

Data Management Plan

We have read and agree to abide by NSF guidelines and award conditions for scientific conduct and data management. Results, data, and collections will be made available to qualified researchers upon request, provided that the quantities and time requirements for compliance do not compromise our research objectives. We will also address data sharing issues in annual and final reports.

Lead Investigator: Jennifer R. Brum (PI)

Institution: Louisiana State University

Collaborators: Bonnie B. Hurwitz (University of Arizona), Virginia Edgcomb (WHOI), Steven Hallam (UBC), Frank Stewart (GA Tech), Gordon Taylor (SUNY Stonybrook), Osvaldo Ulloa (Universidad de Concepcion, Chile), Christopher Algar (Dalhousie University), Matthew Sullivan (Ohio State University), William Wischusen (Louisiana State University)

Project: Ecology and biogeochemical impacts of viruses in marine oxygen minimum zones

Solicitation Info: NSF PD 98-1650

Submission Date: Aug. 15, 2016

Project overview: Sampling at seven major and well-studied oxygen minimum zones has been completed and data generation, curation and management has largely been funded elsewhere. This proposal seeks funding to support analyses of viral and relevant contextual data to look at ecological drivers and biogeochemical impacts of ocean viral community structure, as well as minimal data augmentation relevant towards these ends to cover key missing data points and types.

General data management and sharing overview: Since this project is largely built from data already collected and generated, most of the data needed for the project are already archived and shared with collaborative partners through the NSF-funded CyVerse cyberinfrastructure via iVirus. Upon publication, all raw sequence data will be made available through the ENA repository for molecular data (<http://www.ebi.ac.uk/ena/submit/tara-oceans-checklist>), and all sequence data that is annotated and curated through this project will be made publicly available through iVirus and MetaVir. The iVirus project represents a set of 38 apps and related public datasets dedicated to viral ecology recently developed by collaborator Bonnie Hurwitz and the NSF-funded CyVerse team (<http://www.cyverse.org/>). Bonnie Hurwitz will be directly working with the PI and collaborators on this proposal to enable data accessibility and ensure apps are made publicly available (see LOC-Hurwitz).

Data archives: Viral metagenome raw data will all be stored in NSF-funded CyVerse facilities through the iVirus portal. Viral metagenomes will also be submitted and archived at MetaVir (<http://metavir-meb.univ-bpclermont.fr/>). Curated viral genomes and their associated annotation (e.g. predicted host, detection in microbial metagenomes, etc) will be made available as a separate database at iVirus, in both fasta and genbank file formats so that a user can easily use these as new references. Other curated data products (e.g. viral abundance, morphology, infection frequency) will also be made available through iVirus and the Biological and Chemical Oceanography Data Management Office (<http://www.bco-dmo.org/>). All records will be cross-referenced to sample data, and will follow the naming guideline of each expedition.