

Plan Overview

A Data Management Plan created using DMPTool

Title: CDEBI project Investigating a Mysterious Ammonium Flux and Its Relationship to the Microbial Community

Creator: John Kirkpatrick - **ORCID:** [0000-0002-5741-7005](https://orcid.org/0000-0002-5741-7005)

Affiliation: University of Rhode Island (ww2.uri.edu)

Principal Investigator: John Kirkpatrick

Data Manager: John Kirkpatrick

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 29141

Template: BCO-DMO NSF OCE: Biological and Chemical Oceanography

Last modified: 10-03-2017

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

CDEBI project Investigating a Mysterious Ammonium Flux and Its Relationship to the Microbial Community

Data Policy Compliance

Identify any published data policies with which the project will comply, including the NSF OCE Data and Sample Policy as well as other policies that may be relevant if the project is part of a large coordinated research program (e.g. GEOTRACES).

Question not answered.

Pre-Cruise Planning

If the proposed project involves a research cruise, describe the cruise plans. (Skip this section if it is not relevant to your proposal.) Consider the following questions: (1) How will pre-cruise planning be coordinated? (e.g. email, teleconference, workshop) (2) What types of sampling instruments will be deployed on the cruise? (3) How will the cruise event log be recorded? (e.g. the Rolling Deck to Repository (R2R) event logger application, an Excel spreadsheet, or paper logs) (4) Will you prepare a cruise report?

Question not answered.

Description of Data Types

Provide a description of the types of data to be produced during the project. Identify the types of data, samples, physical collections, software, derived models, curriculum materials, and other materials to be produced in the course of the project. Include a description of the location of collection, collection methods and instruments, expected dates or duration of collection. If you will be using existing datasets, state this and include how you will obtain them.

Experimental data: DNA sequences (16 rRNA) extracted from marine sediment and sequenced on a 454 GS-FLX pyrosequencer.

Data and Metadata Formats and Standards

Identify the formats and standards to be used for data and metadata formatting and content.

Where existing standards are absent or deemed inadequate, these formats and contents should be documented along with any proposed solutions or remedies. Consider the following questions: (1) Which file formats will be used to store your data? (2) What type of contextual details (metadata) will you document and how? (3) Are there specific data or metadata standards that you will be adhering to? (4) Will you be using or creating a data dictionary, code list, or glossary? (5) What types of quality control will be used? How will data quality be assessed and flagged?

Data is archived with and publicly available at the Visualization and Analysis of Microbial Population Structures (VAMPS) website at <https://vamps.mbl.edu/index.php> , under project ID JBK_IO_Bv6v4.

Data Storage and Access During the Project

Describe how project data will be stored, accessed, and shared among project participants during the course of the project. Consider the following: (1) How will data be shared among project participants during the data collection and analysis phases? (e.g. web page, shared network drive) (2) How/where will data be stored and backed-up? (3) If data volumes will be significant, what is the estimated total file size?

Data is stored and accessible at vamps.mbl.edu .

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

Describe mechanisms for data access and sharing, and describe any related policies and provisions for re-use, re-distribution, and the production of derivatives. Include provisions for appropriate protections of privacy, confidentiality, security, intellectual property, or other rights or requirements. Consider the following: (1) When will data be made publicly available and how? Identify the data repositories you plan to use to make data available. (2) Are the data sensitive in nature (e.g. endangered species concerns, potential patentability)? If so, is public access inappropriate and how will access be provided? (e.g. formal consent agreements, restricted access) (3) Will any permission restrictions (such as an embargo period) need to be placed on the data? If so, what are the reasons and what is the duration of the embargo? (4) Who holds intellectual property rights to the data and how might this affect data access? (5) Who is likely to be interested in re-using the data? What are the foreseeable re-uses of the data?

Data was made publicly available in 2016 and is freely available for download and re-analysis by all

interested parties. Data is to be linked to BCO-DMO as well as CDEBI websites (<http://www.darkenergybiosphere.org/>). All sequence data and information collected under this project can and has been made publically available without restriction.

Plans for Archiving

Describe the plans for long-term archiving of data, samples, and other research products, and for preservation of access to them. Consider the following: (1) What is your long-term strategy for maintaining, curating, and archiving the data? (2) What archive(s) have you identified as a place to deposit data and other research products?

The primary archive for this data is the VAMPS website <https://vamaps.mbl.edu/index.php> . If and when data migration become necessary the data will be moved to the National Center for Biotechnology Information (NCBI).

Roles and Responsibilities

Describe the roles and responsibilities of all parties with respect to the management of the data. Consider the following: (1) If there are multiple investigators involved, what are the data management responsibilities of each person? (2) Who will be the lead or primary person responsible for ultimately ensuring compliance with the Data Management Plan?

In cooperation with the PI (Steven D'Hondt), the postdoctoral fellowship awardee, John Kirkpatrick, will be responsible for ensuring data accessibility.
