

Data Management Plan

Collaborative Research: Dispersal depth and the transport of deep-sea methane-seep larvae around a biogeographic barrier

Data Collection

Sample collections will include larval forms obtained with the SyPRID plankton sampler on AUV Sentry, adult and juvenile methane seep animals collected from the ocean floor with Alvin or Jason, and samples of larvae from integrated year-long deployments of larval traps on the sea floor. In addition, we will obtain sea-floor current meter data from 6 sites. Between cruises, samples will be analyzed for microchemistry, identified with genetics, and imaged with SEM. The data from these analyses will be documented in metadata files and archived according to agency policies.

The Deep Submergence Group will generate navigation data and the ships will generate routine oceanographic measurements while underway and on station. Alvin or Jason will generate sea-floor imagery and sensor data.

Underway data and dive data will be generated by the ship or undersea assets. Metadata for collections will be generated by the scientists, under the supervision of the Chief Scientist on cruises and the collaborating PI's between cruises. At the beginning of the project, the PI's will create standardized metadata forms and spreadsheets for most of the data products expected.

Documentation and Metadata

Sample information will be recorded in spreadsheet files. At the beginning of the project, P.I.'s will meet by teleconference to discuss standardized protocols for data collection, specimen disposition, and data management. Metadata forms will be designed in advance for recording information on mooring deployments and sampling stations. We will also design data spreadsheets for standard kinds of data (e.g., larvae from SyPRID), and these will be available on cruises in both digital and paper form. Metadata from SyPRID deployments and ROV or Alvin dives will be recorded in standard *National Deep Submergence Facility* (NDSF) formats. The P.I.'s will be responsible for keeping metadata about experiments, collections, and analyses, including names of individuals who obtained the data, collection and processing dates, sample numbers, sample i.d.'s, collection sites, instructions for accessing data, and final disposition of samples. **The metadata will be uploaded to the BCO-DMO website prior to the submission of each annual report, and all data will be submitted within two years of collection, as required by the Data and Sample Policy of the NSF Division of Ocean Science.** Where the data are derived from genetic or chemical analyses, the metadata will include basic information, including literature citations where appropriate, documenting the methods used.

Ethics and Legal Compliance

Our sample protocols and data management will conform to all rules and policies governing the disposition and sharing of data and samples. We do not anticipate any ethical issues associated with the identities of participants or institutional policies. The data we generate are not expected to be sensitive. As data are published, copyright of the derived products (graphs, etc.) will generally be transferred to the publisher, and we will conform with all publisher policies concerning copyright laws and fair use.

Storage and Backup

During cruises, the Chief Scientist will be responsible for safe storage and back-up of sample and research data and metadata, and will assure that all participants have immediate access to the full set

of shipboard data. Specifically, individual P.I.'s at the three collaborating institutions will all be given a full set of data during the cruises. Analyses of samples and modeling results from the participating labs will be shared among all PI's soon after they become available.

BCO-DMO ensures that the data are archived securely at the National Centers for Environmental Information (e.g., *National Oceanographic Data Center*) for long-term archive preservation. Access to the data and metadata is managed by BCO-DMO through their website.

Genetic Sequence data will be submitted to a publicly accessible data repository as soon as they become available (e.g., [National Center for Biotechnology Information](#)).

Selection and Preservation

Many samples of adults will be used as broodstock for obtaining embryo and larval cultures and will be destroyed in the process. A small number of voucher specimens will be retained for later verification of identifications. Any type specimens of new species described will be deposited in the Smithsonian Museum of Natural History. Adults collected for histological analysis will be destroyed, but the slides of gonads will be retained in the PI's lab and made available for sharing. The larvae and juveniles will be photographed, but will mostly be destroyed in the course of geochemical and genetic analyses. All imagery from submersible or ROV dives will be archived at Woods Hole Oceanographic Institution, where it is available to the PI's from this grant and to others through their standardized sharing protocols.

All data and metadata are deemed to have long-term value and will be archived for long-term access in BCO-DMO at Woods Hole Oceanographic Institution. Voucher specimens not appropriate for deposition in the Smithsonian will be retained at the respective institutions of the PI's.

Data Sharing

All data will be archived in the BCO-DMO system at Woods Hole, and the project metadata will accompany the data. Those few samples that survive the experiments and analyses will be retained in our respective laboratories and made available to other users who request them. We will endeavor to publish results as soon as possible and will make the raw analytical data from those publications available immediately after publication. We will add sample metadata, including locations, species and numbers, to the BCO-DMO archives as soon after cruises as we are able to process them and will not impose any restrictions on data sharing.

Responsibilities

The PIs are responsible for ensuring compliance with the data management plan. The PIs will follow-up post-cruise to assess progress, and where necessary facilitate compliance. Status of data management will be noted in the standard report(s) to NSF for this grant.

The Rolling Deck to Repository program (R2R) is responsible for archiving routine underway data at the appropriate national archive, including the National Centers for Environmental Information. Raw multibeam swath and CT data and will be released within 2 yrs of acquisition, or earlier if all participants agree to an immediate release of some datasets upon completion of the cruise. Derived products such as NetCDF grids of local bathymetry at study sites will be submitted to MGDS in conjunction with acceptance of peer-reviewed papers.

WHOI is responsible for archiving NDSF data. We anticipate that *Alvin/Jason* 'framegrabber' data for each dive will be loaded onto the WHOI NDSF site, served within 30 days after the end of the cruise, and will be publicly available on line.

