

DATA MANAGEMENT PLAN

Our data management practices will be consistent with the requirements for PIs funded through the Biological Oceanography (BioOce) program. Overall, we will utilize the “Biological and Chemical Oceanography Data Management Office” (BCO-DMO; <http://bco-dmo.org/>) as the primary repository for project specific data.

1. Expected Data

This project will generate experimental data at two broad levels: 1) molecular (gDNA sequence, transcriptome profiles, and DNA methylation profiles), and 2) organismal (physiological rates of metabolic performance). The molecular data will be heavily dominated by genetic sequences that can be contained within flat text files (fasta). The physiological data will be better accommodated by storing in matrix formats within text files.

2. Data Format

Ultimately, all data sources will be quantitatively analyzed in some fashion requiring formatted input/retrieval via R stats. These R files (rda) would be the most expedient format for data storage because it provides compression and can be accessed by anyone using R in the future.

3. Access to Data and Data Sharing Practices and Policies

All low-level sequence data will be made available through the BCO-DMO repository. Downstream synthesis of transcriptome profiles, methylation profiles, and pathway identification will also reside with BCO-DMO, but also likely be made publicly available on the academic websites of the PIs.

Physiological data in R data frames will be available for “data sharing” at the discretion of the PIs for not more than 1 year after the project ends. This time window is designed to allow the PIs sufficient opportunity to publish results without fear of direct competition. After this 1 year period, these data will be publicly available through the BCO-DMO or by direct request sent to the PIs.

Software tools and products developed during the course of this project may have restricted distribution limits based on the commercial licensing of any IP by the University of Delaware’s Technology Transfer Office. All attempts will be made to establish commercial vs. academic software releases, but ultimately, UD’s TTO will have the final determination about the scope of distribution in which there may be a commercial conflict.

4. Policies for Re-Use, Re-Distribution

Our “Terms of Use” for any project data or software would simply require an acknowledgment and citation of our primary publications related to that information. We would assist any researcher in extracting as much information as possible from these data sets and would encourage this practice as much as possible.

5. Archiving of Data

Legacy storage/access of critical DNA sequence data (digital files) would be handled on our primary departmental server at the marine campus in Lewes, DE. Storage on rack-mounted drives under the control of UD’s IT services will ensure that as this critical server system is routinely updated, our data would be periodically propagated to the new storage devices handling legacy storage on this system. This digital storage also includes a routine back-up to a server storage system on main campus (Newark, DE).