Metagenome and metatranscriptome sequences from deep-sea hydrothermal vent microbial communities collected on cruises AT42-22, TN405, and NA108 from May 2019 to Jun 2022

Website: https://www.bco-dmo.org/dataset/936069
Data Type: Cruise Results, Other Field Results
Version: 1
Version Date: 2024-10-10

Project
» Collaborative Research: Microbes need frenemies: unveiling microbial relationships with protists and viruses that support deep-sea hydrothermal vent food webs (frenemies)

Contributors	Affiliation	Role
Hu, Sarah K.	Woods Hole Oceanographic Institution (WHOI)	Principal Investigator
Anderson, Rika	Carleton College	Co-Principal Investigator
Huber, Julie	Woods Hole Oceanographic Institution (WHOI)	Co-Principal Investigator
Mickle, Audrey	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This dataset is a collection of sample collection metadata, sample identifier information, and NCBI accession information for samples and sequence runs produced as part of the frenemies project. This project examines trophic interactions among microbial eukaryotes, viruses, bacteria, and archaea at deep-sea hydrothermal vents using metagenomics and metatranscriptomics and characterizes these ecologically-significant interactions, such as mutualism, predator-prey, or virus-host. We sequenced samples collected during the 2020 expedition AT42-22 to the Mid-Cayman Rise hydrothermal vent fields, as well as from the 2019 expedition NA108 to the Gorda Ridge and the 2022 expedition TN405 to the Axial seamount. Sequencing targeted archaea, bacteria, and viruses with metagenomics and microbial eukaryotes with metatranscriptomics. We plan to use these data to identify ecologically-significant interactions among protists, viruses, bacteria, and archaea, with a specific emphasis on microbial mortality via viral lysis and eukaryotic grazing.

Table of Contents

Coverage

Location: Mid-Cayman Rise, Axial Seamount, Gorda Ridge
Spatial Extent: **N**:45.9 **E**:-81.7 **S**:18.13 **W**:-126.7
Temporal Extent: 2013-09 - 2022-07

Methods & Sampling

All samples were collected via ROVs, where an intake valve was positioned at the location of active diffuse flow (<100C) and collected into bags, bottles, or a filter directly. If collected into a bag or bottle, fluid was filtered immediately upon recovery. All filters had a pore size of 0.2µm and were composed of PES material.

Samples collected at the Mid-Cayman Rise and the Axial Seamount locations utilized the ROV Jason HOG sampler. Samples collected at the Gorda Ridge used the ROV Hercules SUPR sampler.

Metatranscriptomics

Messenger RNA (mRNA) was subset from previously extracted RNA from all sites and grazing experiments (Qiagen RNA mini kit). Then metatranscriptome libraries were prepared from the mRNA. Library was prepped with NEBNext Ultra II Directional RNA Library Prep Kit for Illumina #E7760S; with a Poly(A) mRNA magnetic isolation module first. For fragmentation, samples were incubated at 94C for 10 minutes. PCR cycles done until enough was isolated, range of 16-28. Mainly 20 or 22 cycles total. Input total RNA for metatranscriptome libraries was either 100 ng or 10 ng. However, with mRNA content ranging from 1-5% of the total RNA, the PCR amplification steps needed to be modified for the estimated input mRNA.

Libraries were sequenced with NovaSeq at the Northwest Genomics Sequence Center (Seattle, WA). Sequenced with NovaSeq S2 300 cycles. An average of 100 million sequences per sample were recovered from the metatranscriptome sequencing.

Metagenomics

Duplicate samples (other half of filter) from in situ sites at Mid-Cayman Rise were extracted for DNA (n = 12), then prepared as metagenome libraries, and sent for NovaSeq sequencing (data received January 2023). DNA extracted with MasterPure Complete DNA and RNA Purification Kit (Lucigen MC85200). Input DNA diluted to 50 ng total, or all DNA input (For low concentration samples) and sheared to target 400 bps with Covaris M220 focused-ultrasonicator, utilizing SonoLab 7.2. For 70 sections, 200 cycles (bursts) at 50 watts (Peak incident power), duty factor = 10%, and average incident power was 5 watts. Min temp: 18C, setpoint: 20C, and max: 22C.

Libraries were prepped with Ovation Ultralow System V2 (from Nugen). Library amplification was done at 15 cycles, and pooled. Sequencing was done with a NovaSeq S2 300 Cycle (UW Genomics NWGC). Over 80 million sequences per sample were recovered, to ensure sufficient sequencing depth.

BCO-DMO Processing Description

- Imported submitted file "bco-dmo_Frenemies.csv" into the BCO-DMO system.
- Split lat_lon field into separate fields
- Converted lon to lon W, (West is negative)
- Created YEAR field in format YYYY by pulling the year from FIELDYR
- Added a CRUISE_ID field with corresponding cruise IDs given in submission by year
- Exported dataset as 936069_v1_frenemies_accession_metadata.csv

[table of contents | back to top]

Related Datasets

IsRelatedTo

Hu, S. K., Huber, J. (2023) **18S rRNA amplicon sequencing of microbial eukaryotes from the Mid-Cayman Rise acquired Jan-Feb, 2020.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2023-11-06 doi:10.26008/1912/bco-dmo.914399.1 [view at BCO-DMO]

Hu, S. K., Huber, J., Smith, A. R. (2021) **High throughput tag-sequencing data from Gorda Ridge Hydrothermal vent field, including 16S and 18S rRNA gene sequences, and environmental metadata from Gorda Ridge Seamount, May/June 2019.** Biological and Chemical Oceanography Data

Management Office (BCO-DMO). (Version 1) Version Date 2020-11-09 doi:10.26008/1912/bco-dmo.828392.1 [view at BCO-DMO]

References

Huber Lab @ WHOI. Microbial communities of deep-sea hydrothermal vents Raw sequence reads. 2023/10. In: BioProject [Internet]. Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information; 2011-. Available from: http://www.ncbi.nlm.nih.gov/bioproject/PRJNA1029509. NCBI:BioProject: PRJNA1029509. https://www.ncbi.nlm.nih.gov/bioproject/1029509

[table of contents | back to top]

Parameters

Parameter	Description	Units
SAMPLE_ID	Complete sample ID associated with fastq files	unitless
SHORT_SAMPLE_ID	Short name sample identifier associated with cruise ID	unitless
SAMPLE_NAME	Related sample name associated with vent site	unitless
LAB_NUM	Internal number for sample inventory	unitless
CRUISE_ID	Identifer of cruise during which sample was collected	unitless
FIELD_REGION	Region of hydrothermal vent field	unitless
YEAR	Year of sample collection	unitless
FIELD_YEAR	Hydrothermal vent field with year of collection	unitless
VENT	Name of individual vent site or sample origin	unitless
LAT	Latitude of site where samples were collected. Negative values indicate South.	Decimal Degrees
LON	Longitude of site where samples were collected. Negative values indicate West.	Decimal Degrees
ORIGIN_TYPE	Origin of the sample, either in situ or grazing experiment	unitless
ORIGIN_DESCRIPTION	Collection method description	unitless

FRENEMIES_PROJ	Associated project	unitless
LIBRARY	Type of fastq sequence file	unitless
RUN	SRR number in NCBI	unitless
BIOSAMPLE	Biosample ID number in NCBI	unitless
BASES	Total number of bps in giga base pairs (Gb)	Giga base pairs (Gb)
BYTES	Total size of the sequence file in Gigabytes (GB)	Gigabytes
EXPERIMENT	Experiment ID for locating sequences in NCBI	unitless
LIBRARY_NAME	Sample identifer for fastq files	unitless
LIBRARY_SELECTION	Amplication used in library prep	unitless
geo_loc_name_country	Geo location name for SRA metadata	unitless
geo_loc_name_country_continent	Geo location associated with continent	unitless
geo_loc_name	Geo location name for site of vent field	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina NovaSeq 6000
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	Libraries were sequenced with NovaSeq at the Northwest Genomics Sequence Center (Seattle, WA).
Generic Instrument Description	General term for a laboratory instrument used for deciphering the order of bases in a strand of DNA. Sanger sequencers detect fluorescence from different dyes that are used to identify the A, C, G, and T extension reactions. Contemporary or Pyrosequencer methods are based on detecting the activity of DNA polymerase (a DNA synthesizing enzyme) with another chemoluminescent enzyme. Essentially, the method allows sequencing of a single strand of DNA by synthesizing the complementary strand along it, one base pair at a time, and detecting which base was actually added at each step.

Dataset-specific Instrument Name	ROV Jason, HOG sampler
Generic Instrument Name	Hydrothermal Organic Geochemistry Sampler
Dataset-specific Description	All samples were collected via ROVs, where an intake valve was positioned at the location of active diffuse flow (
Generic Instrument Description	The Hydrothermal Organic Geochemistry (HOG) sampler is designed to collect large volume (2–9 L) fluid samples with minimal introduction of organic or microbial contamination, and to be powered and deployed in real time from a submersible. Additional design constraints include utilizing materials appropriate for sampling fluids with elevated temperatures, fitting the sampler into the space available on the submersible, and minimizing the time needed to remove samples and prepare the sampler for re-deployment between dives.

Dataset-specific Instrument Name	Covaris M220 focused-ultrasonicator
Generic Instrument Name	ultrasonic cell disrupter (sonicator)
Dataset-specific Description	Input DNA diluted to 50 ng total, or all DNA input (For low concentration samples) and sheared to target 400 bps with Covaris M220 focused-ultrasonicator
Generic Instrument Description	Instrument that applies sound energy to agitate particles in a sample.

[table of contents | back to top]

Deployments

AT42-22

Website	https://www.bco-dmo.org/deployment/914418
Platform	R/V Atlantis
Start Date	2020-01-14
End Date	2020-02-06

TN405

Website	https://www.bco-dmo.org/deployment/938749
Platform	R/V Thomas G. Thompson
Start Date	2022-07-08
End Date	2022-07-12

NA108

Website	https://www.bco-dmo.org/deployment/828612
Platform	E/V Nautilus
Start Date	2019-05-24
End Date	2019-06-09
Description	See also https://www.rvdata.us/search/cruise/NA108

[table of contents | back to top]

Project Information

Collaborative Research: Microbes need frenemies: unveiling microbial relationships with protists and viruses that support deep-sea hydrothermal vent food webs (frenemies)

Non-technical abstract:
Ecological interactions among microbes (bacteria and archaea), viruses, and eukaryotic microorganisms are critical junctions in marine food webs. These interactions range from mutually beneficial relationships to sources of microbial mortality. Interactions between viruses-microbes and eukaryotes-microbes at deep-sea hydrothermal vents impact local carbon cycling. This project aims to identify these microbial interactions, specifically those related to cell death by protistan grazing or viral lysis, and explore how they vary across different hydrothermal vent habitats. By providing a better understanding of the composition and nature of these relationships, the investigators aim to build a better food web model of deep-sea hydrothermal vents and improve our understanding of how climate change and other human activities impact the ecosystem. Outcomes from this project include the generation of new microbiology, oceanography, and computer science curricula targeted at community college students. In addition, it involves research with undergraduate students at all stages of the research process and provides opportunities for professional development and peer-to-peer mentoring.

Technical abstract:
This project examines trophic interactions among microbial eukaryotes, viruses, bacteria, and archaea at deep-sea hydrothermal vents using metagenomics and metatranscriptomics and characterizes these ecologically-significant interactions, such as mutualism, predator-prey, or virus-host. The investigators are sequencing samples collected to target archaea/bacteria, viruses, and eukaryotic grazers during a 2020 expedition to the Mid-Cayman Rise hydrothermal vent field to accomplish these goals. Specific aims of this project are to 1) Investigate the microbial, viral, and protistan assemblages and determine how lifestyle, community composition, and metabolism vary across venting fluids of the Mid-Cayman Rise; and 2) Identify ecologically-significant interactions among protists, viruses, bacteria, and archaea and incorporate these interactions into a model for turnover and exchange of carbon in the vent-associated food web. By modeling how trophic

interactions influence microbial mortality, the proposed project substantially contributes to our understanding of the fate of carbon in one of the most productive ecosystems of the deep sea.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-2327203

[table of contents | back to top]