

Data Management Plan of the RAPID-NSF project: The impact of Hurricane Maria on the mesophotic reefs of Puerto Rico

Data Policy Compliance

- The project investigators 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*.

Description of Data Types

The project will produce several observational and experimental datasets, described in the list below. Observational data will be collected during >60 shallow and mesophotic dives in the southwest Puerto Rico during Fall 2017-Summer 2018). Experimental data will consist of DNA sequences from corals and dinoflagellate symbionts collected during these dives and isotope chemistry from sediment cores.

(1) Observational Datasets

(a) Photographic Data

These include the photographic records from phototransects and photo/video records from the broad qualitative surveys (e.g. species and frequency of overturned corals or other noticeable disturbances). The phototransects and broad surveys will take place from December 2017 to June 2018. For the phototransects and video records, metadata will include (e.g., date, time, depth, location, photographer, area within (for transects), etc.). Repository: Biological and Chemical Oceanography Data Management Office (BCO-DMO).

(b) Samples/Specimens

Specimen samples consist of coral fragments, which will be processed for coral and *Symbiodinium* genetic analyses. Samples will be collected during the Fall 2017 to early Spring 2018. Voucher specimens will be deposited for all individuals for which sequences are submitted in the Museum of Marine Invertebrates, Isla Magueyes Marine Laboratories, an NSF-funded facility. Accession numbers will be deposited to BCO-DMO.

(c) Physical Measurement Data

These data will be obtained from the instrument array provided by the USGS Pacific Science Laboratory, and the ADCPs already in place. Data will be recorded from July 2017 to December 2017. The USGS laboratory is a federal laboratory and all of the oceanographic data will be deposited in ScienceBase, which is part of Data.gov. The data will be hosted there in perpetuity, will be assigned a DOI, and all links for the data will be deposited to BCO-DMO.

(d) Photographic and Physical Measurement Data from Previous Work.

These will be submitted to the National Oceanographic Data Center's Coral Reef Information System (CoRIS) and all links for the data will be deposited to BCO-DMO.

(2) Experimental Datasets

DNA sequences and SNP data will be generated to analyze the connectivity of two coral species (*A. lamarcki* and *M. cavernosa*) and the clade composition of *Symbiodinium* in these two host coral species, respectively. Sequencing will be performed at Duke University Genomics Center or other facilities that offer competitive prices. Sequences and SNP data will be deposited in the National Center for Biotechnology Information's GenBank and Dryad. Accession numbers will be deposited to BCO-DMO.

Data and Metadata Formats and Standards

Field observation data will be stored in .jpg (phototransect data), in .mp4 files (video files), and flat ASCII files (ADCP data), and XL sheets will be generated from the community analysis, which can be read easily by different software packages. XL sheets will also be generated for the stable isotope work. DNA files will be stored as .txt files. Field data will include date, time, latitude, longitude, and depth. Metadata will include descriptions of collection and analysis procedures.

Data Storage and Access During the Project

The investigators will store project data (including spreadsheets, ASCII files, images, and DNA sequences) on desktop computers that are backed up by external hard drives. PI has accounts with the commercial cloud storage companies Carbonite (unlimited storage) and Dropbox for 1 TeraByte for back up storage. Lead PI Schizas will coordinate with the University of Puerto Rico Computing facilities for storing externally all produced data.

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

Data produced by this project may be of interest to geological, biological oceanographers, coral reef scientists and management agencies (Department of Natural Resources Puerto Rico) interested in the role of major disturbances shelf edge ecosystems. We will adhere to and promote the standards, policies, and provisions for data and metadata submission, access, re-use, distribution, and ownership as prescribed by the BCO-DMO Terms of Use (<http://www.bco-dmo.org/terms-use>). DNA sequences will be deposited in the National Center for Biotechnology Information (NCBI) database GenBank and Dryad, including bioinformatics pipelines upon submission of manuscripts.

Here are examples of publicly accessible data sets from various types of data produced in the lab of lead PI Schizas: Morphbank Image Repository (specimen Images 853168- 853511), GenBank (accession numbers KX140056–KX140702), a BOLD dataset (DS-PARG2016), transcriptome Shotgun Assembly (DDBJ/EMBL/GenBank accession No. GFAS00000000), SNP and microsatellite data deposited in Dryad Digital Repository (<http://dx.doi.org/10.5061/dryad.8gg2p>). Here is an example of a publicly available data deposited and archived by USGS personnel Stollarzi: <https://www.sciencebase.gov/catalog/item/58c07cc8e4b014cc3a3bf868>

Plans for Archiving

The PIs will work with BCO-DMO staff to ensure that genetic, biological, physical, geological data are archived appropriately and that proper and detailed documentation are archived along with the data. Stollarzi from USGS will archive the Oceanography data to ScienceBase.

Roles and Responsibilities

Each PI will be responsible for sharing his data with other PIs in a timely fashion, especially for a short period project such as this one. Sherman will be responsible for coordinating all mesophotic dive activities, photo transects, and retrieval of oceanographic instruments. Weil will be responsible for coordinating and making all data available from shallow dives, photo transects, and retrieval of oceanographic instruments. Senior personnel Ruiz will be responsible for the analysis of phototransect data from past effort as well new phototransect data with the chainlink method. Ruiz and Weil will coordinate the distribution of the phototransect analysis to the other PIs. Cruz Motta will be responsible for the analysis and the availability of analysis pipelines for the ecological data. Stollarzi and Schmidt will be responsible for the availability of all physical oceanographic data. Schizas will oversee the genetic work and will submit the resulting sequences to the GenBank database and Dryad upon submission of manuscripts. All bioinformatic pipelines will be made available to Dryad and BCO-DMO. Schizas, will coordinate the overall data management and will submit with the help of each coPI, the project data and metadata to BCO-DMO. Stollarzi will coordinate the submission of the Oceanography data to ScienceBase and assist Schizas with the submission of the appropriate links to BCO-DMO.