
Collaborative Research: Defining contributions of diverse phototrophic microbes to carbon cycling in two Mid-Atlantic estuaries

A Data Management Plan created using DMPTool

Creator: Barbara Campbell

Affiliation: Clemson University

Template: National Science Foundation (NSF)

Last modified: 02-11-2017

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

Collaborative Research: Defining contributions of diverse phototrophic microbes to carbon cycling in two Mid-Atlantic estuaries

Data Policy Compliance

The PIs 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*.

Pre-Cruise Planning

We will provide all available metadata from our two requested cruises (Summer 2018 and Spring 2019) to the Delaware and Chesapeake Bays. We expect to sample at 6 or 7 sites during each of the cruises. Each site will have approximately 3 depths where data will be collected. In addition, data will be collected after microcosm and cell sorting experiments.

Description of Data Types

Data that will be collected include:

- 1) Observational/underway data of salinity and fluorescence from the ships underway system.
- 2) CTD data from depth profiles, including temperature, depth, salinity, fluorescence, oxygen, turbidity. We expect to sample between 6-8 sites at at least 3 depths per site during each 8 day cruise.
- 3) Data from a spectral Light Profiler (Biospherical Par BIS 2101) to measure PAR and light attenuation and from a Secchi disc will be recorded at each station where CTD casts are performed.
- 3) Data from chemical assays from samples collected with the CTD at different depths; including pH, DIC, DOC, nutrients (nitrogen, phosphorus, ammonium, silicate), chlorophyll a.
- 4) Biological data including: cell counts of various types (phytoplankton, cyanobacteria, rhodopin-containing microbes, other phototrophic microbes, all heterotrophic bacterioplankton), bacterial production.
- 5) Sequence data, including: 16S rRNA gene tags (from at least 18 samples, in duplicate), in situ metagenomic and metatranscriptomic data from approximately 18 samples (in duplicate= 36), (meta)transcriptomic data from 60 microcosm samples, including controls and experiments with different wavelengths, as well as sorted cells.

Data and Metadata Formats and Standards

Data (in defined units) will be submitted to BCO-DMO as tab-separated ASCII files, as with prior submitted data by the col Campbell.

All data collected will be in compliance with the current Minimum Information about a Genome Sequence (MIGS) specifications.

Data Storage and Access During the Project

All primary cruise data will be recorded in laboratory notebooks and results compiled in files stored on local hard drives. The data will be routinely backed up to external hard drives and stored in cloud-based systems (Dropbox, box) for sharing between the four PIs.

Secondary data (primarily sequence data and analyzed sequences) will be stored on at least two physically separated NAS servers at Clemson (in different buildings) with internal redundant drives (RAID 5 or 6) in case of physical damage to one of the servers or internal hard-drive failures. The estimated total file size of raw data is between 4-8 TB; and analyzed data between 20-40 TB. Secondary analyzed data will be shared between labs via cloud-based systems (Dropbox, box, etc.).

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

All primary cruise data will be submitted to the BCO-DMO website within 3 months of the completed cruise.

All secondary cruise data (chemical and biological analyses) will be submitted to the BCO-DMO website within 3 months of final analysis of the data.

All sequence data (raw and appropriate genome bins) will be submitted to NCBI SRA at or before publication or within 1 months after the completion of the grant, whichever comes first.

Plans for Archiving

The data that is online and documented (BCO-DMO and NCBI websites) will be properly archived by the appropriate data center.

Roles and Responsibilities

Primary responsibilities for cruise data submission will be by the chief scientist of each cruise (as of now, Barbara Campbell).

Primary responsibilities for secondary data submission will be with the appropriate laboratories (Hanson and Maresca for chemical and biological data, Campbell for sequence data).

The person ultimately responsible for all submissions will be Hanson.