DATA MANAGEMENT PLAN

PIs: Petra Lenz, Andrew Christie and Daniel Hartline (U. Hawaii at Manoa), Russell Hopcroft (U. Alaska Fairbanks)

The P.I.s agree to support the NSF data management and data dissemination policies as described in the NSF Award and Administration Guide (AAG, Chapter VI.D.4) and to comply with the NSF Division of Ocean Sciences Sample and Data Policy. Outlined below are the types of data and samples that will be collected and how they will be managed.

BCO-DMO
A project page will be maintained at BCO-DMO. This project page will provide links to the oceanographic field and molecular data, and micrographs. It will also be used to provide links to publications describing experiments, and provide additional information that might be too detailed for the methods descriptions in a journal article, yet may be of interest to others doing similar research. The goal is to use the BCO-DMO page as a central repository for original data and to provides links to all data associated with this project that have been deposited into public databases like NCBI, Dryad Data Repository.

Oceanographic field data
Fieldwork for this project will take place along the Seward Line. Samples will be collected in collaboration with the existing 20-year Seward Line monitoring program. These programs already include standard sampling protocols including station locations, instrument deployments, and sample collection and processing. The proposed project will have access to these data through data management portals such as the Alaska Ocean Observing System (AOOS; http://www.aoos.org), which participates in the integrated Ocean Observing System (IOOS, NOAA sponsored site) and Earth Observing Laboratory (https://www.eol.ucar.edu), and archives current data from monitoring along the Seward Line. The LTER network also has extensive data management for core measurements, and will include links to any datasets associated with the LTER activities. Summary data graphs are also made available through the
Seward Line website (www.sfos.uaf.edu).

In addition, during years 1 and 2, plankton collections will be prepared for molecular analysis through preservation in RNAlater and maintained live for controlled experiments to be completed at UAF. Additional data will be generated from these samples (survival, growth, RNAseq, gene sequences), which will be archived in appropriate data repositories (see below). These different datasets will be cross-linked by adding information on web addresses, citations, and providing details in "data comment" sections.

During years 2 and 3, additional field collections are planned. Field data from these trips will be archived through the same data management portals (AOOS, EOL, LTER) and cross-linked to the existing sampling programs, and to other data that will be generated as part of the current proposal.

Molecular Data
The proposed effort will generate large volumes of high throughput DNA and RNA sequencing

data. Raw sequencing data (RNASeq), assembled data, and sequences produced through Sanger sequencing technologies will be submitted to the National Center for Biotechnology Information, a public repository of genome, expression and transcriptome data. The project will have one or more Bioprojects, which accommodate both raw and analyzed sequence data, and expression data. RT-qPCR data will be submitted to the Gene Expression Omnibus (GEO) database following NCBI guidelines. RNA-Seq data will be submitted through the Sequence Read Archive (SRA) database. Links will be provided to experimental and collection metadata archived at BCO-DMO, LTER, AOOS and EOL, providing easy cross-referencing between data types and associated publications. Data deposited at NCBI are reviewed for quality control and standard format making it easily accessible by other researchers. Thus, NCBI databases are heavily accessed by researchers for data analysis and used to answer research questions that are beyond the original study that generated the resource – we expect that the data generated here will be used by others in the biological oceanography community.

Micrographs
Images will be used to record: 1) live plankton for both scientific and outreach purposes; 2) the progression of oogenesis in *Neocalanus flemingeri* females; and 3) confocal microscopy to measure cell division. Some but not all of these images will be used in publications and public web pages, however, all of them will be stored in a database that is accessible to all project participants. The standards for the image and confocal metadata will be adapted from the standards we have developed for electron microscope images stored and accessed through our specialized imaging software (Podbase).

Experimental Data
Links between the experimental data and the field and transcriptomic/genetic data will be established through publications in peer-reviewed journals, which will include detailed descriptions of experimental protocols and results as well as web links to data repositories. In addition, citations to publications resulting from the work will be listed in the databases. The original data on the experiments and protocols will be kept in laboratory notebooks that will be stored with the PIs for a minimum of five years after the completion of the project.

Dissemination of Data and Publication of Results
Dissemination of the data will occur through presentation at national and international meetings by project participants, including the P.I.s, the post-doctoral trainee and undergraduate student participants. Research results will be published in peer-reviewed journals.

Outreach Materials
The Gulf of Alaska Large Marine Ecosystem (LME) harbors a rich and diverse community including important fisheries. Materials generated through this project will be made available to a broad audience through UHM, UAF (www.sfos.uaf.edu/sewardline/), Gulf Watch Alaska, North Pacific Research Board, Alaska Ocean Observing System and the LTER network. Materials will focus on photos, diagrams and short explanatory write-ups tailored for a general audience.