Data Management Plan

Lead PI Hutchins will lead the data management effort with support from the coPIs, technician, and graduate students. All data collected during this laboratory-based experimental evolution culturing and experimental effort 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.

Four general types of data will be produced from this project: 1. Culture experimental physiological and biogeochemical data, 2. Genomic data, including DNA sequences from our cultures, 3. Transcriptomic data, including RNA sequences from our cultures, and 4) Proteomic data from our cultures.

Standards that would be applied for format, metadata content, etc.: We will work closely with the Biological-Chemical Oceanography Data Management Office (BCO-DMO: http://www.bco-dmo.org/) to ensure that data used in our analyses and outcomes from our experiments are publicly available according to NSF guidelines. Further, all data made available will be accompanied by compliant metadata. DNA and RNA sequences generated from this work will be deposited in Genbank at the National Center for Biotechnology for Information (NCBI) within three months of the end of this project.

Unlike genomic data, there is not a single global location for proteomics data archival currently. In lieu of this, proteomics datasets will be made available through multiple means. First, we will work with BCO-DMO as mentioned above. Second, datasets will be published as journal supplementary materials as we have done for four recent publications. Third, datasets will be posted at our Proteomics Laboratory's website data archive (http://www.whoi.edu/sites/ProteomicsLab). Finally, we are working with the Joint Genome Institute to deposit our quantitative proteomic results into their popular Integrated Microbial Genomes Portal site, and will request to do so with datasets from this project.

Provisions for archiving and preservation*:* All biochemical and genomic material produced will be stored at -20 or -80 °C, as appropriate for the sample type. Phytoplankton samples will be stored in 4% buffered glutaraldehyde. A database containing the location of all materials will be maintained for use by all PIs and submitted with the data to BCO-DMO. Access policies and provisions*:* Access to data will be given once the data is quality controlled and published. Availability will be in accordance with NSF guidelines for data accessibility.

As we have always done in the past, we will continue to make the results of our work available to the marine science community through timely peer-reviewed publications and professional meeting presentations. Our budgets contain publication and travel funds that will cover the costs of free access publication and oral and poster presentations at professional venues like ASLO/Ocean Sciences, and Gordon Research Conferences.