Experimental analysis of development, lipid accumulation and gene expression in a high-latitude marine copepod

Vittoria Roncalli[1], Lauren N. Block[2], Jeanette L. Niestroy[2†],
Matthew C. Cieslak[2], Ann M. Castelfranco[2], Daniel K. Hartline[2] and Petra H. Lenz[2*]

[1] Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Vila Comunale, 80121 Naples, Italy

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

*corresponding author

[†] Current address: Institute of Pathology, Klinikum Chemnitz (Hospital of Chemnitz), Flemmingstraße 2, Germany

Running head: Lipid accumulation & gene expression in a copepod

Keywords: zooplankton, *Neocalanus flemingeri*, biomass, Gulf of Alaska, transcriptomics, physiology, lipid accumulation window, development

Abstract

The high-latitude copepod *Neocalanus flemingeri* exploits the spring phytoplankton bloom to accumulate lipid stores to survive food-limited periods and to fuel reproduction. At some point during development lipid-accumulation ends, pre-adults (stage V) molt into adults and descend to depth and enter a state of dormancy termed "diapause." How and when they determine to make this transition is unresolved. According to one hypothesis, the trigger is their attaining a threshold amount of "lipid fullness." Alternatively, they are on a fixed program, entering diapause within a narrow developmental window. To better understand the decision, a 5-week laboratory experiment was conducted to assess the effect of food quantity and type on lipid accumulation, biomass and gene expression in stage CV *N. flemingeri*. In fed individuals, the initial rate of lipid accumulation slowed by the end of the experiment, as a portion of CVs began to molt into adults. While changes in gene expression common to all fed individuals between weeks 1 and 3 were consistent with a developmental program, the duration of the CV stage was variable. Individuals deprived of food maintained lipid stores initially, suggesting physiological acclimatization to conserve energy. A comparison with gene expression profiles of field-collected individuals suggests similar responses to resource in the environment.

Introduction

Many organisms that inhabit highly cyclical environments have evolved life histories characterized by a period of resource accumulation followed by a period of dormancy. This optimizes survival during times of limited resources and enables timing subsequent reproduction to coincide with the reappearance of resources needed to ensure growth and survival of the next generation (Dahms, 1995, Denlinger & Armbruster, 2014, Hahn & Denlinger, 2011). In many high-latitude marine ecosystems this strategy has been adopted by mesozooplankton, the spring biomass of which is dominated by large lipid-rich calanoid copepods having life cycles that are synchronized to the spring phytoplankton bloom (Conover & Huntley, 1991, Coyle & Pinchuk, 2003, Søreide *et al.*, 2010). For these copepods, the lipid-resource accumulation phase is efficiently carried out in the larger, developmentally more advanced stages, followed by dormancy at depth in the pre-adult or adult stages (Baumgartner & Tarrant, 2017, Hirche, 1996). However, it is unclear how the decision is made to cease accumulation and enter dormancy. Furthermore, spatial and temporal variability in bloom dynamics subjects individual copepods to a range of food conditions during this phase of their life history (Mackas & Coyle, 2005). Food resources during the spring can be patchy not only in overall abundance, but also in size and species distribution (Strom *et al.*, 2016, Strom *et al.*, 2019, Strom *et al.*, 2006, Waite & Mueter, 2013). Knowing how resource availability affects an individual's ability to accumulate lipid stores is key to understanding their resilience and ability to cope with non-optimal environmental conditions.

The copepods' ability to store lipids allows them to transform a brief pulse in primary production into a long-term energy source (Conover & Corner, 1968, Darnis *et al.*, 2012, Jónasdóttir *et al.*, 2015, Kaartvedt, 2000, Kattner *et al.*, 2007, Record *et al.*, 2018). These lipid-rich calanoid copepods play a key role in high-latitude pelagic food webs as an important food source for higher trophic levels, so their persistence is critical to ecosystem health (Kattner & Hagen, 2009, Lee *et al.*, 2006). While some species start to build lipid stores during the early copepodite stages, lipids in most accumulate primarily during the pre-adult copepodite CV stage (Miller & Nielsen, 1988, Tsuda *et al.*, 2001). Lipid stores in stage CV individuals increase over the season and maximum lipid sac volumes are reported just prior to or during the early stages of diapause (Coleman, 2022, Miller *et al.*, 2000, Miller *et al.*, 1998). One hypothesis for the cue that ends the lipid build-up phase is the "lipid-accumulation-window hypothesis." This proposes that individuals enter diapause after they have accumulated sufficient lipid stores (Irigoien, 2004, Johnson *et al.*, 2008), suggesting that if resources are sparse diapause is delayed to lengthen the period of lipid accumulation. Alternatively, timing of the initiation of diapause could be controlled by environmental factors such as photoperiod and temperature or an endogenous program (Denlinger,

2002, Häfker *et al.*, 2018, Marcus & Scheef, 2010). While in diapause, the copepods are mostly inactive, their metabolic rates are depressed and development is arrested (Auel *et al.*, 2003, Hirche, 1983, Saumweber & Durbin, 2006). The stored lipids fuel diapause, development during post-diapause, and contribute energy to reproduction (Hirche, 2013, Lenz & Roncalli, 2019, Niehoff *et al.*, 2002, Saumweber & Durbin, 2006).

To address the issues of lipid accumulation, resource availability and the transition to diapause, we designed an experiment to monitor the developmental and physiological conditions during the lipid accumulation phase in the calanoid, *Neocalanus flemingeri*. Females of this species diapause as non-feeding adults, allowing us to restrict examination of the effect of food quantity and type on lipid accumulation in the pre-adult copepodite stage CV. With a single generation per year, the non-feeding adult depends on its lipids to fuel both diapause and winter reproduction (Lenz & Roncalli, 2019, Miller & Clemons, 1988, Tsuda *et al.*, 2001). As a CV, the species grazes on a variety of organisms, preferentially on large ciliates, but also on other phytoplankton and lipid-rich diatoms (Dagg *et al.*, 2009, Liu *et al.*, 2008). We collected stage CIV *N. flemingeri* in mid-April from the Gulf of Alaska, allowed them to molt into the CV stage and then incubated them in four different food treatments to track biomass, lipid accumulation and gene expression. We report on their response to starvation and examine the effects of two different food levels and diet compositions.

Materials and Methods

Zooplankton collection, live sorting and experimental set up

Zooplankton were collected using a QuadNet CalVET net (paired nets with mesh sizes of 150 and 53 µm) towed vertically from 100 m depth to surface at station GAK1 (Lat: 59º 50.7', Long: 149º28', depth: 270 m) on April 15, 2019. Zooplankton collections were immediately diluted in surface seawater, transferred into 20 L containers in coolers and transported to the laboratory (Seward Marine Center, Seward, Alaska). The collection was sorted under the microscope to identify *N. flemingeri* stage CIV individuals, which were transferred into five 5 L containers with ambient seawater with ca. 50 to 75 individuals each. Copepods from the containers were checked under the microscope every other day and newly molted CV individuals were removed and transferred into an experimental flask with ~ 700 ml GF/C filtered seawater (750 ml Falcon flasks) to a full complement of 3 CV individuals per flask that had molted within the past 48 hours.

The experiment was set up over a 4-week period as CIVs molted into CVs in the laboratory. Because the number of recently molted CV individuals were not sufficient to start the experiment in a shorter time span (i.e., one week) the CIV individuals were maintained in a water bath at 6.2-6.3

ºC under dim blue light (see below) and *Chaetoceros muelleri* was added 3x per week to supplement the ambient phytoplankton concentrations. The strain of *C. muelleri* used was one developed for aquaculture of marine plankton and it was readily available from the Alutiiq Pride Shellfish Hatchery. Treatment flask cohorts were started in sets of 4, one for each treatment to assure an even contribution of early and late molting CVs to each treatment.

Experimental flasks were incubated at constant temperature 6.2-6.3 ºC and dim blue light 12:12 L:D (LED light strips, peak wavelength: 460-465 nm, peak spectra irradiance of 0.1 to 0.2 mW m^{-2} nm^{-1} measured with a Qmini Spectrometer, RGB Lasersystems). The custom incubator system was constructed using four coolers (Coleman 316 Series 150-Quart Hard Ice Chest Cooler) that were plumbed to an aquarium cooling pump (EcoPlus 1/10 HP chiller) and a reservoir to produce a temperature-controlled water bath in each cooler. Daily feedings of flasks and experimental harvests (see below) were done at mid-day between 10:30 am and 3 pm.

Experimental design

Using four treatments, we tested the effect of amount and type of food on survival, lipid fullness, dry weight, carbon, nitrogen and relative gene expression in stage CV *N. flemingeri* over a 5-week period (Table 1). Recently molted stage CV individuals were isolated into incubation flasks and fed daily for up to 5 weeks. The four treatments consisted of one with no food (NF [black color code]: starved), and three "fed" ones: 1) low food, flagellates only (Low C [blue]: 75 µgC/flask/day); 2) high food, flagellates + heterotrophs (High C [green]: 225 µgC/flask/day); and 3) high food, flagellates + heterotrophs + diatoms (High C D+, termed "diatom" diet for convenience [orange]: 225 µgC/flask/day) (Table 1). The low and high food carbon concentrations were chosen based on reported phytoplankton carbon and chlorophyll *a* concentrations in the upper 25 m in Prince William Sound during the spring bloom (mid-April to mid-May) in the mid-90s (Tamburello, 2005). During the spring integrated chlorophyll *a* concentrations range from a low of 80 to 400 mg m^{-2} (0-25 m) and carbon to phytoplankton cell carbon ratio is around 25, giving an average range of 80 – 400 µg C L^{-1}.

We set up 56 flasks: 14 flasks per treatment x 4 treatments and with 3 individuals per flask (168 individuals total). Copepod survival was good and similar across flasks with the exception of two flasks in the High C treatment where all individuals died within the first week of incubation. Those flasks/individuals were treated as artifacts, they were removed from the experiment and not included in the survival statistics. For weekly measurements, all individuals from a set of flasks from each treatment were harvested, assessed for survival and sub-sets of the individuals were processed for either carbon, nitrogen and dry weight (C, N, DW) or gene expression (RNA-Seq).

Subsets of individuals were imaged through the microscope for size and lipid fullness. Time points for different measurements and sample sizes are summarized in Table 1. A subset of recently molted individuals (designated Wk0, 04/24) was processed for C, N, DW or RNA-Seq and imaged without being incubated. Details follow.

Phytoplankton cultures and feeding

Phytoplankton species used in the experiment are summarized in Table 2. Starter phytoplankton were obtained from the National Center for Marine Algae and Microbiota, Bigelow Laboratory for Ocean Sciences, East Boothbay, Maine, USA (*Odontella aurita*, *Ditylum brightwellii* and *Thalassiosira rotula*) and from Dr. Suzanne Strom's collection at Western Washington University (*Rhodomonas baltica*, *Dunaliella tertiolecta*, *Isochrysis galbana*, *Oxyrrhis marina*, *Heterocapsa triquetra*). Cultures were grown in F/2 medium in sterile seawater (with added silicate for diatoms) in polycarbonate or glass Erlenmeyer flasks with LED grow lights (Sun Blaze T5 LED 2) on a 16:8 L:D cycle. Heterotrophs (*O. marina*, *H. triquetra*) were raised on *I. galbana.* Prior to feeding, cell culture densities were assessed using a hemocytometer to calculate daily food rations. Variability in culture density required occasionally an adjustment in the proportion of the species within a category. Phytoplankton were harvested, mixed and the appropriate ration was added to each flask using either a pipettor or a small measuring cylinder. Prior to feeding and as needed, seawater was removed from flasks with minimal disturbance to the copepods to maintain experimental volumes near 700 mls.

Sample Processing

Imaging

Individuals were removed from each harvested flask, placed individually into chilled embryo dishes, checked under a Leica MZ16 microscope for stage and condition and imaged in lateral view at 25 or 32x with a Jenoptik 8Mpx or Spot 12Mpx Insight digital camera. Calibration scales were imaged for each magnification setting to convert from pixels to mm or mm^2. Images were analyzed with ImageJ (Schneider *et al.*, 2012) for measurements of prosome length (mm) and prosome and lipid sac areas (mm^2) using the line and area tools, respectively. We used the ImageJ function to calculate the area of the prosome and the lipid sac. A relative fullness index was calculated as the ratio of lipid sac area : prosome area as described previously (Monell *et al.*, 2023).

Carbon, nitrogen and dry weight

For the dry weight, carbon and nitrogen analysis, copepods were imaged under the microscope, rinsed in distilled water, transferred to pre-combusted and pre-weighed tin cups and

immediately frozen at -20ºC. A total of 34 CV individuals were processed for dry weight, carbon and nitrogen at the SOEST Laboratory for Analytical Biogeochemistry (S-LAB). At the S-Lab, copepods were dried at 65ºC for overnight before weighing on an ultramicrobalance (Mettler Toledo UMX2). For the carbon and nitrogen analysis, samples were processed in an elemental analyzer (Exeter Analytical model CE 440).

Statistical analysis

Prosome length, lipid fullness, dry weight, carbon content, and nitrogen content across the four food treatments and incubation times were tested for significant differences using ANOVA. Models were run with an interaction term between food treatments and incubation time. WK0 samples were excluded from this analysis. Dry weight, carbon, and nitrogen data were normalized using a natural log (ln) transformation. Residual plots for the models showed no violation of homoscedasticity assumption of linear models. Post-hoc pairwise comparisons were used to identify the effect of food treatment within timepoints and the effect of incubation time within food treatments using a t-test on estimated marginal means with a Tukey adjustment. Differences between WK0 and the experimental samples were assessed using a Student's t-test. The relationship between carbon content and C:N ratios was analyzed using a correlation analysis (Pearson correlation coefficient). Analysis was conducted in the R statistical and programming environment.

Gene expression analysis

RNA-Seq

For the gene expression studies, harvested individuals were preserved in RNALater Stabilization Reagent (QIAGEN) and frozen at -40ºC. Within 1 month of the conclusion of the experiment total RNA was extracted from individual CVs using the QIAGEN RNeasy Plus Mini Kit (catalog # 74134) in combination with a Qiashredder column (catalog # 79654) following the instructions of the manufacturer, checked for quality and quantity in a Nanodrop 2000 (ThermoScientific) and stored at -80ºC. Total RNA was extracted from four recently molted CVs (WK0) and 4-7 experimental CVs for each treatment/time point combination. Those with the highest quality and RNA yields (n=33, Table 1) were selected for RNA-Seq and shipped on dry ice to the University of Georgia Genomics Facility (dna.uga.edu), where total RNA concentrations and quality were checked using an Agilent Model 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). Double-stranded cDNA libraries were prepared using the Kapa Stranded mRNA-Seq kit (KK8420) following manufacturer's instructions. After enrichment of mRNA with oligo-dT beads, samples were fragmented, reverse transcribed into double-stranded complementary DNA and checked for fragment size and quality (Agilent DNF-474 HS NGS Fragment Kit). Each sample

was tagged with an indexed adapter and paired-end sequenced (PE75 bp) using the Illumina NextSeq 500 platform using High-Output Flow Cell. After quality assessment (FASTQC v1.0.0), any remaining Illumina adapters were removed and the first 9 bp were trimmed from each read using Trimmomatic (v. 0.36). Additionally, reads with low quality ("Phred" score < 30, matched pairs) and length < 50 bp were removed from each library. Each library resulted in 7 to 15 million high-quality reads per sample and an average of 9 million across all samples.

Relative gene expression and downstream analyses

Ribosomal RNA was removed from each RNA-Seq library (SortMeRNA) (Kopylova *et al.*, 2012) prior to mapping reads to a standard *N. flemingeri* reference transcriptome (NCBI: BioProject PRJNA496596, TSA GHLB01000000) (Roncalli *et al.*, 2019). Reads were mapped against the reference using kallisto software (default settings; v.0.43.1) (Bray *et al.*, 2016) and Bowtie2 software(v2.3.5.1) (Langmead *et al.*, 2009). Counts generated by the Bowtie2 mapping, were normalized using the RPKM method (reads per kilobase of transcript length per million mapped reads) (Mortazavi *et al.*, 2008), followed by $\log_2$ transformation of the relative expression data (Log_2[RPKM+1]). The log-transformed normalized expression data were used in the dimensionality-reduction analysis (see below) and to calculate z-scores for each transcript and sample.

For gene expression analysis, kallisto-mapped transcripts with low expression (< 1 count per million in all treatments [1cpm]) were removed leaving 46,416 transcripts (90%) that were tested for differential gene expression using the generalized linear model (Bioconductor package EdgeR, R v. 3.12.1). Sets of pairwise likelihood tests were performed to identify differentially-expressed genes (DEGs) in: 1) the progression from post-molt Week 0 (Wk0) to Week 3 individuals on the three food treatments; 2) the response to starvation in Week 1 individuals; and 3) differences between the fed treatments. For the first set, in each week (Wk1-Wk3) pairwise tests included Wk0 vs blue, green and yellow; for the second set, pairwise tests were run for no-food individuals vs each of the three fed treatments from Wk1. For the third set, we ran pairwise tests between the three fed treatments (Low C vs. High C, Low C vs. High C D+ and High C vs. High C D+) for each week. In all pairwise likelihood ratio tests ($p < 0.05$), p-values were adjusted for false discovery rate (FDR) using the Benjamini-Hochberg correction (default algorithm weighWk01) (Robinson *et al.*, 2010). The pairwise likelihood ratio test results were further analyzed across time and treatment using Venn diagrams and the web-based portal Venny (v2.1.0) to find common DEGs among different paired comparisons (Oliveros, 2007).

Clustering analysis (t-SNE)

The dimensionality reduction method, t-distributed Stochastic Neighbor Embedding (t-SNE) was used to cluster individuals by expression similarity (Cieslak *et al.*, 2020, Van Der Maaten & Hinton, 2008). t-SNE was applied to the normalized expression data (Log_2[RPKM+1]) for the entire set of transcripts (n=51,743) using the R package Rtsne (Krijthe, 2015). For the analysis, we set perplexity to 10 with a maximum number of iterations of 50,000. The algorithm was run multiple times to ensure that the output was representative (Cieslak *et al.*, 2020, Van Der Maaten & Hinton, 2008). Clusters were identified using the density-based clustering algorithm, DBSCAN (with *MinPts*=3), which was applied to the coordinates of the two-dimensional t-SNE representation of the sample (Ester *et al.*, 1996)**.** The clustering cut-off (*Eps* parameter) was chosen to maximize the Dunn index score (Dunn, 1974). Both the DBSCAN algorithm and the Dunn index were run in R using the package clusterCrit: v 1.2. 6. 2015) (Desgraupes & Desgraupes, 2018, Hahsler *et al.*, 2017).

Functional annotation

The differentially expressed genes were reviewed for predicted functions by searching accession numbers in the *N. flemingeri* reference that had been annotated against SwissProt, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases following established protocols (Roncalli *et al.*, 2019). Over-represented biological processes were identified using an enrichment analysis of the DEGs by comparing the GO distribution between DEGs and all annotated transcripts using the R-package topGO (v. 2.88.0) that employs a Fisher exact test with a Benjamini-Hochberg correction (p-value < 0.05; using the default algorithm weight01) (Alexa & Rahnenfuhrer, 2010). Downstream analysis, included further analysis of DEGs annotated with GO terms that were overrepresented (enriched), and target genes annotated to lipid metabolism. Relative expression of these transcripts was visualized in heatmaps as z-scores calculated from the log-transformed expression levels (Log_2[RPKM+1]) using the R-package heatmaply (Galili *et al.*, 2018).

Results

Survivorship, duration of the CV stage and prosome length

Overall survivorship in the experimental flasks averaged 87% (137/162) with the highest mortality recorded in flasks harvested during weeks 4 and 5. There was no significant difference among treatments (range: 82% – 91% survivorship; chi-square = 1.337, d.f. = 3, p = 0.72). It was noteworthy that mortality in the no-food treatment was not higher than in the ones with food. The duration of the stage CV for most individuals under low and high-food conditions was ≥ 5 weeks. Between May 25 and 30, eight individuals molted into the CVI stage (7 females and 1 male)

(incubation time: Wk3 [n=2], Wk4 [n=4], Wk5 [n=2]) out of 72 individuals) and were distributed among three treatments: High C D+ (diatom, n=4), High C (n=3) and Low C (n=1) treatments.

Average prosome length was similar across treatments (Wks 1 - 5; mean=3.4 mm, S.D. 0.4 mm, n=111, ANOVA Type II test, p=0.28, Table 3). Prosome lengths of Wk3 individuals were longer than Wk1 ones, and this difference drove an overall significant effect of incubation time (p = 0.03; Table 3), but not in comparison with the other weeks. In addition, prosome lengths of the Wk0 individuals were smaller with a mean of 2.7 mm (S.D.= 0.2; n=3), which is significantly different from the experimental animals (p= 0.0002, Student's t-test). While the Wk0 individuals were taken from the same collection and it is possible that there is some expansion in size after molting (Skinner, 1985), we cannot discount the possibility that these individuals were indeed smaller.

Biomass: lipid fullness, dry weight, carbon and nitrogen

The ratio of lipid area to prosome area was used as an indicator for lipid fullness as individuals progressed through the CV stage on the four treatments (Fig. 1). In the fed copepods lipid fullness increased from 12% to > 20% over the 5-week period in the low and as well as the high food treatments, while it decreased under no-food conditions. Lipid fullness of the eight individuals that molted into adults were highly variable and ranged from 9 to 28% with an outlier on either side. Significant differences were found among food treatments ($p<<0.0001$), but not for time (p=0.16; ANOVA Type II test, Table 3). In addition, the interaction between time and treatment was significant ($p< 0.002$; Table 3), given the loss of lipid fullness in the no-food treatment and the opposite trend in the fed treatments. However, lipid fullness did not decrease substantially in the no-food treatment until after Wk2, suggesting that the copepods were able to preserve their lipid stores initially. Thereafter, lipid stores decreased and approached zero by Wk4.

At Wk1, lipid fullness in High C D+ (diatom) treatment was higher than in the no-food (p=0.04), but not the other two treatments. At Wk2, lipid fullness in both high-food treatments were higher than the no-food treatment ($p<0.001$). Thereafter, all three fed treatments were significantly different from the no-food treatment (Wks 3 – 5, $p\leq0.01$). Lipid fullness comparisons across the fed treatments at different time points found no significant differences between the Low C and the High C D+ (diatom) treatments. However, High C copepods had the fullest lipid sacs over time (High C vs. Low C: Wk3, p=0.02 and 3, p=0.02; High C vs. High C D+ (diatom): Wk4, p=0.04 and Wk5, p=0.004).

Copepod biomass (dry weight, carbon and nitrogen) increased between Wk0 and Wk1 in all fed treatments (Fig. 2). While we expected growth during the first week, the observed increase in

dry weight, carbon and nitrogen may not accurately represent the actual increase in biomass due to the smaller average size of the Wk0 individuals (see above). Thus, we focused our analysis on the changes in biomass between Wk1 and Wk5. Initially, there was considerable overlap in dry weight, carbon and nitrogen measurements among all treatments including the no-food treatment, although average weights differed (Fig. 2A-C). At Wk2, overlap was still present, but averages were further apart. It was intriguing that the no-food individuals largely maintained their C, N, DW values in the first two weeks before declining, whereas individuals in the Low C did not gain significantly until after this time-period, despite having food available. Starting in Wk3, there was a significant treatment effect across all biomass indicators (ANOVA Type II Test; dry weight, carbon and nitrogen $p \leq 0.0001$); however, this effect was primarily due to the no-food treatment: biomass of individuals in this treatment declined steadily after Wk2. In two food treatments (Low C and High C D+ [diatom]), average dry weight, carbon and nitrogen increased by approximately a factor of 2 (Low C: 2.1, 2.6, 2.1, respectively; High C D+ [diatom]: 1.6, 1.5, 2.3, respectively). The increase in the High C treatment was higher (4.8, 5.6, 4.1, respectively), however, this was based on a single individual from the High C treatment at Wk4. Carbon to nitrogen ratios were highly variable with the C:N ratio ranging from below 3 to over 9 with the highest C:N ratios were typically found in individuals with high carbon content independent of food treatment (Fig. 2D, Pearson correlation coefficient: 0.71, $p<0.001$, n=33, recently-molted [Wk0] individuals were excluded).

Gene expression analysis

Overview

Based on the expression of all genes (n = 51,743) the t-SNE algorithm separated the individuals (n = 33) primarily by incubation time (Fig. 3). This pattern was likely driven by changes in gene expression associated with progression through the CV stage, including the accumulation of lipid, starting with post-molt individuals. While most individuals loosely aggregated by week, a few individuals were more similar to those from a different week. No diet-specific clusters were observed among the individuals from the three fed treatments. At Wk3, the t-SNE plot showed some segregation by fed treatment, although this was not significant. The generalized linear model (GLM) identified over 12,000 differentially expressed genes (DEGs) across the four time points and four food treatments. Downstream, pairwise tests were used to identify DEGs and regulated biological processes that characterized: 1) the progression from post-molt Wk0 individuals to Wk3 individuals, including expression of genes involved in lipid metabolism; 2) gene expression patterns in the three fed treatments; and 3) the response to starvation in Wk1 individuals.

Week 0 to Week 3 – temporal changes in gene expression

Differences in gene expression associated with progression through the CV stage were characterized using pairwise comparisons between the Wk0 individuals and Wk1, Wk3, or Wk3 individuals from each food treatment, but excluding the no-food treatment. Temporal changes in expression were identified using Venn diagrams to find the DEGs that were common among the fed treatments at each time point. The number of shared genes for each week increased over time with a total of 488 DEGs in Wk1, 633 DEGs in Wk2 and 1121 DEGs in Wk3 (Supp. Fig. S1). There was some overlap in these shared DEGs between weeks, and a number were not annotated. Overall, across the three weeks there were 565 unique DEGs with functional annotations. Relative expression of these genes is shown in a heatmap (Fig. 4; Supp. Table S1). Based on functional annotation these DEGs are involved in broadly-conserved biological processes such as development, energy metabolism and muscle function. However, none of these processes were overrepresented based on the enrichment analysis. In Wk1 and Wk2, the up-regulated DEGs included genes associated with muscle contraction (e.g. *paladin*, *myosin*), calcium transport and oxidative phosphorylation (e.g. *cytochromes b-c1 complex subunit 8*, *cytochrome c oxidase subunit 6C-1*, *cytochrome c oxidase subunit 7A1*). In Wk2 and Wk3, there was an increase in the number of up-regulated genes that were annotated to development such as those encoding the proteins Toll, Aubergine, Sonic hedgehog and Speckle, as well as several nuclear hormone receptors (nhr), retinoic receptors and cuticle proteins. In Wk3 there were several DEGs annotated to "response to stimulus" such as *DNAJ*, *CYP450*, *ferritin*, *superoxide dismutase*, *peroxiredoxin* and *glutathione S-transferase* (classes: sigma, mu, and microsomal).

Gene expression differences among the three fed treatments

Visualization of the multidimensional data set of expression-levels of all transcripts using t-SNE did not separate out individuals by diet. This pattern was confirmed in the GLM analysis. Diet-specific differences in gene expression were small, and only 33 DEGs characterized the differences between the Low C and High C diets at all time points (Supp. Fig. S2). The few that were functionally annotated included several vitellogenins, a transporter (major facilitator superfamily), a cuticle protein and an immune-associated protein (*GTPase IMAP family member 4*) (Supp. Table S2). Interestingly, these 33 DEGs were high responders and expression differences between the two diets ranged from +/- 4 to 13 fold-change (FC). Only 24 DEGs were shared between Low C and High C D+ (diatom) for all three time points.

Consistent differences in relative expression between the High C and High C D+ diets were limited to 31 genes across all time points (Supp. Table S2). Only a few of these DEGs were functionally annotated, and these included a NADH dehydrogenase, a transmembrane protease serine and a cuticle protein. Interestingly, at Wk1 we found evidence for a stress response in

individuals on the High C D+ (diatom) diet. Genes involved in response to stimulus (GO:0050896) were up-regulated in comparison with either Low C or High C diets (Fig. 5, Supp. Table S1). The up-regulated DEGs that were specific to the High C D+ diet included *cytochrome P450* (8 transcripts), *multidrug resistance-associated protein* (4 transcripts), genes associated with glutathione pathway (e.g., *GST mu*, *GDH*) and genes involved in response to oxidative stress (*nitric oxide synthase*, *aldehyde dehydrogenase*) (Fig. 5). This response to stress was not present in the later weeks.

Relative expression of genes involved in lipid accumulation

The three fed treatments were similar in the relative expression of genes that are involved in lipid accumulation (Fig. 6, Supp. Table S1). The nine genes shown in Fig. 6 are predicted to be involved in fatty acid synthesis and transport based on their annotation and thus involved in lipid accumulation (Tarrant *et al.*, 2021). While this analysis focused on the three fed treatments, Fig. 6 includes the relative expression in newly molted individuals (Wk0) and the no-food treatment for comparison. We found no significant differences in expression in six of the nine genes that were specific to a fed treatment (across time) or time point (across food treatment) (Fig. 6B-D, 6F-G, 6I). In two of the remaining genes, *acyl-CoA synthase* and *ELOV 4* (Fig. 6A, E) the primary difference was between the Low C and High C treatments at the Wk3 time point. The third gene (*FABP 6*) showed the largest differences between both treatments and time points (Fig. 6H). Relative expression of this gene increased with incubation time and this increase was significant. Furthermore, significant differences in expression were observed between Low C and High C D+ (diatom) in Wk1 and Wk2, but not Wk3. Significant differences in expression between Low C and High C were limited to Wk2.

Response to starvation: extreme physiological acclimatization

The largest differences in gene expression were in response to starvation. It was characterized by a large number of DEGs (from 4,685 to 6,262) between the no food individuals and CVs from the other treatments including the recently molted individuals (Wk0). Nearly 3,000 DEGs were shared among the three food treatments as shown in a Venn diagram (Fig. 7A). Enrichment analysis identified two biological processes that were overrepresented among the shared DEGs: metabolic process (GO:0008152) and muscle contraction (GO:0006936) (Fig. 7B, Supp. Table S3). Genes annotated to metabolic process were down-regulated, including those involved in glycolysis (10/10), citric acid cycle [TCA] (4/8, *aconitase*, *malate dehydrogenase*, *malate dehydrogenase*, *succinate-CoA ligase*), fatty acid [FA] synthesis (e.g. *acyl-CoA reductase*, *delta-9 desaturase*, *ELOV4*, *DGAT1*), FA transport (e.g. *FABP5*, *FATP4*), and digestion (e.g., trypsins) (Figs. 6, 7B). Interestingly, genes involved in lipid catabolism were not up-regulated,

instead we found low expression of genes involved in beta oxidation (Fig. 7B). Furthermore, many genes associated with the cellular stress response were down-regulated in the no-food treatment (Figs 5, 7B). These genes included peroxidases, heat shock proteins [*HSP*], multi xenobiotic response transporters [*MXR*] and genes involved in the glutathione pathway. Genes that were up-regulated under no-food conditions were involved primarily in two biological processes: protein degradation and muscle function. Up-regulated genes included those involved in proteolysis (e.g., *serine proteinase*, *calpain*, *cathepsin*, *matrix metalloproteinase*), protein ubiquitination (e.g., *E3 ubiquitin-protein ligase*, *mitochondrial ubiquitin ligase activator*, *ubiquitin-conjugating enzyme E2*), and muscle contraction (GO:0006936) such as *actin*, *myosin* and *tropomyosin* (Fig. 7B).

Discussion

This study was focused on obtaining a better understanding of development and lipid accumulation in pre-adult (CV) copepods during their preparation for seasonal dormancy (diapause). The goal was to gain insight into the signal(s) that initiates their descent to depth to enter diapause ("lipid-accumulation hypothesis") as well as to assess mechanisms of adaptation and resilience of developmental stages in the face of food insecurity in their habitat. Because the target species, *N. flemingeri*, has non-feeding adults, lipid resources accumulated during the pre-adult stage must suffice to fuel mating, diapause and post-diapause reproduction. The experimental approach included complementary measurements on lipid processing by recently-molted CV stages incubated under different food conditions. Results are discussed below in five areas: 1) the lipid accumulation window hypothesis and its relationship to the duration of the CV stage; 2) lipid accumulation and expression of genes related to fatty acid synthesis in fed animals; 3) evidence for a transient detoxification signal in the first week for animals on the diet including diatoms; 4) evidence for a physiological acclimatization during starvation; and 5) how the results contribute to the development of environmental transcriptomics in zooplankton research.

Lipid accumulation window hypothesis and development duration of the CV stage

The lipid -accumulation-window hypothesis makes two key predictions: 1) lipid fullness at the beginning of diapause will be similar across individuals; and 2) the duration of the CV stage will vary among individuals depending on an individual's rate of lipid accumulation. The average duration of the pre-diapause CV stage in *N. flemingeri* in this study exceeded one month, which is long compared with the earlier copepodite stages (CI to CIV, ≤ 15 days) (Liu & Hopcroft, 2006, Slater, 2004) and duration of the CV stage in direct developing *Calanus finmarchicus* (< 18 days) (Tarrant *et al.*, 2014). Nevertheless, a few individuals molted as early as 3 weeks after their molt

from the CIV to the CV stage, suggesting flexibility in the duration of the CV stage. A study by Coleman found that average lipid fullness in *N. flemingeri* adult females during early diapause was similar across three years (Coleman, 2022), as would be predicted by the lipid-accumulation-window hypothesis. However, while mean lipid fullness was similar, Coleman also found high individual variability, which is not consistent with the lipid accumulation hypothesis.

The developmental signal in the gene expression was the single most pronounced signal from post-molt to week 3. The developmental progression was clear in the t-SNE analysis that reduced the high-dimensional expression data into two dimensions, and in the large number of differentially expressed genes that were shared across the three fed treatments. In the future, this signal might be used to find transcripts that would serve as indicator sets to estimate progress through the CV stage in natural populations (Cieslak *et al.*, 2020, Lenz *et al.*, 2021). Furthermore, the changes in gene expression during the incubation period are consistent with an endogenous developmental program that places at least some constraints on lipid accumulation depending on food history. The picture that emerges is a hybrid hypothesis: while the copepod may be able to delay maturation and diapause to optimize lipid accumulation, the underlying developmental program constrains this flexibility and not all individuals are fully provisioned before entering diapause. Further studies are needed to understand the relationships between development, lipid accumulation and a variable food supply in *N. flemingeri* and other calanid copepods.

Lipid accumulation and fatty acid synthesis in fed animals

Biomass and lipid fullness indicators were highly variable among the fed CV individuals in our experiment. This was similar to the variability reported in *C. finmarchicus* mesocosm experiments (Hygum *et al.*, 2000). Increase in biomass measured as carbon and nitrogen across incubation time was significant during the CV stage, albeit, there was overlap among the three fed treatments. We also observed significant overlap in lipid accumulation among the fed treatments, despite the fact that only one such treatment included diatoms, which are high in lipids (Jónasdóttir, 2019). By week 5, only the individuals on one of the high-food diets, the one without diatoms, had accumulated significantly more lipid. Lipid accumulation is controlled by a large number of genes, and the physiology of fatty acid and wax ester synthesis in copepods is poorly understood (Boyen *et al.*, 2023). In calanids, a number of genes have been proposed as indicators of processes that contribute to active lipid storage (Lenz *et al.*, 2021, Lenz *et al.*, 2014, Tarrant *et al.*, 2016, Tarrant *et al.*, 2021). In the current experiment, as expected, the relative expression of these genes was typically higher in the fed individuals than in the starved or Wk0 ones. We observed few diet-specific differences among these genes. High expression of genes involved in lipid storage under

low food conditions (90 µg C per liter) was also reported for direct developing *C. finmarchicus*, under conditions that were comparable to ours (75 µg C per 0.7 liter flask per day) (Skottene *et al.*, 2020). Interestingly, in this species the beta-oxidation genes were up regulated under low food conditions compared with the high food treatment (200 µg C per liter), something we did not observe in *N. flemingeri* under either low- or even no-food conditions after one week of starvation. While many questions remain on the regulation of lipid storage in copepods, these two studies suggest that substantial lipid accumulation can occur over a range of food conditions.

Transient detoxification signal at Wk1

Diatoms are an important food source for copepods and being high in nutritious lipids (Jónasdóttir, 2019) are considered to be a primary food for lipid-rich copepods. Species in the genera *Calanus* and *Neocalanus* inhabiting high latitude environments synchronize growth, development and lipid accumulation to spring diatom blooms (Conover & Huntley, 1991, Kaartvedt, 2000, Mackas & Tsuda, 1999, Record *et al.*, 2018). The selection of three diatom species for the current experiment was based on their occurrence in the Gulf of Alaska. Diatoms contributed 1/3 of the total carbon in one of the two high-carbon diets. At Wk2, biomass, C and N of individuals in this treatment were higher than in the other two food treatments (low and high C). However, many diatom strains also produce toxic substances such as aldehydes and oxylipins, which are detrimental to copepod growth and reproduction (Ianora *et al.*, 1995, Ianora *et al.*, 2003, Miralto *et al.*, 1999, Paffenhöfer *et al.*, 2005). Furthermore, phytoplankton toxicity can vary depending on growth conditions (Van De Waal *et al.*, 2014). Copepods on the High C D+ (diatom) treatment were characterized by the high expression of genes involved in response to stimulus, including the up-regulation of an *aldehyde dehydrogenase*, *glutathione S-transferase* and *cytochrome P450* genes, which are typically up-regulated in copepods in response to toxic diatoms (Lauritano *et al.*, 2012). Here, this response was observed at Wk1, but not in weeks 2 and 3, a pattern that is consistent with a cellular stress response (CSR) followed by homeostasis (Kültz, 2005, Roncalli *et al.*, 2016). Notwithstanding this possible acclimatization, the individuals with diatoms in their diet did not accumulate as much lipid as did those fed the same amount of carbon without diatoms.

Physiological acclimatization during starvation

Survival under no-food conditions over the 5-week experiment was high and did not differ from the other three treatments. Interestingly, starved individuals were able to maintain both their biomass and most of their lipid fullness for the first two weeks, before falling to very low levels by

the end of the 5-week experiment (Figs. 2A and 1A respectively). Large numbers of genes were down-regulated in response to starvation. These included genes involved in digestion (as might be expected), metabolic processes (e.g. glycolysis, TCA cycle, fatty acid synthesis) and the CSR, an energetically costly process (Kültz, 2005, Kültz, 2020). Another common response of organisms to starvation is to increase access to stored resources in order to maintain homeostasis. This can include mobilization of stored fats, carbohydrates and/or proteins (Finn & Dice, 2006). In *N. flemingeri* CVs, potential store-mobilization genes involved in proteolysis and ubiquitination were up regulated. However, genes involved in fatty acid metabolism were down regulated, including those involved in storage-related beta oxidation. Conserving lipid stores is consistent with the life history of *N. flemingeri*, which depends on stored lipids to complete its life cycle and reproduce. In contrast, the up-regulation of beta-oxidation genes under low-food conditions was reported in *C. finmarchicus*, a species that depends on post-diapause feeding for reproduction (Niehoff *et al.*, 2002, Skottene *et al.*, 2020).

Environmental transcriptomics

A central question is how the results from this controlled study can be used to assess gene expression patterns in field-collected individuals to determine if environmental conditions are sufficient for the copepods to acquire the necessary lipid stores for diapause and post-diapause reproduction. Spatial and temporal variability in bloom dynamics means that individual copepods will experience a range of food conditions during this critical developmental phase (Mackas & Coyle, 2005). Large regional differences in relative gene expression have been correlated with temporal and spatial differences in food resources (Roncalli *et al.*, 2019, Roncalli *et al.*, 2022). A three-year comparison in gene expression patterns in CV individuals concluded that the copepods experienced significant food limitation during one of the years (2016) (Roncalli *et al.*, 2022). In early May, phytoplankton abundances in 2016 were an order of magnitude lower than in 2015 in the western region of Prince William Sound. When we compare the results from the no-food treatment in the current study with the gene expression patterns observed in 2016, we find many similarities. These include the low expression of genes annotated to digestive enzymes, metabolic processes such as the TCA cycle and glycolysis and lipid synthesis. For the TCA cycle and glycolysis pathway we found overlap in the specific genes that were differentially expressed in our experiment and the field. Of particular note, is the low expression of the beta-oxidation genes in both field and experimental individuals. This is consistent with the experimental results demonstrating the copepods' ability to maintain their biomass and lipid fullness even under complete starvation for at least one week. The experimental results strengthen the conclusion that in

2016 *N. flemingeri* CV in Prince William Sound were showing physiological acclimatization to low food conditions by conserving energy without up-regulating the beta-oxidation pathway.

Conclusions

1. The stage CV of *N. flemingeri* is characterized by its resistance to starvation and a duration (≥ 5 weeks) that exceeds that of the other copepodite stages (~ 2 weeks).
2. Under no-food conditions, CVs conserve energy by down-regulating the expression of genes involved metabolism and digestion.
3. While the duration of the pre-adult stage is flexible, the CV undergoes a developmental program that is likely to place constraints on the length of the lipid accumulation window.

Acknowledgements

We would like to thank Russell R. Hopcroft, Caitlin Smoot and Pete Shipton (U. Alaska Fairbanks) for assistance with zooplankton collections, sorting and species identification, Javier Cruz-Hernández (U. Hawai‘i at Mānoa) for assistance with the experiment and the imaging, Suzanne Strom and Kerri Fredrickson (U. Western Washington) and Michael Mahmood (Alutiiq Pride Marine Institute) for providing phytoplankton cultures, Ted Murphy (U. Hawai‘i at Mānoa) for assistance in building the incubation set-up, Lynn Hata (U. Hawai‘i at Mānoa) for administrative support, Russell R. Hopcroft and Megan Porter (U. Hawai‘i at Mānoa) for the loan of equipment, the SOEST S-Lab for analytical services and the Georgia Genomics and Bioinformatics Core for high-throughput sequencing. We would also like to thank two anonymous reviewers for their thoughtful comments on an earlier version of this manuscript. This is the University of Hawai‘i at Mānoa School of Ocean and Earth Science and Technology contribution number XXXXX.

Funding

The work was supported by the National Science Foundation (NSF) (grant numbers: OCE-1756767, PHL, DKH; OCE-1756859, R. R. Hopcroft). Additional support was provided by National Center for Genome Analysis Support (NCGAS) under National Science Foundation (grant numbers: DBI-1062432 and ABI-1062432 to Indiana University). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the National Center for Genome Analysis Support, or Indiana University.

Data availability

Short-sequence read data from this experiment are publicly available through the National Center Biotechnology Information data archives (BioProject PRJNA807352). The shotgun assembly used as a reference is available through BioProject PRJNA496596 (Genbank accession: GHLB00000000. Data tables that summarize gene expression, list all differentially expressed genes (DEGs) and provide a cross-reference between the Trinity-generated DN numbers and the individual NCBI accession numbers are available in BCO-DMO as datasets (https://www.bco-dmo.org/project/720280). The annotation file for the reference transcriptome is also available through BCO-DMO (https://www.bco-dmo.org/project/720280).

Author contributions

All listed authors have: 1) made a substantial contribution to the study in terms of conception, data acquisition, or analysis; 2) contributed substantially to drafting the manuscript; and 3) approved the final submitted manuscript. Specific contributions have been made as follows: PHL, DKH, VR conceived the study design; DKH, PHL, JN, LNB collected and identified specimens, designed and constructed the temperature-control apparatus, performed the experiment and completed the initial analysis; PHL, VR, LNB, AMC, MCC analyzing the data and interpreted the statistical results; PHL, VR, LNB wrote major sections of the manuscript. All authors were involved in reviewing and editing the manuscript.

Competing interests

The authors declare that there are no competing interests

References

Alexa, A. and Rahnenfuhrer, J. (2010) topGO: enrichment analysis for gene ontology. *R package version,* **2,** 2010.

Auel, H., Klages, M. and Werner, I. (2003) Respiration and lipid content of the Arctic copepod Calanus hyperboreus overwintering 1 m above the seafloor at 2,300 m water depth in the Fram Strait. *Mar Biol,* **143,** 275-282.

Baumgartner, M. F. and Tarrant, A. M. (2017) The physiology and ecology of diapause in marine copepods. *Annu Rev Mar Sci,* **9,** 387-411.

Boyen, J., Ribes-Navarro, A., Kabeya, N., Monroig, O., Rigaux, A., Fink, P., Habluetzel, P. I., Navarro, J. C. and De Troch, M. (2023) Functional characterization reveals a diverse array of metazoan fatty acid biosynthesis genes. *Mol Ecol,* **32,** 970-982.

Bray, N. L., Pimentel, H., Melsted, P. and Pachter, L. (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol,* **34,** 525-527.

Cieslak, M. C., Castelfranco, A. M., Roncalli, V., Lenz, P. H. and Hartline, D. K. (2020) t-Distributed Stochastic Neighbor Embedding (t-SNE): A tool for eco-physiological transcriptomic analysis. *Mar Genomics,* **51,** 100723.

Coleman, D. M. (2022) *Lipid Accumulation in Three Species of Neocalanus Copepod in the Northern Gulf of Alaska.* University of Alaska Fairbanks.

Conover, R. and Huntley, M. (1991) Copepods in ice-covered seas—distribution, adaptations to seasonally limited food, metabolism, growth patterns and life cycle strategies in polar seas. *J Mar Syst,* **2,** 1-41.

Conover, R. J. and Corner, E. D. S. (1968) Respiration and nitrogen excretion by some marine zooplankton in relation to their life cycles. *J Mar Biol Assoc UK,* **48,** 49-75.

Coyle, K. O. and Pinchuk, A. I. (2003) Annual cycle of zooplankton abundance, biomass and production on the northern Gulf of Alaska shelf, October 1997 through October 2000. *Fish Oceanogr,* **12,** 227-251.

Dagg, M., Strom, S. and Liu, H. (2009) High feeding rates on large particles by *Neocalanus flemingeri* and *N. plumchrus,* and consequences for phytoplankton community structure in the subarctic Pacific Ocean. *Deep Sea Research Part I: Oceanographic Research Papers,* **56,** 716-726.

Dahms, H. U. (1995) Dormancy in the Copepoda - an Overview. *Hydrobiologia,* **306,** 199-211.

Darnis, G., Robert, D., Pomerleau, C., Link, H., Archambault, P., Nelson, R. J., Geoffroy, M., Tremblay, J.-É., Lovejoy, C. and Ferguson, S. H. (2012) Current state and trends in Canadian Arctic marine ecosystems: II. Heterotrophic food web, pelagic-benthic coupling, and biodiversity. *Climatic Change,* **115,** 179-205.

Denlinger, D. L. (2002) Regulation of diapause. *Annu Rev Entomol,* **47,** 93-122.

Denlinger, D. L. and Armbruster, P. A. (2014) Mosquito diapause. *Annu Rev Entomol,* **59,** 73-93.

Desgraupes, B. and Desgraupes, M. B. (2018) Package 'clusterCrit'. R-Proj.

Dunn, J. C. (1974) Well-separated clusters and optimal fuzzy partitions. *J Cybern,* **4,** 95-104.

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise*KDD.* Vol. 96. pp. 226-231.

Finn, P. F. and Dice, J. F. (2006) Proteolytic and lipolytic responses to starvation. *Nutrition,* **22,** 830-844.

Galili, T., O'callaghan, A., Sidi, J. and Sievert, C. (2018) heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics,* **34,** 1600-1602.

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

Hahn, D. A. and Denlinger, D. L. (2011) Energetics of insect diapause. *Annu Rev Entomol,* **56,** 103-121.

Hahsler, M., Piekenbrock, M., Arya, S. and Mount, D. (2017) dbscan: Density Based Clustering of Applications with Noise (DBSCAN) and related algorithms. *R package version,* 1.0-0.

Hirche, H. J. (1983) Overwintering of *Calanus finmarchicus* and *Calanus helgolandicus*. *Mar Ecol Prog Ser,* **11,** 281-290.

Hirche, H. J. (1996) Diapause in the marine copepod, *Calanus finmarchicus* - A review. *Ophelia,* **44,** 129-143.

Hirche, H. J. (2013) Long-term experiments on lifespan, reproductive activity and timing of reproduction in the Arctic copepod *Calanus hyperboreus*. *Mar Biol,* **160,** 2469-2481.

Hygum, B., Rey, C., Hansen, B. W. and Tande, K. (2000) Importance of food quantity to structural growth rate and neutral lipid reserves accumulated in Calanus finmarchicus. *Mar Biol,* **136,** 1057-1073.

Ianora, A., Poulet, S. and Miralto, A. (1995) A comparative study of the inhibitory effect of diatoms on the reproductive biology of the copepod Temora stylifera. *Mar Biol,* **121,** 533-539.

Ianora, A., Poulet, S. A. and Miralto, A. (2003) The effects of diatoms on copepod reproduction: a review. *Phycologia,* **42,** 351-363.
Irigoien, X. (2004) Some ideas about the role of lipids in the life cycle of *Calanus finmarchicus*. *J Plankton Res,* **26,** 259-263.
Johnson, C. L., Leising, A. W., Runge, J. A., Head, E. J. H., Pepin, P., Plourde, S. and Durbin, E. G. (2008) Characteristics of *Calanus finmarchicus* dormancy patterns in the Northwest Atlantic. *ICES J Mar Sci,* **65,** 339-350.
Jónasdóttir, S. H. (2019) Fatty acid profiles and production in marine phytoplankton. *Mar Drugs,* **17,** 151.
Jónasdóttir, S. H., Visser, A. W., Richardson, K. and Heath, M. R. (2015) Seasonal copepod lipid pump promotes carbon sequestration in the deep North Atlantic. *Proc Natl Acad Sci USA,* **112,** 12122-12126.
Kaartvedt, S. (2000) Life history of Calanus finmarchicus in the Norwegian Sea in relation to planktivorous fish. *ICES J Mar Sci,* **57,** 1819-1824.
Kattner, G. and Hagen, W. (2009) Lipids in marine copepods: latitudinal characteristics and perspective to global warming. In: M. Kainz, M. T. Brett and M. T. Arts (eds) *Lipids in Aquatic Ecosystems.* Springer, New York, pp. 257-280.
Kattner, G., Hagen, W., Lee, R. G., Campbell, R., Deibel, D., Flak-Petersen, S., Graeve, M., Hansen, B. W., Hirche, H. J., Jonasdottir, S. and Madsen, M. L. (2007) Perspectives on marine zooplankton lipids. *Can J Fish Aquat Sci,* **64,** 1628-1639.
Kopylova, E., Noé, L. and Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics,* **28,** 3211-3217.
Krijthe, J. H. (2015) Rtsne: T-distributed stochastic neighbor embedding using Barnes-Hut implementation. *R package version 013, URL https://github com/jkrijthe/Rtsne*.
Kültz, D. (2005) Molecular and evolutionary basis of the cellular stress response. *Annu Rev Physiol,* **67,** 225-257.
Kültz, D. (2020) Evolution of cellular stress response mechanisms. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology,* **333,** 359-378.
Langmead, B., Trapnell, C., Pop, M. and Salzberg, S. L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol,* **10,** 1-10.
Lauritano, C., Carotenuto, Y., Miralto, A., Procaccini, G. and Ianora, A. (2012) Copepod population-specific response to a toxic diatom diet.
Lee, R. F., Hagen, W. and Kattner, G. (2006) Lipid storage in marine zooplankton. *Mar Ecol Prog Ser,* **307,** 273-306.
Lenz, P. H. and Roncalli, V. (2019) Diapause within the context of life-history strategies in calanid copepods (Calanoida: Crustacea). *Biol Bull,* **237,** 170-179.
Lenz, P. H., Roncalli, V., Cieslak, M. C., Tarrant, A. M., Castelfranco, A. M. and Hartline, D. K. (2021) Diapause vs. reproductive programs: transcriptional phenotypes in a keystone copepod. *Commun Biol,* **4,** 1-13.
Lenz, P. H., Roncalli, V., Hassett, R. P., Wu, L. S., Cieslak, M. C., Hartline, D. K. and Christie, A. E. (2014) *De novo* assembly of a transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda)--the dominant zooplankter of the North Atlantic Ocean. *PloS One,* **9,** e88589.
Liu, H., Dagg, M. J., Napp, J. M. and Sato, R. (2008) Mesozooplankton grazing in the coastal Gulf of Alaska: *Neocalanus* spp. vs. other mesozooplankton. *ICES J Mar Sci,* **65,** 351-360.
Liu, H. and Hopcroft, R. R. (2006) Growth and development of *Neocalanus flemingeri/plumchrus* in the northern Gulf of Alaska: validation of the artificial-cohort method in cold waters. *J Plankton Res,* **28,** 87-101.

Mackas, D. and Coyle, K. (2005) Shelf–offshore exchange processes, and their effects on mesozooplankton biomass and community composition patterns in the northeast Pacific. *Deep Sea Rs Part II Top Stud Oceanogr,* **52,** 707-725.

Mackas, D. L. and Tsuda, A. (1999) Mesozooplankton in the eastern and western subarctic Pacific: community structure, seasonal life histories, and interannual variability. *Prog Oceanogr,* **43,** 335-363.

Marcus, N. H. and Scheef, L. P. (2010) Photoperiodism in copepods. *Photoperiods–the biological calendar Oxford University Press, New York,* 193-217.

Miller, C. B. and Clemons, M. J. (1988) Revised life history analysis for large grazing copepods in the subarctic Pacific Ocean. *Prog Oceanogr,* **20,** 293-313.

Miller, C. B., Crain, J. A. and Morgan, C. A. (2000) Oil storage variability in *Calanus finmarchicus*. *ICES J Mar Sci,* **57,** 1786-1799.

Miller, C. B., Morgan, C. A., Prahl, F. G. and Sparrow, M. A. (1998) Storage lipids of the copepod *Calanus finmarchicus* from Georges Bank and the Gulf of Maine. *Limnol Oceanogr,* **43,** 488-497.

Miller, C. B. and Nielsen, R. D. (1988) Development and growth of large, calanid copepods in the ocean subarctic Pacific, May 1984. *Prog Oceanogr,* **20,** 275-292.

Miralto, A., Ianora, A., Poulet, S. A., Romano, G., Buttino, I. I. and Scala, S. (1999) Embryonic development in invertebrates Is arrested by inhibitory compounds in diatoms. *Mar Biotechnol (NY),* **1,** 401-402.

Monell, K. J., Roncalli, V., Hopcroft, R. R., Hartline, D. K. and Lenz, P. H. (2023) Post-Diapause DNA Replication during Oogenesis in a Capital-Breeding Copepod. *Integrative Organismal Biology,* **5,** obad020.

Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods,* **5,** 621-628.

Niehoff, B., Madsen, S., Hansen, B. and Nielsen, T. (2002) Reproductive cycles of three dominant *Calanus* species in Disko Bay, West Greenland. *Mar Biol,* **140,** 567-576.

Oliveros, J. C. (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. *http://bioinfogp cnb csic es/tools/venny/index html*.

Paffenhöfer, G. A., Ianora, A., Miralto, A., Turner, J. T., Kleppel, G. S., D'alcala, M. R., Casotti, R., Caldwell, G. S., Pohnert, G., Fontana, A., Muller-Navarra, D., Jonasdottir, S., Armbrust, V., Bamstedt, U., Ban, S., Bentley, M. G., Boersma, M., Bundy, M., Buttino, I., Calbet, A., Carlotti, F., Carotenuto, Y., D'ippolito, G., Frost, B., Guisande, C., Lampert, W., Lee, R. F., Mazza, S., Mazzocchi, M. G., Nejstgaard, J. C., Poulet, S. A., Romano, G., Smetacek, V., Uye, S., Wakeham, S., Watson, S. and Wichard, T. (2005) Colloquium on diatom-copepod interactions. *Mar Ecol Prog Ser,* **286,** 293-305.

Record, N. R., Ji, R., Maps, F., Varpe, Ø., Runge, J. A., Petrik, C. M. and Johns, D. (2018) Copepod diapause and the biogeography of the marine lipidscape. *J Biogeogr,* **45,** 2238-2251.

Robinson, M. D., Mccarthy, D. J. and Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* **26,** 139-140.

Roncalli, V., Cieslak, M. C., Germano, M., Hopcroft, R. R. and Lenz, P. H. (2019) Regional heterogeneity impacts gene expression in the sub-arctic zooplankter *Neocalanus flemingeri* in the northern Gulf of Alaska. *Commun Biol,* **2,** 1-13.

Roncalli, V., Niestroy, J., Cieslak, M. C., Castelfranco, A. M., Hopcroft, R. R. and Lenz, P. H. (2022) Physiological acclimatization in high-latitude zooplankton *Mol Ecol,* **31,** 1753-1765.

Roncalli, V., Turner, J. T., Kulis, D., Anderson, D. M. and Lenz, P. H. (2016) The effect of the toxic dinoflagellate *Alexandrium fundyense* on the fitness of the calanoid copepod *Calanus finmarchicus*. *Harmful Algae,* **51,** 56-66.

Saumweber, W. J. and Durbin, E. G. (2006) Estimating potential diapause duration in *Calanus finmarchicus*. *Deep Sea Rs Part II Top Stud Oceanogr,* **53,** 2597-2617.

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods,* **9,** 671.

Skinner, D. M. (1985) Molting and regeneration. In: D. E. Bliss and L. H. Mantel (eds) *The Biology of Crustacea Integument, Pigments, and Hormonal Processes.* Vol. 9. Academic Press, Inc., Orlando, Florida, pp. 43-146.

Skottene, E., Tarrant, A. M., Altin, D., Olsen, R. E., Choquet, M. and Kvile, K. Ø. (2020) Lipid metabolism in Calanus finmarchicus is sensitive to variations in predation risk and food availability. *Sci Rep,* **10,** 22322.

Slater, L. M. (2004) *Development, growth, and egg production of Centropages abdominalis and Neocalanus flemingeri from the eastern subarctic Pacific.* MS, University of Alaska Fairbanks.

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

Strom, S. L., Frederickson, K. A. and Bright, K. J. (2016) Spring phytoplankton in the eastern coastal Gulf of Alaska: Photosynthesis and production during high and low bloom years. *Deep Sea Rs Part II Top Stud Oceanogr,* **132,** 107-121.

Strom, S. L., Fredrickson, K. A. and Bright, K. J. (2019) Microzooplankton in the coastal Gulf of Alaska: Regional, seasonal and interannual variations. *Deep Sea Rs Part II Top Stud Oceanogr,* **165,** 192-202.

Strom, S. L., Olson, M. B., Macri, E. L. and Mordy, C. W. (2006) Cross-shelf gradients in phytoplankton community structure, nutrient utilization, and growth rate in the coastal Gulf of Alaska. *Mar Ecol Prog Ser,* **328,** 75-92.

Tamburello, K. R. (2005) *Interannual variations in the carbon to chlorophyll a ratios during the spring bloom in Prince William Sound, Alaska.*

Tarrant, A. M., Baumgartner, M. F., Hansen, B. H., Altin, D., Nordtug, T. and Olsen, A. J. (2014) Transcriptional profiling of reproductive development, lipid storage and molting throughout the last juvenile stage of the marine copepod *Calanus finmarchicus*. *Front Zool,* **11,** 1.

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

Tarrant, A. M., Eisner, L. B. and Kimmel, D. G. (2021) Lipid-related gene expression and sensitivity to starvation in Calanus glacialis in the eastern Bering Sea. *Mar Ecol Prog Ser,* **674,** 73-88.

Tsuda, A., Saito, H. and Kasai, H. (2001) Life history strategies of subarctic copepods *Neocalanus flemingeri* and *N. plumchrus*, especially concerning lipid. *Plankton Biology and Ecology,* **48,** 52-58.

Van De Waal, D. B., Smith, V. H., Declerck, S. A., Stam, E. C. and Elser, J. J. (2014) Stoichiometric regulation of phytoplankton toxins. *Ecology letters,* **17,** 736-742.

Van Der Maaten, L. and Hinton, G. (2008) Visualizing data using t-SNE. *J Mach Learn Res,* **9,** 2579-2605.

Waite, J. N. and Mueter, F. J. (2013) Spatial and temporal variability of chlorophyll-a concentrations in the coastal Gulf of Alaska, 1998–2011, using cloud-free reconstructions of SeaWiFS and MODIS-Aqua data. *Prog Oceanogr,* **116,** 179-192.

Tables

Table 1. Summary of the experimental design and individual measurements. Experimental treatments included a no food treatment (NF) and three fed treatments (75 [Low C] and two 225 [High C and High C D+] µgC/flask/day) with different combinations of food types (flagellates [Low C], flagellates + heterotrophs [High C], flagellates + heterotrophs + diatoms [High C D+]) in the amounts indicated in the table. Flasks were color coded by food treatment. The food species used in the experiment are summarized in Table 2. For each time point, flasks from each treatment were harvested and all individuals were assessed for survival, with sub-sets of individuals (numbers in parentheses) imaged under the microscope. A small number of individuals were processed for either carbon, nitrogen and dry weight (C,N,DW) or gene expression (RNA-Seq).

Treatment	Wk0	NF	Low C	High C	High C D+
Color code		Black	Blue	Green	Orange
Feeding rate (µgC/flask/day)					
Total		0	75	225	225
Flagellates		0	75	150	75
Heterotrophs		0	0	75	75
Diatoms		0	0	0	75
Recently molted	Imaging (6) C,N,DW (3) RNA (3)				
Incubation time (wks)					
1		Imaging (6) C,N,DW (3) RNA (3)	Imaging (6) C,N,DW (2) RNA (3)	Imaging (6) C,N,DW (2) RNA (3)	Imaging (6) C,N,DW (2) RNA (3)
2		Imaging (6) C,N (1) DW (2)	Imaging (6) C,N,DW (4) RNA (3)	Imaging (6) RNA (3)	Imaging (6) C,N,DW (3) RNA (3)
3		Imaging (6) C,N,DW (2)	Imaging (6) C,N,DW (3) RNA (3)	Imaging (6) RNA (3)	Imaging (6) C,N,DW (2) RNA (3)
4		Imaging (4)	Imaging (5) C,N,DW (2)	Imaging (5) C,N,DW (1)	Imaging (4)
5		Imaging (6) C,N,DW (2)	Imaging (4) C,N,DW (2)	Imaging (4)	Imaging (7) C,N,DW (2)

Table 2. Phytoplankton and heterotrophs used in the feeding experiment. Starter cultures were obtained from either the National Center for Marine Algae and Microbiota, Bigelow Laboratory for Ocean Sciences, or from Dr. Suzanne Strom's collection (see materials and methods). Phytoplankton per cell carbon content was estimated using cell size (Menden-Deuer and Lessard 2000).

Species	Strain	Taxonomic Class	pgC/cell
Flagellates			
Rhodomonas baltica	CCMP755	Cryptophyceae	45-52
Dunaliella tertiolecta		Chlorophyceae	49
Isochrysis galbana	CCMP1323	Prymnesiophyceae	10
Heterotrophs			
Oxyrrhis marina		Dinophyceae	470
Heterocapsa triquetra	CCMP448	Dinophyceae	650
Diatoms			
Dytilum brightwellii	CCMP357	Bacillariophyceae	600
Thalassiosira rotula	CCMP3096	Bacillariophyceae	200
Odontella aurita	CCMP595	Bacillariophyceae	220

Table 3. Summary of the global two-way ANOVA carried out on copepod biomass parameters incubated under four different food treatments and measured at five time points. Incubation time: weeks 1-5; df: degrees of freedom; SS: sum of squares. F-values and significance (p value) are also listed (significant effects in bold). Parameters tested: prosome length (mm), lipid fullness (%), dry weight (ln[µg]), carbon (ln[µg C]) and nitrogen (ln[µg N]).

	df	SS	F-value	p value
Prosome length				
Incubation time	4	1.25	2.86	**0.03**
Food treatment	3	0.426	1.30	0.28
Incubation:Food	12	1.23	0.94	0.51
Residuals	91	9.97		
Lipid fullness				
Incubation time	4	455	1.70	0.16
Food treatment	3	10756	53.5	**<0.0001**
Incubation:Food	12	2376	2.96	**0.002**
Residuals	91	6097		
Dry weight				
Incubation time	4	2.31	4.71	**0.008**
Food treatment	3	4.48	12.2	**0.0001**
Incubation:Food	7	3.11	3.63	**0.01**
Residuals	19	2.32		
Carbon				
Incubation time	4	2.80	3.11	**0.04**
Food treatment	3	12.5	18.6	**<0.0001**
Incubation:Food	7	4.49	2.85	**0.03**
Residuals	18	4.05		
Nitrogen				
Incubation time	4	1.92	4.43	**0.01**
Food treatment	3	6.80	20.9	**<0.0001**
Incubation:Food	7	2.92	3.84	**0.01**
Residuals	18	1.95		

Figure captions

Figure 1. *Neocalanus flemingeri* CV percent lipid fullness as a function of incubation time and food treatment based on measurements of the area of the lipid sac as a percent of the area of the prosome. Incubation times measured as weeks since the molt from the CIV into the CV stage. Panels: A) NF: no food, filled black circles; B) Low C: low food treatment, flagellates, filled blue squares; C) High C: high food, flagellates + heterotrophs, filled green triangles; and D) High C D+: high food, flagellates + heterotrophs + diatoms, filled orange diamonds. Sample sizes for each time point and treatment are listed in Table 1. Error bars are standard deviations. Average lipid fullness of recently molted individuals (Wk0; open circle, black) is shown in every panel for each food treatment.

Figure 2. *Neocalanus flemingeri* CV biomass measurements as a function of incubation time and food treatment. Incubation times measured as weeks since the molt from the CIV into the CV stage. Panels: A) dry weight (µg); B) Carbon content (µg); C) Nitrogen content (µg); D) Carbon to Nitrogen ratios for all individuals. Symbols for time zero and food treatments – Wk0: open circle black; NF: no food, filled black circles; Low C: low food treatment, flagellates, filled blue squares; High C: high food, flagellates + heterotrophs, filled green triangles; and High C D+: high food, flagellates + heterotrophs + diatoms, filled orange diamonds. Sample sizes for each time point and treatment are listed in Table 1. Error bars are standard deviations.

Figure 3. t-SNE analysis of gene-expression profiles in *Neocalanus flemingeri* CVs. The analysis included log-transformed relative expression (Log_2 [RPKM +1]) of all genes (n=51,743) (perplexity=10 and # of iterations=50,000). Treatments and time points are indicated by different symbols as shown in the inset in the graph. Dashed circles group clusters identified using DBSCAN (MinPts=3 and Eps= the value that maximized the Dunn index).

Figure 4. Heatmap of relative expression of genes that were regulated over time and were shared across food treatments. For each week, paired comparisons between WK0 and each food treatment were analyzed for shared DEGs. Relative expression for each DEG was computed as a z-score and clustered based on their similarity in the expression pattern (left cluster). Each column represents the average of three biological replicates and time points and treatments are indicated on top (timepoints: Wk0, Wk1, Wk2 and Wk3; treatments: NF=no food, blue=Low C; green=High C, orange=High C D+). Complete list of transcripts are found in Supp. Table S1.

Figure 5. Heatmaps of relative expression of a select group of genes annotated to the GO term response to stimulus (GO:0050896) in *Neocalanus flemingeri* that were differentially expressed. Expression is shown as z-scores of relative gene expression of DEGs with each column representing the average of three biological replicates. Time points and treatments are indicated on top (timepoints: Wk0, Wk1, Wk2 and Wk3; treatments: NF=no food, blue=Low C; green=High C, orange=High C D+). On the right, annotation for the DEGs as CYP450: cytochrome P450; detoxification genes including glutathione S-transferase (GST); glutamate dehydrogenase (GDH) and multidrug resistance-associated proteins (MXR); and NOS: nitric oxide synthase. Arrow: up-regulated genes in the High C D+ (orange) treatment at Wk1. Complete list of transcripts are found in Supp. Table S1.

Figure 6. Relative expression of nine genes involved in fatty acid accumulation in *Neocalanus flemingeri*. Relative gene expression (y-axis) is shown as $\log_2(\text{RPKM} + 1)$ (mean, n=3, error bars: standard deviations). Incubation time (x-axis) indicates time since molt from stage CIV to CV in weeks. A) Acyl-CoA synthetase; B) β-ketoacyl-ACP synthase; C) Delta(5) fatty acid desaturase; D) Delta(9) fatty acid desaturase; E) Elongation of very long chain fatty acids protein 4 (ELOV 4); F) Elongation of very long chain fatty acids protein 6 (ELOV 6); G) Fatty acid-binding protein homolog 5 (FABP 5); H) Fatty acid-binding protein homolog 6 (FABP 6); I) Long-chain fatty acid transport protein 4 (FATP 4). Symbol and color coding - open circle: time zero (Wk0); black circle: no food; blue square: low C; green triangles: high C; orange diamonds: high C D+. Complete list of transcripts are found in Supp. Table S1.

Figure 7. Relative gene expression of genes that were differentially expressed in the no food (starvation) treatment in *Neocalanus flemingeri* pre-adults. A) Venn diagram of DEGs between no food individuals (NF) and individuals on the three food treatments: blue=Low C, green=High C, orange: High C D+). Number of DEGs shared among the paired comparisons in the center in white letters. B) Heatmaps of DEGs associated with energy metabolism, cellular stress response and muscle contraction (left label). Labels on the left indicate general functional categories with more specific biological processes involved within each category labeled on the right. Expression is shown as z-scores of relative gene expression of DEGs with each column representing the average expression of three biological replicates. Time points and treatments are indicated on top (timepoints: Wk0, Wk1, Wk2 and Wk3; treatments: NF=no food, blue=Low C; green=High C, orange=High C D+). Complete list of transcripts are found in Supp. Table S3.

Figure 1

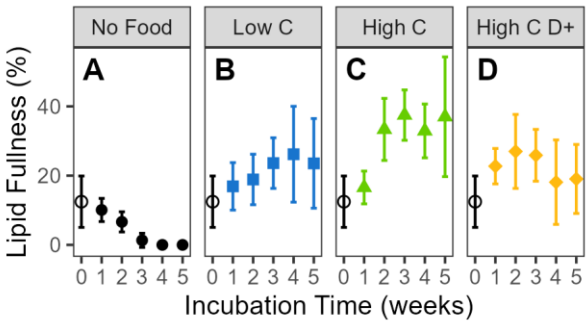

Figure 2

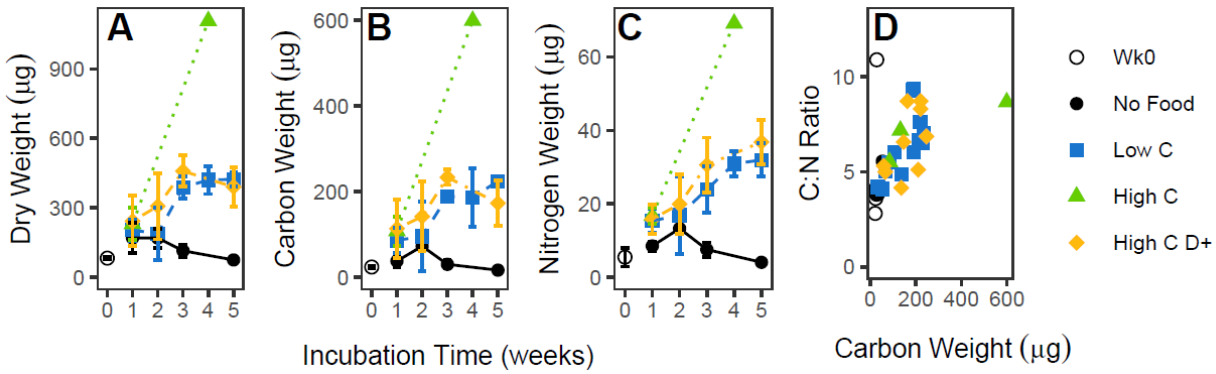

Figure 3

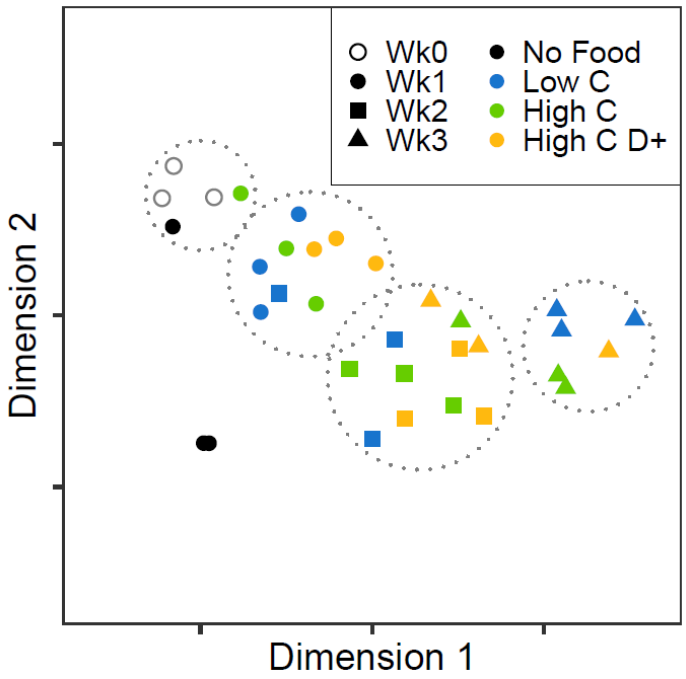

Figure 4

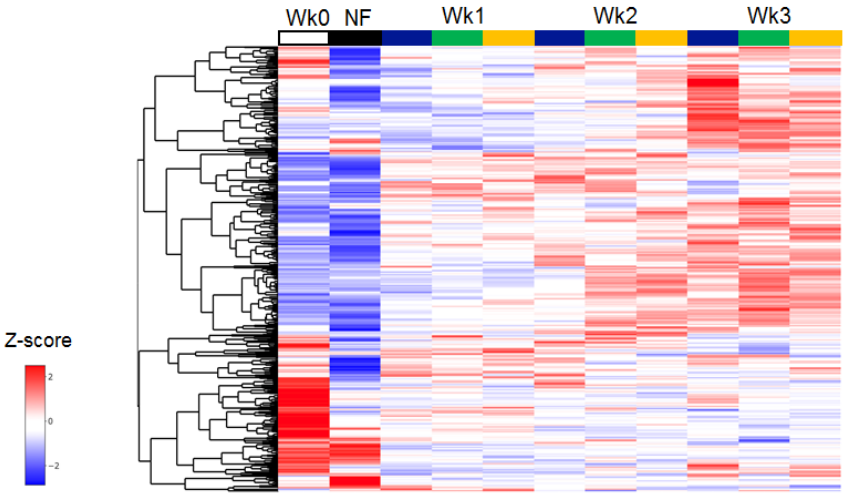

Figure 5

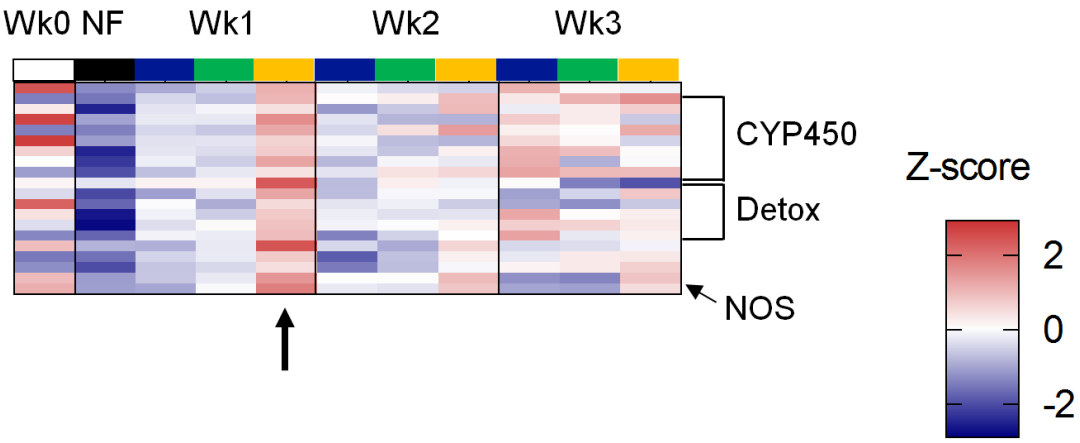

Figure 6

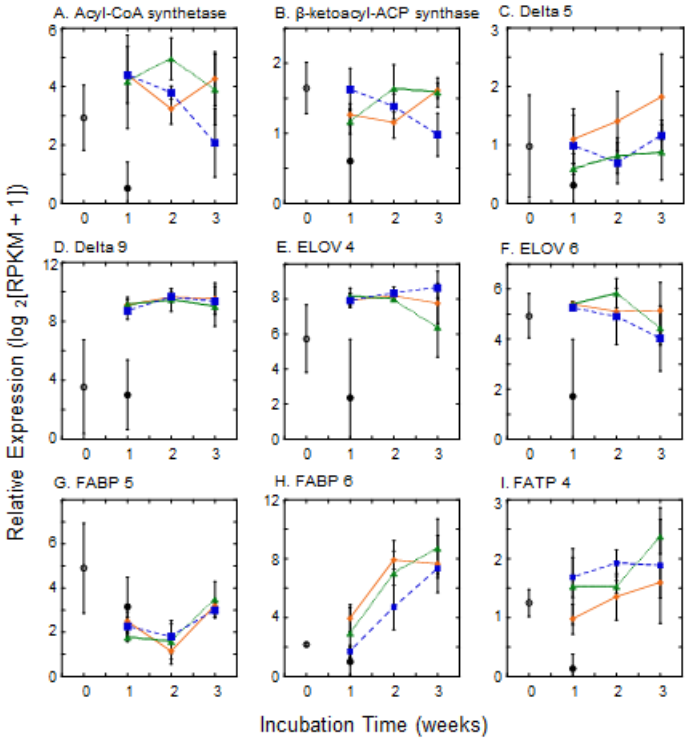

Figure 7