Metadata and accessions for comparative single-cell genomics of Chloroflexi from the Okinawa Trough deep subsurface biosphere from DV/Chikyu OIDP stations, Sept-Oct. 2010 (Subsurface FeOBs project)

Website: https://www.bco-dmo.org/dataset/683021
Data Type: Cruise Results
Version: 2
Version Date: 2021-03-03

Project
» Iron-oxidizing Bacteria from the Okinawa Trough Deep Subsurface Biosphere (Subsurface FeOBs)

Program
» Center for Dark Energy Biosphere Investigations (C-DEBI)

Contributors	Affiliation	Role
Moyer, Craig L.	Western Washington University (WWU)	Principal Investigator
Fullerton, Heather	Western Washington University (WWU)	Scientist
Copley, Nancy	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Metadata and accessions for comparative single-cell genomics of Chloroflexi from the Okinawa Trough deep subsurface biosphere from DV/Chikyu OIDP stations, Sept-Oct. 2010; with links to NCBI and IMG repositories.

Table of Contents

Coverage

Spatial Extent: **N**:27.78333 **E**:126.91667 **S**:27.78333 **W**:125.88333
Temporal Extent: 2010-09 - 2010-09

Methods & Sampling

Methodology from Fullerton and Moyer (2016). See paper for references cited below.

Sample collection. Subsurface sediments were collected on IODP expedition 331 (Deep Hot Biosphere) from 1 September through 4 October 2010 (Fig. 1). Onboard contamination testing of sites C0015 (126°53=E, 27°47=N; hole B; section 1H-5; 5.6mbsf) and C0017 (126°55=E, 27°47=N; hole C; section 1H-7; 26.6 m bsf) found no indication of interior-core contamination using fluorescent microspheres (both holes C0015B and C0017C) and perfluorocarbon tracer (hole C0017C only). The sample from hole C0017C was also verified by

PCR-generated phylotype comparisons based on 97% similarity to phylotypes obtained from drilling mud at a contamination level of 1% or less (26). Subsamples were aseptically collected from the interiors of whole-round cores and stored in cryovials with 27% (vol/vol) glycerol at -80°C.

Single-cell source. Core depths were chosen from sites C0015 and C0017, which were characterized as weakly oxidized pumiceous gravels with no detected sulfide mineralization and less than 0.1 wt% total organic carbon, total nitrogen, and total sulfur (25). The selected samples for single-cell genomes were from subsurface depths of 5.6 m bsf from hole C0015B and from 26.6 m bsf from hole C0017C. Temperatures were estimated at ~10.5°C for C0015B and ~8.1°C for C0017C at these depths. Details of geochemistry and lithography have been previously described (12, 24, 25).

Single-cell sorting, amplification, sequencing, and annotation. Samples from sites C0015 and C0017 (Fig. 1) were diluted with 1 ml of filter-sterilized artificial seawater (27), making a slurry, and then passed through a 90-um nylon mesh filter twice and centrifuged at ~500 ug for 2 min to produce a particle-free cell suspension. The suspension was then processed using fluorescence-activated single-cell sorting at the Single Cell Genome Center (SCGC) at Bigelow Laboratory for Ocean Sciences. Single-cell sorting and multiple displacement amplification (MDA) have been previously described (28). The amplified SSU rRNA gene sequences (27F/907R) were classified using the Ribosomal Database Project (RDP) online classifier (28, 29). Based on their SSU rRNA gene identities, nine Chloroflexi SAGs (of the total 29 unique MDA reactions identified after cell sorting) were chosen for whole-genome sequencing. These SAGs were sequenced and assembled, and contamination was checked by the SCGC, using previously well-described parameters (28, 29). Assembly was done using SPAdes v.3.0.0 (30). All contigs were compared to ensure no cross contamination among SAGs and the NCBI nt database, which was followed by tetramer principal-component analysis as previously described (31-33). These analyses revealed no contamination. The full name for each of the SAGs was shortened, e.g., Anaerolineales bacterium SCGC AC-711-B22 was shortened to An-B22. Phylogeny was abbreviated as follows: Anaerolineales to An, Dehalococcoidales to De, and Thermoflexales to Th. The assembled genomes were annotated using RAST (34). Gene annotations were compared to NCBI GenBank via BLASTn, and the results can be found in Tables S2 to S4 in the supplemental material.

The Anaerolineales SAGs were compared to the genome of Anaerolinea thermophila UNI-1 (GenBank accession number NC_014960) and the single Thermoflexales SAG to that of Thermoflexus hugenholtzii JAD2 (NCBI BioProject PRJNA195829), as they were determined to be their closest respective relatives. The type strain A. thermophila UNI-1 was isolated from an anaerobic granular sludge reactor treating fried soybean curd manufacturing wastewater in Japan (35), while the type strain T. hugenholtzii JAD2 was isolated from the sediment of Great Boiling Spring in Nevada (36). Both are considered thermophilic, Gram-negative, nonspore-forming, heterotrophic bacteria that grow in multicellular filaments (36, 37).

Phylogenetic analysis. SSU rRNA gene sequences and phylogenetic relatives were aligned using the Silva SINA aligner (38). For the rdhA analysis, amino acids were aligned using ClustalW within Geneious (39, 40). The resulting alignments were manually screened and then used to create a phylogenetic consensus tree using MrBayes within Geneious (41). Parameters included using the HKY85 substitution model, the chain length set at 1,100,000, and a subsampling frequency of 200. Priors were set with an unconstrained branch length. The average nucleotide identity (ANI) was calculated for the SAGs and selected genomes, with the BLAST parameters as previously described (42).

Genome completeness estimates. Genome completeness estimates were determined with BLASTP using predicted amino acid sequences against a set of single-copy core genes (43). To be considered valid, all proteins must have at least 30% identity over at least 30% of the length of the core gene (44). The core gene group is made up of 66 previously established genes belonging to a nonredundant list as examined by gene ontology (GO) annotations (44, 45).

Accession numbers. The SSU rRNA gene sequences obtained from MDA have been submitted to the NCBI GenBank database (accession numbers KT119838 to KT119846). All the SAGs have been made public in the Integrated Microbial Genomes (IMG) database (IMG submission identifiers [IDs] 68650, 69642 to 69645, 69647 to 69649, and 69684).

Data Processing Description

BCO-DMO Processing Notes:
Created a servable data file from metadata and data provided by the PI and created links to accession pages at NCBI and IMG, when available.

version 2 (2021-03-01), added IMG accessions links, replaces version 1(2017-03-21).
* on 2022-03-03 archival version 2 (2021-03-03) reimported into the BCO-DMO data system to fix a data access issue.

[table of contents | back to top]

Data Files

File
accessions.csv (Comma Separated Values (.csv), 5.34 KB) MD5:fcf6ea08052ab47e05080ad1f4b1bd84 Primary data file for dataset ID 683021

[table of contents | back to top]

Supplemental Files

File
Supplementary Information for Comparative single-cell genomics of Chloroflexi from the Okinawa Trough deep subsurface biosphere filename: Appl.Environ.Microbiol.2016-Fullerton-3000-8_SUPPL.pdf (Portable Document Format (.pdf), 359.19 KB) MD5:7ff3cb83dbe500e53a2f8c48de4557be Table S1. Phylogeny and metabolism of select organisms in the phylum Chloroflexi Table S2. One Carbon Pool Genes Table S3. Glycolysis Genes Table S4. TCA Cycle Genes Table S5. Estimated genome size based percent genome recovery as determined by tRNAs of nearest neighbor. Figure S1. Pairwise comparisons of Chloroflexi genomes ANI vs. SSU percent identity. Dashed lines represent species level cutoff values. Figure S2. N-terminal amino acid alignment of RdhA sequences, including the one found in SAG An-B22 with boxes highlighting the Fe-S cluster binding motif. This is a subset of the amino acid alignment that was used to create Figure 4.

[table of contents | back to top]

Related Publications

Fullerton, H., & Moyer, C. L. (2016). Comparative Single-Cell Genomics of Chloroflexi from the Okinawa Trough Deep-Subsurface Biosphere. Applied and Environmental Microbiology, 82(10), 3000–3008. doi:10.1128/aem.00624-16 https://doi.org/10.1128/AEM.00624-16
Results

[table of contents | back to top]

Parameters

Parameter	Description	Units
cruise_id	cruise identifier	unitless
cruise_description	description of site	unitless
site	site identification	unitless
site_description	description of site	unitless
station	station identifier	unitless
lat	latitude at which sample was collected	decimal degrees
lon	longitude at which sample was collected	decimal degrees
hole	drill hole identifier	unitless
section	drill core sample section	unitless
depth_bsf	depth below seafloor	meters
gene	gene that was sequenced	unitless
service	method: Single Amplified Genome	unitless
taxon_name	description of taxon	unitless
NCBI_BioSample	NCBI BioSample identifier	unitless
BioSample_link	link to NCBI BioSample page	unitless
NCBI_accession_number	NCBI accession number	unitless
NCBI_accession_link	link to NCBI accession number page	unitless
IMG_taxon_id	IMG taxon id	unitless
IMG_link	link to IMG taxon id page	unitless
date_collected	collection date formatted as Mon-yyyy	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	Genetic analyses performed at Single Cell Genome Center (SCGC) at Bigelow Laboratory for Ocean Sciences
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	
Generic Instrument Name	Thermal Cycler
Dataset-specific Description	Genetic analyses performed at Single Cell Genome Center (SCGC) at Bigelow Laboratory for Ocean Sciences
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

Chikyu-331

Website	https://www.bco-dmo.org/deployment/682963
Platform	Chikyu
Start Date	2010-09-01
End Date	2010-10-03
Description	IODP cruise for Deep Hot Biosphere expedition. For more information, including cruise reports, visit http://publications.iodp.org/proceedings/331/331title.htm

[table of contents | back to top]

Project Information

Iron-oxidizing Bacteria from the Okinawa Trough Deep Subsurface Biosphere (Subsurface FeOBs)

Website: http://www.jamstec.go.jp/chikyu/eng/Expedition/okinawa/exp331.html

Coverage: Iheya North hydrothermal system in the Mid-Okinawa Trough (126 53.80'E, 27 47.45'N)

Description from NSF award abstract:
Communities of Fe-oxidizing Bacteria (FeOB) are common at sites of hydrothermal venting and are known to form complex communities in microbial mats, hydrothermal sediments, oceanic crustal basalts and borehole fluids (among others) that are iron-rich and low in oxygen. Studies from sites around the Pacific Ocean have found the Zetaproteobacteria to be ubiquitous members of these FeOB communities. Previously, borehole fluids from the Southern Mariana Backarc have been shown to support several novel and distinct lineages of endemic Zetaproteobacteria. Sampling from the deep subsurface at the Iheya North hydrothermal field (IODP Expedition 331) explores this deep subsurface FeOB biodiversity and has resulted in multiple enrichments using both microaerophilic and anaerobic culturing conditions. Zetaproteobacteria have been detected at levels up to 13% of the total bacterial community from these subsurface core samples. This project expands post-

cruise analyses to focus on assessing these FeOB subsurface communities using a combined single cell genomics, community-level metagenomics and a FeOB directed cultivation approach. This approach will allow insights into the exclusive physiology and metabolism of these Zetaproteobacteria thereby demonstrating key features as to how they survive, compete and grow within these complex subsurface microbial communities. A comparative analysis (from an evolutionary standpoint) to determine the unique differences between subsurface Zetaproteobacteria and those living above the seafloor surface is an integral part of this effort.

A better understanding of Zetaproteobacteria is of central interest to scientists interested in areas of earth science and oceanography because they illustrate how microbes can influence geochemical cycling and mineral deposition. Furthermore, morphological structures similar to those produced by Zetaproteobacteria can still be identified 100's of millions (and possibly billions) of years back in the rock record, making them of paleontological (and potentially of exobiological) interest. As knowledge of extant populations grows, it is possible they will also help to inform us of environmental change in past Earth history. From a practical standpoint, these organisms might be thought of as 'micro-machines', spinning out threads of iron oxyhydroxide that coalesce in unknown ways. These oxides are known to be highly reactive with a range of other metals, organic compounds, and nutrients, thus impacting many other biogeochemical cycles. In particular, this study allows the opportunity to compare and contrast a known monophyletic class of Proteobacteria, to study the differences between subsurface and supersurface lineages of Zetaproteobacteria, thereby yielding fundamental clues into microbial biogeography. A wealth of educational outreach opportunities will be made possible by this work, including graduate education, research experiences for undergraduates, and teacher training.

Related Publications:

Fullerton, H. and C. L. Moyer. 2016. Comparative single-cell genomics of Chloroflexi from the Okinawa Trough deep subsurface biosphere. Appl. Environ. Microbiol. 82:3000-3008. doi: 10.1128/AEM.00624-16

Takai, K., Mottl, M. J., Nielsen, S. H. H., and the IODP Expedition 331 Scientists: IODP Expedition 331: Strong and Expansive Subseafloor Hydrothermal Activities in the Okinawa Trough, Sci. Dril., 13, 19-27, doi:10.5194/sd-13-19-2012, 2012.

Yanagawa K, Nunoura T, McAllister SM, Hirai M, Breuker A, Brandt L, House CH, Moyer CL, Birrien J-L, Aoike K, Sunamura M, Urabe T, Mottl MJ and Takai K (2013) The first microbiological contamination assessment by deep-sea drilling and coring by the D/V *Chikyu* at the Iheya North hydrothermal field in the Mid-Okinawa Trough (IODP Expedition 331). Front. Microbiol. 4:327. doi:10.3389/fmicb.2013.00327.

[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.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1260710

[table of contents | back to top]