

RESEARCH ARTICLE

Open Access



Post-diapause transcriptomic restarts: insight from a high-latitude copepod

Vittoria Roncalli^{1,2*} , Matthew C. Cieslak¹, Ann M. Castelfranco¹, Russell R. Hopcroft³, Daniel K. Hartline¹ and Petra H. Lenz¹

Abstract

Background: Diapause is a seasonal dormancy that allows organisms to survive unfavorable conditions and optimizes the timing of reproduction and growth. Emergence from diapause reverses the state of arrested development and metabolic suppression returning the organism to an active state. The physiological mechanisms that regulate the transition from diapause to post-diapause are still unknown. In this study, this transition has been characterized for the sub-arctic calanoid copepod *Neocalanus flemingeri*, a key crustacean zooplankton that supports the highly productive North Pacific fisheries. Transcriptional profiling of females, determined over a two-week time series starting with diapausing females collected from > 400 m depth, characterized the molecular mechanisms that regulate the post-diapause trajectory.

Results: A complex set of transitions in relative gene expression defined the transcriptomic changes from diapause to post-diapause. Despite low temperatures (5–6 °C), the switch from a “diapause” to a “post-diapause” transcriptional profile occurred within 12 h of the termination stimulus. Transcriptional changes signaling the end of diapause were activated within one-hour post collection and included the up-regulation of genes involved in the 20E cascade pathway, the TCA cycle and RNA metabolism in combination with the down-regulation of genes associated with chromatin silencing. By 12 h, females exhibited a post-diapause phenotype characterized by the up-regulation of genes involved in cell division, cell differentiation and multiple developmental processes. By seven days post collection, the reproductive program was fully activated as indicated by up-regulation of genes involved in oogenesis and energy metabolism, processes that were enriched among the differentially expressed genes.

Conclusions: The analysis revealed a finely structured, precisely orchestrated sequence of transcriptional changes that led to rapid changes in the activation of biological processes paving the way to the successful completion of the reproductive program. Our findings lead to new hypotheses related to potentially universal mechanisms that terminate diapause before an organism can resume its developmental program.

Keywords: Copepod, Gulf of Alaska, *Neocalanus flemingeri*, Transcriptomics, Diapause, Respiration

* Correspondence: vittoria.roncalli@szi.it

¹Pacific Biosciences Research Center, University of Hawai'i at Mānoa, 1993 East-West Rd, Honolulu, HI 96822, USA

²Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Naples, Italy

Full list of author information is available at the end of the article



© The Author(s). 2021, corrected publication 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Diapause is a type of dormancy that is genetically controlled and widespread among arthropods. It has evolved independently in multiple taxa [1–3]. It is an alternative developmental program that allows organisms to survive during periods of unfavorable environmental conditions and synchronize reproduction and growth to optimize survival [1–3]. The diapause program extends the lifespan of the organism by introducing a period of suspended development [4]. This program has been divided into three major phases: pre-diapause, diapause and post-diapause [4]. During pre-diapause resources are actively accumulated and maturation is postponed. During the diapause phase the organism is in a state of “suspended animation,” characterized by low metabolic rates and increased stress resistance. In post-diapause biological and cellular processes return to an active state and the organism completes its life cycle. While different phases and sub-phases of diapause were recognized early on [2], the physiological progression through the phases remains poorly characterized [4]. The molecular basis of such major physiological transitions is fundamentally interesting, in particular the transition from diapause to post-diapause, which requires the restart of a diverse set of biological processes such as development, metabolic activation, muscle function, cell division and digestion [5, 6]. Here, we investigated this transition in an organism with adult-female diapause, where the primary biological process during post-diapause is the completion of the reproductive program [7, 8].

The sub-arctic calanoid copepod *Neocalanus fleminigeri* [9] is a good model for studying emergence from diapause. With one generation per year, diapause is a critical component of *N. fleminigeri*'s life cycle. Between March and May, the juvenile stages (copepodids CI to CV) are found in the upper 100 m, where they are actively growing and accumulating lipids in preparation for the diapause phase [10–12]. Starting in May, the population is dominated by pre-adults (copepodid stage CV) that continue to accumulate lipid stores. By June, the pre-adults migrate to depth, molt into the non-feeding adult stage (both males and females) and mate [10, 11]. Adult males die, while the females migrate deeper in the water column (≥ 400 m), and females enter the diapause phase becoming inactive for ca. 5–7 months (from summer until December/January) [10]. As part of the post-diapause phase, females initiate the reproductive program that takes ca. seven weeks [13]. As eggs near maturity, females ascend to shallower depth (250–500 m), start spawning and newly recruited young copepodids (CI) appear in the upper 100 m in March [10, 11].

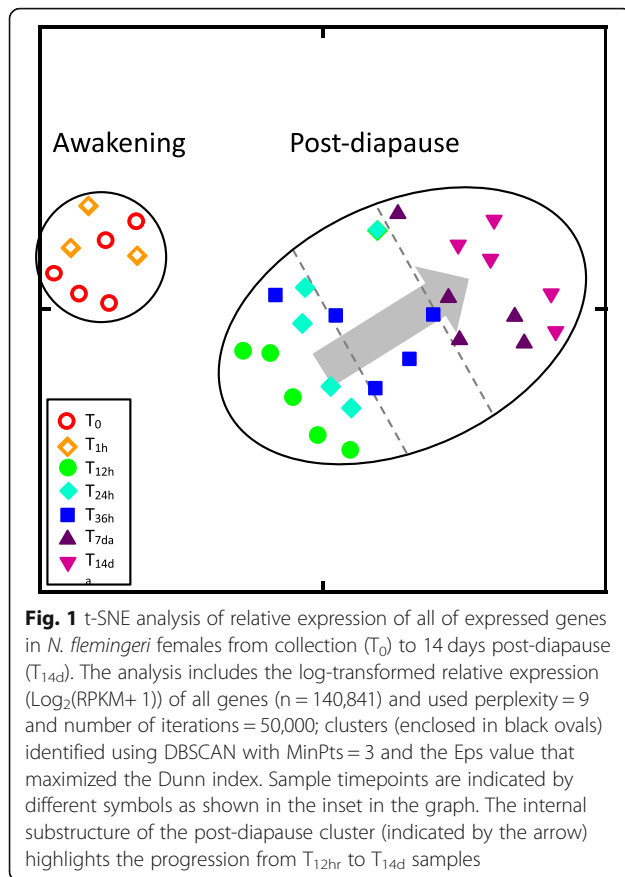
Stimuli associated with the collection of diapausing females from depth with a plankton net were found to set in motion the transition from diapause to post-diapause

in *N. fleminigeri* [13]. By one-week post collection the reproductive program had started as revealed by weekly transcriptomic profiling of females between time of collection and end-of-life. Completion of the reproductive program involves sequential activation of genes involved in germline development, meiotic cycle and oogenesis [13–15]. The program is paired with regulation of genes involved in other biological processes associated with metabolism and homeostasis. The up- and down-regulation of genes associated with catabolism (e.g. glycolysis, β -oxidation, proteolysis), homeostasis and cell maintenance suggests shifts in energy source and energy reallocation among processes as oocytes mature and spawning begins [14]. From these studies it was clear that by week one, all females were in the post-diapause phase and the transition from diapause to post-diapause had been missed. Thus, the goal of this study was to understand the transcriptional changes the organism has to go through to complete the transition from the “suspended” diapause state to the fully “active” post-diapause state. Dimensionality-reduction analysis, gene expression profiling and correlation network analysis were applied to identify how and which genes/processes were sequentially regulated as *N. fleminigeri* females terminated diapause and entered the post-diapause period.

Results

Diapause and post-diapause states have distinct transcriptional phenotypes

The existence of distinct transcriptional phenotypes within a set of samples can be inferred by the approaches described in methods (below) and in particular by the occurrence of clusters in a t-distributed Stochastic Neighbor Embedding (t-SNE) dimensionality-reduction plot. Clustering of the *N. fleminigeri* females ($n = 33$) based on the expression of all genes ($n = 140841$), separated them into two distinct clusters; the first cluster included those from the time of collection (T_0) and those harvested one hour later (T_{1hr}), while the second cluster included females from all other samples (T_{12hr} to T_{14d}) (Fig. 1). By the 12 h timepoint, the females' transcriptional phenotype was distinct from the first two timepoints, suggesting that the shift from the “diapause” state to the “post-diapause” transcriptional phenotype had occurred by 12 h post-collection. While 25 females clustered together, a substructure was observed, with samples aligning by timepoint along a diagonal from T_{12hr} to T_{14d} (arrow, Fig. 1). Because t-SNE attempts to retain as much as possible in its 2D representations the proximities found in higher dimensions [16], the non-overlapping pattern of segregated points within this cluster is evidence of progressive changes in gene expression.



Gene expression analysis confirmed the presence of large differences in transcriptional profiles from T_0 to T_{14d} as indicated by more than 14,000 differentially expressed genes (DEGs) identified across all samples ($p \leq 0.05$ after FDR correction). Pairwise comparison between T_0 females and all other timepoints ($T_{1hr} - T_{14d}$) showed a steep up-ramping of the number of DEGs between the 1 hr. and 36 h timepoints (Table 1). Although T_{1hr} females clustered with T_0 (see Fig. 1) nearly 900 DEGs were identified between these first two timepoints, suggesting that the transition from diapause to post-diapause had already begun at 1 hr (Table 1). Consistent with the clear separation between clusters observed by 12 hr post-collection (see Fig. 1), a 2-fold increase in the number of DEGs was observed comparing T_0 with T_{12hr} ($n = 2523$) with a similar number of DEGs between T_0 and T_{24hr} and T_0 and T_{36hr} (Table 1). The largest number of DEGs was found between T_0 and T_{14d} ($n = 4874$, Table 1).

One-hour time point: transient expression of genes associated with diapause termination

The plankton net collection (~ 40 min long net-tow from 400 m depth) generated significant mechanical, chemical and photic stimulation, which induced the termination

of diapause. An indication that females were already terminating diapause at T_{1hr} came from the observed down-regulation of genes that are up-regulated during diapause and considered to be biomarkers for the diapause state. Among those, we found pro-longevity genes (Embryonic lethal abnormal vision (ELAV) and Cheerio) and several Phosphoenolpyruvate carboxykinases (PEPCK) genes that are responsible for anaerobic metabolism during dormancy.

The T_{1hr} timepoint was also characterized by transient transcriptional changes: 97% of the DEGs were uniquely regulated in 1 hr. females. Only 23 DEGs at T_{1hr} were shared between the two comparisons (T_0 vs T_{1hr} and T_0 vs T_{12hr}) (Fig. 2). In the comparisons with T_0 , the majority of DEGs were up-regulated (Table 1) at the later timepoints, as would be expected in an organism that is emerging from a state of developmental arrest, low metabolic rates and depressed levels of mRNA.

Genes encoding peptides known to be associated with the reactivation of physiological and developmental processes were up-regulated in the T_{1hr} females. At this timepoint, the transient up-regulation was observed for genes involved in the ecdysteroid signaling pathway. This signaling cascade begins with the conversion of the ecdysone (E) to the steroid 20-hydroxyecdysone (20E) mediated by the action of cytochromes P450 (e.g. CYP315, Shade (Shd)). The 20E binds the Ecdysone receptor (EcR) and its dimerization partner Ultraspiracle (USP) triggering the downstream activation of the gene Broad and of the Ecdysone-inducible proteins E74 and Eip78/79. The gene Shd, the receptor genes (EcR, USP) and of the genes Broad, E74 and Eip78/79 were up-regulated at T_{1hr} (Fig. 3). Up-regulation of all genes was transient and only observed at T_{1hr} (Fig. 3).

Concurrent with the regulation of genes involved in the 20E pathway, the immediate and transient up-regulation at T_{1hr} included two pathways associated with energy: the tricarboxylic acid (TCA) cycle and the oxidative phosphorylation pathway. The genes encoding the eight enzymes involved in the TCA cycle (Citrate synthase, Aconitase, Isocitrate dehydrogenase, α -ketoglutarate dehydrogenase, Succinate thiokinase, Succinate dehydrogenase, Fumarate, Malate dehydrogenase) were uniquely up-regulated at T_{1hr} (Fig. 4a). Only one gene, encoding the enzyme Succinate dehydrogenase, maintained high expression during the remaining 14 days of the experimental period (Fig. 4a). Genes encoding enzymes involved in the oxidative phosphorylation pathway were also up-regulated at T_{1hr} (Fig. 4b); however, in contrast to the TCA, the high expression of these genes was not exclusive to this timepoint. For NADH dehydrogenase, Cytochrome C oxidase and Nucleosome remodeling factor 38kD (Nurf-38), the relative expression decreased at T_{12hr} but then

Table 1 Summary of differential gene expression analysis

	DEGs		
	Total	Up-regulated	Down-regulated
Generalized linear model (GLM)	14608		
Likelihood ratio tests			
T ₀ vs T _{1hr}	883	502	381
T ₀ vs T _{12hr}	2523	2026	497
T ₀ vs T _{24hr}	3346	2524	822
T ₀ vs T _{36hr}	3633	2646	987
T ₀ vs T _{7d}	2798	1907	891
T ₀ vs T _{14d}	4874	2678	2106

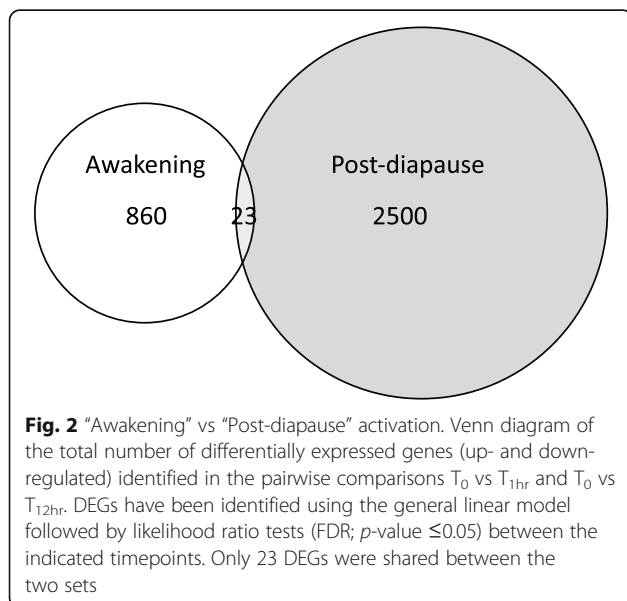
Differentially expressed genes (DEGs) were identified using a general linear model (GLM) followed by likelihood ratio tests (FDR; p-value ≤ 0.05) between T₀ females and females from T_{1hr}, T_{12hr}, T_{24hr}, T_{36hr}, T_{7d} and T_{14d} time points. For each pairwise likelihood test, the number of total DEGs, up- and down-regulated is listed. The pairwise comparisons between other samples are shown in Table S1

increased again at T_{24hr} remaining high to T_{14d} (Fig. 4b). Interestingly, genes encoding enzymes involved in glycolysis and β-oxidation were not among the up-regulated DEGs at T_{1hr}. These genes did not become significantly up-regulated until later (see below).

In addition to the specific pathways regulated at T_{1hr}, we observed reactivation of transcription processes, as suggested by the up-regulation of genes involved in histone acetylation and by the down-regulation of genes associated with chromatin silencing which is one of the processes that prevent RNA metabolism during dormancy. These results are discussed below as part of the WGCNA analysis.

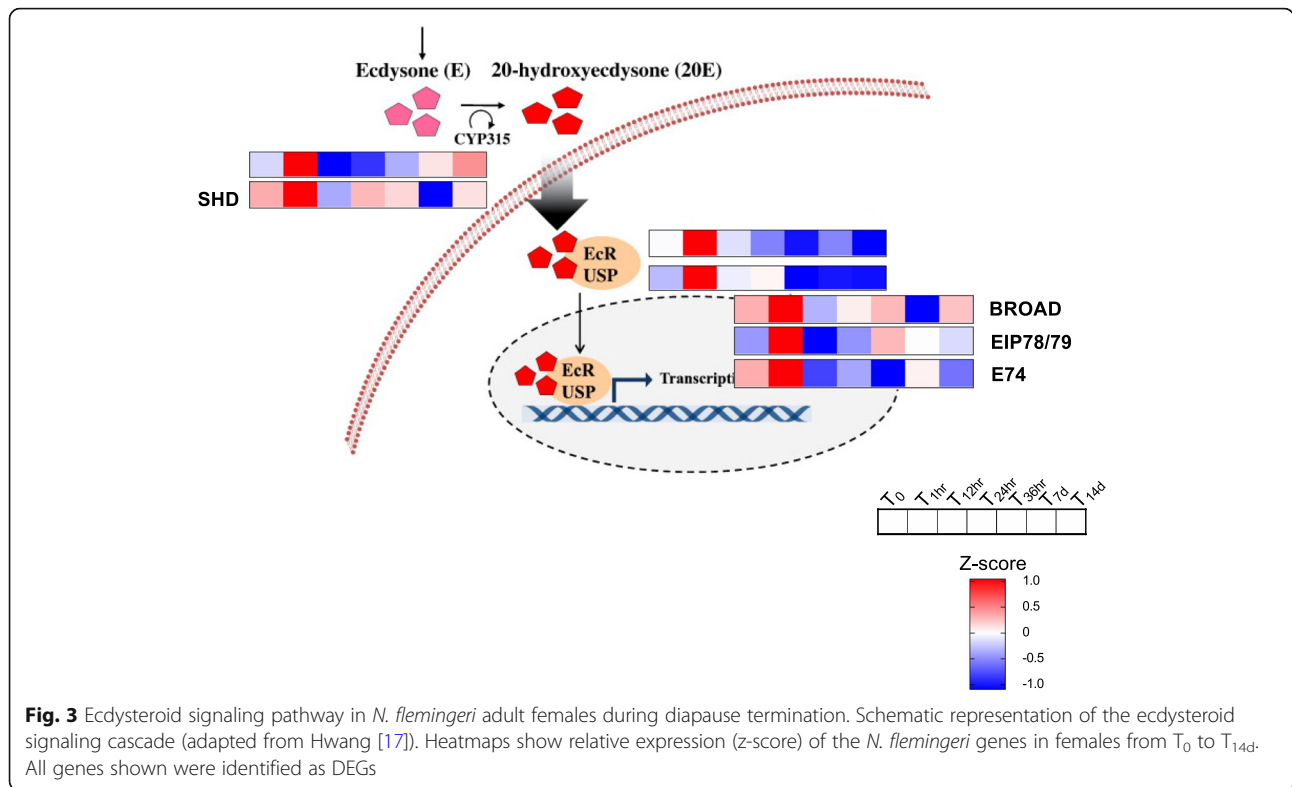
From collection to 14 days post-diapause – an orchestrated progression of transcriptional changes

Transcriptional changes during the 14 days were further analyzed using weighted gene correlation network analysis



(WGCNA) on the DEGs. Based on expression patterns, the differentially expressed genes across samples (n = 14608) were grouped into seven gene network modules that were positively correlated with one or more timepoints (Fig. 5). Six of the seven modules grouped the first two timepoints (T₀ and T_{1hr}) together as would be expected from the t-SNE clusters. However, the black module, with the smallest number of DEGs (n = 174), differentiated T₀ samples from the others. Approximately 10% of the DEGs were placed into the gray module that collects the genes that did not aggregate into a specific gene correlation network (Fig. 5). Enrichment analysis of the DEGs in each module identified overrepresented GO terms belonging to nine biological processes of which five were unique to single modules (Fig. 6).

RNA metabolic process was an enriched process that characterized early changes in gene expression. It was overrepresented for the DEGs identified in two modules that were positively correlated with the T₀ and T_{1hr} timepoints (red and yellow, Fig. 6). This process was also enriched for the DEGs characterized by a more gradual change in relative expression over the time series (green, Fig. 6). This signal was preceded by a single enriched process (chromatin organization) in the black module that differentiated the T₀ from all the other timepoints including T_{1hr}. These four modules were all positively correlated with early timepoints. In contrast, the blue and brown modules were positively correlated with intermediate and later timepoints starting at T_{12hr} and T_{24hr} respectively (Fig. 5). The blue and brown modules, which together contained more than 5000 DEGs, shared three enriched processes (multicellular organism development, signaling, and immune system process). Two additional enriched GO terms were unique to the blue module (cell cycle, cell differentiation) (Fig. 6). The turquoise module, with the largest number of DEGs (n > 3000), showed positive correlation with the last two timepoints (T_{7d} and T_{14d}) and included two enriched processes that were unique to this module (oogenesis, metabolic process) (Fig. 6).



Early responding biological processes: chromatin organization and RNA metabolism

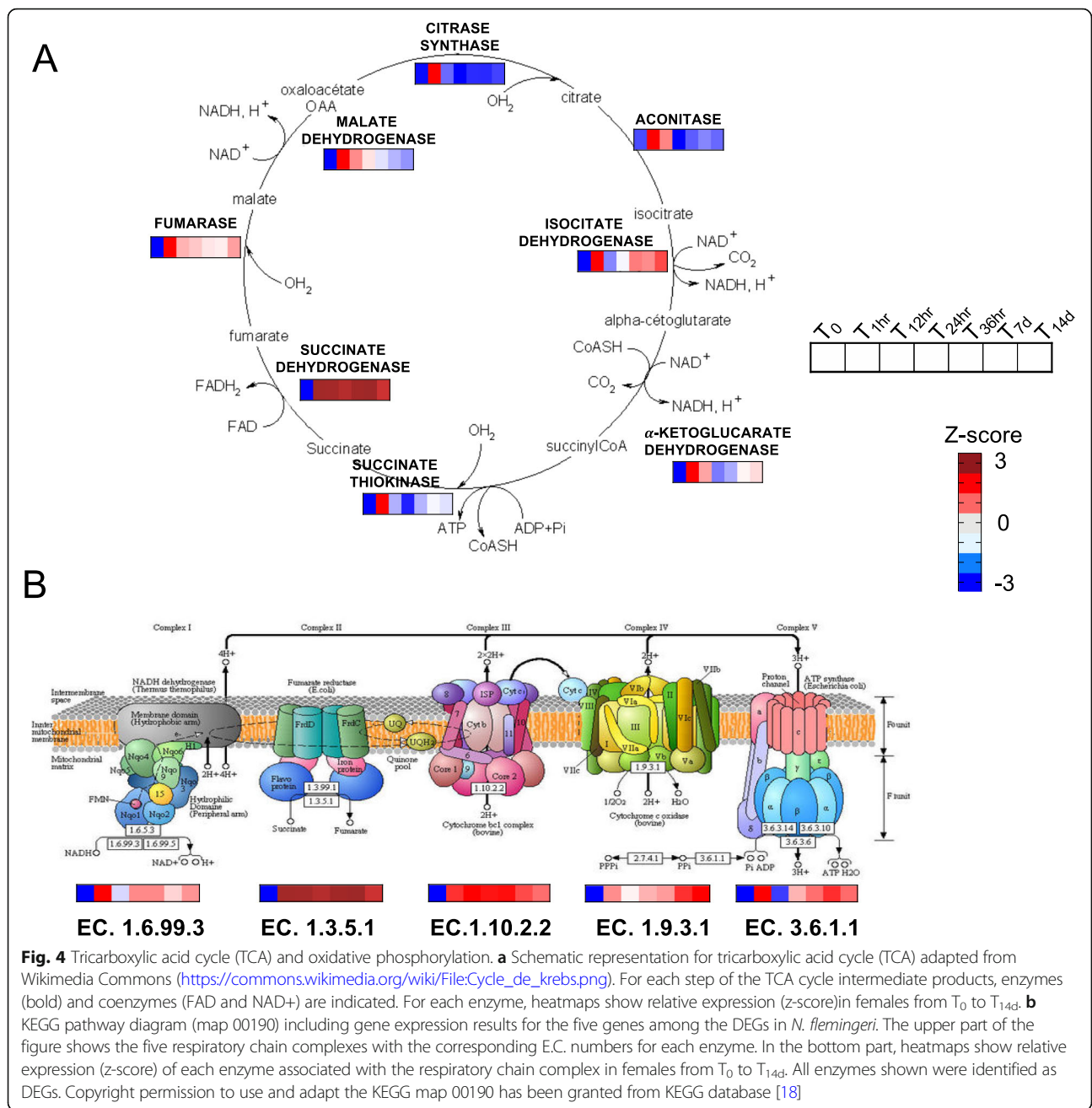
Chromatin organization and in particular chromatin silencing were the processes overrepresented for the DEGs within the black module (Figs. 5, 6). Up-regulation of chromatin silencing is not surprising since suppression of transcriptional activities are characteristic of the dormant state. Considering all DEGs (n = 198) involved in chromatin organization and RNA metabolism (Fig. 7) 34 were exclusively up-regulated at T_0 with respect to later timepoints and annotated as Longitudinal lacking proteins (LOLA), Proliferation-associated SNF2, Histone-lysine N-methyltransferase and Lamins. For some LOLA a significant increase in expression was observed again at T_{14d} consistent with an earlier study [14].

DEGs annotated to RNA metabolism, a process that was overrepresented in three modules (red, yellow, green), were characterized by a complex pattern of changes in gene expression suggesting an organized progression of up- and down-regulation of specific genes during the 14 days after collection (Fig. 7). Seven different GO terms were enriched for this process and these include RNA modification, rRNA processing, regulation of transcription and histone mRNA catabolic process (Fig. 6). Several transcription factors (Sp1, Sp4, Pur-alpha), antioxidants (Peroxiredoxin-5) and genes associated with protein turnover (e.g. E3 ubiquitin-protein ligase Siah1, E3 ubiquitin-protein ligase) and histone

acetylation (e.g. KAT8 regulatory NSL complex, PolyA-RNA polymerase, Histone demethylation 1) were included in the yellow module, with high expression restricted to the T_0 and T_{1hr} timepoints. The red and green modules included transcription related genes (e.g. Purine-rich single stranded DNA binding protein, tRNA pseudouridine) and genes involved in RNA binding (e.g. Eukaryotic translation initiator factor 4, La-related protein CG11505). Although RNA metabolism was not enriched in the other modules (blue, brown, turquoise) some DEGs associated with this process were up-regulated at later timepoints. These DEGs were annotated with additional GO terms indicating that they were also involved in other biological functions such as the immune system (e.g. Embryonic polarity protein dorsal, Toll protein, Spaetzle protein) which is consistent with the enrichment of this biological process in both blue and brown modules (Fig. 6).

Early post-diapause phase: from 12 to 36 h, overrepresentation of cell cycle and cell differentiation

Enrichment of cell cycle and cell differentiation was unique to the DEGs in the blue module, which included genes that were positively correlated with the beginning of the post-diapause period (T_{12hr} to T_{36hr} ; Figs. 5, 6). Two specific processes within cell cycle were enriched: centrosome duplication and nuclear division (Fig. 6). Consistent with a putative increase in cell proliferation, cell



differentiation was also overrepresented during these timepoints (Fig. 6). Among the genes involved in cell cycle, there were several Cyclins and Cyclin-dependent kinases (Cdk) associated with G1/S (CycA, Cdk12, Cdk17, CycB, CycM2) phases (Fig. 8a). Genes involved in germline development, which also includes ovarian follicle development, were up-regulated between T_{24hr} and T_{36hr} and they included Neurogenic delta, Innexins 2 and Septins-7 (Fig. 8a). The expression of these genes suggests that activation of the oogenesis signal already occurs at T_{12hr}. Additional genes involved in cellular differentiation were assigned to

the turquoise module and these were most highly expressed in the last two timepoints (T_{7d} and T_{14d}). Among these genes, were several cAMP kinases and Crumbs proteins (95F) and some Bicaudal D, Aurora kinases and Innexins 2.

Post-diapause phase: from 12 h to 14 days, enrichment of signaling processes, multicellular organism development and immune system

Signaling was an overrepresented process for the DEGs within the blue and the brown modules although different GO terms were enriched in each module (Fig. 6). Of

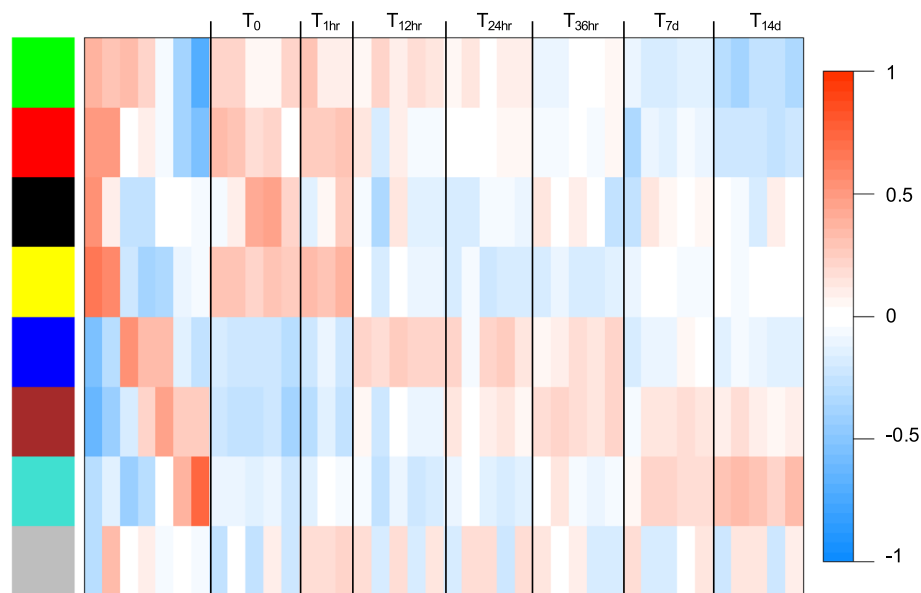


Fig. 5 Correlation of WGCNA modules for the DEGs to sample traits. Heatmap shows correlation of module eigengenes (rows labeled by color) to samples either grouped by preservation time point (first seven columns) or individually. The right most columns ($n = 33$) present the correlations of the eigengene expression for each module with the individual samples as labeled on the top. The color of each cell represents the direction and strength of the correlation (blue = negative and red = positive; color scale on right). Number of DEGs in each module: green ($n = 1409$), red ($n = 1240$), black ($n = 174$), yellow ($n = 1454$), blue ($n = 2909$), brown ($n = 2391$) and turquoise ($n = 3620$). The grey module ($n = 1411$) includes DEGs that did not aggregate with a specific gene correlation network

particular interest was the gene expression pattern associated with the MAPK cascade and the regulation of I-kappaB kinase/NF-kappaB pathway. These GO terms were enriched for the blue module, which was positively correlated with first stages of post-diapause, T_{12hr} to T_{36hr} timepoints. Activation of the MAPK cascade is essential for many cellular processes, such as cell division and cell differentiation. Furthermore, the MAPK pathway is responsible for the downstream activation of I-kappaB kinase/NF-kappaB pathway. Here, several mitogen-activated protein kinases (MAPK3, MAPK2, MAPK1, p38b) and genes associated with the NF pathway (e.g. PI3K, NF-kappa-B-activating protein) were up-regulated in *N. flemingeri* females between T_{12hr} and T_{36hr} (Fig. 8b). Expression of these genes was significantly lower at T_{7d} and T_{14d} timepoints (Fig. 8b).

A large group of genes in the blue and brown modules were enriched for GO terms associated with tissue homeostasis such as regulation of muscle contraction (which was shared between the two modules), locomotor rhythm and myoblast fusion (Fig. 6). These GO terms, which shared the same GO parent (multicellular organism development), were positively correlated with T_{12hr} - T_{14d} timepoints (Fig. 9a). Up-regulated genes included Myosins (light chain, muscle), Titin and Troponin T. Even if muscle GO terms were not enriched in other modules, some DEGs associated with this process, annotated as Myosins, Irregular chiasm C-roughest protein

abd Muscle LIM proteins, were up-regulated at earlier timepoints (T_0 - T_{1hr}).

Another biological process overrepresented for both the blue and brown modules was immune system process, which included Toll signaling pathway as an enriched GO term (Fig. 6). Most of the up-regulated genes were annotated as Toll protein, Spatzle proteins and Embryonic polarity protein dorsal. Relative expression for most of these genes was high in the later timepoints (from T_{12hr} to T_{14d}) (Fig. 9b). Some transcripts encoding Spatzle proteins were also significantly regulated at the earlier timepoints (T_0 - T_{1hr}) compared with the other timepoints, which is not surprising considering that regulation of immune system is also required for maintaining tissue homeostasis (Fig. 9b).

One to two-week post-diapause: overrepresentation of oogenesis and metabolic processes

By T_{7d} and T_{14d} the reproductive program was in progress as suggested by the enrichment of two GO terms associated with oogenesis (Fig. 6). Of the 54 DEGs annotated to those GO terms, 84% had transient up-regulation at T_{7d} and T_{14d} (turquoise module) (Fig. 10a). The majority of these genes were involved in processes occurring during the early phase of oogenesis such as maintenance and division of germ-line stem cells (e.g. Armadillo segment polarity, Dicer protein, Scrawny protein), oocyte determination and formation of the

anterior-posterior axis (e.g. Ovo protein, Egghead protein). The remaining DEGs were grouped into the blue and the brown modules with high expression at the earlier timepoints (Fig. 10a). The blue module included four DEGs annotated as Crumbs protein that are involved in cell differentiation. The brown module (high expression from T_{36hr}) included DEGs annotated as Innexin 2, Cueball protein, Brain tumor protein and Maternal protein staufen. These genes are associated with early processes such as ovarian follicle development and oocyte localization that occur at the beginning of oogenesis; this suggests that the reproductive machinery was set in motion starting at T_{36hr} but became overrepresented only at T_{7d} and T_{14d}.

The turquoise module was characterized by an overrepresentation of DEGs associated with specific metabolic processes (Fig. 6). Among the enriched GO terms, two terms were associated with carbohydrate metabolism, one with lipid metabolism and one with digestion (Fig. 6). High expression of all nine genes encoding enzymes involved in the glycolytic pathway was found at T_{14d} compared with all other timepoints (Fig. 10b). Six of the nine genes (Hexokinase, Phosphofructokinase, Glyceraldehyde3-phosphate dehydrogenase, Phosphoglycerate kinase, Enolase and Pyruvate kinase) were also up-regulated at T_{7d} (Fig. 10b). A similar pattern was observed in the expression of genes involved in lipid catabolism. Relative expression of all genes involved in β -oxidation as well as several lipases (e.g. Monoacylglycerol lipase, Phospholipase) was significantly higher in T_{7d} and T_{14d} compared with all other timepoints (Fig. 10b). Consistent with the enrichment result, up-regulation of several Cytochromes P450 J, involved in epoxigenase activity, and multiple digestive enzymes (Trypsins) was also observed (Fig. 10b). Eleven out of the 20 Trypsins were even more highly expressed in T_{14d} than in T_{7d} (Fig. 10b). Although Trypsins are usually associated with food digestion, in decapods alteration of their enzymatic activity has been reported during post-larvae development [20]. In their entirety, the gene expression pattern suggests that by the end of the experimental period, the females were entering a period of higher metabolic demands.

Respiration rates in pre-diapause and early post-diapause females

Differences in respiration rates were recorded between active pre-diapause and post-diapause females. Active *N. flemingeri* females had respiration rates that averaged 1 $\mu\text{g O}_2$ per individual per hour (range: 0.66 to 1.2, $n = 6$; Fig. 11). In contrast, respiration rates in females collected from depth and incubated in the laboratory averaged 0.3 $\mu\text{g O}_2$ per individual per hour (range: 0.1 to 0.55, $n = 14$) and was similar for all three timepoints

post-collection which included days 2, 2–3 and 5 post-collection (Fig. 11).

Discussion

Diapause, a genetically-controlled suspension of growth and development, can be divided into three main phases: pre-diapause, diapause and post-diapause [4]. The post-diapause phase starts with the reactivation of biological and cellular processes after a period of arrested development and low metabolic rate. Thus, the organism must transition from a state of “suspended animation” to fully active. The post-diapause state varies depending on the next step in the life cycle: the brine shrimp (*Artemia franciscana*) completes embryogenesis; fly pupae (*Sarcophaga crassipalpis*, *Rhagoletis pomonella*) metamorphose into adults; pre-adult stages of some organisms (e.g. the copepod *Calanus finmarchicus*) complete development by molting into adults; other organisms diapause in the adult stage and complete the reproductive program during post-diapause (mosquito *Aedes aegyptii*, copepod *Neocalanus flemingeri*). From a molecular point of view, post-diapause is characterized by large changes in transcription. Within 1 h of a diapause-terminating stimulus, expression levels of genes encoding heat shock proteins and enzymes involved in signaling and respiration pathways are altered in some organisms [21–24]. However, these studies give an incomplete picture, since changes in gene expression were measured in a small number of genes (< 10), providing limited insights into the extensive changes in state that occur during the transition from the diapause to the post-diapause phenotype. The transcriptomic analysis here revealed that the transition from diapause to post-diapause in *N. flemingeri* females involves the finely-orchestrated sequential regulation of thousands of genes.

Transcriptional profiling supports the existence of a diapause and a post-diapause phenotype in the copepod. Based on the relative expression of all genes ($n = 140841$), females separated into two distinct transcriptional phenotypes with the post-diapause phenotype appearing within 12 h after the termination trigger, a combination of photic, chemical, thermal, pressure and mechanical stimuli during net collection. However, in addition to the two distinct transcriptional states, the t-SNE plot (Fig. 1), designed to preserve local relationships, showed sub-structure within the clusters which is indicative of progressive changes in gene expression over time. Differential gene expression and functional analysis confirmed this: the termination trigger initiated a complex program of gene expression regulation that included transient changes in transcription starting with nearly 900 DEGs between the first two timepoints taken 1 h apart (T₀ and T_{1hr}).

Module	Parent GO term	Term description	GO term ID	E value
Black	chromatin organization GO:0006325	chromatin organization involved in negative regulation of transcription	GO:0097549	1.8e-07
		chromatin silencing	GO:0006342	1.9e-06
Yellow	RNA metabolic process GO:00160070	RNA modification	GO:0009451	1.2e-07
		histone mRNA catabolic process	GO:0071044	7.3e-07
Red	RNA metabolic process GO:00160070	RNA secondary structure unwinding	GO:0010501	1.1e-06
		valyl-tRNA aminoacylation	GO:0006438	4.7e-06
		maturation of SSU-rRNA	GO:0000462	3.0e-07
Green	RNA metabolic process GO:00160070	regulation of transcription	GO:0006355	2.4e-07
		rRNA processing	GO:0006364	1.7e-07
Blue	multicellular organism development GO:0007275	regulation of striated muscle tissue development	GO:0016202	1.8e-12
		regulation of muscle contraction	GO:0006937	1.1e-12
		larval somatic muscle development	GO:0007526	5.5e-15
		myoblast fusion	GO:0007520	2.9e-17
	signaling GO:0023052	regulation of I-kappaB kinase	GO:0043122	2.5e-23
		MAPK cascade	GO:0000165	1.0e-07
	immune system process GO:0002376	regulation of Toll signaling pathway	GO:0008592	1.0e-18
		immune system process	GO:0002376	3.1e-13
	cell cycle GO:0007049	nuclear division	GO:0000280	5.8e-06
		centrosome duplication	GO:0051298	9.6e-08
cell differentiation GO:0030154	cell differentiation	GO:0030154	7.7e-06	
Brown	multicellular organism development GO:0007275	regulation of muscle contraction	GO:0006937	3.0e-09
		locomotor rhythm	GO:0045475	1.8e-08
	signaling GO:0023052	integrin-mediated signaling pathway	GO:0007229	1.2e-07
	immune system process GO:0002376	immune system process	GO:0002376	3.1e-13
Cyan	metabolic process GO:0008152	digestion	GO:0007586	3.4e-15
		epoxygenase P450 pathway	GO:0019373	1.9e-07
		glucose metabolic process	GO:0006006	1.2e-05
		mannose metabolic process	GO:0006013	1.4e-05
	Oogenesis GO:0048477	oocyte development	GO:0048599	3.3e-07
	oocyte differentiation	GO:0007311	1.6e-06	

Fig. 6 Overrepresented process in females from collection to 14 days post-diapause. List of GO terms enriched for the differentially expressed genes (DEGs) in seven WGCNA modules as shown in the module correlation heatmap (see Fig. 5). Enrichment analysis was performed independently for the DEGs in each WGCNA module against the annotated reference transcriptome (n=59544). Module colors refer to Fig. 5. For each enriched GO term, parent GO terms (based on Gene Ontology hierarchical assignment), term description, GO ID and p-value adjusted for FDR are listed. Parent GO terms that were enriched exclusively in one module are in bold

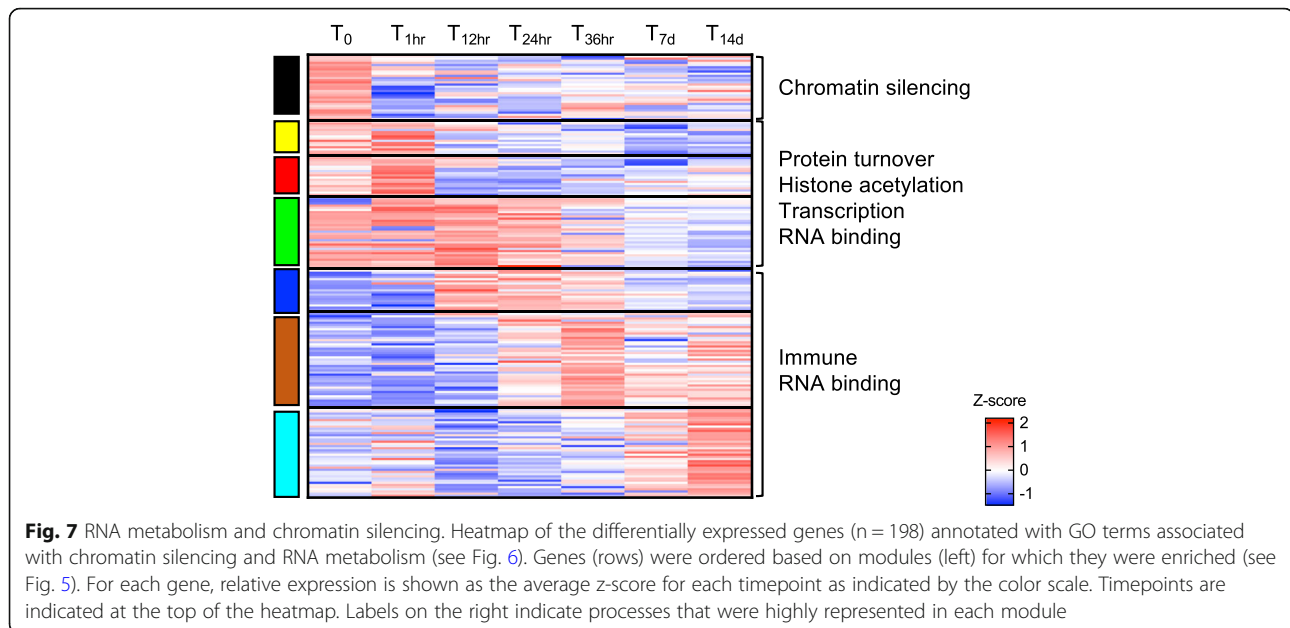
One-hour timepoint: up-regulation of genes within the 20E-cascade pathway

Before the females can switch from diapause to the post-diapause phenotype, they have to “break” dormancy by regulating a specific set of genes. The one-hour timepoint was characterized by the transient up-regulation of genes involved in the 20E-cascade pathway, a pathway which, in insects, is typically silenced (down-regulated) during dormancy to suppress larval growth [25]. In insects, the 20E transcription cascade promotes direct development and is linked to molting and/or metamorphosis [26], although it may also play a role in diapause termination [27]. Up-regulation of the ecdysone receptor was reported as a double peak in the pupae

of the flesh fly *S. crassipalpis* at 1 h and 9 h after the diapause-termination trigger [28]. This may reflect activation of two different processes, the first one linked to the diapause-to-post-diapause transition, the second to the initiation of metamorphosis. The transient up-regulation of these genes in *N. flemingeri* seen only at 1 h post-collection is consistent with a specific role in signaling diapause termination (Fig. 3).

One-hour timepoint: resumption of metabolism

Another transient signal at 1 h was the up-regulation of genes encoding enzymes involved in two major energetic pathways: oxidative phosphorylation and the tricarboxylic acid (TCA) cycle (Fig. 4). In arthropods, metabolic



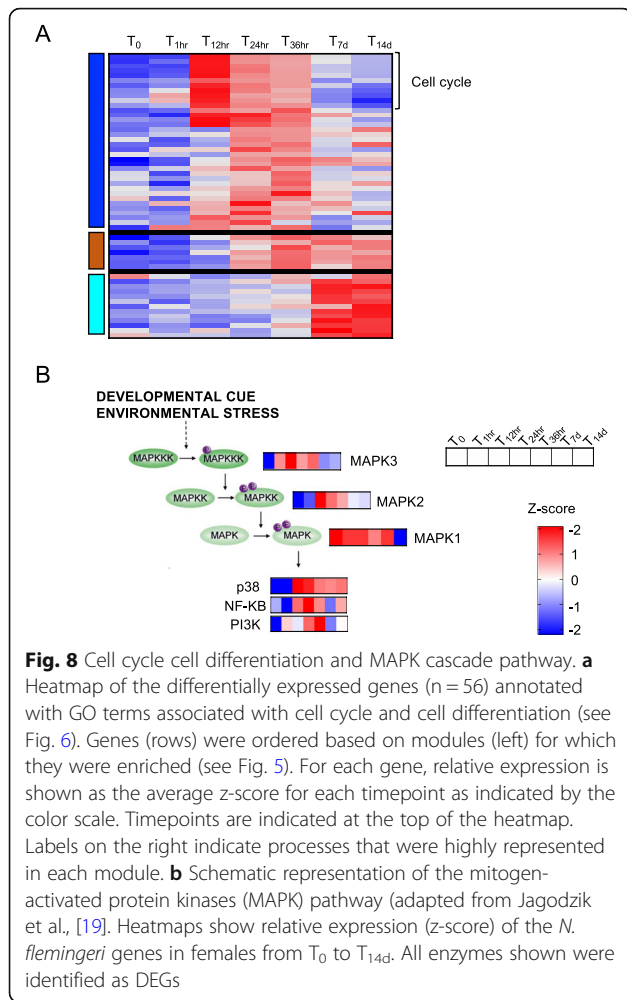
depression is ubiquitous during diapause. This repression can be achieved in different ways: blockage in the delivery of carbon fuel to the mitochondrion (e.g. *A. franciscana*), repression in the expression of genes associated with energetic pathways (e.g. mitochondrial enzymes involved in oxidative phosphorylation), or covalent modification of existing proteins [5]. In diapausing embryos of *A. franciscana*, which are in a deep state of arrest, inhibition of the enzymatic steps within mitochondrial activity occurs during entry into diapause [5]. Hydration-induced termination of diapause in gastrula-stage embryos triggers a transient up-regulation in the expression of two mitochondrial genes (COXI and COXII) in *A. franciscana* [24]. In *N. flemingeri*, genes associated with all five complexes in the oxidative chain were up-regulated at 1 h. In the T_{1hr} females, the peak was transient for genes associated with complexes I, IV and V. Expression of these genes was significantly lower at 12 h, although levels remained above T_0 . For the other two complexes (II and III) gene expression remained high for the remainder of the study (14 days).

An increase in metabolic rate between the diapause and post-diapause phases is universal in arthropods [6, 29]. In T_0 *N. flemingeri* females, expression levels of the eight transcripts encoding TCA enzymes are very low in comparison with the post-diapause levels reported in females starting early oogenesis but prior to spawning [14]. Surprisingly, we found that the increase in expression of these genes occurred before the females had transitioned from the diapause to the post-diapause phase based on t-SNE clustering. At 1 h, we observed a sharp, high peak in expression for all eight transcripts encoding TCA enzymes; at least 3-fold increases in

expression were found between T_0 and T_{1hr} followed by a decrease in expression at 12 h in seven of the eight transcripts. The initial up-regulation of the TCA genes at T_{1hr} preceded any changes in the expression of genes involved in other energetic pathways such as glycolysis and β -oxidation that were observed at T_{7d} , consistent with earlier observations [14].

Respiration rates remain low into post-diapause

Respiration rates are depressed during diapause in all organisms in which they have been measured. However, return to basal levels may not occur immediately, instead, increase of respiration rates can be biphasic as shown in insect pupae during post-diapause [29]. Respiration rates in *N. flemingeri* females averaged about 25% of fully active rates during the first 4 days post-collection (Fig. 11). This is similar to what has been reported for *C. finmarchicus* collected from depth [30, 31]. While these studies suggested that the 75% reduction represented respiration rates during diapause, our transcriptomic data indicated that in *N. flemingeri* the lowered values extend to post-diapause rates, after the up-regulation of genes involved in the TCA cycle and oxidative phosphorylation. Thus, respiration rates in copepods during diapause would be expected to be much lower, possibly approaching the 90% depression reported for diapausing pupae [29]. In addition, it is unclear whether *N. flemingeri* post-diapause respiration rates ever return to the high pre-diapause levels, or whether post-diapause respiration rates are biphasic, as reported for insect pupae.



Post diapause: release from transcriptional repression

Maintenance of low RNA metabolism as well as low protein turnover and low RNA:DNA ratio define the dormant state in diapausing copepods [7, 30, 31]. Not surprisingly, termination of diapause in *N. flemingeri* females included many DEGs involved in RNA metabolism, including several enriched GO child terms in three WGCNA modules (yellow, red and green). Resumption of transcription appeared to occur through two mechanisms: 1) down-regulation of repressors; and 2) up-regulation of genes involved in processes associated with transcription and protein turnover. Chromatin silencing is used in eukaryotes as one mechanism that counteracts transcriptional activity of RNA polymerases to control specific gene expression [32]. *N. flemingeri* is released from a state of transcriptional repression by an immediate down-regulation of genes associated with chromatin silencing (Fig. 7). The expression of several Longitudinal lacking proteins and Laminins, highly expressed in T_0

females, were down-regulated by 1 h. Expression remained low for the remainder of the experiment of this study and on through mid-oogenesis [14].

In addition to chromatin silencing, release from transcriptional repression may be occurring through the down-regulation of a group of 21 “diapause specific genes”, mostly annotated as Nuclear receptors, Retinoic acid receptors and the Fork head protein P. These are highly expressed during diapause preparation in *C. finmarchicus* pre-adults (CV) (Lenz et al. [33]) as well as in *N. flemingeri* T_0 females. In the latter, high expression during diapause is followed by immediate down-regulation at T_{1hr} . While we know little about the specific function of these genes in copepods, they are likely to be involved in keeping the organism in a state of transcriptional repression. The second mechanism involves restarting transcription and translation, which in the copepod was characterized by up-regulation of genes associated with protein turnover, histone acetylation and RNA binding. The regulation of these genes was immediate (T_{1hr}) and complex, characterized by the sequential and sustained up-regulation of multiple genes during the 14-day experimental period.

Cell cycle is reactivated early during post-diapause

Cell division is arrested during insect diapause [1, 33] and a significant increase in expression of Cyclins B (CycB) has been identified as signaling diapause termination. In the maggot *R. pomonella*, this signal occurred within 24 h of the initial ramp up in metabolic rate during post-diapause [6]. In *N. flemingeri* females, up-regulation of CycB genes had been observed at the one-week timepoint by Roncalli et al. [13]. In the current study, we found that cell cycle genes are already up-regulated in post-diapause T_{12hr} females, suggesting that this process may be one of the first to be activated (Fig. 8). Enrichment of two child GO terms associated with cell cycle was found for the blue module, which included six Cyclins A (CycA) genes with peak expression during early post-diapause in females from the T_{12hr} to T_{36hr} timepoints. By the 7 and 14-day timepoints, relative expression levels had decreased back to levels observed in the diapause phenotype (T_0 and T_{1hr} females). In *Drosophila melanogaster*, both CycA and CycB and proteins are highly abundant during the G2 phase of the cell cycle. During meiosis I, CycA peaks first during prophase I, while CycB reaches its maximum expression during metaphase I [34]. The sequential change in gene expression observed in *N. flemingeri* suggests that cell division is synchronized and coordinated, and activated soon after the termination of diapause. This is consistent with histological observations of the S-phase of the cell cycle commencing in the ovary of *N. flemingeri* females within 24 h of collection, and in some females in as little as six to 12 h [35].

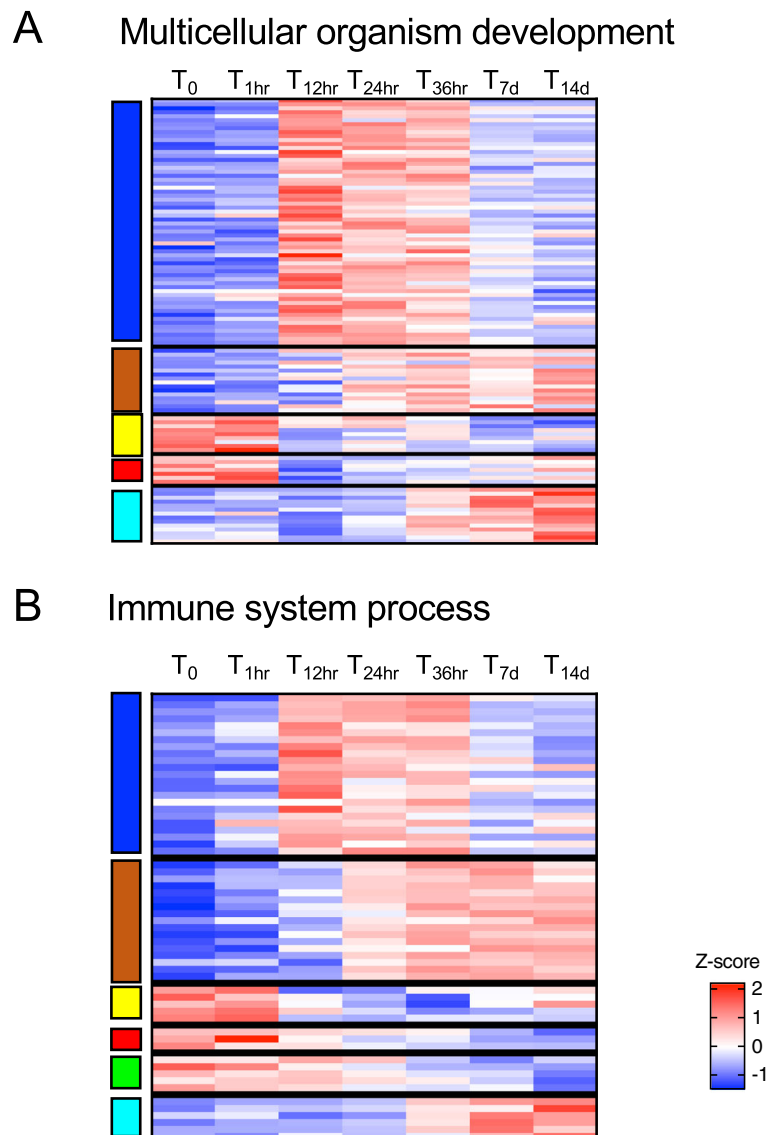


Fig. 9 Multicellular organism development and immune system process. Heatmap of the differentially expressed genes annotated with GO terms associated with: **a** multicellular organism development (n = 108) and **b** immune system process (n = 61) (see Fig. 6). Genes (rows) were ordered based on modules (left) for which they were enriched (see Fig. 5). For each gene, relative expression is shown as the average z-score for each timepoint as indicated by the color scale. Timepoints are indicated at the top of the heatmap

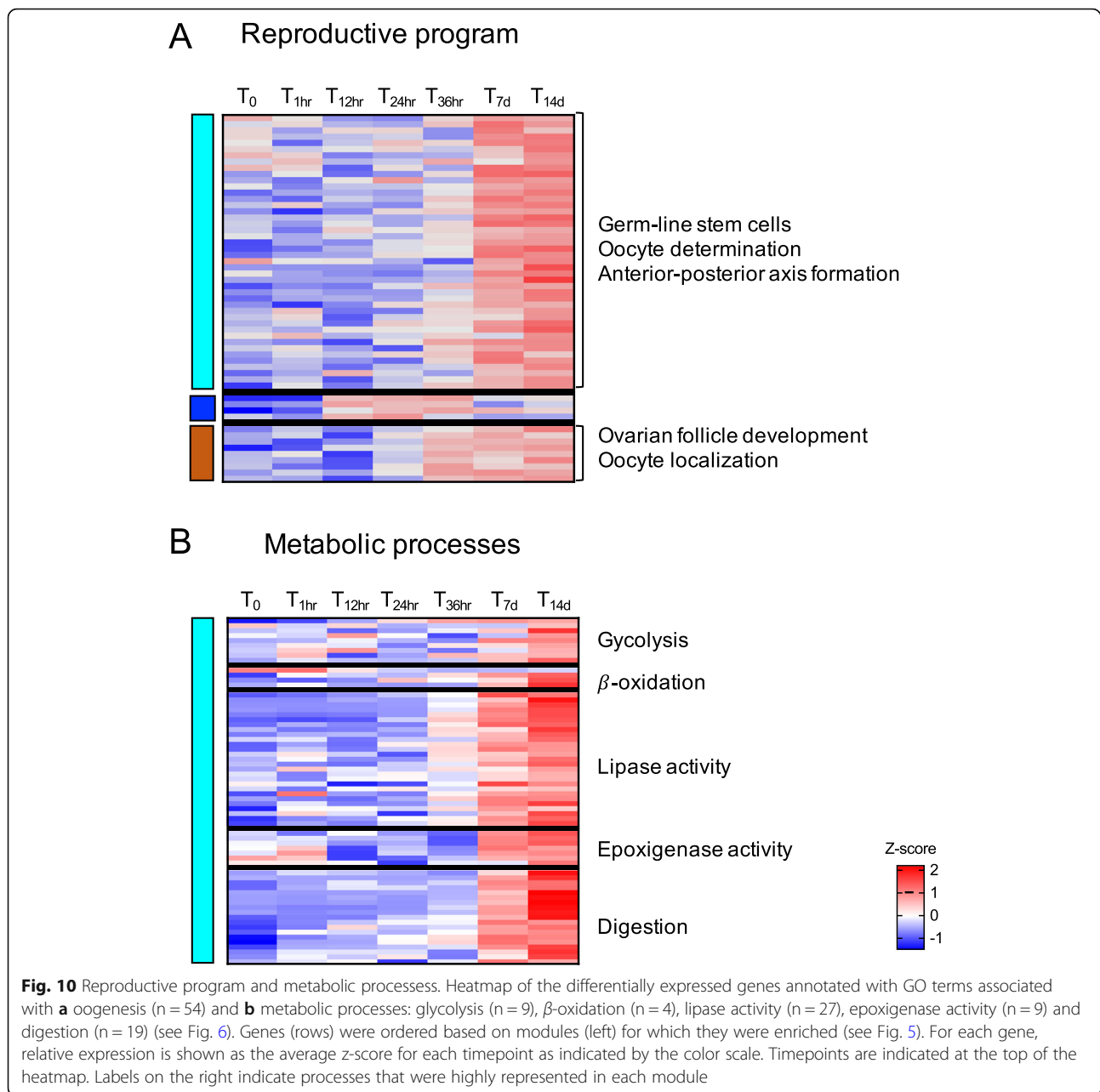
Cell growth and differentiation, the next processes activated during the post-diapause sequence

The transcriptional signal suggests that activation of cell growth and differentiation follows soon after the up-regulation of genes involved in cell cycle. In *N. flemingeri* overrepresentation of genes involved in cell differentiation had been reported to coincide with the enrichment of oogenesis GO terms [13]. Several transcripts encoding for Bicaudal D, Neurogenic delta, Innexins 2 and Aurora kinase showed a 2-fold expression increase at 1 week from collection compared with T₀ females [13]. In the current study, differential expression of genes involved in cell differentiation

started even earlier at T_{12hr} and was characterized by the sequential regulation of different sets of genes (Fig. 8). Some of the genes, up-regulated in females from T_{12hr} to T_{36hr} timepoints were annotated as Septins-7, Neurogenic Delta and Innexin 2. As reported earlier, Bicaudal D and Aurora kinases were only up-regulated in T_{7d} and T_{14d} females, with low expression at the earlier timepoints (this study) and at 3 weeks as reported previously [13].

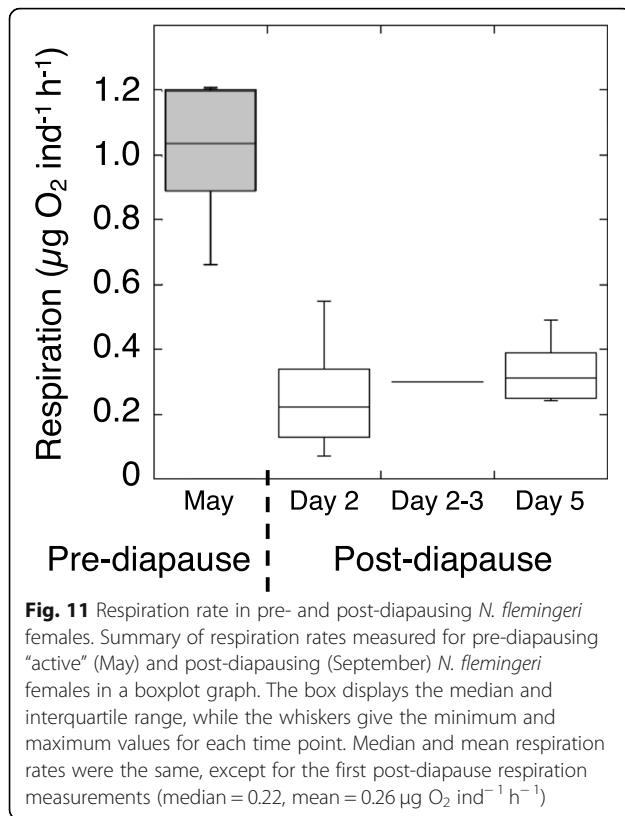
From post-diapause to reproductive program

In *N. flemingeri*, females mature and mate in the late spring and summer, prior to entering diapause. Upon



emergence during the post-diapause phase, they complete their lifecycle and spawn multiple clutches of eggs during the winter [10]. A previous study focused on post-diapause, reported the females' progression through the different stages of oogenesis from one-week post-collection to just prior to spawning, which starts at 7 weeks after diapause termination [13, 14]. Transcriptional profiling found that the reproductive program involves the sequential regulation of genes involved in multiple processes, such as cell differentiation, meiotic cell cycle, oogenesis and oocyte localization [13]. The current study, focused on the transition from diapause to post-diapause, included two time points, T_{7d} and T_{14d}

that overlapped with the earlier study. At those timepoints, we found enrichment of oocyte differentiation, a process that characterizes the early phases of oogenesis (Fig. 10). Approximately 80% of the up-regulated DEGs involved in those enriched processes were specific to T_{7d} and T_{14d} females; many of these genes had been also reported in the earlier study, as differentially expressed between diapausing females and females from 1 to 3-weeks post-collection. The expression of those DEGs returned to lower levels by three to 4 weeks post-collection, confirming their role in the early oogenesis phase [13]. However, it is also clear that the earlier study failed to capture the initial activation of the reproductive



program in the post-diapause phase. Prior to the T_{7d} time point, several genes involved in germline processes and cell differentiation, were up-regulated early in post-diapause, starting at the T_{24hr} to T_{36hr} timepoints. Genes like Armadillo segment polarity, Slowmo protein and Dicer peaked at T_{24hr} and T_{36hr} with a return to low expression at T_{7d} and T_{14d} at similar levels to T_0 . Thus, in this study, while the genes annotated to oogenesis may not have been enriched among the DEGs until T_{7d} , the reproductive program was activated almost immediately after diapause termination, a conclusion that is consistent with histological observations.

Activation of maintenance processes during post-diapause

Post-diapause includes a reactivation of basal organismal processes associated with cellular maintenance and homeostasis. In *N. flemingeri*, investment in cellular maintenance during early to mid-oogenesis phase is suggested by the up-regulation and enrichment of DEGs involved in intracellular signal transduction, muscle activity and immune response [13, 14]. Many of these genes were not up-regulated until the two-week timepoint [13, 14], which was confirmed in the present study. However, the earlier time points indicate that some genes involved in cellular maintenance became up-regulated early in post-diapause (Fig. 9). Several GO terms associated with multicellular organismal development and immune system

were overrepresented in two WGCNA modules (blue, brown), that included genes positively correlated with females from T_{12hr} to T_{14d} . Noteworthy was the transient up-regulation of several Myosins, Titins and genes associated with Toll and MAPK signaling pathways, which were only highly expressed from T_{12hr} to T_{36hr} (Fig. 9).

Conclusions

Among arthropods, post-diapause in *N. flemingeri* is relatively simple in terms of development, physiology and behavior. During post-diapause, females complete the reproductive program while maximizing survival to assure spawning of mature and fertilized eggs. Other biological processes are limited: females are not growing nor molting, they are developmentally mature, they are neither searching for a mate nor engaging in mating behavior, nor are they feeding. Females are nearly neutrally buoyant and swimming activity is very limited, although they are capable of escape responses to a mechanical stimulus. Nevertheless, the transcriptomic program that ends diapause and transitions to post-diapause is highly orchestrated starting within an hour of an effective termination stimulus.

How the diapause program of lipid-rich copepods will be affected by global climate change remains a major question. Changes in phenology and geographic distribution are expected, but heretofore predictions have not incorporated a physiological understanding of the diapause program. Studies have been hampered by the copepods' remote habitat and the inability to duplicate their life cycle in the laboratory. The transcriptional profiling of *N. flemingeri* females during diapause has generated a framework for further experimentation and testing of hypotheses related to female reproductive success.

Methods

Field collection and laboratory incubation

Neocalanus flemingeri adult females were collected from depth on September 21, 2017 in Prince William Sound (near station “KIP2”; depth:~ 589 m; Latitude 60° 27.13' N; Longitude 147° 99.13' W) at night during the fall cruise of the Seward Line Long-term Observation Program (LTOP) (<http://www.sfos.uaf.edu/sewardline/>). Zooplankton were collected from 400 to 550 m depth with a multiple opening and closing plankton net (0.25 m² cross-sectional area; 150 μm mesh nets; Multinet-Midi, Hydro-Bios) towed vertically. Collections were immediately diluted, and *N. flemingeri* females were live sorted with five individuals preserved individually in RNAlater Stabilization Reagent QIAGEN within 15 min of the retrieval of the net (T_0). At collection females were vertically oriented, head up with folded first antennae (A1) and pereopods remoted. They were unresponsive to mechanical stimulation; however, after ca. 1 h, many females had their A1s deployed. Additional undamaged females were sorted and transferred

into 2 L jars containing seawater collected from below 400 m and covered in aluminum foil and incubated at 5 °C (~the temperature of collection depth). After checking for survival, females (n = 5) were preserved at each of the following timepoints: T_{1hr} (n = 3), T_{12hr}, T_{24hr}, T_{36hr}, T_{7d} (one week) and T_{14d} (2 weeks) for RNA-Seq (Table S1). All samples were stored at -80 °C until further processing.

RNA extraction, RNA-Seq, expression profiling and clustering analysis

Total RNA was extracted from individual females using QIAGEN RNeasy Plus Mini Kit (catalog # 74134) in combination with a Qias shredder column (catalog # 79654) and shipped on dry ice to the Georgia Genomics Bioinformatics Core (<https://dna.uga.edu>). Double-stranded cDNA libraries (NGS Stranded) were multiplexed and sequenced using Illumina Next-Seq 500 instrument (High Output Flow Cell, 75 bp, paired end). From each library, reads < 50 bp long and with an average Phred score of < 30 were removed prior to downstream analysis (FASTX Toolkit v.2.0.0; Illumina Basespace Labs). For each individual, expression level was quantified by mapping each quality-filtered library (n = 33) against an existing *N. flemingeri* female reference transcriptome (NCBI BioProject: PRJNA32445; GFUD00000000) [36] using bowtie2 software (default settings; v.2.1.0) [37] and kallisto software (default settings; v.0.43.1) [38]. Kallisto software is expressly designed to reduce potential errors associated with ambiguous mapping, thus it is a better fit for gene expression analysis [37, 38]. Hence, counts generated by kallisto were used for the identification of differentially expressed genes. For the counts generated by the bowtie2 mapping, relative expression of the mapped reads was normalized using the RPKM method (reads per kilobase of transcript length per million mapped reads) [39]. This was followed by log₂ transformation of the relative expression data after adding a pseudocount of 1 to the RPKM value for each transcript (Log₂ [RPKM+1]). These log-transformed relative expression data were used for the dimensionality-reduction analysis (see below) and to calculate z-scores for each transcript and sample.

The dimensionality reduction method, t-distributed Stochastic Neighbor Embedding (t-SNE) [40] was applied to the normalized expression data (Log₂ [RPKM+1]) for all 33 samples. This clustering and visualization tool, which emphasizes local similarities among transcriptional profiles, is widely used in transcriptomics. The t-SNE as implemented in the R package Rtsne (Rtsne URL: <https://github.com/jkrijthe/Rtsne>) [41] was used to analyze the entire set of transcripts (n = 140841). After testing several values for the controlling parameters, as is standard practice [16, 40], a perplexity parameter of 9 and a maximum number of iterations of 50000 were

settled on, and the algorithm was run multiple times to ensure that the output was representative [16]. To provide an objective method of identifying clusters, the density-based clustering algorithm, DBSCAN (with *MinPts* = 3) was applied to the coordinates of the two dimensional t-SNE representation of the samples [16, 42]. The clustering cut-off (*Eps* parameter) was chosen to maximize the Dunn index score [43]. Both the DBSCAN algorithm and the Dunn index were run in R (dbscan: <https://CRAN.R-project.org/package=dbscan>) [44] (clusterCrit: <https://CRAN.R-project.org/package=clusterCrit>). In addition, two other widely used dimensionality reduction algorithms were tested on the dataset and the results (Supplementary Fig. 1).

Differential gene expression and weighted gene correlation network analysis (WGCNA)

EdgeR [45] was used to assess differential gene expression across samples, based on the counts generated from Kallisto mapping (see above). Transcripts with low expression levels (< 1 count per million (1 cpm)) were removed from each sample. As implemented by EdgeR, libraries (n = 33) were normalized by a TMM method (trimmed means of M values). Differentially expressed genes across samples were identified using the negative binomial generalized linear model (GLM) using the glmFit function with adjusted p-value for false discovery rate (Benjamini–Hochberg procedure) (FDR; p-value ≤ 0.05). To identify differences between each timepoint, pairwise comparisons were performed using likelihood ratio tests (glmLRT) (p-value ≤ 0.05, corrected for FDR). A Venn diagram of the DEGs from the pairwise comparisons T₀ vs T_{1hr} and T₀ vs T_{12hr} was generated using the R package Euler [46].

Patterns of differential gene expression among samples were further explored using weighted gene correlation network analysis (WGCNA) [47]. The WGCNA analysis was performed on the log-transformed (Log₂ [RPKM+1]) gene expression of all DEGs identified by the GLM (n = 14608). The WGCNA analysis used a signed network type (networkType = “signed”) with a soft threshold power of 6 and minimum module size of 100 transcripts. A signed network type means the construction of the network was based on the strength of positive correlations among the DEGs and negative correlations were ignored.

Functional analysis

Functional analysis of the DEGs was obtained by searching the *N. flemingeri* reference transcriptome that had been annotated (E-value cut-off = 1 e-05) against SwissProt (n = 62126) and Gene ontology (GO) databases (n = 59544) [36]. To identify overrepresented biological processes, a gene ontology (GO) enrichment analysis was performed on the DEGs grouped in each WGCNA module (network).

Briefly, DEGs from each module with GO annotation (> 50% per module) were compared against the annotated transcripts in the reference transcriptome (n = 59544) using TopGO software [48] which implements a Fisher exact test with a Benjamini-Hochberg correction (p-value ≤ 0.05 v. 2.88.0, default algorithm weight01). Heatmaps of the DEGs annotated within enriched GO terms (see Fig. 6) were generated using relative expression as z-scores of the average of the samples for each timepoint.

Respirometry

Respiration rates were measured for *N. flemingeri* “pre-diapausing” (May) and “post-diapausing” (September) adult females. In May, “pre-diapausing” females were recently molted individuals from pre-adults collected in May 2017 and readily responded with escape swims when mechanically stimulated. Post-diapausing females were individuals collected from depth (September 2016) with their first antennae (A1) folded immediately after collection, a behavioral sign of the diapause state. By 12 h post collection, all females had their A1 deployed perpendicular to the main body axis; even after deployment of the A1s, females were inactive and relatively non responsive to stimulations. These females were maintained in the laboratory during the post-diapause program.

Respiration rate experiments were run in the laboratory at the Seward Marine Center or in Fairbanks. In early May 2017, pre-adult copepodid stage CV were collected using a CalVet net (53 μm) towed vertically through the upper 100 m, individuals were live sorted under the microscope, transferred into carboys and maintained at ambient temperatures (6–7 °C). Recently molted females, which had not been mated, were used in respiration experiments on May 18 and 22 (experiment duration: 24–26 h, pre-diapause). Females in diapause were collected from depth on September 19, 2016 between 19:00 and 20:00 local time and processed as described above. Respiration experiments were conducted on the evening of September 21 (experiment duration: 5.5 h; time point: 2 days), and overnight on September 21–22 (experiment duration: 10.3 h; time point: 2–3 days) and 24–25 (experiment duration: 13.3 h; time point: 5 days). Incubation temperatures were between 4 and 6.5 °C. Each experiment included four blank (control) experiments containing only filtered seawater.

Oxygen consumption was measured by incubating either one (active) or one to six (post-diapause) individuals in 20 mL scintillation vials filled with freshly GFC-filtered sea water from the collection station that had been equilibrated overnight in an incubator, while being stirred gently with a magnetic stirrer to achieve uniform oxygenation. At the start of the experiment, animals were gently drawn into wide-bore pipettes and transferred to vials in the same order as subsequent

measurement. A mini-stir bar was added, bubbles were purged and vials sealed with plastic scintillation-vial caps or cover slips, then submerged in the stirred water bath and placed back in the incubator. At the end of the experiment, each vial was magnetically stirred to distribute deoxygenated water evenly (checked with dye), then uncapped before the oxygen measurement. Percent saturation of O₂ in sea water was determined with a pre-calibrated PreSens TX-3 “optode” (non-destructive optical probe: PreSens Precision Sensing GmbH, Regensburg, Germany) while measuring temperature with a ganged probe lowered to a tip-depth of 3 cm in each vial in turn. The TX-3 was set to record data every 30s and the first reading was discarded to allow settling; thereafter 3–5 readings were taken until 3 fell within 0.5% of each other and the mean of the 3 was recorded. All measurements were taken at the same temperature (± 0.2 °C), as the optode was unexpectedly sensitive to temperature despite its temperature-compensating probe. The difference in percent saturation at the run temperature for each vial from the mean value of 2–4 blanks run simultaneously was converted to $\mu\text{g O}_2 \text{ ind}^{-1} \text{ h}^{-1}$ using standard tables for sea water saturation (www.EngineeringToolBox.com).

Abbreviations

20E: 20-hydroxyecdysone; Genes: Shd: shade; EcR: Ecdysone receptor; USP: Ultraspiracle; Nurf-38: Nucleosome remodeling factor-38kD; cdk: Cyclin-dependent kinases; cyc: Cyclins; MAPK: Mitogen-activated protein kinases; DEGs: Differentially expressed genes; GLM: Generalized linear model; RPKM: Reads per kilobase of transcript length per million mapped reads; WGCNA: Weighted gene correlation network analysis

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07557-7>.

Additional file 1:

Acknowledgements

We would like to thank L. Hata of the University of Hawai‘i at Mānoa for administrative support, M. Belanger and R. Nilsen from the Georgia Genomics Facility at the University of Georgia for sequencing and assistance with samples. We would also like to thank the science party and the crew of the USFWS *RV Tiglax* for at sea support and assistance with collection. This is the University of Hawai‘i at Mānoa School of Ocean and Earth Science and Technology contribution number 11291.

Authors' contributions

VR, PHL and RRH conceived the study. VR, PHL, DKH performed the experiments. VR, PHL, MCC, AMC analyzed the data. VR, in collaboration with PHL, and with advice and suggestions from AMC, DKH, interpreted the data and wrote the manuscript. VR, PHL, DKH, AMC and RRH reviewed the manuscript. All authors approved the final manuscript.

Funding

This work was supported by the National Science Foundation Grant OCE-1459235 to PHL and A. E. Christie and Grant OCE-1459826 to RRH. Additional support was provided by the National Science Foundation NCGAS under Grant Nos. DBI-1458641 and ABI-1062432 to Indiana University. Seward Line core activities were supported by a consortium of the North Pacific Research Board, the Alaska Ocean Observing System, and the Exxon Valdez Oil Spill

Trustee Council (through Gulf Watch Alaska). The views expressed herein are those of the authors and do not reflect the views of the funding agencies.

Availability of data and materials

The datasets for this study can be found as raw short sequences on NCBI; BioProject: PRJNA324453. The shotgun assembly used as the reference transcriptome in the article is available under the accession GFDU00000000 (PRJNA324453). The list of all differentially expressed genes (GLM), their relative expression as Log₂ [RPKM+1] and their functional annotation is included in the Table S2. The list of DEGs from each pairwise likelihood test with their fold change expression, *P*-value and FDR are included in supplementary information files (Files S3, S4, S5, S6, S7, S8).

Declarations

Ethics approval and consent to participate

No special approvals or permissions were required for the collections and experiments on the marine organisms used in this study. This work did not include human subjects or vertebrate animals.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Pacific Biosciences Research Center, University of Hawai'i at Mānoa, 1993 East-West Rd, Honolulu, HI 96822, USA. ²Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Naples, Italy. ³Institute of Marine Science, University of Alaska, Fairbanks, 120 O'Neill, Fairbanks, AK 99775-7220, USA.

Received: 21 October 2020 Accepted: 25 March 2021

Published online: 03 June 2021

References

- Denlinger DL. Regulation of diapause. *Annu Rev Entomol.* 2002;47(1):93–122. <https://doi.org/10.1146/annurev.ento.47.091201.145137>.
- Mansingh A. Physiological classification of dormancies in insects. *Can Entomol.* 1971;103(7):983–1009. <https://doi.org/10.4039/Ent103983-7>.
- Tauber CA, Tauber MJ. Insect seasonal cycles: genetics and evolution. *Ann Rev Ecol Evol Syst.* 1981;12(1):281–308. <https://doi.org/10.1146/annurev.es.12.110181.001433>.
- Košťál V. Eco-physiological phases of insect diapause. *J Insect Physiol.* 2006; 52(2):113–27. <https://doi.org/10.1016/j.jinsphys.2005.09.008>.
- Podrabsky JE, Hand SC. Physiological strategies during animal diapause: lessons from brine shrimp and annual killifish. *J Exp Biol.* 2015;218(Pt 12): 1897–906. <https://doi.org/10.1242/jeb.116194>.
- Ragland GJ, Egan SP, Feder JL, Berlocher SH, Hahn DA. Developmental trajectories of gene expression reveal candidates for diapause termination: a key life-history transition in the apple maggot fly *Rhagoletis pomonella*. *J Exp Biol.* 2011;214(Pt 23):3948–59. <https://doi.org/10.1242/jeb.061085>.
- Hirche H-J. Diapause in the marine copepod, *Calanus finmarchicus*—a review. *Ophelia.* 1996;44(1–3):129–43. <https://doi.org/10.1080/00785326.1995.10429843>.
- Baumgartner MF, Tarrant AM. The physiology and ecology of diapause in marine copepods. *Annu Rev Mar Sci.* 2017;9(1):387–411. <https://doi.org/10.1146/annurev-marine-010816-060505>.
- Miller CB. *Neocalanus flemingeri*, a new species of Calanidae (Copepoda: Calanoida) from the subarctic Pacific Ocean, with a comparative redescription of *Neocalanus plumchrus* (Marukawa) 1921. *Prog Oceanogr* 1988;20:223–273. *Neocalanus flemingeri*, a new species of Calanidae (Copepoda: Calanoida) from the subarctic Pacific Ocean, with a comparative redescription of *Neocalanus plumchrus* (Marukawa) 1921, 4. DOI: [https://doi.org/10.1016/0079-6611\(88\)90042-0](https://doi.org/10.1016/0079-6611(88)90042-0).
- Miller CB, Clemons MJ. Revised life history analysis for large grazing copepods in the subarctic Pacific Ocean. *Prog Oceanogr.* 1988;20(4):293–313. [https://doi.org/10.1016/0079-6611\(88\)90044-4](https://doi.org/10.1016/0079-6611(88)90044-4).
- Tsuda A, Saito H, Kasai H. Life histories of *Neocalanus flemingeri* and *Neocalanus plumchrus* (Calanoida: Copepoda) in the western subarctic Pacific. *Mar Biol.* 1999;135(3):533–44. <https://doi.org/10.1007/s002270050654>.
- Liu H, Hopcroft RR. Growth and development of *Metridia pacifica* (Copepoda: Calanoida) in the northern Gulf of Alaska. *J Plankton Res.* 2006; 28(8):769–81. <https://doi.org/10.1093/plankt/fbl009>.
- Roncalli V, Sommer SA, Cieslak MC, Clarke C, Hopcroft RR, Lenz PH. Physiological characterization of the emergence from diapause: a transcriptomics approach. *Sci Rep.* 2018;8:1.
- Roncalli V, Cieslak MC, Hopcroft RR, Lenz PH. Capital breeding in a diapausing copepod: a transcriptomics analysis. *Front Mar Sci.* 2020;7:56.
- Lenz PH, Roncalli V. Diapause within the context of life-history strategies in Calanid copepods (Calanoida: Crustacea). *Biol Bull.* 2019;237(2):170–9. <https://doi.org/10.1086/705160>.
- Cieslak MC, Castelfranco AM, Roncalli V, Lenz PH, Hartline DK. t-Distributed Stochastic Neighbor Embedding (t-SNE): A tool for eco-physiological transcriptomic analysis. *Mar Genomics.* 2020;51:100723.
- Hwang DS, Han J, Won EJ, Kim DH, Jeong CB, Hwang UK, et al. BDE-47 causes developmental retardation with down-regulated expression profiles of ecdysteroid signaling pathway-involved nuclear receptor (NR) genes in the copepod *Tigriopus japonicus*. *Aquat Toxicol.* 2016;177:285–94. <https://doi.org/10.1016/j.aquatox.2016.06.004>.
- Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30. <https://doi.org/10.1093/nar/28.1.27>.
- Jagodzik P, Tajdel-Zielinska M, Ciesla A, Marczak M, Ludwikow A. Mitogen-activated protein kinase cascades in plant hormone signaling. *Front Plant Sci.* 2018;9:1387. <https://doi.org/10.3389/fpls.2018.01387>.
- Lovett DL, Felder DL. Ontogenetic change in digestive enzyme activity of larval and postlarval white shrimp *Penaeus setiferus* (Crustacea, Decapoda, Penaeidae). *Biol Bull.* 1990;178(2):144–59. <https://doi.org/10.2307/1541973>.
- Rinehart JP, Denlinger DL. Heat-shock protein 90 is down-regulated during pupal diapause in the flesh fly, *Sarcophaga crassipalpis*, but remains responsive to thermal stress. *Insect Mol Biol.* 2000;9(6):641–5. <https://doi.org/10.1046/j.1365-2583.2000.00230.x>.
- Tammariello SP, Denlinger DL. G0/G1 cell cycle arrest in the brain of *Sarcophaga crassipalpis* during pupal diapause and the expression pattern of the cell cycle regulator, proliferating cell nuclear antigen. *Insect Biochem Mol Biol.* 1998;28(2):83–9. [https://doi.org/10.1016/S0965-1748\(97\)00082-9](https://doi.org/10.1016/S0965-1748(97)00082-9).
- Fujiwara Y, Denlinger DL. High temperature and hexane break pupal diapause in the flesh fly, *Sarcophaga crassipalpis*, by activating ERK/MAPK. *J Insect Physiol.* 2007;53(12):1276–82. <https://doi.org/10.1016/j.jinsphys.2007.07.001>.
- Wang W, Meng B, Chen W, Ge X, Liu S, Yu J. A proteomic study on postdiapaused embryonic development of brine shrimp (*Artemia franciscana*). *Proteomics.* 2007;7(19):3580–91. <https://doi.org/10.1002/pmic.200700259>.
- Denlinger D, Yocum G, Rinehart J. Hormonal control of diapause. *Insect Endocrinol*; 2012. p. 430–463. Hormonal Control of Diapause, DOI: <https://doi.org/10.1016/B978-0-12-384749-2.10010-X>.
- Hill RJ, Billas IM, Bonneton F, Graham LD, Lawrence MC. Ecdysone receptors: from the Ashburner model to structural biology. *Annu Rev Entomol.* 2013; 58(1):251–71. <https://doi.org/10.1146/annurev-ento-120811-153610>.
- Kostal V, Stetina T, Poupardin R, Korbelova J, Bruce AW. Conceptual framework of the eco-physiological phases of insect diapause development justified by transcriptomic profiling. *Proc Natl Acad Sci U S A.* 2017;114(32): 8532–7. <https://doi.org/10.1073/pnas.1707281114>.
- Rinehart JP, Cikra-Ireland RA, Flannagan RD, Denlinger DL. Expression of ecdysone receptor is unaffected by pupal diapause in the flesh fly, *Sarcophaga crassipalpis*, while its dimerization partner, USP, is downregulated. *J Insect Physiol.* 2001;47(8):915–21. [https://doi.org/10.1016/S0022-1910\(01\)00064-6](https://doi.org/10.1016/S0022-1910(01)00064-6).
- Ragland GJ, Fuller J, Feder JL, Hahn DA. Biphasic metabolic rate trajectory of pupal diapause termination and post-diapause development in a tephritid fly. *J Insect Physiol.* 2009;55(4):344–50. <https://doi.org/10.1016/j.jinsphys.2008.12.013>.
- Ingvarsdottir A, Houlihan DF, Heath MR, Hay SJ. Seasonal changes in respiration rates of *Calanus finmarchicus* (Gunnerus) at copepodite stage V. *Fish Oceanogr.* 1999;8(Suppl 1):73–83. <https://doi.org/10.1046/j.1365-2419.1999.00002.x>.
- Saumweber WJ, Durbin EG. Estimating potential diapause duration in *Calanus finmarchicus*. *Deep-Sea Res II Top Stud Oceanogr.* 2006;53(23–24): 2597–617. <https://doi.org/10.1016/j.dsr2.2006.08.003>.

32. Beisel C, Paro R. Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet.* 2011;12(2):123–35. <https://doi.org/10.1038/nrg2932>.
33. Lenz PH, Roncalli V, Cieslak MC, Tarrant AM, Castelfranco AM, Hartline DK. Diapause vs. reproductive programs: transcriptional phenotypes in a keystone copepod. *Commun Biol.* 2021;4(1):1–13.
34. Whitfield W, Gonzalez C, Maldonado-Codina G, Glover DM. The A-and B-type cyclins of *Drosophila* are accumulated and destroyed in temporally distinct events that define separable phases of the G2-M transition. *EMBO J.* 1990;9(8):2563–72. <https://doi.org/10.1002/j.1460-2075.1990.tb07437.x>.
35. Monell, K. Characterization of cell division in the tissues of the calanoid copepod *Neocalanus flemingeri* from diapause through early oogenesis. MS Thesis University of Hawai'i at Manoa. <http://hdl.handle.net/10125/73356>.
36. Roncalli V, Cieslak MC, Sommer SA, Hopcroft RR, Lenz PH. *De novo* transcriptome assembly of the calanoid copepod *Neocalanus flemingeri*: a new resource for emergence from diapause. *Mar Genomics.* 2018;37:114–9. <https://doi.org/10.1016/j.margen.2017.09.002>.
37. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25. <https://doi.org/10.1186/gb-2009-10-3-r25>.
38. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* 2016;34(5):525–7. <https://doi.org/10.1038/nbt.3519>.
39. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5(7):621–8. <https://doi.org/10.1038/nmeth.1226>.
40. van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* 2008;9:2579–605.
41. Krijthe JH. Rtsne: T-distributed stochastic neighbor embedding using Barnes-Hut implementation. R package version 013, URL <https://github.com/jkrijthe/Rtsne>. 2015.
42. Ester M, Kriegel H-P, Sander J, Xu X, editors. A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*; 1996.
43. Dunn JC. Well-separated clusters and optimal fuzzy partitions. *J Cybernetics.* 1974;4(1):95–104. <https://doi.org/10.1080/01969727408546059>.
44. Hahsler M, Piekenbrock M, Arya S, Mount D. dbscan: Density Based Clustering of Applications with Noise (DBSCAN) and related algorithms. R package version. 2017:1.0.
45. Schaeffer L, Pimentel H, Bray N, Melsted P, Pachter L. Pseudoalignment for metagenomic read assignment. *Bioinformatics.* 2017;33(14):2082–8. <https://doi.org/10.1093/bioinformatics/btx106>.
46. Larsson J. Eulerr: area-proportional Euler and Venn diagrams with ellipses. R package version. 2018;4(0).
47. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 2008;9(1):559. <https://doi.org/10.1186/1471-2105-9-559>.
48. Alexa A, Rahnenfuhrer J. Gene set enrichment analysis with topGO. *Bioconductor Improv.* 2009;27:1–26.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

