Sub-seafloor sediment eukaryotic rRNA collected on R/V JOIDES Resolution Legs 201 and 204, R/V Maria Merian cruise MSM11-1 at North Pond, and R/V Meteor cruise M76-1a at the Benguela Upwelling System from 2002-2009

Website: https://www.bco-dmo.org/dataset/626184
Version: 06 November 2015
Version Date: 2015-11-06

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
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 Hydrate Ridge (IODP Leg 204 Site 1244a; 44° 35' 17" N, 125° 07' 19" W), Peru Margin (IODP Leg 201 Site 1227a; 79° 57' 349" W, 08° 59' 463" S), and Eastern Equatorial Pacific (IODP Leg 201 Site 1225a; 110° 43' 289" W, 02° 46' 247" N) were obtained from the Gulf Coast Core Repository (University of Texas A&M). Gravity core subsurface samples from North Pond near the Mid-Atlantic Ridge (22° 48' 04" E, 46° 06' 30" N) and Benguela Upwelling System (14° 15' 04" E, 27° 44' 40" S) were collected on March 3rd 2009 onboard the R/V Maria Merian and April 21st 2008 onboard the R/V Meteor, respectively and were provided by Andreas Teske (University of North Carolina, Chapel Hill, NC). Careful precautions were taken during sampling to avoid contamination during the sampling process. For 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 RNA was extracted. Sediment samples at a sediment depth of 0.01 and 0.08 mbsf from Little Sippewissett Salt Marsh were taken November 13th 2011 using a sterile syringe. Sulfide was detectable in both samples and thus samples were presumed anoxic. No specific permits were required for the described field studies. The locations sampled are not privately owned or protected and field studies did not involve endangered or protected species.

RNA extraction and purification: RNA was extracted from 25 grams of sediment using the FastRNA Pro Soil-Direct Kit in a laminar flow hood to reduce contamination from aerosols. Extractions were performed at Woods Hole Oceanographic Institution. Several modifications were made to the protocol provided with the kit to increase RNA yield from low biomass subseafloor samples. It was necessary to scale up the volume of sediment that is typically extracted with the kit (~0.5 grams) due to the expected low biomass of subsurface eukaryotes. Four 15 ml Lysing Matrix E tubes (MP Biomedicals, Solon, OH) were filled with 5 g sediment and 5 ml of Soil Lysis Solution. 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 seconds at a setting of 4.5 on the FastPrep-24 homogenizer. Contents of the 15 ml tubes were combined into two RNAse free 50 ml falcon tubes and centrifuged for 30 minutes at 4000 RPM. The supernatants were combined in a new 50 ml RNAse-free falcon tube and 1/10 volume of 2M Sodium Acetate (pH 4.0) was added. An equal volume of phenol-chloroform (pH 6.5) was added and vortexed for 30 seconds, incubated for 5 minutes at room temperature, and centrifuged at 4000 RPM for 20 minutes at 4 degrees C. The top phase was carefully transferred to a new 50 ml falcon tube and 2.5x volumes 100% ethanol and 1/10 volume 3M Sodium Acetate were added and incubated overnight at −80 degrees C. After incubation, tubes were centrifuged at 4000 RPM for 60 minutes at 4 degrees C and the supernatant removed. Pellets were washed with 70% ethanol, centrifuged for 15 minutes 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 RNAse-free tube. 1/10 volume of 2M Sodium Acetate (pH 4.0) and an equal volume of phenol:chloroform (pH 6.5) were added, the tube was vortexed for 1 minute, and incubated for 5 minutes at room temperature. The tube was then centrifuged for 10 minutes at 4 degrees C, the top phase removed into a new RNAse free 1.5 ml tube, and 0.7 volumes of 100% isopropanol was added and incubated for 1hour at −20 degrees C. After incubation tubes were centrifuged for 20 minutes at 14000 RPM at 4 degrees C and the supernatant was removed. Pellets were washed with 70% ethanol and centrifuged at 14000 RPM for 5 minutes at 4 degrees C. Ethanol was removed and the pellets air-dried. Pellets were resuspended with 200 ul of RNAse free sterile water and DNA was removed using the Turbo DNA-free kit (Life Technologies, Grand Island, NY). DNAse incubation times were increased to 1 hour to ensure removal of contaminating DNA. Samples were then 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 optional column purification step to remove residual humic acids. To further purify the RNA, we used the MEGA-Clear RNA Purification Kit. Extraction blanks were performed (adding sterile water instead of sample) to identify aerosolized contaminants that may have entered sample and reagent tubes during the extraction process. To reduce contamination, all RNA extractions were performed in a laminar flow hood.

RT-PCR amplification of eukaryotic rRNA: To amplify the V4 hypervariable region of eukaryotic rRNA, we used PCR primers targeting this region: EukV4F (5' – CGTATCGCCTCCCTCGCGCCATCAGxxxxxxxxxxxCCAGCASCYGCGGTAATTCC – 3') and EukV4R (5' – CTATGCGCCTTGCCAGCCCGCTCAGACTTTCGTTCTTGATYRA – 3'), where the x region represents the unique MID barcode used for each sample, the linker primer sequence is underlined, and the 18S rRNA eukaryotic primer is bold. These primers were chosen because they target a wide range of eukaryotic taxa. RT-PCR was performed using the SuperScript One-Step RT-PCR with Platinum Taq kit. Individual reactions consisted of 2 ul RNA template, 25 ul buffer, 1 ul of forward Primer, 1 ul of reverse primer, 2 ul of the Platinum RT-Taq enzyme mix, and 18 ul RNAse free sterile water. The cDNA step was performed at 55 degrees C and cDNA was amplified in 40 cycles of PCR with an annealing temperature of 65 degrees C (55 degrees C for 30 minutes, 95 degrees C for 5 minutes, [95 degrees C for 15 seconds, 65 degrees C for 30 seconds, 68 degrees C for 1 minute]x40, 68 degrees C for 5 minutes). To check for DNA carryover during the RNA extraction protocol, a separate PCR reaction (at the same number of cycles) was included in which Taq polymerase was substituted for the reverse-transcriptase/platinum Taq enzyme mix. For each sample, 5–10 RT-PCR reactions were performed and extracted using the Zymo Research Gel Extraction Kit. A gel volume of 100% isopropanol was added to each dissolved gel slice before addition to the DNA collection column. Dissolved gel slices from each sample were pooled by centrifuging them all through the same DNA collection column. cDNA was quantified fluorometrically prior to 454 sequencing using the Qubit 2.0. To identify contaminants we performed additional RT-PCR amplifications at 55 cycles using RNAse free sterile water and RNA extraction blanks (resulting from RNA extractions in which no sample was added) as template. Contaminants were amplified with primers containing a unique MID in 55 cycles of PCR.

Quality control, clustering, and taxonomic assignment of 454 data: cDNA amplicons were sequenced on a GS-FLX Titanium 454 sequencer at EnGenCore (University of South Carolina, Columbia, SC), which resulted in ~37000 reads. To reduce homopolymer errors inherent to 454 sequencing, the dataset was put through the denoise protocol as described in the QIIME software package using the denoise_wrapper.py command. After denoising, chimeric sequences were identified and removed using ChimeraSlayer with the blast_fragments method in QIIME. The data were subjected to quality score filtering using the split_libraries.py command and clustered at various levels of sequence identity (80%, 85%, 90%, 93%, 95%, 97%) in QIIME using the uclust method of all-to-all pair-wise comparisons via the pick_otus.py command.

The QIIME taxonomy classification pipeline was not able to accurately classify the majority of eukaryotic OTUs. Thus, we used Jaguc, a program developed specifically for classification of eukaryotic rRNA sequence data, to classify our sequence reads. 90% of eukaryotic OTUs were classified to genus using this approach. OTU tables were created using the make_otu_table.py command in QIIME and the Jaguc taxonomy for each OTU was amended onto this table using a custom perl script developed by the authors for this purpose. This perl script is available from the authors upon request.

Terminal Restriction Fragment Length Polymorphism (TRFLP) analysis of fungal rRNA: To further investigate the fungal diversity in our samples, we used a TRFLP approach using PCR primers specific to fungal 18S rRNA. The fungal primers used were EF3 (5' – TCCTCTAAATGACCAAGTTTG – 3') and Fung5 (5' – GTAAAAGTCCTGGTTCCCC – 3'). The forward primer, EF3, was labeled with the phosphoramidite dye 6-Carboxyfluorescein (6-FAM) at the 5'-end (Integrated DNA Technologies, Coralville, Iowa). Fungal rRNA was amplified using a cDNA incubation step at 50 degrees C followed by 40 cycles of PCR with an annealing temperature of 53 degrees C. Three RT-PCR reactions were performed for each sample, gel extracted, and pooled using the same protocol as above. Fungal rRNA amplicons were digested with three different restriction enzymes: MspI, RsaI, and HhaI (New England Biolabs, Ipswich, MA), for 1 hr at 37 degrees C. These restriction enzymes were chosen because they have been shown to provide statistically significant TRFLP data for interpreting fungal community structure across different samples. Digests were mixed with the Applied Biosystems size marker GS600LIZ and HiDi Formamide in the ratio 1:1:9 and run on an Applied Biosystems 3730 DNA analyzer. Electropherograms were analyzed using the PeakScanner software package (Applied Biosystems, Carlsbad, CA) to identify the size, height, and peak area of each T-RF. T-REX was used to filter out noise from true peaks and to align peaks.

Related references:
The manuscript is at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0056335

Data Processing Description

Statistical Analyses: Canonical Correspondence Analysis (CCA) was used to elucidate relationships between eukaryotic community structure and concentrations of dissolved oxygen (O_2), nitrate (NO_3-) dissolved inorganic carbon (DIC), total organic carbon (TOC), and sulfide. Multi-response Permutation Procedure (MRPP) was used to test for a statistically significant influence of sediment depth, DIC, sulfide, TOC, and oxygen on the observed OTU distributions. All ordination and multivariate statistical analyses were performed on the TRFLP and pyrosequenced datasets as a whole, as well as the five major eukaryotic subgroups that dominated our 454 dataset: Metazoa, Viridiplantae, Diatoms, Alveolates, and Fungi. Analyses were performed on sequences affiliated with these groups clustered at 80, 85, 90, 93, and 97% sequence identity thresholds as well as the fungal TRFLP dataset. MRPP and CCA were implemented using the PC-ORD software package (MjM Software Design). Weighted UniFrac analysis was performed in QIIME. Prior to UniFrac and alpha-diversity comparisons, the number of sequences per sample were normalized to the sample with the least number of sequences by randomly selecting a subset of sequences from each sample using the multiple_rarefactions.py script in QIIME.

Quality-filtered reads and raw reads are publicly available through the NCBI SRA at http://www.ncbi.nlm.nih.gov/sra?term= SRA052670

Data Files

File
deep_biosphere_euks.csv(Comma Separated Values (.csv), 214 bytes) MD5:8e2ab1b663aba54b1a6e57d24ba1b59d
Primary data file for dataset ID 626184

Parameters

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

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	GS-FLX Titanium 454 sequencer
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	cDNA amplicons were sequenced on a GS-FLX Titanium 454 sequencer at EnGenCore (University of South Carolina, Columbia, SC), which resulted in ~37000 reads.
Generic Instrument Description	A DNA sequencer is an instrument that determines the order of deoxynucleotides in deoxyribonucleic acid sequences.

Dataset-specific Instrument Name	PCR
Generic Instrument Name	Thermal Cycler
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

JRES-204

Website	https://www.bco-dmo.org/deployment/626215
Platform	R/V JOIDES Resolution
Report	http://dmoserv3.whoi.edu/data_docs/C-DEBI/cruise_reports/204PREL.PDF
Start Date	2002-07-07
End Date	2002-09-02
Description	Leg 204 Drilling Gas Hydrates on Hydrate Ridge, Cascadia Continental Margin Sites 1244-1252 7 July-2 September 2002 Cruise report obtained from http://www-odp.tamu.edu/publications/pubs.htm

MSM11-1

Website	https://www.bco-dmo.org/deployment/626221
Platform	R/V Maria S. Merian
Report	http://dmoserv3.whoi.edu/data_docs/C-DEBI/cruise_reports/msm11-1-scr.pdf
Start Date	2009-02-16
End Date	2009-05-12
Description	Objectives from the cruise report: North Pond, an isolated sediment pond located on the Western flank of the mid-Atlantic Ridge, offers the opportunity to study microbial communities and their activities in deeply-buried sediments and the underlying basement. One important argument for choosing North Pond is that the geochemistry, hydrology, and geologic setting of North Pond have been previously studied by a series of DSDP and ODP drill holes and seafloor observatories. However the existing site survey data are not sufficient for the approved North Pond IODP drilling expedition, scheduled to take place not before FY 2010. The goal of the proposed investigations on MSM 11/1 is therefore to map North Pond in a detailed way in order to be able to position planned IODP holes precisely. These surveys compromise a detailed seismic mapping of the sediment-basement interface of North Pond, additional heat flow data and geochemical and microbial sampling of the sediments. In addition dredging of the upper crust, surrounding North Pond, will help to better define the petrological setting of North Pond.

M76-1a

Website	https://www.bco-dmo.org/deployment/626226
Platform	R/V Meteor
Report	http://dmoserv3.whoi.edu/data_docs/C-DEBI/cruise_reports/awi_doi_10.2312_cr_m76.pdf
Start Date	2008-04-12
End Date	2008-04-25

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.

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

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-0939564
WHOI Ocean Life Institute (WHOI OLI)	OLI-27071359