Data Management Plan

This plan is presented to comply with NSF policy and directions on the dissemination and sharing of research results as described in Grant Proposal Guide (GPG) Chap. II.C.2.j.

Data Policy Compliance

Our bone/wood lander studies will generate a diverse set of data to be managed by the two Principal Investigators and their collaborators. The two PIs (Smith and Halanych) have worked extensively in biodiversity and molecular connectivity studies and will build upon their experience to successfully manage and integrate the data. Here we present a summary of the process to be implemented during the project.

Data Management

Pre-Cruise Planning: Pre-cruise conference calls will be use to discuss our sampling strategies and allocation of samples among the PIs and collaborators. Organizational schemes and approaches to data formatting, recording, storage, and integration, as well as sample handling protocols, will be discussed in detail.

On-Cruise Management - Event Management: Data management and event management during cruises will be tightly-coupled. In the past, we have worked closely to produce data fusion using event information and the e-log (free access on line). We anticipate developing and managing three broad classes of metadata: (1) event (e.g., station location and time), (2) instrument (bone/wood lander or multiple corer) and sample characteristics, and (3) combined data for submission to public repositories. Event metadata (1) will be produced in a shared spreadsheet during the cruise. This is the primary means of identifying, ordering, and cataloging sampling activities aboard ship, and ensures traceability of samples across activities. The investigators will share this responsibility and designate an individual to maintain the spreadsheet. The event log will also be published in the science party's onboard web-site. Instrument and sample metadata (2) will depend on the characteristics of the instrument and will include georeferenced location (lat., long., depth), replicate number, and type of data to be generated (e.g., substrate type, depth into bone/wood/sediment, macrofaunal species structure, etc.). Combined metadata for submission (3) will be produced from the composite of event and the instrument/sample metadata to support the submission of our datasets to appropriate long-term archives.

Post-Cruise Management: In the third year of the project, a data integration workshop will be conducted to (1) inform all PIs and collaborators about the nature of the data generated, (2) conduct synthetic analyses to evaluate patterns (e.g., covariance) across different types of biota (microbes versus macrofauna) and data types (molecular versus morphological), (3) plan collaborative papers integrating our results for ecosystem-level syntheses, and (4) plan and coordinate distribution of data to public data repositories.

Data Archiving

Public Data Repositories: We will employ a few different NSF-recognized repositories according to the nature of the data. This will provide public access to our research data and ensure that metadata are provided within current policy guidelines when depositing the data with the archive. DNA sequence data will be deposited in GenBank, and alignment and reconstructed phylogenies submitted to TreeBase. Type and voucher specimens (especially from species descriptions) will be deposited in the Field Museum of Natural History (FMNH) in Chicago (usually holotypes), and/or at the Museum of Natural History in London (usually paratypes). All specimens catalogued in FMNH collections will be included in the online searchable databases at http://fieldmuseum.org/explore/department/zoology/invertebrates/collections and will be available for loan. Species lists generated from different substrates, depths and locations will be made available on the C. Smith UH website, and species records, together with metadata (sampling date, substrate type, location, and depth), with be placed in OBIS.

Institutional Sample Repositories: Both Smith and Halanych maintain sample inventories at their respective institutions for samples collected on cruises. Whereas some specimens (e.g., holotype and paratypes) will be placed into museum collections, most will be inventoried for use in subsequent analyses. Both PIs maintain space for both wet morphological and molecular (frozen and EtOH) samples in secured laboratories. Collections are cataloged and maintained with help from technicians. Material from macrofaunal biodiversity studies, stable-isotope analyses, and bone/wood samples will be archived at the University of Hawaii. Molecular material will be archived at Auburn University. Archived material will be listed on the project websites at UH and Auburn, and made available, upon reasonable request, to other investigators.