Data Management Plan: Marine Diatom-Parasite Relationships in Upwelling Systems

---

Data Policy Compliance

We will comply with the requirements of the National Science Foundation Division of Ocean Sciences Sample and Data Policy (May 2011). BCO DMO may be the appropriate repository for some of our data, but most of the data from this project will be molecular sequence data, which is more suitable for submission to NCBI. Metagenomic sequence data will be archived in the Sequence Read Archive (SRA) at NCBI and registered as a BioProject along with the metadata for the sequenced samples at NCBI.

Pre-Cruise Planning

No new field work is planned as part of this project

Description of Data Types

We categorize the data to be collected into the following groups:

- (1) Experimental data: DNA and RNA sequences (18S rDNA tag sequences, metagenomic and metatranscriptomic data)

- (2) Derived data products: publications, preliminary data reports, sequence data.

The experimental and sequence analysis results will require 2 years or more for analysis.

Data and Metadata Formats and Standards

Metagenomic data will be deposited at NCBI (BioProject) in the required format (usually fasta files). Meta data include experiment and sample ID, citations of related publications, contact person.

Data Storage and Access During the Project

Data will be shared among project participants via shared DropBox and GoogleDrive folders managed by Ward. These drives also serve as secure backup for multiple users and devices. The largest files are likely to be sequence database and analysis files, which will be stored on Princeton University servers. The Department of Geosciences at Princeton supports PICSciE (Princeton Institute for Computational Science and Engineering), which in turn provides data storage capacity and high performance computing support for sequencing projects.

Mechanisms and Policies for Access, Sharing, Re-Use, and Re-Distribution

Data availability: Experimental data will be published in peer-reviewed publications and deposited at BCO DMO as appropriate or at NCBI (GenBank, SRA) as appropriate at the time of publication. After publication, the authors will make all data available to others upon request.

Data Sharing via BCO-DMO: We will submit project metadata to BCO-DMO at the time the project is funded. Sequence data generated during the proposed research will be archived at NCBI and a link to the Sequence Read Archive (SRA) will be posted on the BCO-DMO project site.

Underway Shipboard Data: We do not expect to generate new shipboard data for this project.

Plans for Archiving

We do not plan to collect physical samples that require long term archiving. Particulate material collected for the purpose of molecular analysis has been preserved in -80°C Freezers at Princeton University; these are controlled and monitored by a central monitoring and alarm program. Some of these samples will probably be retained after the end of the project, but they do not constitute a formal permanent archive. Long term availability of the observational, experimental and derived data will be assured by its repository at BCO-DMO and NCBI.

Roles and Responsibilities

Ward is responsible for ensuring compliance with the Data Management Plan.