Data Management Plan

We will conform to the NSF and OCE policies on the dissemination and sharing of samples and research results. Details on the various types of data that will be generated through this project, as well how data will be archived and distributed, are described below:

Types of Data, Samples and Physical collections

This project will yield a number of different types of samples and data, including environmental measurements (physical, chemical, and biological parameters), biogeochemical rates, and molecular data (e.g., functional gene and 16S rRNA sequences, qPCR data, etc.). All of the data associated with our samples will be maintained in existing offline PI databases.

Policies for access and sharing (including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements)

There are no legal (copyright, licensing) or ethical (privacy) issues concerning data management for this project.

Data Distribution:

During the funding period of the project and until the data are published, experimental data will be stored in electronic format on laboratory and desktop computers. Data will be tabulated and stored in spreadsheets (e.g., MS Excel, CSV, etc.) for easy manipulation and transfer to other programs for plotting and statistical evaluation. Stored data will be clearly labeled with unique identifiers and dates of collection and updates. Accompanying metadata and documentation explaining experimental conditions and measurements will be created by the project PIs. Following publication, the data will be made publically available.

Data and Sample Archiving:

Samples of genomic DNA, RNA, and unprocessed filters/cores collected during this study will be stored/archived in multiple -80°C freezers at Stanford. Sequence data will be deposited in NCBI as well as other relevant genomic databases, along with appropriate metadata, within 2 years of the completion of the project. Data from this project will also be contributed to the Biological and Chemical Oceanography Data Management Office (BCO-DMO), based at WHOI (http://bcodmo.org) and dedicated to data generated from NSF funding. The BCO-DMO system is particularly useful because it provides access to complex datasets in a consistent manner, with sufficient metadata, so that others can make full use of these data for their own purposes.