BCO-DMO NSF OCE: Biological and Chemical Oceanography
National Science Foundation

Data Policy Compliance

This project will comply with the NSF OCE Data and Sample Policy

Description of Data Types

(1) **Field observational**: Physico-chemical data: Time of collection, date, location, depth, core temperature profile from ROV Alvin's thermoprobe, salinity, dissolved oxygen. To maximize access to these data, data will be stored in BCO-DMO by the end of this project and in PANGAEA (http://www.pangaea.de/about/) by the PI. PANGAEA is a member of the ICSU World Data System, created by the International Council for Science, is jointly hosted by the Alfred Wegener Institute and the University of Bremen, and is established as a repository for the long-term availability and accessibility of archived data and metadata in secure and machine readable formats.

(2) **Hydrocarbon profiles and nutrient analyses**: 3-cm horizons from sampled cores will be analyzed for saturated hydrocarbons, PAHs and alkylated PAHs (~200 compounds) by Alpha Analytical. Porewater concentrations of NO_2+NO_3, NH_4, PO_4, total dissolved organic and inorganic carbon, dissolved organic nitrogen, will be supplemented by total organic carbon, total carbon, and total nitrogen (determined with LSU Wetlands Biogechemistry Labs). All data will be deposited in BCO-DMO and PANGEA immediately after their analysis and made publicly available by the PI.

(3) Experimental:

a. **Sequence data**: MiSeq PE 300bp iTAG data will be generated for Bacteria and Archaea and for Fungi by the end of year 1 for all samples in our collection. These data will be deposited to GenBank and made publicly available by the end of the project period by the postdoc. Links to data will be provided through BCO-DMO. Metatranscriptome data for culture studies of biodegradation will be deposited to MG-RAST, and archived on the NCBI-Gene Expression Omnibus (GEO) database and made publicly available by the end of the project period by the postdoc. Links will be provided through BCO-DMO.

b. **Environmentally relevant fungal strains:** Fungal strains that have complete taxonomic descriptions (pending collaborator effort) and that have been fully characterized will be deposited in the American Type Culture Collection (ATCC) and the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen) by VE. All unique fungal isolates will be directly cryopreserved in 15% glycerol stocks and/or 5% dimethyl sulfoxide and deposited in the UBOCC (University of Brest Culture Collection) by collaborator Burgaud, and duplicate stocks in -80°C freezers in Edgcomb's laboratory (on emergency power backup) by the postdoc. All isolates will be available upon request during or after our project term. Data from taxonomic characterization and physiological experiments will be shared electronically among all collaborators by collaborator Burgaud and the postdoc, and will be posted to a dedicated project page on Basecamp (https://basecamp.com) by the end of the project period by the postdoc.

c. **All hydrocarbon data** (saturated hydrocarbons and alkylated PAHs) for field samples (spreadsheet output and chromatographic profiles) from Alpha Analytical and GC x GC data (chromatographic and spreadsheet data) for culture incubation

studies produced by S. Sylva in C. Reddy's lab will be deposited in BCO-DMO as they are generated by Edgcomb.

d. **Experimental protocols**: DNA and RNA extraction protocols and protocols for all culturing efforts and hydrocarbon degradation / biosurfactant production screening in the study will be made publicly available through protocols.io by the end of the project period by the postdoc.
e. **Educational outreach materials**: Protocols for the laboratory exercise on FISH hybridization for high school students will be posted on Amazon Inspire, the new education resource for teachers dedicated to the free search, discovery and sharing of digital education resources (http://www.amazoninspire.com) by Edgcomb.
f. All data will be shared through peer reviewed publications.

Data and Metadata Formats and Standards

1. Observational data, hydrocarbon profiles, and nutrient analyses will be submitted as spreadsheet files (.xlsx).
2. Genetic data formatting will follow the standards of NCBI, and MG-RAST and links to these data will be provided through BCO-DMO.
3. Fungal isolate lists and outcomes of incubation studies with individual hydrocarbons will be submitted as spreadsheet files.
4. Hydrocarbon profiling data for field sediment samples and for culture incubation studies will be submitted to BCO-DMO as chromatographic data (PDF format) and spreadsheets (.xlsx).

Data Storage and Access During the Project

All data will be shared electronically among all collaborators as data are generated by posting to a dedicated project page on Basecamp (https://basecamp.com) or Google Drive. All experimental data generated at WHOI will be stored on our laboratory computers and associated Network Accessed Storage devices that are backed up daily by WHOI's automated backup service.

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

Our project data will all be publicly available by the end of the project period through NCBI GenBank/Gene Expression Omnibus (GEO) database, BCO-DMO, and MG-RAST. Our data will not involve security concerns, nor does access need to be restricted in any way.

Data Sharing via BCO-DMO and Archiving

BCO-DMO staff will work with us to manage our project data, and data generated during the proposed research project will be contributed to the BCO-DMO system. Links to data repositories for our genomic data will be provided through BCO-DMO (see above). After data contributed to BCO-DMO are online and fully documented, BCO-DMO ensures that the data are archived properly at the appropriate National Data Center (e.g. NODC) for long-term archive preservation.

Roles and Responsibilities

Lead PI Edgcomb will ensure compliance with this Data Management Plan. She will be responsible for deposition of all field observational, nutrient, and hydrocarbon profile data. She will supervise the deposition of all experimental data by the postdoc.