set and at least a twofold log 2 (fpkm+1) change in the expression levels had to be achieved for any comparison. The DE matrices produced by edgeR were used to draw a heatmap showing the behavior of the DE transcripts (DET) throughout all the samples. The transcripts in this heatmap were clustered together based on their behavior through the timepoints and replicates using a similarity tree as the clustering criteria. A cutoff value of 50% was set for defining the clusters. Additionally, volcano plots representing all the transcripts present in each timepoint were produced to observe differences between transcripts between conditions. Subsequently, and according to the volcano plots, six transcripts were selected to be plotted individually to obtain a better understanding of their behavior. The criteria to select these transcripts was twofold: First, the transcripts had to be significant DE according to the previously established criteria and between subitaneous and quiescent stages. Secondly, the transcripts that were going to be further analyzed had to be annotated to genes and or proteins, which are either involved in developmental processes or stress response.

Finally, an enrichment analysis for biological processes was conducted. This was done using the Bioconductor package GOSeq (Young et al., 2010) as well as Uniprot site to retrieve the GOs (Bateman et al., 2021). Once the output was generated, the results were manually curated to discard redundant terms, as well as those terms enriched by only one transcript. The results are all shown together in a bubble plot that represents the most enriched pathways at each embryologic stage according to their adjusted p-value and after applying the Benjamini-Hochberg method to calculate the FDR which was set at 5% level (Benjamini and Hochberg, 1995).

To further analyze the dataset, we performed a correlation network analysis with the 'WGCNA' package for R (Langfelder and Horvath, 2008), searched the GEO and PRIDE databases (Barrett et al., 2013; Perez-Riverol et al., 2022) for other crustaceans, where quiescent embryonic development had been studied to compare our results, and finally analyzed the expression of *ferritin* mRNA in quiescence as suggested in a recent review (Tarrant et al., 2019).

Plotting was done using the RStudio and R-package 'Tidyverse' (Wickham et al., 2019) and GraphPad Prism vs. 7. Fig. 1 and the graphical abstract were created using Biorender.com.

Table 2
Assembly characteristics.

BUSCO score	95.8%
N50 contig score	1,233 nt
GC content	38.34%
Assembled bases	180,251,909 nt

3. Results

The embryonic stages of the initial development program of *Acartia tonsa* are similar to other animals and include the gastrula and organogenesis as the main stages. Gastrula being defined as the moment when blastopore formation occurs in the embryo by an investigation of the blastocyst. Moreover, organogenesis is the stage immediately after gastrulation in which vital structures and organs of *A. tonsa* are formed (Fig. 1A).

The workflow for the present study is illustrated in Fig. 1B. In short, the time points corresponding to gastrulation in both developmental programs (7 h and 26 h) as well as the stage after gastrulation (organogenesis) in subitaneous development (14 h) were analyzed. While the subitaneous embryos were kept under the standard environmental conditions recommended for *A. tonsa* cultures (16.9 °C, normoxia, no light), quiescent eggs were immediately transferred to a cold (4 °C) anoxic environment right after sampling. We focused on embryos in gastrulation stage or that had recently passed this stage, as we expect to find the most important transcriptional changes at these timepoints for comparison with quiescent embryos. Using this approach, we show that embryonic quiescence is transcriptionally very close to subitaneous gastrulation. Even though these stages are transcriptionally close, we also identify a small selection of candidate transcripts for involvement in quiescence, based on their selective expression in quiescence. These transcripts belong to categories such as autophagy, stress response or are factors involved in controlling embryonic development.

3.1. Differential expression

The *A. tonsa* libraries were prepared with batches of approximately 1000 embryos per replicate and time point. The final assembly consisted of more than 180, 000, 000 assembled bases, six time points in

subitaneous development (3 h, 7 h, 14 h, 24 h, 42 h) and one time point for quiescent development (26 h). In this way, we assured that all the stages followed during the 48 h that subitaneous development lasts are represented in the dataset. The results from the differential expression (DE), together with the transcriptome assembly and the RNA libraries are available at GSE210554. Moreover, the analysis shows 2777 putative transcripts in the transcriptome to be DE between any two of the analyzed time points (Supplementary Table 1) (https://doi.org/10.6084/m9.figshare.22215121). Additionally, all replicates from any one time point cluster together. Approximately one third (Supplementary Table 1) of the DE transcripts in the dataset are not annotated to any known features in the analyzed databases, even when less stringent criteria and manual blast searches were performed for a subset of the data. This search consisted of a 70 non-annotated sample of sequences that were significantly DE at cluster 2 between S7 and Q26, and the blast parameters can be found in Table 3. If considering those transcripts DE between subitaneous and quiescent gastrulation, the number of annotated sequences drops to 26%.

The generated heatmap of DE transcripts (Fig. 2), shows that the stages S7 and Q26 have overall similar DE profiles, which corresponds to the time of gastrulation in either developmental program. Based on the heatmap, five clusters of transcripts can be identified. The most pronounced differences in expression patterns occur between the gastrula stages (S7 and Q26) and organogenesis (S14), indicated by Cluster I, III and IV. Inspection of the clusters points towards a clear prominence of organogenesis as the time point characterized by most changes in regulation of transcript levels (Clusters I, III, IV) (Fig. 2). Many of the DET show their highest expression at this stage (2493 transcripts).

Most transcripts expressed during S7 behave in the same way as Q26 throughout most clusters, although differences are observed in Cluster II (Fig. 2), where 283 transcripts have a peak in expression value at Q26, while they have decreased expression at S7 and even lower at S14 (Supplementary Table I). Upregulated transcripts DE between S7 and Q26 in Cluster II includes a homologue to *ADAM29* (ADAM Metallopeptidase Domain 29), involved in developmental processes by mediating cell-cell and cell-matrix interactions and a homologue of *D. melanogaster Endou* (Poly(U)-specific endoribonuclease), involved in controlling the response to cellular stress (Lee et al., 2021).

Moreover, we also observe a difference between S7 and Q26 in Cluster IV (Fig. 2), where 27 transcripts are upregulated at Q26 compared with S7 (Supplementary Table I). Most of these transcripts encode myosins and are further upregulated at the S14 time point. However, a transcript homologous to *Holotrichia diomphalia* (Korean black chafer beetle) *ppaf2* (Phenoloxidase-activating factor 2) is upregulated at Q26 (Fig. 5) versus S7, indicating an increased immune response in quiescence (Zhao et al., 2023). One of the clusters of the heatmap contained only one unannotated sequence with no sequence similarity to known transcripts or blast hits. Therefore, it was not considered to be of interest, and is not displayed in Fig. 2.

The overall similarity between S7 and Q26 is underlined by the similarity matrix (Fig. 3), which differentiates the samples in three clear groups corresponding to gastrulation (S7 and Q26) and organogenesis (S14), but with gastrulation samples (S7 and Q26) being more closely related.

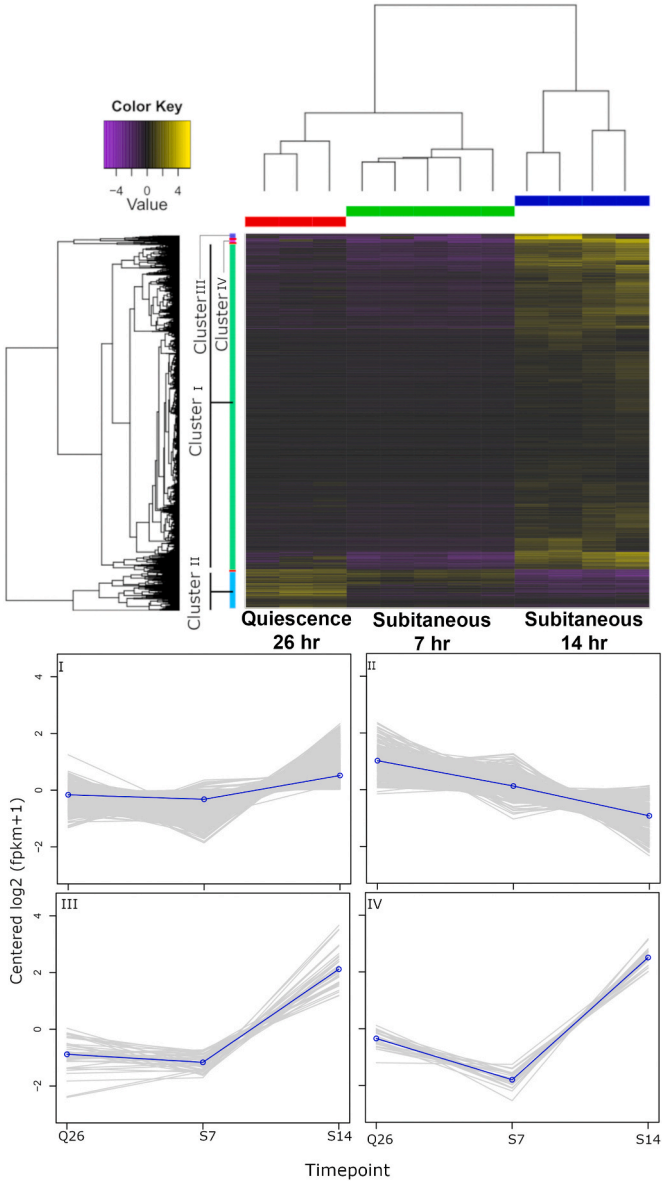

Fig. 2. Developmental analysis of *A. tonsa* eggs. Top: A heatmap displays the 2777 Differentially Expressed Transcripts (DET) across the analyzed time points. The numbers indicate egg age in hours, and the letters denote developmental types - quiescent (Q) or subitaneous (S). Adjacent to the heatmap, a similarity tree is presented, with a 50% similarity cutoff producing a longitudinally plotted dendrogram (I-IV). Plots (I-IV) exhibit grey lines representing transcripts within their respective clusters. The blue line portrays the average expression value per time point. The y-axis represents fold change of sequences in a logarithmic scale (log2), with 0 indicating baseline transcriptome expression. The x-axis signifies the studied developmental stages. The clusters encompass the following transcript counts: I: 2426; II: 283; III: 40; IV: 27.

Table 3
Input parameters for manual BLAST searches.

Input/output parameters	
Max target sequences	100
Expectation threshold	0.1
Word size	3
Max matches in a query range	0
Scoring parameters	
Matrix	BLOSUM62
Gap costs	Existence: 9. Extension 1
Compositional adjustments	Conditional compositional score matrix adjustment
Filters and masking	
Filter	Low complexity regions

For a more comprehensive understanding of the clusters and the transcripts that are likely coregulated, a volcano plot displaying in detail the 28,018 transcripts present at either S7 or Q26 is shown (Fig. 4). Selected genes involved in the stress response or in developmental processes are indicated in the figure.

Among the DET, many are chaperones involved in stress response such as Sequestosome (*sqstm*) and lethal essential for life (*l(2)efl*), which is a homologue for Heatshock protein 20 (HSP20). The selected transcripts contain developmental genes scarface (*scarf*), nucleoredoxin like

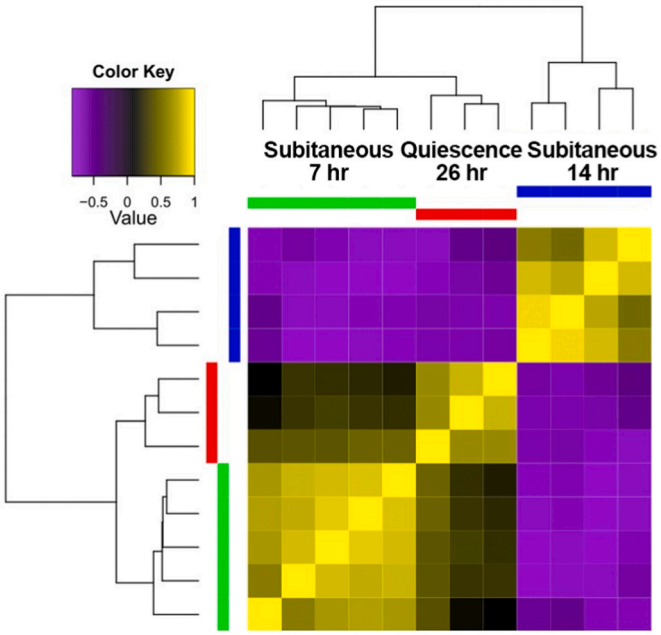

Fig. 3. Sample similarity analysis for *A. tonsa*. This figure presents a sample correlation matrix, illustrating the degrees of similarity between various timepoints and replicates. High similarity is denoted by yellow, while shades of purple indicate increasing dissimilarity. Dendrograms are featured both at the top and left sides, depicting the sample relationships based on their similarity scores. The colors alongside these dendrograms correspond to different sample categories: S7 (Green), S14 (Blue), and Q26 quiescent samples (Red).

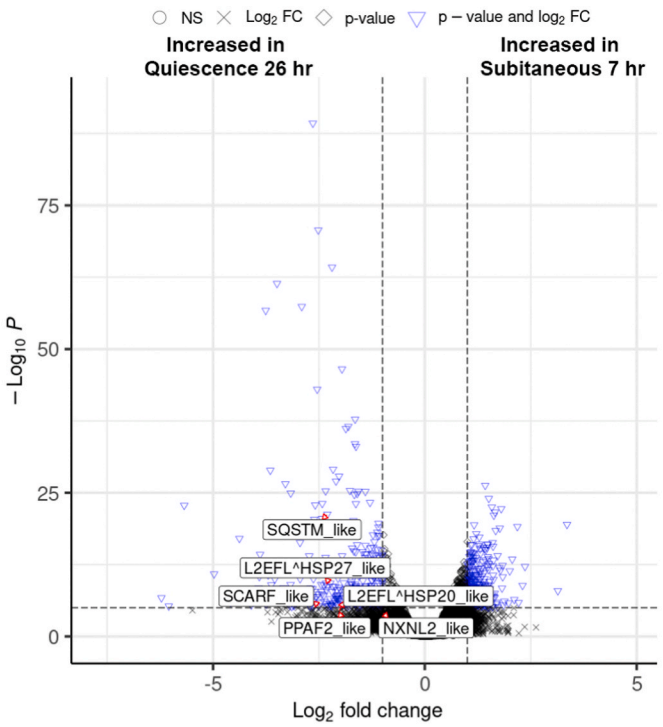

Fig. 4. Transcript analysis in *A. tonsa* during development. Volcano plot illustrating the distinct behavior of the 28,018 transcripts observed in either quiescent (Q26) or subitaneous gastrulation (S7) stages, inclusive of the differentially expressed transcripts. Notably, six sequences in the plot are labeled, representing transcripts annotated to developmental or stress response genes/proteins. Transcripts with increased expression during quiescence are positioned on the left side, while those elevated during subitaneous gastrulation are on the right. The x-axis denotes the fold change in expression, and the y-axis represents the significance of the observation. Values meeting both selection criteria are displayed in blue, while non-significant values are in black.

2 (*nxnl2*)), and stress response mediators (*l2efl/hsp20/hsp27*, *sqstm*, phenoloxidase-activating factor 2 (*ppaf2*)) (Fig. 5A). Some of the selected transcripts (*l(2)efl, sqstm*) contain different isoforms throughout data and these isoforms are regulated similarly except for two forms associated to *l(2)efl*, which are both plotted in Fig. 5. The boxplots illustrate the tendency, similarly to the heatmap that there is no annotated transcript in the data that is completely uniquely expressed in quiescent gastrulation (Q26). The subitaneous organogenesis timepoint (S14) shares similar expression values for some of the transcripts that are upregulated during Q26 but not at S7 (Fig. 5A). In general, the annotated transcripts identified as being upregulated in Q26 vs. S7 more commonly share expression values halfway in between subitaneous gastrulation and organogenesis (S7 and S14, respectively). In some instances, as it happens for *sqstm* and *ppaf2*, transcripts show a clear DE between S7 and Q26 (Fig. 5A), although the number of DET showing this behavior is limited (only 216 transcripts were more than 2-fold DE between S7 and Q26).

In addition to transcripts involved in stress response, the data set also contains mapped transcripts, which are downregulated at S7 and S14, while present at higher levels at Q26 (Fig. 5B). These belong to Cluster II and include the matrix metalloproteinase *adam29*, *endou*, some transcripts of *uaf30* (upstream activation factor subunit UAF30), some transcripts of *lpmo* (lytic polysaccharide monooxygenase), *rtkn2* (anti-apoptotic rhotekin 2 homologue) and *tret1* (facilitated trehalose transporter 1 homologue) as well as other transcripts related to matrix/chitin medication and transcriptional regulation (Fig. 5B, Supplementary Table 1). Similar to the trehalose transporter, also a facilitated glucose transporter homologue (*slc2a2*, Solute carrier family 2, facilitated glucose transporter member 2) is expressed in Q26 and has lower expression at S7 and S14 time points.

3.2. Pathway and correlation network analysis

In the pathway analysis, we focus on pathways in enriched quiescent gastrulation (Q26) when compared against subitaneous gastrulation (S7) (Fig. 6). The enriched processes show an increase in the stress response in the embryo evident at the molecular level (aggrephagy – a form of autophagy), regulation of protein stability as well as the immune response level (regulation of inflammatory pathways are enriched). Moreover, stress pathways associated with hypoxia, such as response to mitochondrial depolarization and response to ischemia, (i.e., a GO term linked to ischemia could reflect hypoxic conditions in the environment) are also enriched in quiescent eggs. The stress response shown in this analysis pictures processes typically activated during hypoxic conditions and reinforces the importance of differentially regulated chaperone (regulation of protein complex stability and aggrephagy). The transcripts annotated to the mammalian gene/protein pair *sqstm1*/p62 are among those that show a DE between S7 and Q26 (Figs. 4 and 5A). At the same time, they are present in most of the enriched processes in Fig. 6. The enriched processes show that a stress response occurs at the transcript level with several pathways indicating responses to wrongly folded proteins. In the analysis, no process was found to be enriched for subitaneous gastrulation.

The weighted gene correlation network analysis (WGCNA) showed four groups of transcripts according to their expression patterns (Fig. 7). Overall, WGCNA use is intended for identifying co-expression patterns across groups of genes. The data is grouped into clusters of genes that show a similar profile and are likely to be co-regulated. Moreover, the analysis selects a “hub gene” that is proposed to play a central role in the co-expression network due to a high number of connections to other genes. In the case of our four groups, module brown showed a cluster of 279 genes to be closely regulated together. The GO analysis of these genes showed biological processes linked to regulation of developmental

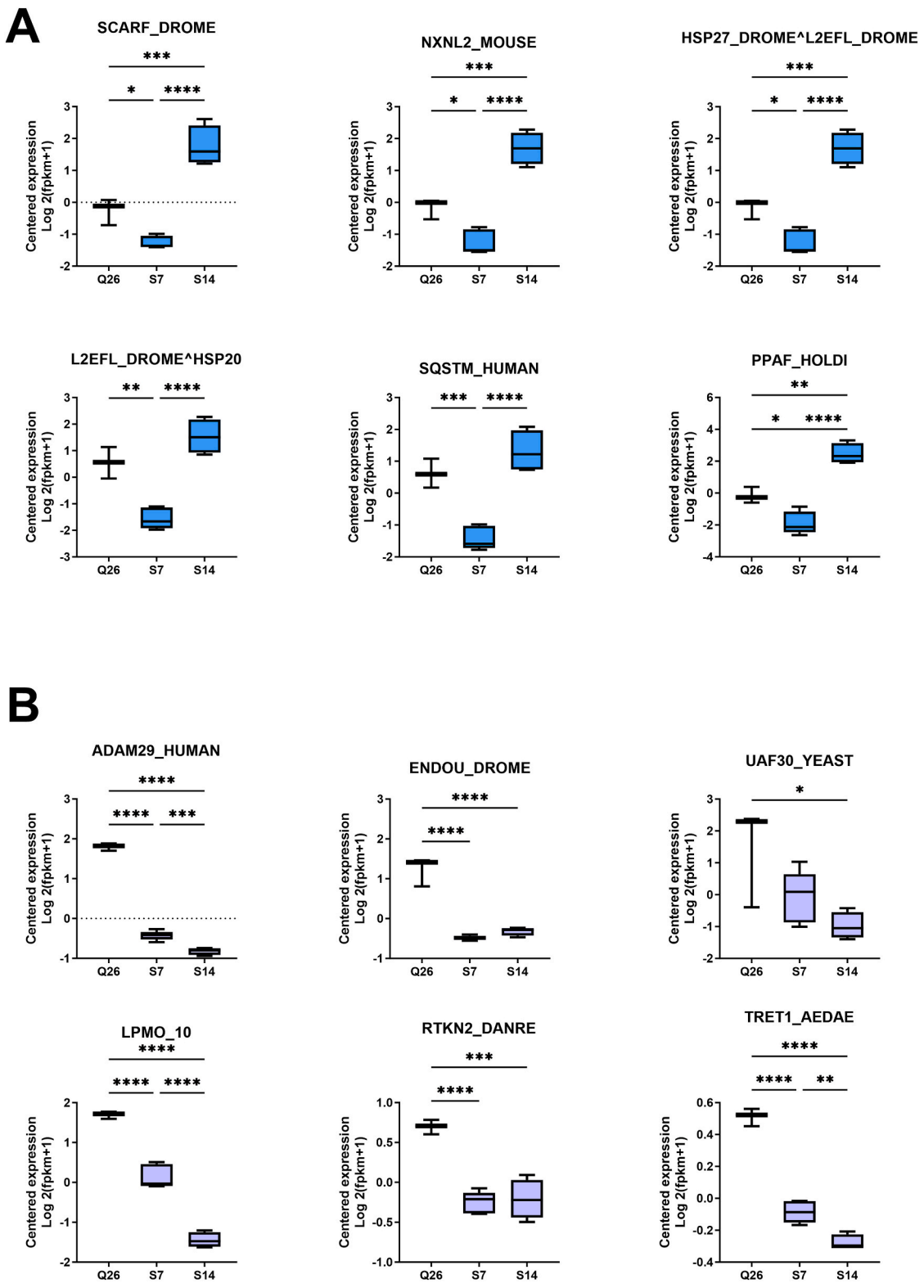

Fig. 5. Gene expression analysis in *A. tonsa.* A) Boxplots illustrating the behavior of six transcripts associated with stress and/or developmental processes across quiescent gastrulation (Q26), subitaneous gastrulation (S7), and organogenesis (S14) stages. B) Boxplots highlighting the behavior of six transcripts that are specifically upregulated during quiescent gastrulation (Q26). These plots depict the expression patterns for Q26, S7, and S14 stages. For both A and B: Expression levels are presented on a log2 scale, centered around fpkm+1, where 0 signifies the transcriptome's baseline expression. The gene name and corresponding species are displayed atop each plot. All displayed transcripts exhibit significant regulation between at least two conditions based on DE edgeR analysis (FDR level of 0.01). To aid interpretations of certainty one-way ANOVA P-values following Tukeys post-hoc test are displayed using asterisks: *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$, ****: $P < 0.0001$.

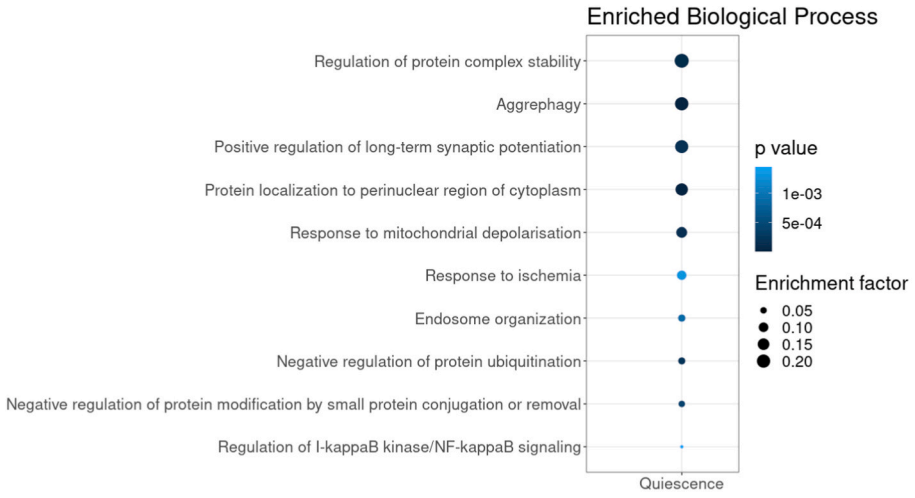

Fig. 6. Enrichment analysis in quiescent development of *A. tonsa*. This bubble plot visualizes the ten most enriched pathways/processes during quiescent development at 26 h compared to subitaneous gastrulation (S7). The size of each bubble corresponds to the enrichment factor, and its color represents the associated p-value. The results are arranged in descending order, with the most enriched pathways at the top.

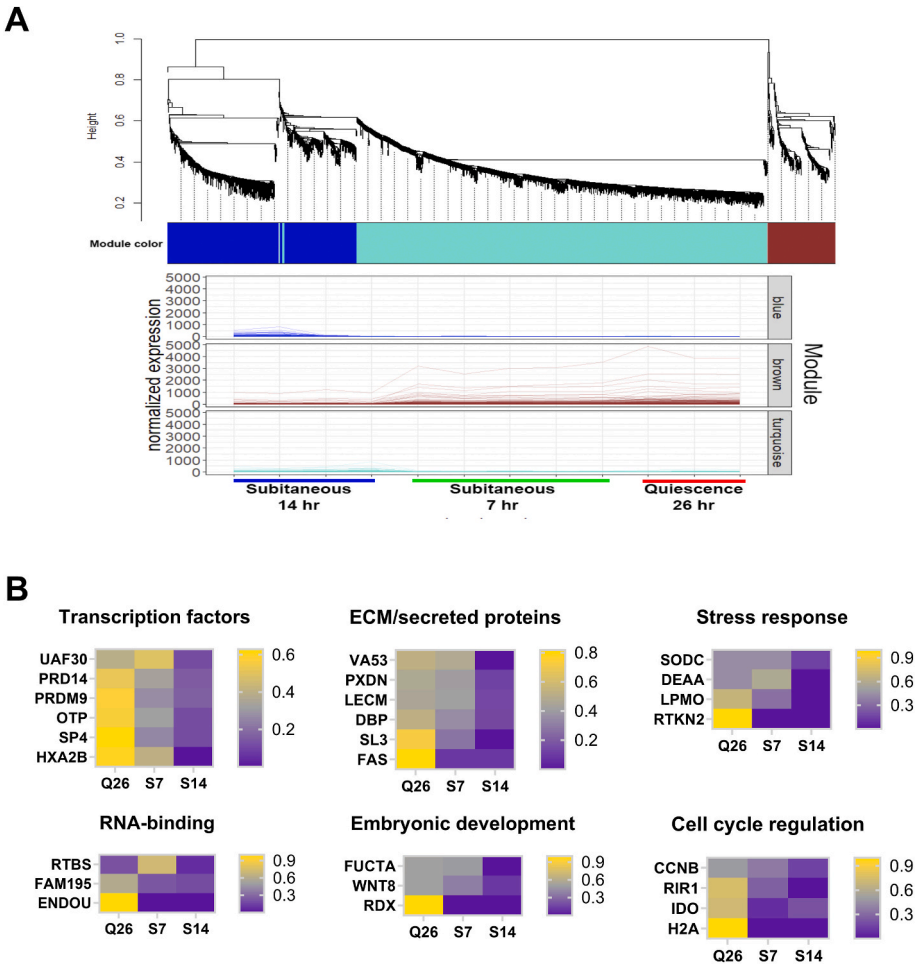

Fig. 7. Weighted network correlation analysis performed with the WGCNA package for R. A) Cluster dendrogram showing the network of correlated genes split in modules. The brown module corresponds to S7 and Q26 samples, indicating enrichment in gastrulation. B) Heatmaps of transcripts from the brown module sorted into functional categories of Transcription factors, ECM/secreted factors, Stress response, RNA-binding, Embryonic development, and Cell cycle regulation.

processes, cell cycle control and apoptosis regulation (Fig. 7B). This module showed an upregulation at timepoints S7 and Q26, and while expressed in both groups, it showed bigger upregulation at Q26, also in concordance with the transcripts displayed in Fig. 7B. Further analysis of the data showed that two annotated transcripts in this module are among the most connected genes. These are *Rtbs* (RNA directed DNA

polymerase from transposon BS; having reverse transcriptase activity) from *D. melanogaster* and a transcript annotated to a cell wall component of a freshwater cyanobacteria; *Fasciclin* (*FAS*). Of these two, *Rtbs* has a higher connectivity with both gastrulation timepoints and lower with organogenesis. This points towards a potential importance RNA-mediated regulation of gastrulation of the embryo. Nevertheless, there is no difference between the connectivity of this transcript and the two developmental programs. This leads us to believe that it may be a transcript linked with gastrulation but not necessarily with quiescent development. Another transcript with high connectivity in both gastrulas is *Uaf30*, also displayed in Fig. 5. The analysis further underlines the possible involvement of this transcript in the development of *A. tonsa*. Finally, the major hub gene from the correlation analysis is a non-annotated sequence. Other transcripts in this module, although with lower connectivity were some of the developmental and stress proteins that we have followed throughout the analysis (*Rdx*, *Wnt8*, *Sodc*) (Fig. 7B), while other transcripts were linked to the extracellular environment and to development. Of interest is the observation of lectin-related transcripts being increased in Q26 (*Lecm, Sl3/Snaclec3*), as well as peroxidasin homologue (*Pxcn*) and vitamin D binding protein (*Dbp*), which are all encoding secreted proteins. Of note, Pxcn has been linked with diapause dormancy in the subarctic copepod *Neocalanus flemingeri* (Roncalli et al., 2020). In the WGCNA, module blue contained 48 transcripts upregulated at the S14 timepoint, the GOs identities of this module were in most cases indicating embryonic development and RNA metabolism. Finally, modules grey (not shown in Fig. 7) and turquoise did not show regulation at any time point and were mainly linked to metabolic processes.

To complement our analysis of *A. tonsa* quiescence mechanisms, we searched for datasets of closely related species in GEO and PRIDE databases, but these did not contain data from any crustaceans in which embryonic dormancy had been studied in a similar fashion as in our study and thus it was not possible to compare expression patterns of genes and or proteins. Finally, ferritin did not show any DE here when compared between subitaneous and quiescent gastrulation.

4. Discussion

Our clustering analysis of the DET showed that most of the transcripts are DE between time points corresponding to the two measured subitaneous embryological stages and surprisingly, the transcripts specific to quiescence are relatively few and often unannotated. Of notice, the differences in transcription between subitaneous and quiescent gastrulation are not very pronounced, even though it is at this moment that the subitaneous and quiescent embryos of *Acartia tonsa* begin to display a different behavior (Nilsson and Hansen, 2018). On the one hand, after reaching gastrulation quiescent embryos will remain in this stage until acceptable environmental conditions are restored. On the other hand, subitaneous eggs will continue normal development and reach the next embryogenesis stage (i.e., organogenesis) within a few hours as observed in the S14 samples.

Nevertheless, and besides the few differences between gastrulation in subitaneous and quiescent development from the transcriptome point of view, our data show that a small number of transcripts are upregulated during the quiescent gastrula phase. However, not all those transcripts are annotated, even when less stringent search criteria were used against the databases. Of the annotated sequences, stress response genes arise as the main DET category, with three chaperones (*hsp27*, *hsp20* and *sqstm/p62*) observed in the present analysis. Transcripts annotated to SQSTM/p62, a gene/protein pair that codes for a chaperone involved in stress response and autophagy in mammals are the most DE transcripts observed in our analysis. The finding of numerous chaperones being actively upregulated during quiescence is supported by previous research that documented the need for hypoxic conditions in the quiescent embryo together with the behavior of two stress proteins: Hsp70 and ferritin in *A. tonsa* quiescence (Jørgensen et al., 2019b; Nilsson et al., 2014).

Here, *ferritin* was not found to be DE and while *hsp70* was identified among the DE chaperones, its behavior was not as clear as for the chaperones here shown. Regarding the un-annotated sequences, it is possible that there are genes unique to *Acartia* and/or other Calanoid copepods that produce resting eggs and that may be key to understanding the complete process. The BUSCO results demonstrate that the current assembly has a greater sequence diversity than previous ones and thus we consider the quality of the transcriptome to be adequate for the identification of annotated sequences that regulate the quiescent stage. However, we should also consider the possibility that post-translational processes, which cannot be detected at the transcriptional level, such as the hypoxia response factor pathway, are acting within the quiescent egg. Thus, it could be relevant to perform shotgun proteomics to investigate differences between quiescent and subitaneous gastrulation at the whole proteome level.

Another interesting observation from the dataset is the confirmation at the molecular level of gastrulation being the stage that controls the start of quiescence as proposed previously (Nilsson and Hansen, 2018). This statement is supported by the overall similarity between the samples that compose the subitaneous gastrulation and those that represent quiescence. Even though these samples are quite close at the molecular transcriptional level, this is not true from a time perspective. There is a 19 h difference between subitaneous gastrulation (7 h) and quiescent samples at the same developmental stage (26 h). As proposed by Nilsson and Hansen, not only will embryogenesis occur at a slower pace during anoxic environments, but ultimately it will remain at the gastrula stage until normoxia is restored. Although the time difference is clear between our samples, almost no progression can be seen in the transcript expression of the samples, concordant with developmental arrest at the gastrulation stage in quiescent embryos. The increase of the stress response molecules could be a potential molecular process that represses normal embryonic development, although this would need to be experimentally verified.

In the case of the organogenesis stage (S14) this shows a clearly different expression pattern from the gastrula samples (as shown in the heatmap in Fig. 2) and it is more dissimilar to quiescent gastrula than to subitaneous gastrula. The identified modules from the correlation network analysis further support this observation and bring forward the possibility that RNA-mediated regulation may play a role in the early stages of copepod embryogenesis. Moreover, we observe pathways linked to cell cycle regulation and apoptosis being upregulated in Q26 but not in S7, which is concordant with the observation of upregulation of cell cycle regulators, transcriptional regulators and antiapoptotic transcripts in quiescent samples (Fig. 5B and 7). Meanwhile no other differences in biological processes were observed between these two timepoints. Moreover, S14 shows developmental processes that were not present at either of the other two timepoints S7 and Q26.

The exact process driving quiescence in copepods remains to be identified, but evidence points towards an increase in stress response proteins, mostly chaperones that could be acting against ROS caused by the adverse environmental conditions i.e., as observed in emergence from conditions of low temperature and hypoxia. Overall, the data indicates that closely regulated stress response processes are being triggered during quiescent development. This observation is clearly remarkable through the presence of many GO terms that share parent terms and indicates that some of these processes may operate in a coordinated manner. Upregulation during quiescence would help survival of the embryo through the adverse environmental conditions that trigger quiescence (Jørgensen et al., 2019b). As for the autophagy activator *Sqstm* we propose it to be an important mediator involved in this process at least during the initial stages of embryogenesis. This is supported by it being shown to play a role in most of the enriched pathways while at the same time is the most DE sequence in quiescent samples when compared with the subitaneous gastrulating embryo. Sqstm activates autophagy and mediates stress responses and induced upregulation of Sqstm in

D. melanogaster prolongs fly lifespan and health (Aparicio et al., 2019), and it is conceivable that Sqstm could induce a similar stress response in the quiescent embryo. In diapause eggs from *Artemia parthenogenetica*, autophagy is increased just prior to the onset of dormancy, and it is possible that autophagy also has a role in quiescence (Lin et al., 2016). Functional analysis focusing on demonstrating the role of Sqstm for quiescence would be a valuable follow up experiment to the present study. Additionally, a metadata analysis looking for Sqstm in other metazoans capable of inducing embryonic dormancy could help elucidate its role in dormancy and therefore be of interest for developmental researchers.

The upregulation of transcriptional regulators and chromatin modifiers in Q26 samples suggests active transcription continues in quiescence, and that it could be possible to elucidate further pathways mediating quiescence by identifying the chromatin targets and thereby the target RNAs of the transcriptional regulators with high expression in quiescence and low expression during subitaneous development.

Furthermore, it would be relevant to investigate mechanisms related to desiccation as a protective mechanism in dormancy, since transcripts encoding transporters of glucose and trehalose are increased specifically in quiescence. Increased osmolality due to sugar uptake could be one of the protective mechanisms of extended quiescence, as also demonstrated for the brine shrimp *Artemia* (MacRae, 2016). This observation, together with the expression of antiapoptotic transcripts such as rhotekin 2, which encodes for an effector protein for the small GTPase Rho, would serve to maintain viability.

The use of RNA sequencing analysis on copepod eggs aids in bringing some light onto their embryonic development. Our data confirmed and considerably extended previous research findings about quiescent development. Based on the transcriptomic analysis the quiescent stage is characterized by a trigger in stress response and the arrest of the developmental processes that occur in the embryo after gastrulation. Surprisingly few genes are being differentially expressed while in quiescence which leads to the hypothesis that posttranslational regulation may be occurring during induction of both gastrulation and quiescence. Further, we propose a homologue of mammalian Sequestosome (*sqstm*) to be a key gene for the quiescent state, but to the extent to which this gene is acting and whether it is responsible for initiation/maintenance of quiescence is still unknown. Moreover, we also observed other DE transcripts, which may have roles in initiating or maintaining quiescence (i.e., *endou*, *uaf30*, *adam29*, *lpmo*, *rtkn2* and *tret1*).

Finally, more research on the topic would be greatly beneficial for the understanding of the mechanisms of quiescence. Following the observations in the present contribution, we propose to focus on Sqstm as well as *A. tonsa* uncharacterized genes in future research, and to carry out a complementary proteomics assays as the next step in research in quiescent development. This would greatly expand the current knowledge on quiescence and the possible role of Sqstm together with the remaining cues on post-translational regulation of quiescence.

Data availability

The gene expression data are available at fighshare:https://doi.org/10.6084/m9.figshare.22215121 and the genome assembly and RNA seq data at GSE210554.

Acknowledgements

We thank Laboratory technician Tina Thane for excellent technical support in generating the RNA-seq libraries. We are indebted to Professor Kim Rewitz for fruitful comments and discussion on an earlier version of our manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ydbio.2023.09.004.

References

Acebal, M.C., Dalgaard, L.T., Jørgensen, T.S., Hansen, B.W., 2023. Embryogenesis of a calanoid copepod analyzed by transcriptomics. Comp. Biochem. Physiol., Part D: Genomics Proteomics 45, 101054. https://doi.org/10.1016/J.CBD.2022.101054.

Andrews, S., 2010. FastQC: a Quality Control Tool for High Throughput Sequence Data, vol. 2010. https://github.com/s-andrews/FastQC.

Aparicio, R., Rana, A., Walker, D.M., 2019. Upregulation of the autophagy adaptor p62/SQSTM1 prolongs health and lifespan in middle-aged drosophila. Cell Rep. 28 (4), P1029–P1040. https://doi.org/10.1016/j.celrep.2019.06.070.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S., Soboleva, A., 2013. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res 41 (Database issue):D991-5.

Bateman, A., Martin, M.J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Silva, A. Da, Denny, P., Dogan, T., Ebenezer, T.G., Fan, J., Castro, L.G., Garmiri, P., Georghiou, G., Gonzales, L., Hatton-Ellis, E., Hussein, A., Ignatchenko, A., Insana, G., Ishtiaq, R., Jokinen, P., Joshi, V., Jyothi, D., Lock, A., Lopez, R., Luciani, A., Luo, J., Lussi, Y., MacDougall, A., Madeira, F., Mahmoudy, M., Menchi, M., Mishra, A., Moulang, K., Nightingale, A., Oliveira, C.S., Pundir, S., Qi, G., Raj, S., Rice, D., Lopez, M.R., Saidi, R., Sampson, J., Sawford, T., Speretta, E., Turner, E., Tyagi, N., Vasudev, P., Volynkin, V., Warner, K., Watkins, X., Zaru, R., Zellner, H., Bridge, A., Poux, S., Redaschi, N., Aimo, L., Argoud-Puy, G., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Blatter, M.C., Bolleman, J., Boutet, E., Breuza, L., Casals-Casas, C., de Castro, E., Echioukh, K.C., Coudert, E., Cuche, B., Doche, M., Dornevil, D., Estreicher, A., Famiglietti, M.L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Hyka-Nouspikel, N., Jungo, F., Keller, G., Kerhornou, A., Lara, V., Le Mercier, P., Lieberherr, D., Lombardot, T., Martin, X., Masson, P., Morgat, A., Neto, T.B., Paesano, S., Pedruzzi, I., Pilbout, S., Pourcel, L., Pozzato, M., Pruess, M., Rivoire, C., Sigrist, C., Sonesson, K., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C., Chen, Y., Garavelli, J.S., Huang, H., Laiho, K., McGarvey, P., Natale, D.A., Ross, K., Vinayaka, C.R., Wang, Q., Wang, Y., Yeh, L.S., Zhang, J., da Silva, A., Denny, P., Dogan, T., Ebenezer, T.G., Fan, J., Castro, L.G., Garmiri, P., Georghiou, G., Gonzales, L., Hatton-Ellis, E., Hussein, A., Ignatchenko, A., Insana, G., Ishtiaq, R., Jokinen, P., Joshi, V., Jyothi, D., Lock, A., Lopez, R., Luciani, A., Luo, J., Lussi, Y., MacDougall, A., Madeira, F., Mahmoudy, M., Menchi, M., Mishra, A., Moulang, K., Nightingale, A., Oliveira, C.S., Pundir, S., Qi, G., Raj, S., Rice, D., Lopez, M.R., Saidi, R., Sampson, J., Sawford, T., Speretta, E., Turner, E., Tyagi, N., Vasudev, P., Volynkin, V., Warner, K., Watkins, X., Zaru, R., Zellner, H., Bridge, A., Poux, S., Redaschi, N., Aimo, L., Argoud-Puy, G., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Blatter, M.C., Bolleman, J., Boutet, E., Breuza, L., Casals-Casas, C., de Castro, E., Echioukh, K.C., Coudert, E., Cuche, B., Doche, M., Dornevil, D., Estreicher, A., Famiglietti, M.L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Hyka-Nouspikel, N., Jungo, F., Keller, G., Kerhornou, A., Lara, V., Le Mercier, P., Lieberherr, D., Lombardot, T., Martin, X., Masson, P., Morgat, A., Neto, T.B., Paesano, S., Pedruzzi, I., Pilbout, S., Pourcel, L., Pozzato, M., Pruess, M., Rivoire, C., Sigrist, C., Sonesson, K., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C., Chen, Y., Garavelli, J.S., Huang, H., Laiho, K., McGarvey, P., Natale, D.A., Ross, K., Vinayaka, C.R., Wang, Q., Wang, Y., Yeh, L.S., Zhang, J., 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 49, D480–D489. https://doi.org/10.1093/NAR/GKAA1100.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57, 289–300. https://doi.org/10.1111/J.2517-6161.1995.TB02031.X.

Berggreen, U., Hansen, B., Kiørboe, T., 1988. Food size spectra, ingestion and growth of the copepod *Acartia tonsa* during development: implications for determination of copepod production. Mar. Biol. 99, 341–352. https://doi.org/10.1007/BF02112126.

Bryant, D.M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M.B., Payzin-Dogru, D., Lee, T.J., Leigh, N.D., Kuo, T.H., Davis, F.G., Bateman, J., Bryant, S., Guzikowski, A. R., Tsai, S.L., Coyne, S., Ye, W.W., Freeman, R.M., Peshkin, L., Tabin, C.J., Regev, A., Haas, B.J., Whited, J.L., 2017. A tissue-mapped Axolotl de novo transcriptome enables identification of limb regeneration factors. Cell Rep. 18, 762–776. https://doi.org/10.1016/j.celrep.2016.12.063.

Diniz, D.F.A., De Albuquerque, C.M.R., Oliva, L.O., De Melo-Santos, M.A.V., Ayres, C.F. J., 2017. Diapause and quiescence: dormancy mechanisms that contribute to the geographical expansion of mosquitoes and their evolutionary success. Parasites Vectors 10, 1–13. https://doi.org/10.1186/s13071-017-2235-0.

Gee, J.M., 1989. An ecological and economic review of meiofauna as food for fish. Zool. J. Linn. Soc. 96, 243–261. https://doi.org/10.1111/J.1096-3642.1989.TB02259.X.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly

from {RNA}-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., Macmanes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 88, 1494–1512. https://doi.org/10.1038/nprot.2013.084.

Hairston, N.G., Kearns, C.M., 2002. Temporal dispersal: ecological and evolutionary aspects of zooplankton egg banks and the role of sediment mixing. Integr. Comp. Biol. 42, 481–491. https://doi.org/10.1093/ICB/42.3.481.

Hand, S.C., Denlinger, D.L., Podrabsky, J.E., Roy, R., 2016. Mechanisms of animal diapause: recent developments from nematodes, crustaceans, insects, and fish. Am. J. Physiol. Regul. Integr. Comp. Physiol. 310, R1193–R1211. https://doi.org/10.1152/ajpregu.00250.2015.

Hand, S.C., Podrabsky, J.E., 2000. Bioenergetics of diapause and quiescence in aquatic animals. Thermochim. Acta 349, 31–42. https://doi.org/10.1016/S0040-6031(99)00511-0.

Hansen, B.W., 2019. Copepod embryonic dormancy: “An egg is not just an egg.”. Biol. Bull. 237, 145–169. https://doi.org/10.1086/705546.

Holm, M.W., Kiørboe, T., Brun, P., Licandro, P., Almeda, R., Hansen, B.W., 2018. Resting eggs in free living marine and estuarine copepods. J. Plankton Res. 40, 2–15. https://doi.org/10.1093/PLANKT/FBX062.

Jørgensen, T.S., Petersen, B., Petersen, H.C.B., Browne, P.D., Prost, S., Stillman, J.H., Hansen, L.H., Hansen, B.W., 2019a. The genome and mRNA transcriptome of the cosmopolitan calanoid copepod *Acartia tonsa* dana improve the understanding of copepod genome size evolution. Genome Biol. Evol. 11, 1440. https://doi.org/10.1093/GBE/EVZ067.

Jørgensen, T.S., Jepsen, P.M., Petersen, H.C.B., Friis, D.S., Hansen, B.W., 2019b. Eggs of the copepod *Acartia tonsa* Dana require hypoxic conditions to tolerate prolonged embryonic development arrest. BMC Ecol. 19, 1–9. https://doi.org/10.1186/S12898-018-0217-5.

Langfelder, P., Horvath, S., 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. 9, 1–13. https://doi.org/10.1186/1471-2105-9-559.

Leandro, S.M., Tiselius, P., Queiroga, H., 2006. Growth and development of nauplii and copepodites of the estuarine copepod *Acartia tonsa* from southern Europe (Ria de Aveiro, Portugal) under saturating food conditions. Mar. Biol. 150, 121–129. https://doi.org/10.1007/S00227-006-0336-Y/.

Lees, A.D., 1956. The physiology and biochemistry of diapause. Annu. Rev. Entomol. 1, 1–16. https://doi.org/10.1146/annurev.en.01.010156.000245.

Lee, H.-C., Fu, C.-Y., Lin, C.-Y., Hu, j.-R., Huang, T.-Y., Lo, K.-Y., Tsai, H.-Y., Sheu, J.-C., Tsai, H.-J., 2021. Poly(U)-specific endoribonuclease ENDOU promotes translation of human CHOP mRNA by releasing uORF element-mediated inhibition. EMBO J. 40, e104123 https://doi.org/10.15252/embj.2019104123.

Lin, C., Jia, S.-N., Yang, F., Jia, W.-H., Yu, X.-J., Yang, J.-S., Yang, W.-J., 2016. The transcription factor p8 regulates autophagy during diapause embryo formation in *Artemia parthenogenetica*. Cell Stress Chaperones 21, 665–675. https://doi.org/10.1007/s12192-016-0692-6.

Lopes, L.F. de P., Agostini, V.O., Moreira, R.A., Muxagata, E., 2021. Acartia tonsa dana 1849 as a model organism: considerations on acclimation in ecotoxicological assays. Bull. Environ. Contam. Toxicol. 106 https://doi.org/10.1007/S00128-021-03175-X, 734–675.

MacRae, T.H., 2016. Stress tolerance during diapause and quiescence of the brine shrimp, *Artemia*. Cell Stress Chaperones 21, 9–18. https://doi.org/10.1007/s12192-015-0635-7.

Marcus, N.H., 1996. Ecological and evolutionary significance of resting eggs in marine copepods: past, present, and future studies. Hydrobiol. (Sofia) 320, 141–152. https://doi.org/10.1007/BF00016815.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., Bateman, A., 2021. Pfam: the protein families database in 2021. Nucleic Acids Res. 49, D412–D419. https://doi.org/10.1093/NAR/GKAA913.

Nilsson, B., Hansen, B.W., 2018. Timing of embryonic quiescence determines viability of embryos from the calanoid copepod, *Acartia tonsa* (Dana). PLoS One 13, 1–16. https://doi.org/10.1371/journal.pone.0193727.

Nilsson, B., Jepsen, P.M., Bucklin, A., Hansen, B.W., 2018. Environmental stress responses and experimental handling artifacts of a model organism, the Copepod *Acartia tonsa* (Dana). Front. Mar. Sci. 5, 156. https://doi.org/10.3389/FMARS.2018.00156.

Nilsson, B., Jepsen, P.M., Rewitz, K., Hansen, B.W., 2014. Expression of hsp70 and ferritin in embryos of the copepod *Acartia tonsa* (Dana) during transition between subitaneous and quiescent state. J. Plankton Res. 36, 513–522. https://doi.org/10.1093/plankt/fbt099.

Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., Walzer, M., Wang, S., Brazma, A., Vizcaíno, J.A., 2022. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 50, D543–D552. https://doi.org/10.1093/NAR/GKAB1038.

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, 139. https://doi.org/10.1093/BIOINFORMATICS/BTP616.

Roncalli, V., Cieslak, M.C., Hopcroft, R.R., Lenz, P.H., 2020. Capital breeding in a diapausing copepod: a transcriptomics analysis. Front. Mar. Sci. 7, 56. https://doi.org/10.3389/FMARS.2020.00056.

Saiz, E., Kiørboe, T., 1995. Predatory and suspension feeding of the copepod *Acartia tonsa* in turbulent environments. Mar. Ecol. Prog. Ser. 122, 147–158. https://doi.org/10.3354/MEPS122147.

Simao, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., Simão, F. A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. {BUSCO}: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

Støttrup, J.G., Richardson, K., Kirkegaard, E., Pihl, N.J., 1986. The cultivation of *Acartia tonsa* Dana for use as a live food source for marine fish larvae. Aquaculture 52, 87–96. https://doi.org/10.1016/0044-8486(86)90028-1.

Tarrant, A.M., Nilsson, B., Hansen, B.W., 2019. Molecular physiology of copepods - from biomarkers to transcriptomes and back again. Comp. Biochem. Physiol., Part D: Genomics Proteomics 30, 230–247. https://doi.org/10.1016/j.cbd.2019.03.005.

Walter, T.C., Boxshall, G., 2022. The World of Copepods Database. https://doi.org/10.14284/356 [WWW Document].

Wassmànn, P., 1997. Retention versus export food chains: processes controlling sinking loss from marine pelagic systems. Hydrobiologia 363, 29–57. https://doi.org/10.1023/A:1003113403096.

Wickham, H., Averick, M., Bryan, J., Chang, W., D, L., Mcgowan, A., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. J. Open Source Softw. 4, 1686. https://doi.org/10.21105/JOSS.01686.

Young, M.D., Wakefield, M.J., Smyth, G.K., Oshlack, A., 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 11, 1–12. https://doi.org/10.1186/GB-2010-11-2-R14/TABLES/4.

Zhao, B.-R., Wang, X.-X., Liu, P.-P., Wang, X.-W., 2023. Complement-related proteins in crustacean immunity. Dev. Comp. Immunol. 139, 104577 https://doi.org/10.1016/j.dci.2022.104577.

Zirbel, M.J., Miller, C.B., Batchelder, H.P., 2007. Staging egg development of marine copepods with DAPI and PicoGreen. Limnol Oceanogr. Methods 5, 106–110. https://doi.org/10.4319/LOM.2007.5.106.