Data Management Plan – Bigelow Laboratory

Our data management plan is based on guidelines established by the National Science Board, and the National Science Foundation, and covers dissemination and sharing of materials and data that are expected to be collected as part of the research described in the project proposal. For this collaborative project, this data management plan is agreed to by both institutions, and may be considered common to both.

I. Types of Samples and Data, and their Distribution.

This project entails coordinated field sampling of sediment samples obtained using a multi-corer system allowing collection of replicate cores. Meta-data will be gathered from the cores with regard to important geochemical parameters, and a subset of cores will be extruded for more detailed laboratory analysis. The geochemical and associated meta-data from cores will be logged in way that is consistent with data acquisition protocols for the biological and chemical oceanography data management office (BCO-DMO) and deposited with BCO-DMO. As appropriate geochemical and related data from laboratory-based experimental efforts will also be deposited with BCO-DMO, the extent of the analysis are described in the project description.

Data from individual microcosm analyses will be processes and analyzed and stored on servers that are routinely backed up at the respective home institutions. The same is true of any imaging that is done of core samples, or experimental apparatus, or photomicrographs. A selection of personal photographs taken during field operations showing different aspects of field activities may be collected, and disseminated publically via social networking sites like Facebook.

The largest amount of data that will be acquired will be genomic. Genomic data generated using nextgen sequencing technologies for RNAseq and SSU amplicon sequencing will be placed in the NCBI's sequence read archive (http://www.ncbi.nlm.nih.gov/sra) along with associated meta data.

II. Individual Scientists Responsibilities.

Each PI is responsible for maintaining data associated with their own research groups activities, and in accordance with any institutional requirements of the PI's home institution. Basic practices for key areas are outlined below.

Lab notebooks. All information connected with initial data collection, analysis, and results will be kept in a lab notebook. When studying genomic data, for example, data are too plentiful to record by hand in a paper notebook. In such cases, digital notebooks may be preferable. Regardless of media, these notebooks will be stored as well. Hard-copy notebooks will be stored and archived to enhance institutional retrieval. All individuals' lab notebooks remain the property of the host institution.

Data security. In addition to data hosted on local computers, it is advised that all important data be stored on off-network mobile devices (e.g. hard disks) or offsite cloud resources. Password protection is strongly encouraged. This data (and associated passwords) must be made available to senior institution officials in the case that any institutional liability issues should arise.

Data backup. Stored data shall be regularly backed up, preferably weekly, and is the responsibility of the overseeing PI.