Data Management Plan

All significant data, findings, and novel software from the proposed work will be promptly prepared and submitted for publication with authorship that accurately reflects the contributions of all those involved.

Types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project:
- *Illumina* RNAseq data for the functional genomics aim as raw data (long-term storage), as processed data, and as assembled cDNA contigs for analysis.
- Assembled cDNAs maintained in a MySQL database and used for comparative analysis as well as in flatfile format for distribution and other uses (e.g., Blast2go annotation).
- Metadata from the culture manipulations will be stored as Excel files and made available at the project web site.
- RNAi sequences and reagents will be made available upon publication.
- All antibodies developed for this project will be made available at the time of initial manuscript publication.
- Crystal structure data and coordinates for the CARPs will be promptly published.
- Online teaching and outreach materials.

Standards to be used for data and metadata format and content:
The **transcriptome data** (individual reads and assembled contigs) will be available at the Falkowski/Bhattacharya project web site and the raw data will be made available at http://www.ncbi.nlm.nih.gov/sra as has been done for other NSF-funded projects (e.g., RedToL [http://dblab.rutgers.edu/redtol/], *Cyanophora paradoxa* genome project [http://dblab.rutgers.edu/cyanophora/]).

Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements:
The *S. pistillata* cDNA contigs and predicted proteins will be made available as soon as they are produced and have been checked for quality at a project web site as well as related project information/accomplishments (e.g., http://dblab.rutgers.edu/redtol/index.html, http://dblab.rutgers.edu/cyanophora/). Project bioinformaticist Udi Zelzion will implement this web site. All data release will coincide with the publication of the initial analyses. An example data release page is shown for a recently published paper in *Current Biology* that was led by post-doc Cheong Xin Chan in the Bhattacharya lab (http://dbdata.rutgers.edu/data/plantae/). This web site will also provide access to all downstream analyses such as gene annotations (e.g., Blast2GO, NCBI) and output of phylogenomic analyses.

All novel computer programs developed for the proposed research to aid processing and analysis of data will be made freely available at the project web sites. Our results will be disseminated in presentations at scientific meetings and peer-reviewed journal articles, including a review paper on coral biomineralization at the end of the project.

Plans for archiving data, samples, and other research products, and for preservation of access to them:
Data generated by the research and related computer programs will be accessible through a link on the project web sites hosted by Falkowski and Bhattacharya. The URLs for the sites will be provided on all publications generated by the proposed work so that interested readers can

immediately go to the sites to access the information. All data, including raw data, will be either included as supplemental information in manuscripts submitted and approved for peer-review publications, or submitted to archival sites including GenBank (genome), discussed above.