

## Data Management Plan

### I. Title and contact information

1- *Project title*: “Investigating Structural Changes In Reef-Associated Biodiversity Along A Natural Gradient In Ocean Acidification”

2- *Points of contact*: Overall data products and molecular data: Laetitia Plaisance, [plaisancel@si.edu](mailto:plaisancel@si.edu) (202-633-0684) and Nancy Knowlton, [knowlton@si.edu](mailto:knowlton@si.edu) (202-633-0668)

### II. Data and metadata standards

Data Policy Compliance: The research products generated in the proposed project include genetic, morphological data in the form of specimen vouchers and photographic data. We will cooperate and collaborate with all appropriate data repositories and the Biological and Chemical Oceanography Data Management Office (BCO-DMO; [www.bco-dmo.org](http://www.bco-dmo.org)) as appropriate, and we will adhere to the data sharing philosophy. The details on the use of these repositories are presented below.

#### A: molecular data:

Laboratory-based DNA sequencing will be conducted. Records will comprise voucher-based DNA sequencing of the cytochrome oxidase subunit I (COI) gene and next generation sequencing of the mini-COI barcode for bulk communities. Molecular data will be stored in different formats (such as tab-delimited, fasta, fastq, biom, xls) to ensure compatibility between different programs and databases. Sequence data of individual vouchers will be submitted to GenBank along with relevant metadata (date and location of collection, including a unique identifier specifying the settlement structure with which the specimen was associated). Metabarcodes from bulk preserved whole communities will be deposited into the Dryad repository along with relevant metadata. The complete data product from the proposed work will be made available through the BCO-DMO.

#### B: Morphological data:

All sequenced individuals will be deposited as vouchers at the Museum Support Center facility of the National Museum of Natural History, Smithsonian Institution. The genomic DNA for both the individuals and whole communities will be deposited in the Biorepository, Smithsonian Institution.

#### C: Metadata:

Metadata associated with the preserved vouchers and photographs (collection information as above, GenBank numbers) will be stored in the museum’s centrally maintained database, Ke-EMu. Contextual data collected prior to the start of this grant (sampling information: sampling time, horizontal and vertical location, oceanographic metadata, technology, pH conditions, along with other metadata standards) will be presented in excel spreadsheets and submitted to the Biological and Chemical Oceanography Data Management Office (BCO-DMO)

### **III. Data access and sharing**

The data produced during this project will be released for public use one year after the end of the grant or with publication of results, whichever is sooner. Specimens and associated metadata will be available as soon as they are accessioned into the NMNH collections before the end of the project's second year. BCO-DMO will make the data publicly available through their system immediately upon receipt and subject to any approved embargo period. Data may be parsed into different subsets that work with the format of the data. The molecular datasets will be split into different files corresponding to the sequencing technology used (sanger sequencing vs next generation sequencing). All datasets for the project will be connected such that anyone looking at the project will be able to link the different information together.

### **IV. Policies and provisions for re-use or re-distribution**

There will be no permission restrictions on data and they will be open for sharing and re-use upon publication.

### **V. Plans for archiving and Preservation of access**

Data and research products will be archived to publically accessible databases and scientific journals. Extracted DNA will be stored in -80°C freezers in the Biorepository of the National Museum of Natural History. Organism vouchers will be stored at the National Museum of Natural History's collections of the Museum Support Center. Metadata (for genomic DNA, vouchers and photographs, as well as GenBank and Dryad numbers) will be stored in the Museum's centrally maintained database, Ke-EMu. Data will also be stored in excel files to be provided as electronic supplementary information associated with publications.