Data Management Plan

Products of Research: A tank experiment will be conducted on visibly bleached and visibly non-bleached corals of *Montipora capitata* and *Porites compressa* from two sites in Oahu, HI for 1 year with corals reared under 4 treatments: ambient $p\text{CO}_2$ without feeding, ambient $p\text{CO}_2$ with feeding, elevated $p\text{CO}_2$ without feeding, and elevated $p\text{CO}_2$ with feeding. Data collected from this experiment will include the following: coral total lipids, lipid classes, total soluble protein, total soluble carbohydrates, biomass, Chlorophyll *a*, endosymbiont cell density, photosynthesis, respiration, feeding rates, and metabolic carbon budget data (i.e., CZAR, CHAR, CTAR). In addition, any leftover sample material will be archived and stored in Grottoli’s laboratory at -80°C. Typically, 25% of each coral fragment, most of the skeleton, and portions of all coral slurries remain after analyses and are archived. Archived materials have been used by Grottoli’s group and in collaboration with others to produce new research findings that were not originally part of the funded project (i.e., Grottoli & Rodrigues 2011; Killbourne et al 2011; Wu & Grottoli 2010; Lavigne et al 2008, 2010; Hoenish et al 2004) and provided material for three masters and two senior thesis projects (i.e., Baumann 2013; Chrystal 2008; Wu 2004; Ringwald 2012; Borg 2010). All remaining samples will be archived and made available for collaboration. Transcriptomics of bleached and unbleached corals are described under genetic data below.

2 & 3. Data Format, Storage, and Preservation: Data will be saved in excel spreadsheets, as comma delimited DAT files. DAT files will be used because they are a standard format that can be used on a variety of database and software programs including Microsoft excel. All data will be stored using standards developed for widespread usage and will include date (month, day, yr., hour, sec), treatment, species, and other relevant details. Data will initially be stored on a 1 TB WB backup drive in Grottoli’s lab and on the School of Earth Sciences secure server. For long-term storage, the data will archived with the Biological and Chemical Oceanography Data Management Office (BCO-DMO) (http://bcodmo.org/).

Genetic data: All transcriptomic datasets generated by this project will be archived with a standardized meta-data table outlining original data generator(s), associated publications, and details about sampling sites, experimental history, etc. As the input file formats for each of the analyses are often different, and will be generated for each species, they will also be linked to each raw dataset. CoPI Toonen will be primarily responsible for creating and maintaining our data storage and sharing system based on our in-lab Synology DiskStation RAID 5 array (DS1513+ 5300 NAS) with 15TB of hot-swappable mirrored disk space that provides biweekly backups of the database. There is currently no universal standard format for storage of genetic data, but Toonen is currently co-leading a NESCent working group to standardize archiving format for population genetic data and develop a shared database for permanent archiving (RCN DIPnet for database development currently under consideration by DEB).

4. Data Dissemination & Policies for Data Sharing and Public Access: All data will be made publicly available through the BCO-DMO site and directly upon request following publication and after the embargo period for respective journal publications. After the project is completed all data will be made publicly available on the open access BCO-DMO repository within 2 years of project completion to allow sufficient time to prepare manuscripts. For genetic data sharing, the DIPNet system is based on three essential aspects: (1) a specimen-based metadata system based on global standards, (2) globally unique and persistent unique keys based on EZIDs (http://ezid.cdlib.org/) - a form of globally unique identifier, and (3) genotypes and metadata stored locally and sequences stored in GenBank, using BioProject, NucleotideDB, and the Short Read Archive to make data useful and publically available.

5. Roles and Responsibilities: The data management will be the responsibility of PI Andrea Grottoli who will work with BCO-DMO to insure that the data submission and archiving plan is followed. Data will be made openly available to all collaborators and responsibility for the data will be transferred to BCO-DMO following completion of the project. Grottoli will also be responsible for the archiving and preserving of any unused sample material and making them available for collaboration. Toonen will be responsible for management, archiving and sharing of the transcriptomic data.