

Data Management Plan

The data that will be produced from this project will include results from both laboratory culture experiments using diatoms, as well as field work from a seasonal study in coastal California. These efforts will be aimed to understand the potential for interactions between temperature, nutrients and carbonate chemistry under the two contrasting climate regime scenarios (upwelling and heatwaves), as they individually and together influence *Pseudo-nitzschia* physiology and toxin production. Measurements in both lab and field studies will include rates of growth, CO₂ fixation, domoic acid production, as well as basic cell count, chlorophyll and elemental quota data (C,N,P,Si). Lab studies will produce transcriptome datasets under the various experimental treatments, and communities in the field incubation experiments will be characterized using 16S and 18S rRNA amplicon sequencing, flow cytometry, and microscopy. This project will thus produce data that will include experimental, biogeochemical, physiological, genomic, and transcriptomic datasets.

Sample Archiving: All biochemical and genomic material produced will be stored at -20 or -80 °C, as appropriate for the sample type. Cell count samples will be stored in 1% buffered glutaraldehyde. A database containing the location of all materials will be maintained for use by the PI, and submitted along with the data to BCO-DMO.

Data Archiving: All data collected during the laboratory- and field-based experiments will be securely stored in multiple redundant formats (hand-written lab notebooks, backed-up files on discs/hard drives) indefinitely in the PI's offices. Upon receipt of the award, I will contact the Biological-Chemical Oceanography Data Management Office (BCO-DMO: <http://www.bco-dmo.org/>) to register our project. I will submit all data collected from the field and laboratory portions to BCO-DMO for archiving per NSF guidelines. All biochemical and genetic material produced will be stored at -20 or -80 °C, as appropriate for the sample type. In addition to my findings, a searchable database containing the location of all processed and non-processed materials in storage will be maintained by the PIs and submitted to BCO-DMO.

Data Release: Access to data will be given once it is quality controlled and published, or at any rate always within 2 years of collection. Availability will be in accordance with NSF guidelines for data accessibility. I will keep NSF abreast of our compliance with data management through my annual reports and all data will be made available to the general public as expeditiously as possible. I will work closely with the BCO-DMO to ensure that data used in my analyses and outcomes from the proposed experiments are publicly available according to NSF guidelines. Further, all data made available will be accompanied by compliant metadata. Raw nucleic acid sequencing reads will be deposited to the National Center for Biotechnology for Information (NCBI) Sequence Read Archive (ncbi.nlm.nih.gov/sra) and Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>) databases within three months of the end of this project, with appropriate links from our project metadata page at BCO-DMO. Furthermore, all raw and processed sequence data will be backed up and stored at the University of Southern California High-Performance Computing Cluster (hpcc.usc.edu).

As I have done in the past, I will continue to make the results of my work available to the marine science community through timely peer-reviewed publications and professional meeting presentations. The budget contains publication funds that will cover the costs of open access publications.