DATA MANAGEMENT PLAN

Data policy compliance
The project PIs will comply with data management and dissemination policies described in the NSF Award and Administration Guide (PAPPG, Chapter XI.D.4) and the NSF Division of Ocean Sciences Sample and Data Policy. They will work with Biological and Chemical Oceanography Data Management Office (BCO-DMO) to properly manage and archive data.

The Data Sharing Plan (DSP) will be used as a guiding document for managing and sharing all types of data in this project. This plan is not only critical for sharing data among collaborators during the project execution, but also for sharing with resource managers and the broader research community after the completion of the project. Data collected and generated during this project will be made available to the community in a timely manner with formats that are widely used and accessible. Stringent Quality Assurance/Quality Control (QA/QC) evaluation will be performed prior to data archiving and sharing. The DSP will be distributed to all project participants at the beginning of the project so that the data will be organized and managed in a consistent manner throughout the project. Our goal for data management is to ensure accurate and precise data collection with associated metadata, careful backup and local archiving of data, and long-term data archiving through publicly available sites.

Description of data types
The data generated through this project include observational and experimental datasets, as well as sequence data, described below. Data will be collected throughout the project period.

1. Water quality data (observational): Data will be collected with YSI EXO2 Multiparameter sonde and PME MiniDOT loggers. Data will include environmental measurements (e.g., depth, salinity, dissolved oxygen, pH, temperature). File types: .csv and tab-delimited text files. Repository BCO-DMO.
2. Event logs (observational): Sampling event log will include start & end times, dates, and locations of instrument deployments and casts. Will be recorded on paper log sheets. File types: .xlsx, .csv, tab-delimited text files, pdf scans of log sheets. Repository BCO-DMO.
3. Benthic field surveys (observational): Will include data on identity and abundance of corals. Will be recorded in situ on underwater paper. File types: .xlsx, .csv, tab-delimited text files, pdf scans of data sheets. Repository BCO-DMO.
4. Coral performance data (observational and experimental): Will include data on photosynthetic rates, pigment concentrations, and zooxanthellae concentrations. Will be recorded in laboratory notebooks or by PAM fluormeter. File types: .xlsx, .csv, tab-delimited text files, pdf scans of notebooks. Repository BCO-DMO.
5. Microbiome amplicon and metagenome libraries (observational and experimental): DNA sequences from microbiome samples collected from corals in field and laboratory experiments. Sequencing will be conducted at the University of Florida Interdisciplinary Center for Biotechnology Research on an Illumina MiSeq or NextSeq. File types: .fastq. Repository: NCBI's Sequence Read Archive; accession numbers to be provided to BCO-DMO.

Data and metadata formats and standards
The majority of the data will be created and stored as plain text files as they are easily portable between data processing software programs (e.g. Matlab, Excel, etc), can be read on any computer platform, and are suitable for uploading to web-based databases. Metadata needed to support our field and laboratory observations include variables such as time and day of year, units of measure, accuracy and precision of measurements, methods of measurement and sampling, investigator, and data processing protocols. Metadata will be entered into a computer spreadsheet program and stored as ASCII text. Metadata standards and practices, including availability of metadata in machine-readable formats (e.g., XML following Ecological Metadata Language- EML standards), for the project will be developed in consultation with the University of Florida metadata standards and discipline-specific liaison librarian (http://guides.uflib.ufl.edu/c.php?g=147756&p=966438).

Data storage and access during the project
The investigators will store project data (including text, spreadsheet, ASCII files, and PDFs of scanned documents) on laboratory computers that are backed up on external hard drives daily and offsite at the University of Florida's shared network drive GatorCloud-OneDrive@UF (https://it.ufl.edu/services/search/gatorcloud) which will also be used for sharing data among project participants during data collection and analysis phase. The estimated data volume for the project is 250 GB, with sequencing data contributing the vast majority of that content.

Mechanisms and policies for access, sharing, re-use, and re-distribution
All data will be made publicly available within two years of collection. The raw data itself, or links to data in publicly available databases, will be made available through the BCO-MDO repository. The project investigators will work with BCO-DMO data managers to make project data available online in compliance with the NSF OCE Sample and Data Policy. Investigators 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).

Data, samples, and other information collected under this project can be made publically available without restriction once submitted to the public repositories.
Data produced by this project may be of interest to biological oceanographers and ecologists studying the role of ocean deoxygenation in tropical ecosystems.

Interpreted data products will be incorporated primarily into peer-reviewed journal articles, and in Supplemental Information to journal articles, and will also appear in conference presentations. In the case of journal articles, pdf versions of the article will be posted in full compliance with publisher. The quality-filtered sequencing reads for 16S rRNA gene sequence data and metagenomes will be submitted to NCBI's Sequence Read Archive and links to these datasets will be published in open access journal articles. Access to R scripts and other bioinformatic pipelines used in the analysis of data will be provided through the Github repository.

Plans for archiving
The investigators will work with BCO-DMO ensure that project data are submitted to the appropriate national data archive, and to ensure data are archived appropriately and that proper and complete documentation are archived along with the data.

Roles and responsibilities
Each PI will be responsible for sharing his/her subset of data among the project participants in a timely fashion. R. Collin will be responsible for collecting and analyzing data from laboratory experiments. J. Meyer will oversee the molecular work and will submit the resulting sequences to the National Center for Biotechnology Information's (NCBI) Sequence Read Archive. The Lead PI, A. Altieri, will coordinate the overall data management and sharing process and will submit the project data, including NCBI accession numbers, and metadata to the BCO-DMO who will be responsible for forwarding these data and metadata to the appropriate national archive.