DATA MANAGEMENT PLAN

Data management and sharing are essential to the goals of the proposed research and the Broader Impacts of this proposal. Thus, PI Apprill is committed to working with the Biological and Chemical Oceanography Data Management Office (BCO-DMO), an NSF-funded data repository, to archive and make all data sources publically available. Apprill has spoken with BCO-DMO and is assured that they will be able to assist with archiving and storing data generated from this proposed project. Part of the budgeted salary time for Apprill during Y3 will be devoted to organizing, managing and publishing data from this project.

This project will produce a large amount of data in the form of:

1) Partial SSU rRNA gene sequence data from coral-bacteria and archaea generated by pyrosequencing.
2) Full-length SSU rRNA gene sequence data from coral-bacteria and archaea generated by Sanger sequencing.
3) Quantitative PCR data on general and specific microbial groups.
4) Micrograph images.

Raw data management and storage

Data will be recorded and managed using the database management software Microsoft Access, which will facilitate access for statistical analysis. The database and raw data from all of these analyses will be stored on computers with daily back-ups to a remotely accessed WHOI server.

Archiving and sharing SSU rRNA gene data

The pyrosequencing data will be submitted to the National Center for Biotechnology Information (NCBI) Sequence Read Archive depository (http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi). This submission will include the raw files and metadata about the sequences that conforms to MIMARKS, the Genomic Standards Consortium (http://gensc.org). The quality-controlled full-length bacterial and archaeal DNA sequences will be submitted to the NCBI environmental database (http://www.ncbi.nlm.nih.gov/genbank/) along with metadata. Additionally, representative quality-controlled sequences will also be made available in a custom coral-microbe ARB software database that Apprill will make available on her WHOI website, and also link this with BCO-DMO. Accession numbers for all sequences will be made available in respective publications.

BCO-DMO archival and integration of project data

PI Apprill will work with BCO-DMO to archive and integrate all data (SSU rRNA gene data, qPCR data and micrograph images) from this project. BCO-DMO will assist with archiving the processed quantitative PCR and processed, labeled micrograph data generated from this project. These data will be linked to metadata about the coral samples, specific experimental conditions (i.e., FISH experimental conditions), and also the NCBI accession numbers from the sequence data. **The idea is that for a single coral sample, all forms of microbial data generated on that sample will be collectively integrated.** Data archived with BCO-DMO are publically available and easily searchable, including by sample type, date of collection, type of data, and project name. Data will be submitted to the BCO-DMO by the completion of project Year 3. Links and references to the data sources will be available in associated publications and presentations.