Gene expression profiles for Neocalanus flemingeri pre adults (CV) exposed to four different experimental food conditions collected from the M/V Dora in the Gulf of Alaska at station GAK1 from April 2019

Website: https://www.bco-dmo.org/dataset/914459
Data Type: experimental, Other Field Results
Version: 1
Version Date: 2024-05-30

Project
» Collaborative Research: Molecular profiling of the ecophysiology of dormancy induction in calanid copepods of the Northern Gulf of Alaska LTER site (Diapause preparation)

Contributors	Affiliation	Role
Lenz, Petra H.	University of Hawaiʻi at Mānoa (PBRC)	Principal Investigator, Contact
Hartline, Daniel K.	University of Hawaiʻi at Mānoa (PBRC)	Scientist
Niestroy, Jeanette L.	University of Hawaiʻi at Mānoa (PBRC)	Scientist
Roncalli, Vittoria	University of Hawaiʻi at Mānoa (PBRC)	Scientist
Block, Lauren N	University of Hawaiʻi at Mānoa (PBRC)	Student
Cieslak, Matthew C.	University of Hawaiʻi at Mānoa (PBRC)	Technician, Data Manager
Merchant, Lynne M.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This experimental dataset includes relative expression of individual Neocalanus flemingeri stage CV individuals incubated for different lengths of time and four different food treatments. The experimental protocol and results are described in detail in Roncalli et al., 2023. Briefly, field-collected N. flemingeri were allowed to molt into stage CV and then sorted into four different treatments: no food, low carbon, high carbon and high carbon with diatoms. After a one-week incubation, individuals from all four treatments were processed individually for RNA-Seq. In addition, following two and three-week incubations, copepods from the three fed treatments were processed individually for RNA-Seq. Short-sequence reads were mapped against a reference transcriptome and normalized gene expression was computed for each transcript. The dataset includes log-transformed relative gene expression in reads per kilobase per million reads (RPKM) (log2[RPKM+1]). The dataset also includes a list of differentially expressed genes and a look-up table that cross-references the hierarchical identifications of transcripts generated by the Trinity assembly software and the corresponding National Center for Biotechnology Information (NCBI) accession number. These data are further described in the following publications: Roncalli, et al. (2023) (DOI: 10.1093/plankt/fbad045) and Roncalli, et al. (2019) (DOI: 10.1038/s42003-019-0565-5)

Table of Contents

- [Deployments]
- [Project Information]
- [Funding]

Coverage

Location: Gulf of Alaska
Spatial Extent: **Lat**:59.8443 **Lon**:-149.4838
Temporal Extent: 2019-04-15

Dataset Description

These data are further described in the following publications:

Roncalli, V., Block, L. N., Niestroy, J. L., Cieslak, M. C., Castelfranco, A. M., Hartline, D. K., & Lenz, P. H. (2023). Experimental analysis of development, lipid accumulation and gene expression in a high-latitude marine copepod. *Journal of Plankton Research*, 45(6), 885–898. https://doi.org/10.1093/plankt/fbad045

Roncalli, V., Cieslak, M. C., Germano, M., Hopcroft, R. R., & Lenz, P. H. (2019). Regional heterogeneity impacts gene expression in the subarctic zooplankter Neocalanus flemingeri in the northern Gulf of Alaska. *Commun Biology*, 2(1). https://doi.org/10.1038/s42003-019-0565-5

Methods & Sampling

Zooplankton were collected on a day-trip to station GAK1 (59º50.7′ N, Long: 149º28′ W, depth 264 m, Gulf of Alaska) (http://research.cfos.uaf.edu/gak1/) aboard the M/V Dora on April 15, 2019. Collections were made using QuadNet with two 150 µm and two 53 µm mesh nets towed vertically from 100 to 0 m. Collection details are provided in Roncalli et al. (2023). Zooplankton samples were diluted, brought back to the laboratory and sorted under a dissection microscope to select stage CIV Neocalanus flemingeri individuals. As individuals molted into CVs, they were removed from the holding containers and transferred into 750 ml Falcon flasks with 3 individuals per flask and assigned to one of the 4 food treatments, as described in detail in Roncalli et al. (2023). Three individuals were preserved upon molting (Wk0). Individuals were harvested at 3 incubation times (Wk1, Wk2 and Wk3) and preserved in RNALater Stabilization Reagent. Preserved copepods were frozen first in -40ºC during the experiment, and then transferred to −80°C until further processing.

Total RNA extraction, library construction, RNA sequencing and quality control: total RNA was extracted from individuals using QIAGEN RNeasy Plus Mini Kit (catalog # 74134) in combination with a Qiashredder column (catalog # 79654). Sequencing was performed on 3 Wk0 individuals and 3 replicate individuals for each time x treatment combination. Total RNA was shipped on dry ice to the Georgia Genomics Bioinformatics Core (https://dna.uga.edu) for RNA-Seq. There, double-stranded cDNA libraries (KAPA Stranded mRNA-Seq Kit, with KAPA mRNA Capture Beads (cat #KK8421]) from each individual were multiplexed and sequenced using an Illumina Next-Seq 500 instrument (High-Output Flow Cell, 75 bp, paired end). Quality of each RNA-Seq library was reviewed with the FastQC software28. From each RNA-Seq library, low quality reads were removed using FASTQ Toolkit (v. 2.2.5 within BaseSpace). Illumina adaptors, reads <50 bp long, reads with an average Phred score <30 and the first 12 bp from each read, were removed from each library.

Data Processing Description

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 log2 transformation of the relative expression data (Log2[RPKM+1]).

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) with p-values were adjusted for false discovery rate (FDR) using the Benjamini-Hochberg correction (default algorithm weight01) (Robinson et al., 2010).

BCO-DMO Processing Description

Steps for processing the main dataset file, the differentially expressed genes file, and the Seq ID and GenBank accession cross reference table.

1. Loaded submitted files into the BCO-DMO laminar processor. Submitted files loaded are 2024-Jan-sra_result-Seward2019-Expt.xlsx, File1-GeneExpression-Log2(RPKM+1).csv, File2-DEGs-GLM Analysis.csv, and File3-CrossReference-Trinity Genbank.csv.
2. Renamed parameters in the cross reference file from Trinity_ID to seq_id to match the parameter names in the other files and Genbank_Accession_number to Genbank_accession
3. Added the version number 1, suffix '.1', to the Genbank_accession values of File3-CrossReference-Trinity Genbank.csv because the NCBI GenBank accession numbers should contain a version number.
4. Joined the cross reference file with the differentially expressed genes file, File3-CrossReference-Trinity Genbank.csv, on the column seq_id to add the corresponding GenBank accession numbers to the file.
5. Joined the cross reference file with the relative gene expression file, File1-GeneExpression-Log2(RPKM+1).csv, on the column seq_id to add the corresponding GenBank accession numbers to the file.
6. Joined the submitted metadata table to the relative gene expression file, File1-GeneExpression-Log2(RPKM+1).csv, to add the metadata to the file.
7. Added a date field of the format %Y-%m-%d created from the day, month, and year values.
8. Reordered the columns to move the metadata columns to the front of the relative gene expression file
9. Renamed parameters in the relative gene expression file to follow the BCO-DMO naming protocol. Renamed column headers that have a period or space in their name to an underscore. Removed '(m)' from the Depth range parameter name since units will be indicated in the parameters section of the dataset page.

—---

Steps to create an unpivoted version of the submitted relative gene expression file and a metadata table

1. Load in submitted files and a data manager metadata file into the BCO-DMO laminar processor.
2. First loaded in the submitted metadata file 2024-Jan-sra_result-Seward2019-Expt.xlsx and the data manager metadata file dm_replicate_experiments_metadata.csv into laminar.

The submitted metadata file has the columns: Experiment Accession, Experiment Title, Organism Name, Year, Month, Day, Station, Latitude, Longitude, Depth range (m), Study Accession, Study Title, Sample Accession, Replicate. The data manager created metadata file has the columns: Replicate, week_after_molting_to_CV, feeding_protocol, BioProject, BioSample
3. Joined the submitted metadata table and the data manager created metadata table on the Replicate field into a new metadata table named metadata_table_with_ncbi_accessions.
4. Renamed the column headers in the new metadata table according to BCO-DMO naming protocols. Replaced spaces with underscores and removed the text "(m)" from the Depth range parameter name since this unit will be included in the units section of parameter definitions on the dataset page.
5. Added a collection date column of the format %Y-%m-%d from the year, month, and day columns.
6. Loaded into laminar the submitted cross reference file named File3-CrossReference-Trinity Genbank.csv.
7. Added the version number text '.1' to the parameter 'Genbank Accession Number' in the lookup table.
8. Loaded in the submitted relative gene expression file named "File1-GeneExpression-Log2(RPKM+1).csv".
9. Applied the laminar process 'unpivot' to the relative gene expression file.
10. Unpivoted on the column names which are of the form T0.1, T0.2, NF.2, GW1.3.
11. Named the unpivoted table "unpivoted_relative_gene_expression" to later save as a csv file.
12. In the unpivoted file, renamed Genbank_Accession_number to Genbank_accession.
13. Joined the metadata table "metadata_table_with_ncbi_accessions" with the unpivoted table "unpivoted_relative_gene_expression" on the column "Replicate" to add metadata to the unpivoted table.
14. Joined the file "File3-CrossReference-Trinity Genbank.csv" with the unpivoted file "unpivoted_relative_gene_expression" on the column "Trinity_ID" in the cross-reference file and "seq_id" in the unpivoted file.
15. Because the main dataset will have the Replicate column names in the metadata table "metadata_table_with_ncbi_accessions" were renamed from T0.1 to T0_1, etc., the replicate columns in the joined table were renamed in the same pattern by replacing the period with an underscore so that the final

metadata table will match the run_id values in the main dataset file. The same renaming was done for the joined table “unpivoted_relative_gene_expression”.
16. Genbank_Accession_number was renamed to match the pattern of the other accession parameter names.
17. The parameter fields in the unpivoted table and metadata table were reordered to group the accession parameters at the end of the tables.
18. Removed the NCBI accession numbers and titles except for the GenBank accession numbers from the unpivoted file to reduce the file size.

[table of contents | back to top]

Data Files

File	
914459_v1_relative_gene_expression.csv Primary data file for dataset ID 914459, version 1 Relative gene expression given in log2(RPKM +1) for all transcripts based on RNA-Seq data mapped against a reference transcriptome.	(Comma Separated Values (.csv), 13.06 MB) MD5:05703749ca48ac3358891cb931c99249

[table of contents | back to top]

Supplemental Files

File
Differentially expressed genes filename: differentially_expressed_genes.csv (Comma Separated Values (.csv), 477.77 KB) MD5:c142ca72eca93cbed14437cf5a795cd4 List of differentially expressed genes (DEGs) calculated using EdgeR GLM analysis. Columns: seq_id, Genbank_accession See the parameters section of this dataset for definitions of seq_id and Genbank_accession parameters used in this file. Blank values in Genbank_accession represent no corresponding value for seq_id
Metadata table including NCBI accessions filename: supplemental_files/metadata_table_with_ncbi_accessions.csv (Comma Separated Values (.csv), 10.55 KB) MD5:9f972fafe716cdaf58d5581fac488dd9 See the supplemental file "Metadata table parameter definitions", metadata_table_parameter_definitions.csv, for definitions of parameters included in this metadata table. Columns: Replicate, Organism_Name, Station, Latitude, Longitude, Collection_date, Year, Month, Day, Depth_range, week_after_molting_to_CV, feeding_protocol, BioProject, Study_Accession, Study_Title, Experiment_Accession, Experiment_Title, BioSample, Sample_Accession
Metadata table parameter definitions filename: metadata_table_parameter_definitions.csv (Comma Separated Values (.csv), 1.54 KB) MD5:14b21267df32633e4585862b913cf479 This file contains parameter definitions for the supplemental file "Metadata table including NCBI accessions", metadata_table_with_ncbi_accessions.csv Columns: suppliedName, description, suppliedUnits, no_data_value, bcodmo_standard_parameter_name, Datatype, Format suppliedName = Supplied Name description = parameter description suppliedUnits = parameter units no_data_value = fill value for no data bcodmo_standard_parameter_name = BCO-DMO standard parameter name DataType = parameter data type Format = date format if the parameter data type is date
Seq ID and GenBank accession numbers cross-reference table filename: seq_id_genbank_accession_cross_reference.csv (Comma Separated Values (.csv), 1.96 MB) MD5:65425ff3133194f7defd2090c4789ef0 Cross-reference between Trinity IDs and Genbank accession numbers Columns: seq_id, Genbank_accession See the parameters section of this dataset for definitions of seq_id and Genbank_accession parameters used in this file.
Species WoRMS taxonomy filename: species_list.csv (Comma Separated Values (.csv), 211 bytes) MD5:0996e3c2eab4418bf250f26afec882aa Species WoRMS taxonomy table with columns: ScientificName, AphiaID, LSID, Authority, Class, Order, Family, Genus, Species
Unpivoted version of the relative gene expression dataset filename: supplemental_files/unpivoted_relative_gene_expression.csv (Comma Separated Values (.csv), 230.86 MB) MD5:1fb45c6c56d0b9edafb19491614aa68d See the supplemental file "Metadata table parameter definitions", metadata_table_parameter_definitions.csv, for definitions of parameters included in this file that are not listed as parameters on this dataset page. Columns: seq_id, Genbank_Accession, Organism_Name, Station, Latitude, Longitude, Collection_date, Year, Month, Day, Depth_range, Replicate, relative_expression, week_after_molting_to_CV, feeding_protocol

[table of contents | back to top]

Related Publications

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
Software

BaseSpace Labs. (n.d.). FASTQ Toolkit (Version 2.2.5) [Computer software]. Illumina. https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/fastq-toolkit.html

Software

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology, 34(5), 525–527. https://doi.org/10.1038/nbt.3519
Methods

FastQC (2015), FastQC [Online]. Available online at: https://qubeshub.org/resources/fastqc.
Software

Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics, 28(24), 3211–3217. https://doi.org/10.1093/bioinformatics/bts611
Software

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology, 10(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25
Methods

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature Methods, 5(7), 621–628. https://doi.org/10.1038/nmeth.1226
Methods

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology, 11(3), R25. https://doi.org/10.1186/gb-2010-11-3-r25
Methods

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616
Methods

Roncalli, V., Block, L. N., Niestroy, J. L., Cieslak, M. C., Castelfranco, A. M., Hartline, D. K., & Lenz, P. H. (2023). Experimental analysis of development, lipid accumulation and gene expression in a high-latitude marine copepod. Journal of Plankton Research, 45(6), 885–898. https://doi.org/10.1093/plankt/fbad045
Methods

Roncalli, V., Cieslak, M. C., Germano, M., Hopcroft, R. R., & Lenz, P. H. (2019). Regional heterogeneity impacts gene expression in the subarctic zooplankter Neocalanus flemingeri in the northern Gulf of Alaska. Communications Biology, 2(1). https://doi.org/10.1038/s42003-019-0565-5
Results

[table of contents | back to top]

Related Datasets

IsRelatedTo

Hartline, D. K., Lenz, P. H., Cieslak, M. C. (2024) **Annotated de novo transcriptomes generated from six co-occurring species of calanoid copepods from the R/V Tiglax TXF18, TXS19, TXF15, TXF17 in the Gulf of Alaska from 2015-2019.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2024-07-02 doi:10.26008/1912/bco-dmo.908689.1 [view at BCO-DMO]

Lenz, P. H., Cieslak, M. C., Roncalli, V., Hartline, D. K. (2025) **Molecular identification of genetic variants of Neocalanus flemingeri in the Gulf of Alaska from samples collected from 2015 to 2023.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2025-02-21 doi:10.26008/1912/bco-dmo.954181.1 [view at BCO-DMO]
Relationship Description: Related datasets to raw sequence accessions at the National Center for Biotechnology Information.

Lenz, P. H., Roncalli, V., Cieslak, M. C. (2024) **Multiyear RNA-Seq of Neocalanus flemingeri stages CV and Adult Female from the R/V Tiglax and R/V Sikuliaq in the Northern Gulf of Alaska from 2015-2022.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2024-07-26 doi:10.26008/1912/bco-dmo.922330.1 [view at BCO-DMO]
Relationship Description: Related datasets to raw sequence accessions at the National Center for Biotechnology Information.

University of Hawaii at Manoa (2018). Neocalanus flemingeri, Neocalanus flemingeri pre adult (CV). 2018/10.

NCBI:BioProject: PRJNA496596 [Internet]. Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information; Available from: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA496596.

University of Hawaii at Manoa (2022). Neocalanus flemingeri, Response to food availability in pre-adult Neocalanus flemingeri. 2022/02. NCBI:BioProject: PRJNA807352.[Internet]. Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information; Available from: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA807352.

[table of contents | back to top]

Parameters

Parameter	Description	Units
seq_id	Sequence identification using Trinity identification of assembled transcripts	unitless
Genbank_accession	NCBI GenBank acession number	unitless
Organism_Name	Species analyzed	unitless
Station	Station	unitless
Latitude	Latitude. Locations south of equator are negative.	decimal degrees
Longitude	Longitude. Locations west of prime meridian are negative.	decimal degrees
Collection_date	Collection date	unitless
Year	Collection year	unitless
Month	Collection month	unitless
Day	Collection day	unitless
Depth_range	Collection depth range	meters (m)
T0_1	Relative gene expression for replicate T0_1. Food protocol:: No food, Week after molting to CV: 0	Log2[RPKM+1]
T0_2	Relative gene expression for replicate T0_2. Food protocol:: No food, Week after molting to CV: 0	Log2[RPKM+1]

T0_3	Relative gene expression for replicate T0_3. Food protocol:: No food, Week after molting to CV: 0	Log2[RPKM+1]
NF_1	Relative gene expression for replicate NF_1. Food protocol:: No food, Week after molting to CV: 1	Log2[RPKM+1]
NF_2	Relative gene expression for replicate NF_2. Food protocol:: No food, Week after molting to CV: 1	Log2[RPKM+1]
NF_3	Relative gene expression for replicate NF_3. Food protocol:: No food, Week after molting to CV: 1	Log2[RPKM+1]
BW1_1	Relative gene expression for replicate BW1_1. Food protocol:: Low Carbon diet, Week after molting to CV: 1	Log2[RPKM+1]
BW1_2	Relative gene expression for replicate BW1_2. Food protocol:: Low Carbon diet, Week after molting to CV: 1	Log2[RPKM+1]
BW1_3	Relative gene expression for replicate BW1_3. Food protocol:: Low Carbon diet, Week after molting to CV: 1	Log2[RPKM+1]
GW1_1	Relative gene expression for replicate GW1_1. Food protocol::High Carbon diet, Week after molting to CV: 1	Log2[RPKM+1]
GW1_2	Relative gene expression for replicate GW1_2. Food protocol::High Carbon diet, Week after molting to CV: 1	Log2[RPKM+1]
GW1_3	Relative gene expression for replicate GW1_3. Food protocol::High Carbon diet, Week after molting to CV: 1	Log2[RPKM+1]
YW1_1	Relative gene expression for replicate YW1_1. Food protocol::High Carbon diet + diatom, Week after molting to CV: 1	Log2[RPKM+1]
YW1_2	Relative gene expression for replicate YW1_2. Food protocol::High Carbon diet + diatom, Week after molting to CV: 1	Log2[RPKM+1]
YW1_3	Relative gene expression for replicate YW1_3. Food protocol::High Carbon diet + diatom, Week after molting to CV: 1	Log2[RPKM+1]
BW2_1	Relative gene expression for replicate BW2_1. Food protocol::Low Carbon diet, Week after molting to CV: 2	Log2[RPKM+1]

BW2_2	Relative gene expression for replicate BW2_2. Food protocol::Low Carbon diet, Week after molting to CV: 2	Log2[RPKM+1]
BW2_3	Relative gene expression for replicate BW2_3. Food protocol::Low Carbon diet, Week after molting to CV: 2	Log2[RPKM+1]
GW2_1	Relative gene expression for replicate GW2_1. Food protocol::High Carbon diet, Week after molting to CV: 2	Log2[RPKM+1]
GW2_2	Relative gene expression for replicate GW2_2. Food protocol::High Carbon diet, Week after molting to CV: 2	Log2[RPKM+1]
GW2_3	Relative gene expression for replicate GW2_3. Food protocol::High Carbon diet, Week after molting to CV: 2	Log2[RPKM+1]
YW2_1	Relative gene expression for replicate YW2_1. Food protocol::High Carbon diet + diatom, Week after molting to CV: 2	Log2[RPKM+1]
YW2_2	Relative gene expression for replicate YW2_2. Food protocol::High Carbon diet + diatom, Week after molting to CV: 2	Log2[RPKM+1]
YW2_3	Relative gene expression for replicate YW2_3. Food protocol::High Carbon diet + diatom, Week after molting to CV: 2	Log2[RPKM+1]
BW3_1	Relative gene expression for replicate BW3_1. Food protocol: Low Carbon diet, Week after molting to CV: 3	Log2[RPKM+1]
BW3_2	Relative gene expression for replicate BW3_2. Food protocol: Low Carbon diet, Week after molting to CV: 3	Log2[RPKM+1]
BW3_3	Relative gene expression for replicate BW3_3. Food protocol: Low Carbon diet, Week after molting to CV: 3	Log2[RPKM+1]
GW3_1	Relative gene expression for replicate GW3_1. Food protocol: High Carbon diet, Week after molting to CV: 3	Log2[RPKM+1]
GW3_2	Relative gene expression for replicate GW3_2. Food protocol: High Carbon diet, Week after molting to CV: 3	Log2[RPKM+1]
GW3_3	Relative gene expression for replicate GW3_3. Food protocol: High Carbon diet, Week after molting to CV: 3	Log2[RPKM+1]
YW3_1	Relative gene expression for replicate YW3_1. Food protocol: High Carbon diet + diatom, Week after molting to CV: 3	Log2[RPKM+1]

YW3_2	Relative gene expression for replicate YW3_2. Food protocol: High Carbon diet + diatom, Week after molting to CV: 3	Log2[RPKM+1]
YW3_3	Relative gene expression for replicate YW3_3. Food protocol: High Carbon diet + diatom, Week after molting to CV: 3	Log2[RPKM+1]

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina Next-Seq 500
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	Desktop sequencer
Generic Instrument Description	A DNA sequencer is an instrument that determines the order of deoxynucleotides in deoxyribonucleic acid sequences.

Dataset-specific Instrument Name	Dissection microscope
Generic Instrument Name	Microscope - Optical
Generic Instrument Description	Instruments that generate enlarged images of samples using the phenomena of reflection and absorption of visible light. Includes conventional and inverted instruments. Also called a "light microscope".

Dataset-specific Instrument Name	QuadNet
Generic Instrument Name	Plankton Net
Dataset-specific Description	Two 150 µm and two 53 µm mesh nets
Generic Instrument Description	A Plankton Net is a generic term for a sampling net that is used to collect plankton. It is used only when detailed instrument documentation is not available.

[table of contents | back to top]

Deployments

Lenz_Gulf_of_Alaska_2019-04-15

Website	https://www.bco-dmo.org/deployment/923842
Platform	M/V Dora
Start Date	2019-04-15
End Date	2019-04-15
Description	location: station GAK1 (latitude: 59º50.7′ N, longitude: 149º28′ W)

[table of contents | back to top]

Project Information

Collaborative Research: Molecular profiling of the ecophysiology of dormancy induction in calanid copepods of the Northern Gulf of Alaska LTER site (Diapause preparation)

Coverage: Northern Gulf of Alaska LTER

NSF Award Abstract:
The sub-arctic Pacific sustains major fisheries with nearly all commercially important species depending either directly or indirectly on lipid-rich copepods (Neocalanus flemingeri, Neocalanus plumchrus, Neocalanus cristatus and Calanus marshallae). In turn, these species depend on a short-lived spring algal bloom for growth and the accumulation of lipid stores in order to complete an annual life cycle that includes a period of dormancy. The intellectual thrust of this project measures how the timing and magnitude of algal blooms affect preparation for dormancy using a combination of field and experimental observations. The Northern Gulf of Alaska - with four calanid species that experience dormancy, steep environmental gradients, well-described phytoplankton bloom dynamics, and a concurrent NSF-LTER program - provides an unusual opportunity to identify the factors that affect dormancy preparation. Education and outreach plans are integrated with the research. Educational efforts focus on interdisciplinary opportunities for undergraduate, graduate and post-doctoral trainees. The project will generate content for existing graduate and undergraduate courses. U. of Alaska Fairbanks and U. Hawaii at Manoa are Alaska Native and Native Hawaiian Serving Institutions, and students from these groups will be recruited to participate in the project. Because fishing is a major industry in the Gulf of Alaska, outreach will communicate the role copepods play in marine ecosystems using the concept of a dynamic food web tied to production cycles.

Diapause (dormancy) and the accompanying accumulation of lipids in copepods have been identified as key drivers in high latitude ecosystems that support economically important fisheries, including those of the Gulf of Alaska. While the disappearance of lipid-rich copepods has been linked to severe declines in fish stocks, little is known about the environmental conditions that are required for the successful completion of the copepod's life cycle. A physiological profiling approach that measures relative gene expression will be used to test two alternative hypotheses: the lipid accumulation window hypothesis, which holds that individuals enter diapause only after they have accumulated sufficient lipid stores, and the developmental program hypothesis, which holds that once the diapause program is activated, progression occurs independent of lipid accumulation. The specific objectives are: 1) determine the effect of food levels during N. flemingeri copepodite stages on progression towards diapause using multiple physiological and developmental markers; 2) characterize the seasonal changes in the physiological profile of N. flemingeri across environmental gradients and across years; 3) compare physiological profiles across co-occurring calanid species (N. flemingeri, Neocalanus plumchrus, Neocalanus cristatus and Calanus marshallae); and 4) estimate the reproductive potential of the overwintering populations of N. flemingeri. The broader scientific significance includes the acquisition of new genomic data and molecular resources that will be made publicly available through established data repositories, and the development of new tools for routinely obtaining physiological profiles of copepods.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

NOTE: Petra Lenz is a former Principal Investigator (PI) and Andrew Christie is a former Co-Principal Investigator (Co-PI) on this project (award #1756767). Daniel Hartline is the PI listed for the award #1756767 and is now a former Co-PI on this project.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1756767
NSF Division of Ocean Sciences (NSF OCE)	OCE-1756859

[table of contents | back to top]