Supplementary Table 4C: Statistics of reads retained through bioinformatic processing of iTAG data for the 11 samples and control samples and metatranscriptome data.

Website: https://www.bco-dmo.org/dataset/813173
Data Type: Other Field Results
Version: 1
Version Date: 2020-05-28

Project
» Collaborative Research: Delineating The Microbial Diversity and Cross-domain Interactions in The Uncharted Subseafloor Lower Crust Using Meta-omics and Culturing Approaches (Subseafloor Lower Crust Microbiology)

Program
» International Ocean Discovery Program (IODP)

Contributors	Affiliation	Role
Edgcomb, Virginia P.	Woods Hole Oceanographic Institution (WHOI)	Principal Investigator, Contact
Soenen, Karen	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Supplementary Table 4C: Metatranscriptome data summary for cellular activities presented and statistics on sequencing and removal of potential contaminant sequences: Statistics of reads retained through bioinformatic processing of iTAG data for the 11 samples and control samples and metatranscriptome data. Samples taken on board of the R/V JOIDES Resolution between November 30, 2015 and January 30, 2016

Table of Contents

Coverage

Spatial Extent: **Lat**:-32.70567 **Lon**:57.278183
Temporal Extent: 2015-11-30 - 2016-01-30

Dataset Description

Supplementary Table 4C: Metatranscriptome data summary for cellular activities presented and statistics on sequencing and removal of potential contaminant sequences: Statistics of reads retained through

bioinformatic processing of iTAG data for the 11 samples and control samples and metatranscriptome data. Samples taken on board of the R/V JOIDES Resolution between November 30, 2015 and January 30, 2016

Acquisition Description

Rock material was crushed while still frozen in a Progressive Exploration Jaw Crusher (Model 150) whose surfaces were sterilized with 70% ethanol and RNase AWAY (Thermo Fisher Scientific, USA) inside a laminar flow hood. Powdered rock material was returned to the -80°C freezer until extraction.

DNA was extracted from 20, 30, or 40 grams of powdered rock material, depending on the quantity of rock available. A DNeasy PowerMax Soil Kit (Qiagen, USA) was used following the manufacturer's protocol modified to included three freeze/thaw treatments prior to the addition of Soil Kit solution C1. Each treatment consisted of 1 minute in liquid nitrogen followed by 5 minutes at 65 °C. DNA extracts were concentrated by isopropanol precipitation overnight at 4°C.

The low biomass in our samples required whole genome amplification (WGA) prior to PCR amplification of marker genes. Genomic DNA was amplified by Multiple Displacement Amplification (MDA) using the REPLI-g Single Cell Kit (Qiagen) as directed. MDA bias was minimized by splitting each WGA sample into triplicate 16 µL reactions after 1 hr of amplification and then resuming amplification for the manufacturer-specified 7 hrs (8 hrs total).

DNA was also recovered from samples of drilling mud and drilling fluid (surface water collected during the coring process) for negative controls, as well as two "kit control" samples, in which no sample was added, to account for any contaminants originating from either the DNeasy PowerMax Soil Kit or the REPLI-g Single Cell Kit.

Bacterial SSU rRNA gene fragments were PCR amplified from MDA samples and sequenced at Georgia Genomics and Bioinformatics Core (Univ. of Georgia). The primers used were: Bac515-Y and Bac926R. Dual-indexed libraries were prepared with (HT) iTruS (Kappa Biosystems) chemistry and sequencing was performed on an Illumina MiSeq 2 x 300 bp system with all samples combined equally on a single flow cell.

Raw sequence reads were processed through Trim Galore [http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/], FLASH (ccb.jhu.edu/software/FLASH/) and FASTX Toolkit [http://hannonlab.cshl.edu/fastx_toolkit/] for trimming and removal of low quality/short reads.

Quality filtering included requiring a minimum average quality of 25 and rejection of paired reads less than 250 nucleotides.

Operational Taxonomic Unit (OTU) clusters were constructed at 99% similarity with the script pick_otus.py within the Quantitative Insights Into Microbial Ecology (QIIME) v.1.9.1 software and 'uclust'. Any OTU that matched an OTU in one of our control samples (drilling fluids, drilling mud, extraction and WGA controls) was removed (using filter_otus_from_otu_table.py) along with any sequences of land plants and human pathogens that may have survived the control filtering due to clustering at 99% (filter_taxa_from_otu_table.py). As an additional quality control measure, genera that are commonly identified as PCR contaminants were removed. Unclassified OTUs were queried using BLAST against the GenBank nr database and further information about these OTUs is provided in the Supplementary Discussion text under the section "Taxonomic diversity information from iTAGs." OTUs that could not be assigned to Bacteria or Archaea were removed from further analysis. For downstream analyses, any OTUs not representing more than 0.01% of relative abundance of sequences overall were removed as those are unlikely to contribute significantly to in situ communities. The OTU data table was transformed to a presence/absence table and the Jaccard method was used to generate a distance matrix using the dist.binary() function in the R package ade4.

Processing Description

BCO-DMO processing notes:

- Reformatted table structure
- Added columns Latitude, Longitude and Depth
- Adjusted column header names to comply with database requirements

[table of contents | back to top]

Related Publications

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. Nature Methods, 7(5), 335–336. doi:10.1038/nmeth.f.303
Methods

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2015). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environmental Microbiology, 18(5), 1403–1414. doi:10.1111/1462-2920.13023
Methods

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., … Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biology, 12(1). doi:10.1186/s12915-014-0087-z
Methods

Sheik, C. S., Reese, B. K., Twing, K. I., Sylvan, J. B., Grim, S. L., Schrenk, M. O., … Colwell, F. S. (2018). Identification and Removal of Contaminant Sequences From Ribosomal Gene Databases: Lessons From the Census of Deep Life. Frontiers in Microbiology, 9. doi:10.3389/fmicb.2018.00840
Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
Sample_ID	Sample ID	unitless
Latitude	Latitude of sample, south is negative	decimal degrees
Longitude	Longitude of samples, west is negative	decimal degrees
Depth	Depth - meters below seafloor (mbsf)	meters (m)
iTAG_Raw	iTAG data - Raw reads	number of reads
iTAG_Paired_QC	iTAG data - paired reads after QC	number of reads
iTAG_Paired_Contmnt_Rem	iTAG data - Paired reads surviving removal of potential contaminants matching sequences in control samples or known contaminants.	number of reads
iTAG_OTU	iTAG data - Number of OTUs at 99% identity	number of OTUs
Metatr_Raw	Metatranscriptome data - Raw reads from sequencing	number of reads
Metatr_Paired_QC	Metatranscriptome data - Paired reads after QC	number of reads
Metatr_Paired_Contmnt_Rem	Metatranscriptome data - Paired reads surviving removal of potential contaminants matching sequences in control samples or known contaminants.	number of reads
Metatr_Reads_Remaining	Metatranscriptome data - Percent of original paired reads remaining	percentage (%)

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina MiSeq 2 x 300 bp platform
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	DNA sequencing performed using the Illumina MiSeq 2 x 300 bp platform (Univ. of Georgia)
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.

[table of contents | back to top]

Deployments

IODP-360

Website	https://www.bco-dmo.org/deployment/810905
Platform	R/V JOIDES Resolution
Report	http://publications.iodp.org/scientific_prospectus/360/index.html
Start Date	2015-11-30
End Date	2016-01-30

[table of contents | back to top]

Project Information

Collaborative Research: Delineating The Microbial Diversity and Cross-domain Interactions in The Uncharted Subseafloor Lower Crust Using Meta-omics and Culturing Approaches (Subseafloor Lower Crust Microbiology)

Coverage: SW Indian Ridge, Indian Ocean

NSF abstract: The lower ocean crust has remained largely unexplored and represents one of the last frontiers for biological exploration on Earth. Preliminary data indicate an active subsurface biosphere in samples of the lower oceanic crust collected from Atlantis Bank in the SW Indian Ocean as deep as 790 m below the seafloor. Even if life exists in only a fraction of the habitable volume where temperatures permit and fluid flow can deliver carbon and energy sources, an active lower oceanic crust biosphere would have implications for deep carbon budgets and yield insights into microbiota that may have existed on early Earth. This is all of great interest to other research disciplines, educators, and students alike. A K-12 education program will capitalize on groundwork laid by outreach collaborator, A. Martinez, a 7th grade teacher in Eagle Pass, TX, who sailed as outreach expert on Drilling Expedition 360. Martinez works at a Title 1 school with ~98% Hispanic and ~2% Native American students and a high number of English Language Learners and migrants. Annual school visits occur during which the project investigators present hands on-activities introducing students to microbiology, and talks on marine microbiology, the project, and how to pursue science related careers. In addition, monthly Skype meetings with students and PIs update them on project progress. Students travel to the University of Texas Marine Science Institute annually, where they get a campus tour and a 3-hour cruise on the R/V Katy, during which they learn about and help with different oceanographic sampling approaches. The project partially supports two graduate students, a Woods Hole undergraduate summer student, the participation of multiple Texas A+M undergraduate students, and 3 principal investigators at two institutions, including one early career researcher who has not previously received NSF support of his own. Given the dearth of knowledge of the lower oceanic crust, this project is poised to transform our understanding of life in this vast environment.

The project assesses metabolic functions within all three domains of life in this crustal biosphere, with a focus on nutrient cycling and evaluation of connections to other deep marine microbial habitats. The lower ocean crust represents a potentially vast biosphere whose microbial constituents and the biogeochemical cycles they mediate are likely linked to deep ocean processes through faulting and subsurface fluid flow. Atlantis Bank represents a tectonic window that exposes lower oceanic crust directly at the seafloor. This enables seafloor drilling and research on an environment that can transform our understanding of connections between the deep subseafloor biosphere and the rest of the ocean. Preliminary analysis of recovered rocks from Expedition 360 suggests the interaction of seawater with the lower oceanic crust creates varied geochemical conditions capable of supporting diverse microbial life by providing nutrients and chemical energy. This project is the first interdisciplinary investigation of the microbiology of all 3 domains of life in basement samples that combines diversity and "meta-omics" analyses, analysis of nutrient addition experiments, high-throughput culturing and physiological analyses of isolates, including evaluation of their ability to utilize specific carbon sources, Raman spectroscopy, and lipid biomarker analyses. Comparative genomics are used to compare genes and pathways relevant to carbon cycling in these samples to data from published studies of other deep-sea environments. The collected samples present a rare and time-sensitive opportunity to gain detailed insights into microbial life, available carbon and energy sources for this life, and of dispersal of microbiota and connections in biogeochemical processes between the lower oceanic crust and the overlying aphotic water column. About the study area: The International Ocean Discovery Program (IODP) Expedition 360 explored the lower crust at Atlantis Bank, a 12 Ma oceanic core complex on the ultraslow-spreading SW Indian Ridge. This oceanic core complex represents a tectonic window that exposes lower oceanic crust and mantle directly at the seafloor, and the expedition provided an unprecedented opportunity to access this habitat in the Indian Ocean.

[table of contents | back to top]

Program Information

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.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1658031

[table of contents | back to top]