GenBank accession numbers for microbial 16S rRNA sequences collected at the East Pacific Rise by DSV/Alvin during R/V Atlantis research cruises (Jan. 2004 - June 2006) (Microbial Communities at Deep-Sea Vents project)

Website: https://www.bco-dmo.org/dataset/687844
Data Type: Cruise Results
Version:
Version Date: 2017-04-20

Project
» An Integrated Study of Energy Metabolism, Carbon Fixation, and Colonization Mechanisms in Chemosynthetic Microbial Communities at Deep-Sea Vents (Microbial Communities at Deep-Sea Vents)

Programs
» Dimensions of Biodiversity (Dimensions of Biodiversity)
» Center for Dark Energy Biosphere Investigations (C-DEBI)

Contributors	Affiliation	Role
Sievert, Stefan M.	Woods Hole Oceanographic Institution (WHOI)	Principal Investigator
Copley, Nancy	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Table of Contents

Coverage

Spatial Extent: **N**:9.84 **E**:-104.29 **S**:9.8383 **W**:-104.2917
Temporal Extent: 2004-02-03 - 2006-07-01

Dataset Description

Bacterial diversity and successional patterns during biofilm formation on freshly exposed basalt surfaces at diffuse-flow deep-sea vents.

This dataset is the microbial community composition based on 16 rRNA gene analysis of biofilms on basalt collected during research cruises AT11-07, AT11-10, AT11-20 and AT15-06 at the 9ºN deep-sea hydrothermal vent field on the East Pacific Rise, Pacific Ocean between Jan. 2004 and June 2006. The dataset contains links to full-length 16S rRNA sequence data deposited at NCBI GenBank with accession numbers KT257735-KT257859. The 16S rRNA tag data were assigned GenBank SRA Bioproject number PRJNA288972.

Methods & Sampling

Basalt panels were deployed and recovered by DSV Alvin at TicaVent on the EPR (9°50.4 N, 104°17.5 W: 2513-m depth) during R/V Atlantis research cruises AT11-07 (2004/02/03 –2004/02/19), AT11-10 (2004/04/07–2004/04/24), AT11-20 (2004/11/11–2004/11/24). Basalt panels were deployed either at a diffuse-flow vent site (Experiment site1: ~15◦C) or a nearby, non-vent site (Control site: ~1.8◦C). Experiment site1 was located within a thriving patch of tubeworms (Riftia pachyptila) and mussels (Bathymodiolus thermophilus). The Control site was located 2.5- m from the experimental site, ~1/2-m from mussels marking the edge of the colonized vent. These panels are considered 'pre-eruption' samples because they were collected 19–28 months before the 2005/2006 EPR eruption. For the purpose of comparison with our experimental results, one post-eruption native basalt biofilm sample (hereafter referred to as 'native Basalt') was collected from Tamtown (9°50.3 N, 104°17.4 W: 2503-m depth; ~0.58 km from TicaVent), during the RESET06 cruise (AT15-06, 2006/06/25-2006/07/01). [official dates: AT15-06 (2006/06/18-2006/07/07)]

Basalt panels (10.2 x 10.2 x 2.5 cm) were constructed of basalt collected in the area of 9oN on the EPR (Supplementary Figure 1D). Before deployment, all panels were sealed in aluminum foil and autoclaved. Each panel was deployed from and recovered into an ethanol-wiped biobox mounted on Alvin's basket, filled prior to the dive with either 0.2-µm filtered seawater or double-distilled water to prevent contamination with surface seawater (as were the Basalt and Trap samples). All subsequent handling of the basalt panels was with sterilized gloves or tools. Control and experimental basalt panels were exposed for 5 time intervals: 4-days, 9-days, 13-days, 76-days, and 9-mos [283-days (control) and 293-days (experimental)]. One panel was lost (76-day, control), leaving a total of 9 panels for analysis, in addition to the native Basalt and Trap samples.

Environmental DNA was extracted from the basalt panel surfaces using a large-volume CTAB (hexadecyltrimethylammonium bromide) extraction. Individual basalt panels were placed into warmed (55oC) DNA extraction buffer, composed of 100 mM TRIS-HCl (pH 8.0), 20 mM EDTA (pH 8.0), 1.4 M NaCl and 2% CTAB, with mercaptoethanol added to 0.2% via syringe filter. Filter-sterilized Proteinase K solution and sodium dodecyl sulfate (SDS) solution were added to final concentrations of 0.1 mg/mL and 0.65% respectively. Covered by this extraction solution, each block was agitated on a shaker table for 2 hrs at 55oC. Biofilm removal was confirmed by inspection of the extracted blocks under a dissecting microscope. DNA was extracted from this solution with an equivalent volume of phenol:chloroform:isoamyl alcohol (25:24:1, pH 8.0), followed by an equivalent volume of chloroform:isoamyl alcohol (24:1). To precipitate the DNA, 0.1 volume of 3M sodium acetate and 2.5 volumes of cold 100% ethanol were added, then placed into -20oC for 3 hours before centrifuging at 10,000 x g for 15 min. After decanting the supernatant, the pellets were covered with cold 70% ethanol and centrifuged at 16,000 x g for 5 min. The ethanol was pipetted off and the pellets were dried. Isolated DNA was resuspended in 50-200 µl sterile water and kept at -80oC until use. Environmental DNA from Trap and native Basalt samples was extracted with using the Ultraclean Soil DNA extraction kit (MoBio laboratories).

The 16S rRNA region of environmental DNA was amplified (27F, 1492R primers) and replicates of 10 PCR amplifications (15 cycles each) were combined, precipitated using a QIAquick PCR purification kit (Qiagen), resuspended in 35 µl of sterile water and purified using the QIAquick gel extraction kit (Qiagen). Replicate PCR reactions were combined in order to minimize PCR bias (Polz and Cavanaugh, 1998) and only 15 cycles were used to decrease the formation of chimeric sequences and Taq error. The combined products were then reamplified with five additional PCR cycles to minimize the formation of heteroduplex molecules (Thompson et al., 2002) and purified using a QIAquick gel extraction kit (Qiagen).

Purified PCR products of the 16S rRNA gene were cloned with the Strataclone PCR cloning kit (Stratagene) for sequencing. Nearly complete, double-stranded sequences of 16S rRNA genes (~1,475 bp) were sequenced on a 96-capillary 3730xl DNA analyzer (Applied Biosystems) with primers M13F and M13R. For all libraries, single strand sequences were grouped by 97% similarity with the program Sequencher (version 4.1.2, Applied Biosystems) and one representative of each group was selected to sequence both forward and reverse strands. These groupings were designated operational taxonomic units (OTUs). Sequences were tested for the presence of chimeras with the program Mallard (Ashelford et al., 2006) and with the Bellerophon server (Huber et al., 2004). The ARB software package was used to analyze sequence data and construct trees (version 2.5b; O. Strunk and W. Ludwig, Technische Universitat Munchen, Munich, Germany).

Primers 517F and 806R were used to target the V4 region of the 16S rRNA gene (Caporaso et al., 2010). For multiplex sequencing, 8 forward primers were synthesized (Eurofin), each with a different tag. PCR reactions (50 µl volume) contained 250 nM of each of forward and reverse primers, 5 ng of template DNA and were performed using Picomaxx taq polymerase (Invitrogen, Carlsbad, CA) using the thermal profile 95°C for 2 min followed by 25 cycles of denaturation at 94°C for 15 s, primer annealing at 61°C and extension at 72°C for 45 s, with final extension of 72°C for 3 min. Amplicons were sequenced at EnGenCore (University of South Carolina) using 454 FLX chemistry. Raw sequences were processed with QIIME 1.8.0 (Caporaso et al., 2010). Sequences were excluded from analysis if they had a mean quality score < 25, were either < 200 or > 1000

bp in length, contained ambiguous nucleotides or had any mismatches in the forward and reverse primers. The sequences were assigned to individual samples by their adapter tags and the 16S rRNA primers were removed prior to analysis. The resulting data sets contained 254 bp of the bacterial V4 region. Potential chimeras were identified and removed using ChimeraSlayer. Trimmed sequences were then classified with RDP classifier. Operational taxonomic units (OTUs) were clustered at 97% similarity based on a distance matrix generated with the program QIIME.

Phylogenetic trees were constructed from aligned clone sequences and closely related environmental clones and cultures using the ARB software package (version 5.5; O. Strunk and W. Ludwig, Technische Universität Munchen, Munich, Germany). Tag sequences found at high abundances (>50 tags detected per sample), as well as those that were identical to a clone sequence or those that represent a unique lineage were inserted into bootstrapped trees using parsimony insertion tool. Alpha and beta diversity estimates were calculated using QIIME. After trimming each sample to an equal number of tags, bacterial diversity was estimated with N obs (observed richness), Chao1 (nonparametric richness estimator), nonparametic (np) Shannon diversity index, Simpson Evenness and Equitability. Structure of the pyrosequenced microbial communities was compared with two different methods. Principal coordinate analysis (PCoA) maps the samples on a set of orthogonal axes in order to explain the maximum amount of variation by the first coordinate and the second largest amount of variation by the second coordinate based on weighted UniFrac values (Lozupone et al., 2010). The UPGMA (unweighted pair group method with arithmetic mean) distance tree is based on a hierarchical clustering in which topological relationships are identified in order of similarity. The robustness of the UPGMA clusters was tested with jackknife analysis, based on 100 randomized subsamples.

Data Processing Description

The ARB software package was used to analyze sequence data and construct trees (version 2.5b; O. Strunk and W. Ludwig, Technische Universitat München, Munich, Germany).

The raw tag sequences were processed with QIIME 1.8.0 (Caporaso et al., 2010). Sequences were excluded from analysis if they had a mean quality score < 25, were either < 200 or > 1000 bp in length, contained ambiguous nucleotides or had any mismatches in the forward and reverse primers. The sequences were assigned to individual samples by their adapter tags and the 16S rRNA primers were removed prior to analysis. The resulting data sets contained 254 bp of the bacterial V4 region. Potential chimeras were identified and removed using ChimeraSlayer. Trimmed sequences were then classified with RDP classifier. Operational taxonomic units (OTUs) were clustered at 97% similarity based on a distance matrix generated with the program QIIME.

Phylogenetic trees were constructed from aligned clone sequences and closely related environmental clones and cultures using the ARB software package (version 5.5; O. Strunk and W. Ludwig, Technische Universität Munchen, Munich, Germany). Tag sequences found at high abundances (>50 tags detected per sample), as well as those that were identical to a clone sequence or those that represent a unique lineage were inserted into bootstrapped trees using parsimony insertion tool. Alpha and beta diversity estimates were calculated using QIIME. After trimming each sample to an equal number of tags, bacterial diversity was estimated with N obs (observed richness), Chao1 (nonparametric richness estimator), nonparametic (np) Shannon diversity index, Simpson Evenness and Equitability. Structure of the pyrosequenced microbial communities was compared with two different methods. Principal coordinate analysis (PCoA) maps the samples on a set of orthogonal axes in order to explain the maximum amount of variation by the first coordinate and the second largest amount of variation by the second coordinate based on weighted UniFrac values (Lozupone et al., 2010). The UPGMA (unweighted pair group method with arithmetic mean) distance tree is based on a hierarchical clustering in which topological relationships are identified in order of similarity. The robustness of the UPGMA clusters was tested with jackknife analysis, based on 100 randomized subsamples.

BCO-DMO Processing Notes:
- added conventional header with dataset name, PI name, version date
- modified parameter names to conform with BCO-DMO naming conventions
- changed n/a (not available) to nd (no data)
- converted latitude and longitude to decimal degrees
- reformatted date from dd-Mon-yy to yyyy-mm-dd
- created links to NCBI accession pages

[table of contents | back to top]

Data Files

File
16S_sequences.csv (Comma Separated Values (.csv), 27.51 KB) MD5:57edb4754e5cabb13c1b81381d6d5c75 Primary data file for dataset ID 687844

[table of contents | back to top]

Related Publications

Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J., & Weightman, A. J. (2006). New Screening Software Shows that Most Recent Large 16S rRNA Gene Clone Libraries Contain Chimeras. Applied and Environmental Microbiology, 72(9), 5734–5741. doi:10.1128/aem.00556-06 https://doi.org/10.1128/AEM.00556-06
Methods

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

Gulmann, L. K., Beaulieu, S. E., Shank, T. M., Ding, K., Seyfried, W. E., & Sievert, S. M. (2015). Bacterial diversity and successional patterns during biofilm formation on freshly exposed basalt surfaces at diffuse-flow deep-sea vents. Frontiers in Microbiology, 6. doi:10.3389/fmicb.2015.00901
Related Research

Huber, T., Faulkner, G., & Hugenholtz, P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics, 20(14), 2317–2319. doi:10.1093/bioinformatics/bth226
Methods

Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2010). UniFrac: an effective distance metric for microbial community comparison. The ISME Journal, 5(2), 169–172. doi:10.1038/ismej.2010.133
Methods

Polz, M. F., & Cavanaugh, C. M. (1998). Bias in Template-to-Product Ratios in Multitemplate PCR. Applied and Environmental Microbiology, 64(10), 3724–3730.
Methods

Thompson, J. R. (2002). Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by “reconditioning PCR.” Nucleic Acids Research, 30(9), 2083–2088. doi:10.1093/nar/30.9.2083
Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
sample_descrip	sample description: either 16S tags or clones	unitless
cruise_id	cruise identification	unitless
date_deployed	date that panels were deployed	yyyy-mm-dd
date_recovered	date that panels were recovered	yyyy-mm-dd
lat	latitude; north is positive	decimal degrees
lon	longitude; east is positive	decimal degrees
depth_w	depth of water	meters
site	sampling site	unitless
treatment	treatment type: experimental, control, or natural basalt (see methodology description for details)	unitless
panel_id	settling panel identifier	unitless
BioProject	NCBI BioProject accession number	unitless
BioProject_link	link to NCBI BioProject accession page	unitless
SRA_experiment	SRA experiment accession number	unitless
SRA_expt_link	link to SRA experiment accession page	unitless
NCBI_accession	NCBI sequence accession number	unitless
accession_link	link to sequence accession page	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Capillary 3730xl DNA analyzer (Applied Biosystems) and 454 FLX Technology
Generic Instrument Name	Automated DNA Sequencer
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

AT11-07

Website	https://www.bco-dmo.org/deployment/616332
Platform	R/V Atlantis
Start Date	2004-01-28
End Date	2004-02-24

AT11-07_Alvin_Dives

Website	https://www.bco-dmo.org/deployment/626019
Platform	Alvin
Start Date	2014-02-03
End Date	2014-02-19

AT11-20

Website	https://www.bco-dmo.org/deployment/654275
Platform	R/V Atlantis
Start Date	2004-11-05
End Date	2004-11-26
Description	More information is available from the Rolling Deck to Repository (R2R).

AT15-06

Website	https://www.bco-dmo.org/deployment/654285
Platform	R/V Atlantis
Start Date	2006-06-18
End Date	2006-07-07
Description	More information is available from Rolling Deck to Repository (R2R).

AT11-10

Website	https://www.bco-dmo.org/deployment/687864
Platform	R/V Atlantis
Start Date	2004-04-06
End Date	2004-04-30
Description	Part of Ridge Interdisciplinary Global Experiments (Ridge2000) Program to study hydrothermal vent area in the East Pacific Rise.

[table of contents | back to top]

Project Information

An Integrated Study of Energy Metabolism, Carbon Fixation, and Colonization Mechanisms in Chemosynthetic Microbial Communities at Deep-Sea Vents (Microbial Communities at Deep-Sea Vents)

Deep-sea hydrothermal vents, first discovered in 1977, are poster child ecosystems where microbial chemosynthesis rather than photosynthesis is the primary source of organic carbon. Significant gaps remain in our understanding of the underlying microbiology and biogeochemistry of these fascinating ecosystems. Missing are the identification of specific microorganisms mediating critical reactions in various geothermal systems, metabolic pathways used by the microbes, rates of the catalyzed reactions, amounts of organic carbon being produced, and the larger role of these ecosystems in global biogeochemical cycles. To fill these gaps, the investigators will conduct a 3-year interdisciplinary, international hypothesis-driven research program to understand microbial processes and their quantitative importance at deep-sea vents. Specifically, the investigators will address the following objectives: 1. Determine key relationships between the taxonomic, genetic and functional diversity, as well as the mechanisms of energy and carbon transfer, in deep-sea hydrothermal vent microbial communities. 2. Identify the predominant metabolic pathways and thus the main energy sources driving chemoautotrophic production in high and low temperature diffuse flow vents. 3. Determine energy conservation efficiency and rates of aerobic and anaerobic chemosynthetic primary productivity in high and low temperature diffuse flow vents. 4. Determine gene expression patterns in diffuse-flow vent microbial communities during attachment to substrates and the development of biofilms.

Integration: To address these objectives and to characterize the complexity of microbially-catalyzed processes at deep-sea vents at a qualitatively new level, we will pursue an integrated approach that couples an assessment of taxonomic diversity using cultivation-dependent and -independent approaches with methodologies that address genetic diversity, including a) metagenomics (genetic potential and diversity of community), b) single cell genomics (genetic potential and diversity of uncultured single cells), c) meta-transcriptomics and -proteomics (identification and function of active community members, realized potential of the community). To assess function and response to the environment, these approaches will be combined with 1) measurement of in situ rates of chemoautotrophic production, 2) geochemical characterization of microbial habitats, and 3) shipboard incubations under simulated in situ conditions (hypothesis testing under controlled physicochemical conditions). Network approaches and mathematical simulation will be used to reconstruct the metabolic network of the natural communities. A 3-day long project meeting towards the end of the second year will take place in Woods Hole. This Data Integration and Synthesis meeting will allow for progress reports and presentations from each PI, postdoc, and/or student, with the aim of synthesizing data generated to facilitate the preparation of manuscripts.

Intellectual Merit. Combining the community expression profile with diversity and metagenomic analyses as well as process and habitat characterization will be unique to hydrothermal vent microbiology. The approach will provide new insights into the functioning of deep-sea vent microbial communities and the constraints regulating the interactions between the microbes and their abiotic and biotic environment, ultimately enabling us to put these systems into a quantitative framework and thus a larger global context.

Broader Impacts. This is an interdisciplinary and collaborative effort between 4 US and 4 foreign institutions, creating unique opportunities for networking and fostering international collaborations. This will also benefit the involved students (2 graduate, several undergraduate) and 2 postdoctoral associates. This project will directly contribute to many educational and public outreach activities of the involved PIs, including the WHOI Dive & Discover program; single cell genomics workshops and Cafe Scientifique (Bigelow); REU (WHOI, Bigelow, CIW); COSEE and RIOS (Rutgers), and others. The proposed research fits with the focus of a number of multidisciplinary and international initiatives, in which PIs are active members (SCOR working group on Hydrothermal energy and the ocean carbon cycle, http://www.scorint. org/Working_Groups/wg135.htm; Deep Carbon Observatory at CIW, https://dco.gl.ciw.edu/; Global Biogeochemical Flux (GBF) component of the Ocean Observatories Initiative (OOI), http://www.whoi.edu/GBF-OOI/page.do?pid=41475)

[table of contents | back to top]

Program Information

Dimensions of Biodiversity (Dimensions of Biodiversity)

Website: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503446

Coverage: global

(adapted from the NSF Synopsis of Program)

Dimensions of Biodiversity is a program solicitation from the NSF Directorate for Biological Sciences. FY 2010 was year one of the program. [MORE from NSF]

The NSF Dimensions of Biodiversity program seeks to characterize biodiversity on Earth by using integrative, innovative approaches to fill rapidly the most substantial gaps in our understanding. The program will take a broad view of biodiversity, and in its initial phase will focus on the integration of genetic, taxonomic, and functional dimensions of biodiversity. Project investigators are encouraged to integrate these three dimensions to understand the interactions and feedbacks among them. While this focus complements several core NSF programs, it differs by requiring that multiple dimensions of biodiversity be addressed simultaneously, to understand the roles of biodiversity in critical ecological and evolutionary processes.

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.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1136727
NSF Division of Ocean Sciences (NSF OCE)	OCE-0452333

[table of contents | back to top]