16S rRNA sequences from subsurface microbial communities (Blackwood Sinkhole)

Website: https://www.bco-dmo.org/dataset/847824
Data Type: Other Field Results
Version: 1
Version Date: 2021-05-25

Project
» Subseafloor prokaryotic and viral communities and interactions in anoxic marine basins: a case study from a 3000-year old stratigraphic succession (Virus-host anoxic sediment)

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

Contributors	Affiliation	Role
Labonté, Jessica M.	Texas A&M, Galveston (TAMUG)	Principal Investigator
van Hengstum, Pete	Texas A&M, Galveston (TAMUG)	Co-Principal Investigator
Rauch, Shannon	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This dataset includes information about 16S rRNA sequences from subsurface microbial communities at Blackwood Sinkhole. Samples were collected in July 2018.

Table of Contents

Coverage

Spatial Extent: **Lat**:26.79 **Lon**:-77.42
Temporal Extent: 2018-07-29

Methods & Sampling

Methods are detailed in Risley et al., submitted. Briefly, a core was recovered from Blackwood Sinkhole, Bahamas, and divided into 24 stratigraphic layers. For each layer, DNA was extracted and the 16S rRNA gene was amplified and sequenced to characterize the microbial communities in relation to the C:N ratio.

Study site
Blackwood Sinkhole is located ~220 m from the shoreline on the northeastern coast of the Great Abaco Island on the Little Bahama Bank (26.79, 77.42) (van Hengstum et al., 2016). Blackwood Sinkhole is 32 m in diameter, and this groundwater-fed basin has a stratified water column. The sinkhole hydrography is characterized by an upper ~10 m of meteoric water (1.4 psu), which transitions to a mixing zone below (10-15 meters below sea level), and this rests above anoxic saline groundwater (~15-40 meters below sea level) (39.9 psu) (van Hengstum et al., 2016). The groundwater stratification and geomorphology of Blackwood Sinkhole has

promoted excellent sediment preservation from a lack of vertical mixing from either invertebrate bioturbation, or physical mixing from wave or wind action. Basal sedimentary deposits in Blackwood Sinkhole are carbonate (karst) gravels, which abruptly transition to laminated sapropel (i.e., high organic matter) interbedded with carbonate horizons that were deposited over the last 3,000 years (van Hengstum et al., 2016). The laminated sapropel has a high total organic carbon (TOC) content ~10% (Tamalavage et al., 2018). However, sedimentation rate and location of sedimentation in the sinkhole has not been laterally or temporally uniform over time. This is because (a) the bottom geometry of the sinkhole basin is complex as inherited from the collapse of an original limestone cave, and (b) the sources and delivery mechanisms of sediment during the last 3000 years are linked to primary productivity, erosion of the vertical sinkhole walls, and deposition of organic matter derived from adjacent wetland habitats (Tamalavage et al., 2018). Using detailed radiocarbon dating on other core samples, constant sedimentation (~0.4 mm yr^{-1}) was initiated 3,000 Calibrated years Before Present (Cal yrs BP) on a core collected from the sinkhole periphery, with evidence of coarse-grained particle deposition during intense hurricane strikes on Abaco Island (van Hengstum et al., 2016). In contrast, sedimentation in the center of the sinkhole was delayed until ~1,600 Cal yrs BP, and sedimentation rates in the center were higher than the periphery (1.2 mm yr^{-1}).

Based on a detailed analysis on the changes in organic matter provenance through time using a 3-endmember mixing model using stable carbon isotopes (δ^{13}Corg) and the C:N ratio, there were three periods of time in Blackwood Sinkhole when organic matter was dominated by a different source (Tamalavage et al., 2018). The oldest part of the record (Group 3, 1520-2990 Cal yrs BP) was characterized by primarily terrestrial organic matter deposition, and the middle part of the record was dominated by inputs from primary productivity (Group 2, 1009-1502 Cal yrs BP). Finally, the last millennium was dominated by organic matter inputs from an adjacent wetland (Group 1, 0-1008 Cal yrs BP). These wetlands were emplaced on the epikarst surface in response to concomitant regional sea-level and groundwater-level rise inundating topographic lows on the landscape adjacent to Blackwood Sinkhole, which created favorable wetland habitat (Tamalavage et al., 2018).

Sample collection and handling

A 7.6 cm push core (BLWD-C7; 26.79°N, 77.42°W) was collected on July 29, 2018 with a polyvinyl chloride pipe, using advanced technical scuba diving procedures following guidelines established by the American Academy of Underwater Sciences. The periphery of the sinkhole bottom was targeted for core sampling in an attempt to re-collect the last 3,000 years of sedimentary deposition (van Hengstum et al., 2016). In the lab, the core was sectioned lengthwise into a working and archive half (subsequently stored at 4°C), photographed, radiographed, and the lithology was qualitatively described (Schnurrenberger et al., 2003). The working half was split into 24 stratigraphic layers, based on sediment color and texture, and stored at -80°C for further analyses.

To validate the age of the core and compare to previous results, five samples (20.1−22.2 cm, 32.9−35 cm, 52.1−54.2 cm, 64.5−67.7 cm, and 79.5−86.9 cm) of terrestrial plant macrofossils (e.g., leaves, twigs) were selected for radiocarbon dating of at the National Ocean Sciences Accelerator Mass Spectrometry facility at Woods Hole Oceanographic Institution (Woods Hole, MA, USA). Conventional radiocarbon ages were calibrated into Calendar years Before Present (Cal yrs BP, present is considered 1950 Common Era) with IntCAL13 (Reimer et al., 2013). A final downcore Bayesian age model for BLWD-C7 was computed using the R program Bacon v2.2 (Blaauw and Christen, 2011) to provide probability estimates at each core depth using three of the five samples (32.9−35 cm, 52.1−54.2 cm, 64.5−67.7 cm) and the surface designated as −68 Cal yrs BP (2018 year of collection). The shallowest radiocarbon result from the base of the prominent sapropel horizon at 20.1 to 22.2 cm in BLWD-C7 provided a calibrated age result that exceeded 500 Cal yrs BP, yet previous results indicate a much younger age for this horizon (291-223 Cal yrs BP, Tamalavage et al., 2018). Similarly, the deepest radiocarbon calibration result from 79.5−86.9 cm in BLWD-C7 caused significant change in sedimentation rate, and as such is suspected to be old and reworked terrestrial plant remains. These dates were rejected from the final radiocarbon age model.

Measurement of carbon and nitrogen content

Total carbon and nitrogen content measurements were performed on subsamples from the 24 sediment horizons from BLWD-C7. First, the subsamples were freeze-dried overnight, homogenized, and 2 to 6 mg of ground sample were placed into tin capsules and measured on a Costech instruments ECS 4010 CHNSO Analyzer (Costech Analytical Technologies) to measure total carbon (TC) and total nitrogen (TN). Data calibration was determined relative to acetanilide and standard reference material for organics in marine sediment according to the National Institute of Standards and Technology (NIST). To measure organic carbon (with mass correction applied), the samples were acidified using 8 mL of 1M HCl for 24 h or until effervescence ceased, then desiccated at 60°C, and re-homogenized. The ground, acidified samples were weighed (0.5 to 1.8 mg) into silver capsules then processed on the CHN analyzer. The potential loss of carbon from the direct acidification process was corrected by multiplying the percent of sample remaining (post-acidification weight subtracted from pre-acidification weight/pre acidification weight) (Tamalavage et al., 2018). The atomic C:N

ratio was determined using the organic carbon (OC) acidified values divided by the Total Nitrogen (TN) values unacidified and multiplied by the molecular weight ratio (14.01/12.01) (Tamalavage et al., 2018). Precision on replicates measurements of C and N was within ±2% weight percent.

16S rRNA gene sequencing and analysis
Total DNA was extracted from separate sediment subsamples (1-2 g) from each of the 24 horizons using the DNeasy PowerSoil kit (Qiagen, USA) and stored at -20˚C until PCR amplification. A negative control using 0.5 mL of 0.2 µm sterile MilliQ water was also extracted to identify possible contamination from the ambient lab and kit reagents. The PCR primers 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and 806R (5'-CCGYCAATTYMTTTRAGTTT-3') were used to target the V4 region of the 16S rRNA gene (Parada et al., 2016). Thermal cycling was performed under the following conditions: initial preheating for 3 min at 94˚C; 35 cycles of denaturation at 94˚C for 1 min, annealing at 50˚C for 1 min, and extension at 72˚C for 1 min 45 s; final extension at 72˚C for 10 min. PCR was completed in triplicates and the products were pooled and cleaned using the MinElute PCR Purification Kit (Qiagen, USA). Amplicons (100-200 ng) were sent to the Texas A&M AgriLife Bioinformatics and Genomics facility for library preparation with the Amplicon library preparation kit (Illumina) and sequencing with Illumina MiSeq with 250 bp paired-ends.

Analyses of the 16S rRNA gene amplicons were completed using the software *mothur* as presented in the MiSeq Standard Operating Procedure (SOP) tutorial (Kozich et al., 2013), which included reducing sequencing and PCR errors, processing the improved sequences, running an alignment using the reference SILVA alignment (v132), removing poorly aligned sequences and undesirables, and pre-clustering the sequences into amplicon sequence variants (ASVs). The remove.lineage command using all the lineages observed in the negative control at taxonomic level 6 (genus) was used to remove any potential contaminants from final analyses. Operational taxonomic unit (OTU)-based analyses using sequences clustered with the split method argument were also performed (i.e., rarefaction curves and heatmap of shared OTUs). Using a Bray-Curtis dissimilarity matrix, the beta-diversity between samples was examined and ordinated by non-metric multidimensional scaling (NMDS) in R (RCore Team, 2013), with overlaying the carbon content parameters (TC, TN, and C:N ratio) applying the 'envfit' function from the 'vegan' package (Dixon, 2003). Analysis of stratigraphically-constrained (i.e., age constrained according to radiocarbon results) hierarchical clustering using the package 'rioja' in R (Juggins, 2017). The sequences are available in GenBank under the BioProject #PRJNA639820.

Data Processing Description

Data Processing
Analyses of the 16S rRNA gene amplicons were completed using the software *mothur* as presented in the MiSeq Standard Operating Procedure (SOP) tutorial (Kozich *et al.*, 2013), which included reducing sequencing and PCR errors, processing the improved sequences, running an alignment using the reference SILVA alignment (v132), removing poorly aligned sequences and undesirables, and pre-clustering the sequences into amplicon sequence variants (ASVs). The remove.lineage command using all the lineages observed in the negative control at taxonomic level 6 (genus) was used to remove any potential contaminants from final analyses. Operational taxonomic unit (OTU)-based analyses using sequences clustered with the split method argument were also performed (*i.e.*, rarefaction curves and heatmap of shared OTUs). Using a Bray-Curtis dissimilarity matrix, the beta-diversity between samples was examined and ordinated by non-metric multidimensional scaling (NMDS) in R (RCore Team, 2013), with overlaying the carbon content parameters (TC, TN, and C:N ratio) applying the 'envfit' function from the 'vegan' package (Dixon, 2003). Analysis of stratigraphically-constrained (*i.e.*, age constrained according to radiocarbon results) hierarchical clustering using the package 'rioja' in R (Juggins, 2017). The sequences are available in GenBank under the BioProject #PRJNA639820.

BCO-DMO Processing:
- renamed fields to conform with BCO-DMO naming conventions;
- created separate fields for latitude and longitude;
- corrected year to 2018;
- removed empty/unnecessary columns.

[table of contents | back to top]

Data Files

File
BWLD-C7_16S_rRNA.csv (Comma Separated Values (.csv), 4.44 KB) MD5:7b5ea5b1b94cc51fe2150ab8404b5b59
Primary data file for dataset ID 847824

[table of contents | back to top]

Related Publications

Blaauw, M., & Christen, J. A. (2011). Flexible paleoclimate age-depth models using an autoregressive gamma process. Bayesian Analysis, 6(3). doi:10.1214/11-ba618
Methods

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. Journal of Vegetation Science, 14(6), 927–930. doi:10.1111/j.1654-1103.2003.tb02228.x
Methods

Juggins S. (2017). rioja: Analysis of Quaternary Science Data. http://www.staff.ncl.ac.uk/stephen.juggins/.
Software

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. Applied and Environmental Microbiology, 79(17), 5112–5120. doi:10.1128/aem.01043-13
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

RCore Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (http://www.R-project.org/)
Software

Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Ramsey, C. B., … van der Plicht, J. (2013). IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. Radiocarbon, 55(4), 1869–1887. doi:10.2458/azu_js_rc.55.16947
Methods

Schnurrenberger, D. (2003). Journal of Paleolimnology, 29(2), 141–154. doi:10.1023/a:1023270324800
Methods

Tamalavage, A. E., van Hengstum, P. J., Louchouarn, P., Molodtsov, S., Kaiser, K., Donnelly, J. P., … Fall, P. L. (2018). Organic matter sources and lateral sedimentation in a Bahamian karst basin (sinkhole) over the late Holocene: Influence of local vegetation and climate. Palaeogeography, Palaeoclimatology, Palaeoecology, 506, 70–83. doi:10.1016/j.palaeo.2018.06.014
Methods

Van Hengstum, P. J., Donnelly, J. P., Fall, P. L., Toomey, M. R., Albury, N. A., & Kakuk, B. (2016). The intertropical convergence zone modulates intense hurricane strikes on the western North Atlantic margin. Scientific Reports, 6(1). doi:10.1038/srep21728
Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
sample_name	Sample name	unitless
sample_accession	NCBI BioSample accession number	unitless
bioproject_accession	NCBI BioProject accession number	unitless
collection_date	Date of sample collection; format: YYYY-MM-DD	unitless
depth	Sample depth	centimeters (cm) below seafloor
env_local_scale	description of the geographic environmental feature sampled (Blackwood Sinkhole)	unitless
env_medium	environment where the sample was obtained; all subsamples are from the same sediment core	unitless
geo_loc_name	Geolocation name	unitless
latitude	Latitude of sample collection	degrees North
longitude	Longitude of sample collection	degrees East
design_description	how the sample was analyzed	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina MiSeq
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	Amplicons (100-200 ng) were sent to the Texas A&M AgriLife Bioinformatics and Genomics facility for library preparation with the Amplicon library preparation kit (Illumina) and sequencing with Illumina MiSeq with 250 bp paired-ends.
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	Costech instruments ECS 4010 CHNSO Analyzer
Generic Instrument Name	Costech International Elemental Combustion System (ECS) 4010
Dataset-specific Description	Total carbon (TC) and total nitrogen (TN) were measured using a Costech instruments ECS 4010 CHNSO Analyzer (Costech Analytical Technologies).
Generic Instrument Description	The ECS 4010 Nitrogen / Protein Analyzer is an elemental combustion analyser for CHNSO elemental analysis and Nitrogen / Protein determination. The GC oven and separation column have a temperature range of 30-110 degC, with control of +/- 0.1 degC.

Dataset-specific Instrument Name	push core
Generic Instrument Name	Push Corer
Dataset-specific Description	A 7.6 cm push core (BLWD-C7; 26.79°N, 77.42°W) was collected on July 29, 2018 with a polyvinyl chloride pipe, using advanced technical scuba diving procedures following guidelines established by the American Academy of Underwater Sciences.
Generic Instrument Description	Capable of being performed in numerous environments, push coring is just as it sounds. Push coring is simply pushing the core barrel (often an aluminum or polycarbonate tube) into the sediment by hand. A push core is useful in that it causes very little disturbance to the more delicate upper layers of a sub-aqueous sediment. Description obtained from: http://web.whoi.edu/coastal-group/about/how-we-work/field-methods/coring/

[table of contents | back to top]

Project Information

Subseafloor prokaryotic and viral communities and interactions in anoxic marine basins: a case study from a 3000-year old stratigraphic succession (Virus-host anoxic sediment)

Website: https://www.darkenergybiosphere.org/award/subseafloor-prokaryotic-and-viral-communities-and-interactions-in-anoxic-marine-basins-a-case-study-from-a-3000-year-old-stratigraphic-succession/

Coverage: Blackwood Sinkhole, Bahamas

Prokaryotes make up the majority of the biomass in sediment, which covers two thirds of the Earth's surface. In sediment, prokaryotes and viruses play a role in cycling organic carbon and regulate the fluctuation of organic matter. In anoxic sediment, the relationships between geochemical gradients, genomic potential, and virus-host interactions remain understudied and poorly understood. We characterized the prokaryotic and viral diversity along the geochemical gradients of a laminated sediment core from an anoxic sinkhole, Blackwood Sinkhole, Bahamas. We analyzed the pore water chemistry (nutrients, carbon, nitrogen) and identified the various sources contributing to the preserved sedimentary organic matter (surrounding terrestrial inputs or authigenic primary productivity). We characterized the microbial community composition of specific stratigraphic layers using 16S ribosomal RNA gene sequencing and metagenomics to show that there is a microbial succession down the core. Finally, we looked at viral assemblages using metagenomics and viral production through virus-induced microbial experiments to characterize virus-host interactions and better understand the role of viral lysis in this system. Through the characterization of the relationships of microbes between each other and with their environment, we aimed to identify the role organic and inorganic matter availability plays in shaping viral and prokaryotic communities, as well as how microbial communities shape their environment.

[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-0939564

[table of contents | back to top]