

## Data Management Plan

### 1 Data Policy Compliance

The project investigators will comply with the data management and dissemination policies described in the *NSF Award and Administration Guide* (AAG, Chapter VI.D.4) and the *NSF Division of Ocean Sciences Sample and Data Policy*.

### 2 Field Sample Planning

Sample planning will be coordinated among the PIs through email and phone. Detailed plans for field sampling (R/V Ira C cruises & SCUBA operations) and field measurements will be agreed upon three months in advance of the field season. The sampling events and any field data will be recorded on paper logs and scanned into PDF documents.

### 3 Description of Data Types

**Observational Datasets: CTD:** CTD data collected using a SeaBird SBE CTD package; processing to be done using SeaBird's SeaSave software; data will include standard environmental measurements (such as pressure, temperature, salinity, fluorescence) and mPAR data collected with logging software. File types: Raw (.con, .hdr, .hex, .bl) and processed and .cnv, .asc, .bt1) ASCII files. Repository: BCO-DMO. **Event log:** Cruise scientific sampling event log; will include event numbers, start/end dates, times & locations of instrument deployments, as recorded on paper log sheets. File types: Excel file converted to .csv; scanned PDFs. Repository: BCO-DMO

**Experimental Datasets: Biogeochemical data:** We will collect dissolved O<sub>2</sub>, NO<sub>3</sub><sup>-</sup>, <sup>15</sup>NO<sub>3</sub><sup>-</sup>, NH<sub>4</sub><sup>+</sup>, <sup>15</sup>NH<sub>4</sub><sup>+</sup>, and <sup>15</sup>N-N<sub>2</sub> in the overlying water of whole core experiments. We will collect the following data from measurements on discrete sediment samples: TOC, chl a, diatom counts, ICNO<sub>3</sub>, IC<sup>15</sup>NO<sub>3</sub>, porewater NO<sub>3</sub><sup>-</sup>+NO<sub>2</sub><sup>-</sup>, porewater <sup>15</sup>NO<sub>3</sub><sup>-</sup>+<sup>15</sup>NO<sub>2</sub><sup>-</sup>, porewater and extractable NH<sub>4</sub><sup>+</sup> and <sup>15</sup>NH<sub>4</sub><sup>+</sup>, TN, and T<sup>15</sup>N. File types: Excel file(s). Repository: BCO-DMO.

**Nucleic acid sequence data:** 2 x 150 bp sequencing reads will be generated on Illumina NexSeq for metagenomic and metatranscriptomic analyses. File types: short-read archive (.sra) and .fasta files. Repository: NCBI GenBank.

### 4 Data and Metadata Formats and Standards

**Observational Data:** Field observation data will be stored in flat ASCII files. Field data will include date, time, latitude, longitude, cast number, and depth, as appropriate. Quality flags will be assigned according to the ODS IODE Quality Flag scheme (IOC Manuals and Guides, 54, volume 3; [http://www.iode.org/mg54\\_3](http://www.iode.org/mg54_3)). Metadata will be prepared in accordance with BCO-DMO conventions (i.e. using the BCO-DMO metadata forms) and will include detailed descriptions of collection and analysis procedures.

**Experimental Data:** Biogeochemical data will be stored in Excel files. Sequencing reads will be trimmed for quality and filtered to remove primers, adaptors and rRNA sequences using Ribopicker v.0.4.3, with file types: short-read archive (.sra) and .fasta files. Metadata pertaining to sequences generated will be stored as a .txt file.

### 5 Data Storage and Access During the Project

The investigators will store project data (including spreadsheets, ASCII files, images, and PDFs of scanned logs) on laboratory computers, external hard drives, and google drive that are backed up by the University's central IT organization. Personal computers in all laboratories are backed up daily using Apple Time Machine to an onsite external hard drive.

## **6 Mechanisms and Policies for Access, Sharing, Re-Use, and Re-Distribution**

All data collected through the proposed research will be archived and publication related information and material will be made available immediately. Publicly available data will be delivered in an interoperable format that enables wide-spread data sharing and facilitates secondary use. All data analyses source code will be made publicly available upon manuscript submission and accession numbers, project IDs and other digital object identifiers (DOIs) will be provided to BCO-DMO. All SSU rRNA sequence data will be submitted to the NCBI Sequence Read Archive (SRA).

## **7 Plans for Archiving**

We will work with the BCO-DMO to ensure that field observational and biogeochemical data are appropriately archived. Additionally, we will archive all of the data on the University's tape-recorded permanent back up system. All nucleic acid sequence data will be archived at NCBI GenBank.

## **8 Roles and Responsibilities**

Rich and Giblin will be responsible for managing data. The PIs agree to share data with each other in an open and timely manner and adhere to appropriate data storage and backup policies.