

## Data Management Plan

Two research cruises are planned, one in year 1 and another in year 2, both utilizing the ROV *Jason* or *Global Explorer*. Pre-cruise planning will be done via web-based teleconferencing (Zoom). Detailed plans for dive activities including sampling locations and priorities at each vent field and sampling strategy, as well as plans for shipboard experiments with specimens collected on each dive, will be written up as a science implementation plan for each cruise. The actual sampling events will be recorded on paper logs (scanned into PDF documents) and in a digital event log, following the Biological and Chemical Oceanography Data Management Office best practices for data management.

Soon after the completion of the cruise, the original underway data will be contributed by the vessel operator to the UNOLS central data repository at <http://www.rvdata.us/catalog/> managed by the Rolling Deck to Repository (R2R) project. Also, R2R will ensure that the original underway measurements will be archived permanently at NODC and/or NGDC as appropriate for the data type.

Navigation, motion imagery, still imagery, and real-time sensor data from ROV will be archived and available through the National Deep Submergence Facility (NDSF).

Likewise, appropriate measurements made by the science party will be submitted to PANGAEA (<https://www.pangaea.de/>) and the data sets will be freely available online from their earth and environmental science data system.

Data types generated by the science party during and after the cruises will include:

- 1) Trawl data: water column data (e.g., depth, salinity, temperature) collected by net-mounted sensors, and abundance of target taxa from Tucker trawls, will be logged.
- 2) Physiological data: shipboard electrophysiological and emitted luminescence data will be stored as digitized data on the computer hard drives, and as hand-written data in lab notebooks. Digitized data will be backed-up on accessory hard drives. Hand-written data will be transcribed into Microsoft Excel spreadsheets after each experiment and backed-up on accessory hard drives and servers.
- 3) Collections for experiments: metadata for all collected samples (e.g., suction samples, traps, images, etc.) will be incorporated into a Microsoft Excel database containing: sample number, taxonomic classification, dive number, Lat/Long, collection depth, time/date, habitat type, temperature, salinity, image/video numbers, fixative, and final distribution (PI, Museum).
- 4) Specimen archival for molecular analyses: All whole samples and tissue plucks will be identified and cataloged in a digital database in the Florida International Crustacean Collection (FICC) maintained by H. Bracken-Grissom, which currently houses over 10,000 archived crustaceans. The databases currently in place (FileMakerPro) allows for the following data entries: full taxonomic classification, catalog number, photo voucher/digitized images, associated locality information with an interactive map, geographical distribution, fixative, collector field notes, information about associated organisms and a remarks section. The database also includes a molecular component that can be used to track the entire genetic history of the organism/sample. Archival of samples will be in 70-80% ethanol in the Florida FICC or donated to the National Museum of Natural History (NMNH) in Washington D.C.

- 5) Additional collections: animals collected from vent sites but not used for experimental work will be stored in research grade ethanol and used by H. Bracken-Grissom or made available for distribution to laboratories working on genomics for Tree of Life studies. Specimens utilized for visual physiology that are not needed for genomics will be housed in J. Cohen or T. Frank's climate-controlled laboratories. Bioluminescent specimens not utilized for genomics studies will be housed in S. Johnsen's climate-controlled laboratory. Rare or unusual species will be provided to the National Museum of Natural History, Smithsonian Institution to add to their collections. Those requiring expert identifications to complement genomics will be sent to professional taxonomists for identification.
- 6) Histological data: fixation and tissue processing information will be recorded in lab notebooks and transcribed into Microsoft Word files. Images will be stored as digitized files. All digital data will be stored on hard drives and backed-up onto accessory hard drives and servers.
- 7) Representative images of collected animals will be posted immediately on the expedition webpages (cruise blog; the WhaleTimes *Creep into the Deep* webpage, *UndertheScope* webpage). The images will be low resolution images only, but contact information will be provided on the webpages, and high resolution images will be freely provided for any educational purposes.
- 8) All DNA barcoding data (16S and COI) will be uploaded to GenBank with accompanying metadata. Raw illumina reads (from gene expression studies) and output from Trinity assemblies will be uploaded to NCBI's SRA (Sequence Read Archive), which is standard practice for NGS data and allows download in their native format. NCBI primary data (i.e., sequences) are archived in perpetuity, although metadata can be edited. Analysis of large datasets will be processed on the FIU high performance computer (see **FIU Facilities**).

Analyzed data will be presented at yearly conferences and submitted for publication in peer reviewed journals. All techniques/protocols used will be available in the publications. Once published, original data not already available through PANGAEA or as supplementary material with publications will be available to any researcher requesting access to them. Some of the animals collected on the research cruises will be used by the PIs to enhance their deep-sea biology, marine biology, invertebrate biology and sensory physiology lab courses, as well as public lectures and lectures to various school groups.

All original data will be retained in the laboratory of the collaborator collecting the data. For shared experiments, all investigators involved will retain copies of all data pertaining to that experiment. Should key personnel choose to leave the project before its completion, their data will be made available to another investigator involved in the project, or another investigator with expertise in the same field who is interested in completing the project. Any publications coming out of this data transfer will be a shared publication between the labs contributing data to the project.