

DATA MANAGEMENT PLAN

Data policy compliance

The project investigator will comply with the data management and dissemination policies described in the NSF Award and Administration Guide and in the NSF Division of Ocean Sciences Sample and Data Policy.

Types of data and curriculum materials

Data: Expected data types include:

1. Oceanographic data collected during the fieldwork component of the study from six sites located in Oahu, Hawaii (Objective 1). Environmental data will include temperature, salinity, nutrients and irradiance. Temperature, salinity, and irradiance will be recorded every hour for two years. Nutrient samples will be obtained in duplicates four times a year for two years.
2. Microplastic abundance and chemical identity in water, sediments and corals (IR spectroscopic data). Phthalates in water, sediments and corals (Liquid Chromatography-Mass Spectrometry).
3. Biological data for symbiont and coral host. For Objectives 1 and 3, we will obtain histological (gamete development, fecundity), lipid, and fatty acid data. For objective 2, we will obtain ingestion and adhesion rates in bleached and non-bleached corals using different types of plastics. For objective 3, we will measure chlorophyll, dinoflagellate density, photosynthetic and respiration rates in bleached and non-bleached colonies, and carbohydrates and proteins in coral tissue. For objective 4, we will obtain electron microscopy pictures (SEM and TEM) and data on coral synchrony, microplastics and phthalate concentrations in coral eggs, and fertilization rates.
4. Transcriptomic data: Tag-sequencing data will be produced for *Montipora capitata* in lab experiments (Objective 3 and 4).
5. Derived data: Results of statistical analyses (i.e. WGCNA, PCA, ANOVA) will be recorded in R-markdown/spreadsheet form and published in peer-reviewed journals.

Curriculum Materials: Educational materials developed from outreach activities in Hawaii will include (1) materials for volunteers involved in the citizen science project (sampling design and data collection field booklet and online and in-person training materials); (2) materials from workshop with elementary teachers from Hawaiian emergent schools; (3) science kits for Hawaiian teachers that attended the coral reef education workshop; and (4) lesson plans on climate change and plastic pollution and field activities (geo-inquiry) for middle school students from Hawaii.

Data format, access and public release:

Processed experimental data from objectives 1-4 will be stored as Excel spreadsheets or text files. Transcriptomic data will be stored in fastq.files. All data will be stored and backed up on different external hard drives and cloud platforms (google, dropbox, code42). All mass spectrometry raw files collected through this project will be uploaded onto the <https://chorusproject.org/> website under the publicly available database and listed under the "Padilla-Gamino lab". Associated location and preparation details will be provided with the analysis files. All raw files will be accessible and downloadable by anyone at any time. All files will be uploaded within 2 weeks of collection. The Chorus Project was created for the sole purpose of providing research scientists and developers with the ability to store, analyze, and share their mass spectrometry data. The Chorus application is available on the cloud to anyone and can be accessed for free by the scientific and general public, allowing the dissemination and sharing of data no matter the original raw file format or mass spectrometry platform used to generate the data. All mass spectrometry data files relevant to the project will be released in peer-reviewed publications. Microplastics spectral libraries files will be easily accessible (.txt file format) through the Padilla-Gamiño lab website.

Data and metadata will be formatted using the guidelines suggested by Borer et al. (2009) that emphasize the use of a scripted statistics program (i.e., SAS Statistical Software); a read-only file for original data with copies of the file made for further statistical analyses; and an accompanying read.me text file for notes that support the original data file. These procedures will ensure that the data will both be easy-to-access for users and will not become inadvertently modified when used. Coral data will be

standardized using coral ash-free dry weight and surface area, depending on the parameter, to allow for comparison to the published literature related to coral physiology.

Curriculum Materials: All curriculum-resources that can be provided via the internet will be available on Dr. Padilla-Gamiño's website at: <http://padilla-gaminolab.weebly.com>. This lab website will also have the data generated by the citizen science project (.xls files) so that it is available to the UW research team, students, other research groups, decisionmakers, and the public.

Our Policy of Open Science

The PI on this project is well aware of the need for broad sharing of scientific information and will include multiple activities to provide unrestricted access for other researchers and the public as final results are available. All data generated from this project will be contributed to the Biological and Chemical Oceanography Data Management Office (BCO-DMO). 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>). We will upload published and unpublished mass spectrometry data within two weeks of collection onto Chorusproject.org for free public download, viewing, and analysis. We aim to publish our data in peer-reviewed scientific journals following the timeline in the project description. We will make publications "open access" so that they are easy to acquire and accessible to a broader community. Pipelines and scripts used in data analyses will be made available in github and lab website. All members of the Padilla-Gamiño lab will be trained in data management tools as part of their research program.

Archiving data and samples

Data:

The final repository for the oceanographic data (physical, chemical, and biological, as previously described) from the three sites in Kaneohe Bay, Hawaii will be the National Oceanographic Data Center (NODC) in Silver Springs, MD. In addition, biological data collected during the study (coral reproductive, transcriptomic, and physiological data as previously described) will be submitted to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) as the primary data management source for this study. Transcriptomic data will be archived in the NCBI databases.

Samples:

Any remaining phenotypic-trait samples (those not consumed during the analytical procedures) will be archived in the Padilla-Gamiño lab: whole coral samples, eggs, and larvae will be frozen at -80°C suitable for additional physiological analyses, coral skeletons will be dried suitable for additional biogeochemical analyses, and fixed samples will be preserved in ethanol suitable for additional histological analyses. These archived samples will be listed on the PI website and made available, upon reasonable request, to other investigators.

REFERENCE

Borer ET, Seabloom EW, Jones MB, Schildhauer M (2009) Some simple guidelines for effective data management. Bulletin of the Ecological Society of America 90 (2): 205-214.