

DATA MANAGEMENT PLAN

Carly D. Kenkel (Univ. of So. California), Ross Cunning (Shedd Aquarium), John Parkinson (Univ. of South Florida), Andrew Baker (Univ. of Miami), Craig Dahlgren (Perry Institute of Marine Science)

(1) Products of Research

- 1) Environmental data (temperature log files)
- 2) Raw underwater data collection sheets and coral photographs
- 3) Snap-frozen and ethanol preserved tissue samples
- 4) Processed tissue samples and aliquots
- 5) DNA and associated sample derivatives (PCR products, sequencing libraries)
- 6) Raw physiological trait data (.csv spreadsheets and metadata)
- 7) Raw sequencing data (FASTQ files and metadata)
- 8) Processed sequence data (.csv tables of SNP counts)
- 9) Photoscan 3-D model output files
- 10) Annotated statistical analysis scripts (e.g. .R and .py files)
- 11) Results from analyses (summary files and figures)
- 12) Written reports and manuscript drafts

(2) Data Format

In-water and laboratory protocols and field-collected organismal and environment data will be stored in notebooks and/or folders that remain at all times in the respective PI's lab, and also recorded and archived in electronic form (PDFs, text files or .csv spreadsheets). Tissue samples and processed aliquots, including DNA and derivatives, will be stored in labeled boxes at -80C or -20C as appropriate, and storage information will be recorded on the respective lab freezer organization electronic spreadsheets. Raw sequence read data will be stored as compressed FASTQ files. Raw physiological trait data will be stored as .csv spreadsheets. Processed sequence data will be stored as tab-delimited text files and/or .csv spreadsheets. Statistical analyses data will be stored as tab-delimited text or .csv input files and R source files. Results summaries will be stored as Word documents and PDF figure files.

Prior to publication, data stored in notebooks will be kept strictly in labs at the respective institutions. Digital data will be stored in four locations: 1. Personal computers; 2. Backup discs; 3. Respective lab file storage / ftp servers; 4. Cloud file storage through USC's unlimited Google Drive or OneDrive for Business account.

(3) Access to Data, Data Sharing and Data Archival

Tissue subsamples of all experimental corals will be archived in PI Kenkel's lab following best practice recommendations developed at the Coral Bleaching RCN Workshop on sample preservation in July 2020 (OCE-1838667).

Manuscripts will be uploaded to open-access pre-print servers, such as BioRxiv, prior to formal review and publication. All primary data that have been included in the results section of the published manuscripts and statistical analysis scripts sufficient to reproduce reported results will be submitted to the journal as Supplemental Information or archived in open-access databases, such as GitHub. PI's Kenkel (under ckenkel) and Cuning (under jrcuning) are regular contributors to this platform.

Raw sequencing data will be permanently archived and released within one year of sequencing through NCBI's Sequence Read Archive. The release of the data will be announced through the Coral-List email server maintained by NOAA and through social media platforms.

Final peer-reviewed journal manuscripts and supplemental information, such as data tables for graphical information in manuscript figures, will be posted in the USC Digital Repository and will be available on this public website no later than 12 months after publication. These records will be durable, accessible through web protocols and made safe from tampering or falsification. The storage media will be updated as necessary to keep it current.

The Location of all datasets will be registered with the Biological and Chemical Oceanography Data Management Office (BCO-DMO) as soon as they are made publicly available or within one year of project completion, providing links to the locations of specific datasets.

(4) Re-Use, Re-Distribution and Production of Derivatives

Primary data, such as raw coral samples and their derivatives (e.g. DNA/RNA), and other supporting materials will be available for unrestricted use to all academic researchers upon request. USC's policy is to encourage, wherever appropriate, research data to also be shared with the general public through internet access. Terms of use will include requirements of attribution along with disclaimers of liability in connection with any use or distribution of the research data.