Data Policy Compliance: The project investigators will comply with the data management and dissemination policies described in the *NSF Award and Administration Guide* (AAG, Chapter VI.D.4) and the *NSF Division of Ocean Sciences Sample and Data Policy*.

Pre-Cruise Planning: Pre-cruise planning will occur through email, teleconferences, and a planning workshop with all co-PIs and involved students. This cruise will utilize a remotely operated vehicle (ROV JASON, Woods Hole Oceanographic Institution) and an autonomous underwater vehicle (AUV Sentry, Woods Hole Oceanographic Institution). Preliminary dive plans will be written prior to the cruise, and updated and amended prior to each dive based on cruise events.

Cruise event logging: Detailed dive reports for each ROV and AUV dive will be digitally compiled as Microsoft Word documents during the cruise. Each dive report will include site information, times of launch and recovery, sampling events, and watch-stander logs. These, along with each dive plan, participant information, and samples logs will be compiled into a cruise report (PDF document) at the conclusion of the cruise.

Description of Data Types

This project will produce observational datasets through the ship's underway sensors, ROV sensors and cameras, AUV sensors, and derived from analysis of collected biological specimens. Observational data will be collected on a south-western Pacific research cruise planned to take place during the fall, winter or spring months.

Cruise underway data: Standard underway data collected along the ship's track (e.g., sea surface temperature, salinity, etc.). File types: .csv ASCII files. Repository: BCO-DMO and the Rolling Deck to Repository (R2R).

ROV data: Routine sensor data (e.g., temperature), video, and images collected by the ROV Jason, as well as dive event logs, are made available through the Jason Virtual Van and archived at the National Deep Submergence Facility (NDSF) at Woods Hole Oceanographic Institution. File types: .csv ASCII, .jpg, .mpeg Repository: Virtual Van and NDSF archive

AUV data: Routine sensor data (e.g., CTD, dissolve oxygen) collected by the AUV Sentry is archived and made publicly available by the NDSF at Woods Hole Oceanographic Institution. File types: .csv ASCII, .ssc ASCII .mat Matlab. Repository: NDSF archive

Genomic and gene sequences: Microbial genomic, bacterial 16S rRNA gene sequences, and snail mitochondrial CO1 gene sequences will be derived from water/fluid samples, substrate samples, FRIE colonization devices, and metazoan larvae, juveniles, and adults.
16S rRNA gene sequencing will be performed at the Argonne National Laboratory, and genomic sequencing will be performed at the Georgia Genomics Facility at the University of Georgia. FISH probes will be designed for the *Alviniconcha* spp. symbionts: ε-proteobacterial symbiont, γ-1, and γ-Lau proteobacterial symbionts The resulting assembled and raw data will be made publicly available. File types: Short-read archive .sra, .fastq read files, .fasta files. Respository: NCBI; probeBase; accession numbers and associated sample metadata will be submitted to BCO-DMO.

Snail collections: Quantitative data on the abundance, weight, and size of snail species collected via the mussel pot benthic samplers will be collected at sea and recorded in Excel spreadsheets. After the cruise, genetic species ID data will be added to these datasheets after sequencing and analysis. File types: Excel file(s). Repository: BCO-DMO.

Larval samples: Larvae will be collected via the AUV Sentry and the SyPRID, benthic tube traps, and high volume pumps. Taxonomic identification data and photographs derived from fresh and preserved samples will be recorded in Excel spreadsheets. File types: Excel file(s). Respository: BCO-DMO.

Data and Metadata Formats and Standards: Genetic sequence data will be prepared in accordance with the minimum information about a metagenome sequence (MIMS) specifications and minimum information about a marker gene sequence (MIMARKS) developed by the Genomic Standards Consortium, and the Earth Microbiome Project standards and protocols for 16S rRNA gene amplicon sequencing. All other data will be stored in Excel spreadsheets, which can be submitted to the BCO-DMO. In addition, metatdata will be prepared using BCO-DMO metadata forms and will include detailed descriptions of sampling and analysis protocols.

Data Storage and Access During the Project:
The investigators will store project data (including Excel spreasheets, PDFs of dive logs, etc.) on a shared Google Drive directory that will be hosted through the University of Rhode Island's unlimited space Google Suite. In addition, this data will be stored on investigator's laboratory computers that are backed up by their institution's central IT organizations. Genetic sequence data will be kept on laboratory network attached storage and on the Brown University large data storage with automatic off-site backups associated with its high-performance computing cluster, which URI researchers have access to through an EPSCoR agreement.

Mechanisms and Policies for Access, Sharing, Re-Use, and Re-Distribution: Immediately after completion of the research cruise, underway data and metadata will be submitted to the Rolling Deck to Repository. Within four months of the conclusion of the cruise, the final cruise report will be submitted to the BCO-DMO. All other data will be submitted to the appropriate repository (detailed previously) upon submission of manuscripts, or 2 years post-cruise, whichever comes first. Genbank accession numbers will be submitted to the BCO-DMO after sequences are submitted to Genbank.

Data produced by this project may be of interest to chemical and biological oceanographers, particularly deep-sea biologists studying hydrothermal vent and seep microbial and invertebrate communities. 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), the minimum information about a metagenome sequence (MIMS) specifications and Minimum information about a marker gene sequence (MIMARKS) developed by the Genomic Standards Consortium, and the Earth Microbiome Project standards and protocols for 16S rRNA gene amplicon sequencing.

Plans for Archiving: Rolling Deck to Repository will ensure that the original underway measurements are archived permanently at NODC and/or NGDC as appropriate. NDSF maintains permanent archive of ROV and AUV sensor data, images, and video. The NCBI Genbank maintains a permanent archive of sequence data. The PI will work with R2R, NDSF, BCO-DMO 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. S. Arellano and C. Young will be responsible for collecting and analyzing the larval data and microscopy. R. Beinart will oversee the genetic and genomic sequencing work and will submit the resulting sequences to the National Center for Biotechnology Information's (NCBI) GenBank database. The Lead PI, R. Beinart, will coordinate the overall data management and sharing process and will submit the project data, including GenBank accession numbers, and metadata to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) who will be responsible for forwarding these data and metadata to the appropriate national archive.