

## **Data Management Plan of the RAPID-NSF project: Microbiome and population dynamics in SCTLD-infected corals in 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 by photographing tagged colonies will be collected during >40 shallow dives in the southwest Puerto Rico during Fall 2019-Summer 2020. Experimental quantitative data will consist of 16S rDNA sequences from coral microbiomes collected during these dives. Field quantitative data of disease prevalence, incidence and mortality will be collected from surveys of our permanent transects during the project. Mortality rates will be generated from quantitative data collected from time-series photographs of tagged, diseased colonies.

#### **(1) Observational Datasets**

##### **(1a) Photographic Data**

These include the photo/video records from the broad qualitative surveys (e.g. species and frequency of diseased corals) and time series photographs of tagged colonies. The video transects generated from the UW drone represents extended and broad surveys over the reefs and shelf edge off La Parguera. Metadata will include (date, time, depth, location, photographer, area surveyed, etc.). Repository: Biological and Chemical Oceanography Data Management Office (BCO-DMO).

##### **(1b) Genomic Samples/Specimens**

Specimen samples consist of coral fragments, which will be processed for microbiome genetic analyses. Samples will be collected during the Fall 2019 to June 2020. 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.

##### **(1c) Physical Measurement Data (Temperature)**

These data will be obtained from Hobo loggers that will be deployed during the course of this project. Temperature data has been recorded continuously from 2003 to the present and will be augmented until the end of this project. The temperature data will be deposited to BCO-DMO.

#### **(2) Quantitative data sets**

##### **(2a) Genomic experimental data**

16S rDNA sequences of microbiomes will be generated from the targeted scleractinian species. Sequencing will be performed at a Genomic facility negotiated through Genohub that offer competitive prices and outstanding quality and service. Sequences will be deposited in the National Center for Biotechnology Information's GenBank and Dryad. Accession numbers will be deposited to BCO-DMO.

##### **(2b) Epizootic and mortality data**

Disease surveys will be done along our permanent transects (12 per reef). Quantitative data will be generated on disease prevalence and incidence along twelve 20m<sup>2</sup> band transects in each of the six reefs in the LPNR.

### **Data and Metadata Formats and Standards**

Field observation data will be stored in .jpg (photographic data), flat ASCII files (ADCP data). 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 projected data (including spreadsheets, ASCII files, images, and DNA sequences) on desktop computers that are backed up by external hard drives. Co-PI Schizas has accounts with the commercial cloud storage companies Carbonite (unlimited storage) and Dropbox for 1 TeraByte for back up storage. PIs 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 biological oceanographers, coral reef scientists and management agencies (Department of Natural Resources Puerto Rico) interested in the presence and/or progress of SCTL in Puerto Rico. 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 co-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>).

### **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.

### **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. PI Weil will be responsible for coordinating all dive activities, photo transects, and tissue collections. Weil will be responsible for coordinating and making all data available from surveys and individual colony assessments. Weil and Motta will be responsible for the analysis of transect 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. Weil will coordinate the overall data management and will submit with the help of each co-PI, the project data and metadata to BCO-DMO.