DATA MANAGEMENT PLAN: Davies & Meyer-Kaiser

Data policy compliance. The project investigators will comply with the data management and dissemination policies described in the *NSF Award and Administration Guide* and the *NSF Division of Ocean Sciences Sample and Data Policy*.

Description of data types. The project will produce several observational and experimental datasets, described in the list below.
1. Environmental data: We will record light and temperature data from our study sites.
2. *In situ* transplanted recruit growth and survival data: Photos of coral recruits on terra cotta tiles will be recorded before transplantation, at 6 months and at one year to measure growth and survival (Aims 1a, 3a).
3. Growth and survival rate estimates will be made at different time-points for recruits following thermal challenge treatments (Aims 2a, 3b).
4. Genomic data: 2bRAD sequencing data will be produced for *P. lobata* parents and for recruits following transplantation (Aim 1b).
5. Gene expression data: Tag-sequencing data will be generated for corals and their symbionts following transplantation (Aim 2b).
6. Metabarcoding and microsatellite data: 16S Metabarcoding sequencing and fragment analyses of microsatellite loci will be used to characterize the microbial and algal symbiont communities associated with transplanted *P. lobata* (Aim 2c).
7. Coral tissue samples: Some of our analyses will result in ethanol-preserved or frozen *P. lobata* tissue samples. Tissue of coral recruits may be exhausted by the planned analyses.
8. Derived data: The results of statistical analyses will be recorded in text files, spreadsheets, or Rmarkdown documents and published in peer-reviewed journals.

Data and metadata formats and standards. Numerical and observational data (types 1-3 above) will be stored as text files or Excel spreadsheets. Genomic, transcriptomic, and metagenomic data (types 4-6 above) will be stored as fastq or chromatogram files. Photographic data (type 2) will be stored as TIFF files. Coral tissue samples will be frozen or preserved in 95-100% ethanol and stored in the Davies lab at BU. Metadata will be prepared in accordance with BCO-DMO conventions (i.e. using the BCO-DMO metadata forms) and will include detailed descriptions of collection and analysis procedures.

Data storage and access during the project. The investigators will store project data (including spreadsheets and text files) on both individual laboratory computers and institution servers that are maintained by Boston University's Shared Computing Cluster (SCC) and WHOI's Information Services group (PI Meyer-Kaiser has unlimited Google drive space through WHOI). Fastq files from 2bRAD, TagSeq, and 16S metabarcoding sequencing will be stored on the SCC, which has over a petabyte of space for long-term data storage. Adult coral tissue samples will be archived in the Davies Lab Stirling Ultra-low -80C freezer for a period of at least 5 years. All resulting DNA and RNA extracts for all tissue types will also be archived (when not exhausted for analyses) in the Davies lab for a period of at least 5 years.

Mechanisms and policies for sharing, access, re-use, and distribution. The repository for all data is the Biological and Chemical Oceanography Data Management Office (BCO-DMO), unless otherwise noted, following recommendations in the BCO-DMO Data Management Best Practices Guide (bcodmo.org/resources/). Data sets produced by the science party will be made available through the BCO-DMO data system within two years from the date of collection. The Principal Investigators will work with BCO-DMO data managers to make project data available online in compliance with the NSF OCE Sample and Data Policy. Data, samples, and other information collected under this project can be made publically available without restriction once submitted to the public repositories. We 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). We aim to publish our data in peer-reviewed international scientific journals in a timely manner following the proposed time frame in the project description. When possible, we will make publications 'open access' to allow for a broader community of researchers and the public to acquire manuscripts easily. Scripts and pipelines used for data analyses will be made available along with published manuscripts, on co-authors github pages or the lab websites. The interns, undergraduates, technicians, and postdocs in both labs will be trained to use these data management tools as part of their education.

Plans for archiving. The investigators will work with BCO-DMO to ensure that data are archived appropriately and with necessary documentation. Genomic and transcriptomic data will also be archived in NCBI databases (i.e., short read archive (SRA)). Metadata provided will follow the latest directives from the Genomic Standards Consortium (GSC) for "Minimal Information about a Marker Sequence" (MIMARKS). This is a curated standard for the acquisition and display of information associated with sample acquisition, processing, handling, sequencing, and analysis. These are community standards, agreed using consensus and updated where necessary by annual meetings of the GSC (www.gensc.org). In addition, these standards are recognized by the International Nucleotide Sequence Database Collaboration (INSDC) and reported by a keyword (GSC) for compliant sequences.

Roles and responsibilities. PIs Davies and Meyer-Kaiser will oversee data management for the science team and interface with the BCO-DMO staff. Davies will be responsible for genomic, metagenomic, and transcriptomic data, while Meyer-Kaiser will be responsible for experimental data and environmental data including images and derived data. In the case of personnel departure, responsibility for data management will fall to the remaining PI.