

DATA MANAGEMENT PLAN

Our project has 4 main objectives:

- 1) Environmental sampling and urchin collection for microbiome and Metabolomic analyses. At each site, one water, and one sediment sample per visit are to be collected as well as 5 diseased and 5 healthy urchins (expected, depending on the animal density at each area).
- 2) Microbiome/Metabolite comparisons between healthy and diseased animals.
- 3) Demographic data collection.
- 4) Environmental characterization of study sites.

The Principal Investigator (PI) will be responsible for participants' safety on a daily basis and together with the Co-PI and team will guarantee that all permits (including sample collection, as well as management of samples in the lab (IACUC and IBC – biosafety) are acquired and maintained. The PI will evaluate the progress of the study, to review procedures for maintaining the quality of data collection, management, and analyses. Dr. Godoy will be coordinating the biorepository of samples (soil, water and urchins) in her lab and samples will be stored at -80C. Metadata will accompany all the biorepositories and will be available to be submitted to the BCO-DMO project. All data (including genomic, metabolomic, and environmental data) will be available within two years of collection.

Sample collection of Physical Specimens

During the course of our project, we will collect environmental data, water and sediment samples, as well as sea urchin samples at all sample sites.

Demographic measured

We will select four reefs along the northeastern and eastern coast of Puerto Rico. Cerro Gordo reef (CGD; 18°29'05.9"N 66°20'21.0"W), Dorado Beach (DBE; 18°28'36.2"N 66°17'52.4"W) and Playa Azul reef (PAZ; 18°23'04.7"N 65°42'56.2"W) are fringing reefs located in the northern coast of Puerto Rico. At each reef, we will set up eight transects (10 x 2m each) to estimate sea urchin density (# of individuals/m²). We will also measure the test size (in cm) of no less than two sea urchins (yet the final number measured will depend on the specific density at each reef). Each reef will be visited four times (at three-month intervals) for one year.

Coral reef metrics

Coral and algal % cover will be estimated by placing a 1m² quadrant along the 10m transects. Reef rugosity (topographic complexity) will be assessed at each transect by counting the number of links in a 2m chain. Demographic and coral reef measurements will be estimated every visit (at three months intervals). Data will be associated to the metadata and available to the BCO-DMO.

Environmental measurements

At each site, Water Temperature (C°) will be measured by placing a data logger device. The device will be programmed to record the temperature every hour for three months. Likewise, pH, [-log (H⁺)], dissolved O (mg/L), and conductivity (µS/cm) will be measured using devices programmed to record every hour for three-month intervals. Calibration will be performed according to the manufacturer's instructions.

Total measurements per environmental factors, 24 measurements per day for 90 days total. The environmental devices will be retrieved at each visit.

Dr. Godoy will be coordinating the biorepository of samples (soil, water and urchins) in her lab and samples will be stored at -80C. Consistent data collection will be facilitated with the use of detailed operating manuals that will be written for this protocol.

Molecular data

Microbiome data: 16S rRNA sequence data and metagenomic data will be uploaded to QIITA and will also be available in the BCO-DMO Project. This will allow any researcher to replicate the data analyses. Oversight of the microbiome data analyses and metabolomic data integration will be provided by the Co-PI, Dr. Godoy.

Metadata will accompany all the biorepositories and will be available to be submitted to the BCO-DMO project. All data (including genomic and metabolomic, and environmental data) will be available within two years of collection and immediately once a manuscript has been submitted for publication. Field collections of water, sediment, and sea urchin gut specimens and genomic DNA will be made available to the Center for tropical Biodiversity, PR Science Technology and Research Trust, the Microbiota Vault and BCO-DMO. Loans to future collaborating institutions will be obtained based on individual formal agreements. All data is made accessible and open source through collection databases and gene data in public repositories with its metadata like QIITA. Genomic data and its links will be provided directly to the BCO-DMO Project. Metadata and data will be sent to the Biological and Chemical Oceanography Data Management Office, during the project and the most within two years of data collection.

Microbiome data analyses will be done by Dr. Godoy and her graduate student, and will be made public to the NCBI and EBI via QIITA <https://qiita.ucsd.edu/> and BCO-DMO. The graduate student and other students in the project will be responsible for compiling safety reports, which will be sent to the PIs. The content of the report will include a summary of progress including a discussion of issues or problems, a project organizational chart including a projected timetable and schedule.

We follow the earth Microbiome standards for metadata formatting and metadata Ontology (EMPO). The EMP Ontology (EMPO) is built upon existing ontologies and is used to assign samples to habitats in a hierarchical framework that captures major axes of microbial community diversity. The specific sample metadata will also be made public to QIITA and BCO-DMO.

Metabolomic Data

A list of metabolites detected for each sample will be available along with the metadata in the BCO-DMO project with the associated metadata.