Sub-seafloor metatranscriptomes from anaerobic Peru Margin sediments collected on R/V JOIDES Resolution Leg 201 in 2002

Website: https://www.bco-dmo.org/dataset/626150
Version: 05 November 2015
Version Date: 2015-11-05

Project
» World-wide exploration of microbial eukaryote diversity and activity in the marine subsurface (Microbial Euk Div Mar Subsurface)

Programs
» Center for Dark Energy Biosphere Investigations (C-DEBI)
» International Ocean Discovery Program (IODP)

Contributors	Affiliation	Role
Orsi, William D.	Woods Hole Oceanographic Institution (WHOI)	Principal Investigator, Contact
Biddle, Jennifer F.	University of Delaware	Co-Principal Investigator
Christman, Glenn D.	University of Delaware	Co-Principal Investigator
Edgcomb, Virginia P.	Woods Hole Oceanographic Institution (WHOI)	Co-Principal Investigator, Contact
Rauch, Shannon	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Table of Contents

Methods & Sampling

Sample collection and storage: Subsurface sediment samples from the continental shelf of Peru, Ocean Drilling Program (ODP) Site 1229D (77° 57.4590′ W, 10° 58.5721′ S), were obtained during ODP Leg 201 on 6 March 2002. Careful precautions were taken to avoid contamination during the sampling process. For Integrated Ocean Drilling Program (IODP) cores, contamination tests were performed using perfluorocarbon tracers and fluorescent microspheres. Sediment samples were immediately frozen at −80 degrees C after sampling and stored at −80 degrees C until used for mRNA extractions in this study (10-year storage time at −80 degrees C).

RNA extraction and purification: Extraction of sub-seafloor RNA was performed according to the protocol described previously. In brief, RNA was extracted from 25 g of sediment using the FastRNA Pro Soil-Direct Kit (MP Biomedicals). It was necessary to scale up the volume of sediment that is typically extracted with the kit (~0.5 g) owing to the low biomass inherent to marine subsurface samples. All tubes, tips and disposables used were certified RNase free and all extraction procedures were performed in a laminar flow hood to reduce aerosol contamination by bacterial and fungal cells/spores. Five 15-ml Lysing Matrix E tubes (MP Biomedicals) were filled with 5 g sediment and 5 ml of Soil Lysis Solution (MP Biomedicals). Tubes were vortexed to suspend

the sediment and Soil Lysis Solution was added to the tube leaving 1 ml of headspace. Tubes were then homogenized for 60 s on the FastPrep-24 homogenizer (MP Biomedicals) with a setting of 4.5. Contents were pooled into two 50-ml tubes and centrifuged for 30 min at 4000 r.p.m. (3220g) at room temperature (25 degrees C). Supernatants were combined in a new 50-ml tube and 1/10 volume of 2 M sodium acetate (pH 4.0) was added. An equal volume of phenol-chloroform (pH 6.5) was added and vortexed for 30 s, incubated for 5 min at room temperature, and spun at 4000 r.p.m. (3220g) for 20 min at 4 degrees C. The aqueous phase was transferred to a new 50-ml tube. Nucleic acids were precipitated by adding 2.5 and 1/10 volumes 100% ethanol and 3 M sodium acetate, respectively, and incubating overnight at −80 degrees C. The next day, tubes were spun at 4000 r.p.m. (3220g) for 60 min at 4 degrees C and the supernatant removed. Pellets were washed with 70% ethanol, spun for 15 min at 4 degrees C and air-dried. Dried pellets were resuspended with 0.25 ml RNase-free sterile water and combined into a new 1.5-ml tube. 1/10 volume of 2 M sodium acetate (pH 4.0) and an equal volume of phenol-chloroform (pH 6.5) were added, vortexed for 1 min and incubated for 5 min at room temperature. This was necessary to remove residual organic material (that is, humic acids) resulting from the rather large pellet/precipitate. After centrifuging at 14000 r.p.m. (20817g) for 10 min at 4 degrees C, the top phase was removed into a new 1.5-ml tube. 0.7 volumes of 100% isopropanol was added and incubated for 1 h at −20 degrees C (to precipitate nucleic acids). Tubes were then centrifuged for 20 min at 14000 r.p.m. (20,817g) at 4 degrees C and the supernatant removed. Pellets were washed with 70% ethanol and centrifuged at 14000 r.p.m. (20817g) for 5 min at 4 degrees C. After removing ethanol and air-drying, pellets were re-suspended in 0.2 ml of RNase free sterile water. DNA was removed using the Turbo DNA-free kit (Life Technologies), increasing the incubation time to 1 h to ensure rigorous DNA removal. After this step, samples were taken through the protocol supplied with the FastRNA Pro Soil-Direct kit to the end (starting at the RNA Matrix and RNA Slurry addition step), including the column purification step to remove residual humic acids (see FastRNA Pro Soil-Direct Kit manual). Extraction blanks were performed (adding sterile water instead of sample) to ensure that aerosolized contaminants did not enter sample and reagent tubes during the extraction process. Absence of DNA and RNA contamination was confirmed by no visible amplification of small subunit (SSU) ribosomal RNA (rRNA) and rRNA genes from extraction blanks after 35 cycles of PCR and RT–PCR.

After RNA extraction, the MEGA-Clear RNA Purification Kit (Life Technologies) was used to purify the RNA. This kit removes short RNA fragments (mostly produced during the extraction protocol) and residual inhibitors (that is, humics). We followed the protocol all the way through the optional precipitation/concentration step, re-suspending the RNA pellet in 10 ul of RNase-free sterile water. Before cDNA amplification, the removal of contaminating DNA in RNA extracts was confirmed by the absence of visible amplification of SSU rRNA genes after 35 cycles of PCR using the RNA extracts as template.

cDNA amplification and Illumina sequencing: Five microlitres of purified RNA was used as template for whole-cDNA amplification using the Ovation RNA-Seq v2 System (NuGEN technologies, http://www.nugeninc.com/nugen/index.cfm/products/cs/ngs/rna-seq-v2/). We followed the manufacturer's instructions for cDNA amplification, and the resulting quantity of cDNA was checked on a Nanodrop (Thermo Scientific) and Fluorometer (Qubit 2.0, Life Technologies). Quality of the amplified cDNA was checked on a Bioanalyzer (Agilent Biotechnologies) before Illumina sequencing. Illumina library preparation and paired-end sequencing was performed at the University of Delaware Sequencing and Genotyping Center (Delaware Biotechnology Institute).

Related references:
For more sampling information see www-odp.tamu.edu/publications/prelim/201_prel/201toc.html
The manuscript is at http://www.nature.com/nature/journal/v499/n7457/full/nature12230.html

Data Processing Description

Quality control of the data set was performed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), with a quality score cutoff of 28. Approximately 1 billion paired-end reads that passed quality control were imported into CLC Genomics Workbench 5.0 (CLC Bio) and assembled using the paired-end Illumina assembler. Contigs were assembled over a range of k-mer sizes (20, 50, 60, 64) with a minimum contig size cutoff of 300 nucleotides. The k-mer size of 50 resulted in the highest number of contigs and these contigs were chosen for use in downstream analyses. To reduce the formation of chimaeric assemblies, we used a paired-end sequencing approach and performed assemblies without scaffolding. Reads were mapped onto the contigs using the read mapping option in CLC Genomics Workbench to retain information on relative abundance of contigs. Quality-filtered reads and raw reads are publicly available through the NCBI SRA at http://www.ncbi.nlm.nih.gov/sra?term=SRA058813

[table of contents | back to top]

Data Files

File
peru_margin.csv (Comma Separated Values (.csv), 1.52 KB) MD5:12ba39b846e1f63f539b21f0c84e2412 Primary data file for dataset ID 626150

[table of contents | back to top]

Parameters

Parameter	Description	Units
accession_number	NCBI accession number.	dimensionless
SRA_number	NCBI SRA accession number.	dimensionless
description	Brief description of the sequence.	dimensionless
SRA_URL	Hyperlink to NCBI SRA accession.	dimensionless
accession_URL	Hyperlink to NCBI accession.	dimensionless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Bioanalyzer
Generic Instrument Name	Bioanalyzer
Dataset-specific Description	Quality of the amplified cDNA was checked on a Bioanalyzer (Agilent Biotechnologies) before Illumina sequencing.
Generic Instrument Description	A Bioanalyzer is a laboratory instrument that provides the sizing and quantification of DNA, RNA, and proteins. One example is the Agilent Bioanalyzer 2100.

Dataset-specific Instrument Name	Fluorometer
Generic Instrument Name	Fluorometer
Dataset-specific Description	We followed the manufacturer’s instructions for cDNA amplification, and the resulting quantity of cDNA was checked on a Nanodrop (Thermo Scientific) and Fluorometer (Qubit 2.0, Life Technologies).
Generic Instrument Description	A fluorometer or fluorimeter is a device used to measure parameters of fluorescence: its intensity and wavelength distribution of emission spectrum after excitation by a certain spectrum of light. The instrument is designed to measure the amount of stimulated electromagnetic radiation produced by pulses of electromagnetic radiation emitted into a water sample or in situ.

Dataset-specific Instrument Name	Nanodrop
Generic Instrument Name	Spectrophotometer
Dataset-specific Description	We followed the manufacturer’s instructions for cDNA amplification, and the resulting quantity of cDNA was checked on a Nanodrop (Thermo Scientific) and Fluorometer (Qubit 2.0, Life Technologies).
Generic Instrument Description	An instrument used to measure the relative absorption of electromagnetic radiation of different wavelengths in the near infra-red, visible and ultraviolet wavebands by samples.

Dataset-specific Instrument Name	PCR
Generic Instrument Name	Thermal Cycler
Dataset-specific Description	Absence of DNA and RNA contamination was confirmed by no visible amplification of small subunit (SSU) ribosomal RNA (rRNA) and rRNA genes from extraction blanks after 35 cycles of PCR and RT–PCR.
Generic Instrument Description	A thermal cycler or "thermocycler" is a general term for a type of laboratory apparatus, commonly used for performing polymerase chain reaction (PCR), that is capable of repeatedly altering and maintaining specific temperatures for defined periods of time. The device has a thermal block with holes where tubes with the PCR reaction mixtures can be inserted. The cycler then raises and lowers the temperature of the block in discrete, pre-programmed steps. They can also be used to facilitate other temperature-sensitive reactions, including restriction enzyme digestion or rapid diagnostics. (adapted from http://serc.carleton.edu/microbelife/research_methods/genomics/pcr.html)

[table of contents | back to top]

Deployments

JRES-201

Website	https://www.bco-dmo.org/deployment/626163
Platform	R/V JOIDES Resolution
Report	http://dmoserv3.whoi.edu/data_docs/C-DEBI/cruise_reports/201PREL-1.pdf
Start Date	2002-01-27
End Date	2002-03-29
Description	Leg 201 Controls on Microbial Communities in Deeply Buried Sediments, Eastern Equatorial Pacific and Peru Margin Sites 1225-1231 27 January-29 March 2002 Cruise report obtained from http://www-odp.tamu.edu/publications/pubs.htm

[table of contents | back to top]

Project Information

World-wide exploration of microbial eukaryote diversity and activity in the marine subsurface (Microbial Euk Div Mar Subsurface)

Coverage: Peru Margin

Project description obtained from C-DEBI:
Practically nothing is known about microbial eukaryotes (mEuks) in the marine subsurface. mEuks are pivotal members of microbial communities because they regenerate nutrients and modify or remineralize organic matter through grazing on prokaryotic and other eukaryotic prey. Thus, mEuks help determine metabolic potentials of microbial communities and influence elemental cycling. Only one study has addressed mEuk diversity in the marine subsurface (Edgcomb et al. 2010), which suggested Fungi dominate the eukaryotic subsurface community and are active in sediments 35 mbsf at the Peru Margin. Thus, some mEuks may be specifically adapted to the deep subsurface and may play significant roles in the utilization and regeneration of organic matter and nutrients in deep-sea sediments.

One objective of this study will be to further investigate whether Fungi are consistently the dominant group of mEuks in the marine subsurface by examining mEuk diversity in a broad range of subsurface samples from ODP expeditions spanning the world's oceans. Deep sequencing of SSU rRNA in these samples will provide a proxy for mEuk diversity and activity in the marine subsurface. A second objective will be to 'ground truth' an mRNA isolation protocol for mEuks in marine subsurface sediments. Once established, this protocol will enable the third objective, which is the creation of a eukaryotic metatranscriptome from ODP site 1229. This metatranscriptome will provide insights into the functional role of mEuks in the marine subsurface and perhaps new insights into microbial evolution.

This project was funded by a C-DEBI Postdoctoral Fellowship.

[table of contents | back to top]

Program Information

Center for Dark Energy Biosphere Investigations (C-DEBI)

Website: http://www.darkenergybiosphere.org

Coverage: Global

The mission of the Center for Dark Energy Biosphere Investigations (C-DEBI) is to explore life beneath the seafloor and make transformative discoveries that advance science, benefit society, and inspire people of all

ages and origins.

C-DEBI provides a framework for a large, multi-disciplinary group of scientists to pursue fundamental questions about life deep in the sub-surface environment of Earth. The fundamental science questions of C-DEBI involve exploration and discovery, uncovering the processes that constrain the sub-surface biosphere below the oceans, and implications to the Earth system. What type of life exists in this deep biosphere, how much, and how is it distributed and dispersed? What are the physical-chemical conditions that promote or limit life? What are the important oxidation-reduction processes and are they unique or important to humankind? How does this biosphere influence global energy and material cycles, particularly the carbon cycle? Finally, can we discern how such life evolved in geological settings beneath the ocean floor, and how this might relate to ideas about the origin of life on our planet?

C-DEBI's scientific goals are pursued with a combination of approaches:
(1) coordinate, integrate, support, and extend the research associated with four major programs—Juan de Fuca Ridge flank (JdF), South Pacific Gyre (SPG), North Pond (NP), and Dorado Outcrop (DO)—and other field sites;
(2) make substantial investments of resources to support field, laboratory, analytical, and modeling studies of the deep subseafloor ecosystems;
(3) facilitate and encourage synthesis and thematic understanding of submarine microbiological processes, through funding of scientific and technical activities, coordination and hosting of meetings and workshops, and support of (mostly junior) researchers and graduate students; and
(4) entrain, educate, inspire, and mentor an interdisciplinary community of researchers and educators, with an emphasis on undergraduate and graduate students and early-career scientists.

Note: Katrina Edwards was a former PI of C-DEBI; James Cowen is a former co-PI.

Data Management:
C-DEBI is committed to ensuring all the data generated are publically available and deposited in a data repository for long-term storage as stated in their Data Management Plan (PDF) and in compliance with the NSF Ocean Sciences Sample and Data Policy. The data types and products resulting from C-DEBI-supported research include a wide variety of geophysical, geological, geochemical, and biological information, in addition to education and outreach materials, technical documents, and samples. All data and information generated by C-DEBI-supported research projects are required to be made publically available either following publication of research results or within two (2) years of data generation.

To ensure preservation and dissemination of the diverse data-types generated, C-DEBI researchers are working with BCO-DMO Data Managers make data publicly available online. The partnership with BCO-DMO helps ensure that the C-DEBI data are discoverable and available for reuse. Some C-DEBI data is better served by specialized repositories (NCBI's GenBank for sequence data, for example) and, in those cases, BCO-DMO provides dataset documentation (metadata) that includes links to those external repositories.

International Ocean Discovery Program (IODP)

Website: http://www.iodp.org/index.php

Coverage: Global

The International Ocean Discovery Program (IODP) is an international marine research collaboration that explores Earth's history and dynamics using ocean-going research platforms to recover data recorded in seafloor sediments and rocks and to monitor subseafloor environments. IODP depends on facilities funded by three platform providers with financial contributions from five additional partner agencies. Together, these entities represent 26 nations whose scientists are selected to staff IODP research expeditions conducted throughout the world's oceans.

IODP expeditions are developed from hypothesis-driven science proposals aligned with the program's science plan *Illuminating Earth's Past, Present, and Future*. The science plan identifies 14 challenge questions in the four areas of climate change, deep life, planetary dynamics, and geohazards.

IODP's three platform providers include:

- The U.S. National Science Foundation (NSF)
- Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT)
- The European Consortium for Ocean Research Drilling (ECORD)

More information on IODP, including the Science Plan and Policies/Procedures, can be found on their website at http://www.iodp.org/program-documents.

A summary table with links to IODP datasets currently hosted on Zenodo (https://zenodo.org/communities/iodp) can be accessed using the following link: https://iodp.tamu.edu/database/zenodo.html

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-0939564
NSF Division of Integrative Organismal Systems (NSF IOS)	IOS-1238801

[table of contents | back to top]