Microbiology Resource Announcements

| Environmental Microbiology | Announcement

Sediment and groundwater metagenomes from subsurface microbial communities from the Oak Ridge National Laboratory Oak Ridge Reservation, Oak Ridge, Tennessee, USA

Lauren M. Lui,[1] Torben N. Nielsen,[1] Heidi J. Smith,[2] John-Marc Chandonia,[1] Jennifer V. Kuehl,[1] Fangchao Song,[1] Andrew Sczesnak,[3] Andrew Hendrickson,[1] Terry C. Hazen,[4,5] Matthew W. Fields,[2,6] Adam P. Arkin[1,3]

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT We report 26 subsurface sediment and 9 groundwater metagenomes from the Oak Ridge Reservation at Oak Ridge, TN, USA. Samples were collected from various depths and phases (attached vs planktonic) to study subsurface microbial metabolism, the effect of contamination on microbial communities, and differences across groundwater and sediment microbial communities.

KEYWORDS metagenomics, sediment, groundwater, subsurface

The terrestrial subsurface is a critical biome for the Earth's biogeochemical cycling and sources of consumable freshwater; however, we still know little about the microbial communities that drive the biogeochemical processes and how they are impacted by environmental conditions (1). Environmental sequencing has provided the best avenue for studying these microbial communities, especially since many species have unknown conditions for lab culturing (2). In this study, we generated metagenomes to study the subsurface microbiome from the centimeter scale to the field scale (3) and to study the relationship between the sediment and water communities at the Oak Ridge Reservation, a site that has been impacted by heavy metal, nitrate, and radionuclide contamination due to leaching from waste disposal ponds (4, 5). We generated shotgun metagenomics data from 77 sediment samples and 33 groundwater samples.

Sediment samples were taken from the boreholes EB106 and EB271, whose drilling, geology, location, and biogeochemistry are described in reference 6. Each borehole was cut into approximately 22.86 cm cores. From each core, we took 5 g replicates for DNA extraction using a modified protocol of the Qiagen PowerMax Soil kit to reduce bead-beating (and thus DNA fragmentation) as described in reference 7 and documented on Protocols.io (dx.doi.org/10.17504/protocols.io.kqdg3kepqv25/v1). For EB271, cores had 1–3 replicates. For EB106, for 12 cores, we were able to obtain enough DNA for metagenomics sequencing. The first five cores had three replicates. The other cores only have one replicate, either due to poor DNA yield or failed metagenomics library prep, and some required ~20 g of sediment to obtain enough DNA for sequencing, as indicated in Table 1.

Water samples were collected from adjacent groundwater wells (FW106, FW115-24, GW271) to the boreholes 1 week after drilling. Well construction, location, and biogeochemistry are described in reference 6. On the day of sampling, 5 L of water was filtered sequentially in an anaerobic chamber through 0.2 and 0.1 um PES filters. Water was also shipped to Lawrence Berkeley Lab overnight, and biomass from 2 L of water was collected on 0.2 um PES filters. DNA was extracted from filters using the same method as the sediment.

Editor Irene L. G. Newton, Indiana University, Bloomington, Indiana, USA

Address correspondence to Lauren M. Lui, lmlui@lbl.gov, or Adam P. Arkin, aparkin@lbl.gov.

The authors declare no conflict of interest.

See the funding table on p. 4.

Ed. Note: A conflict of interest was identified after the acceptance of this paper. The editor-in-chief provided a final review.

Received 6 February 2025
Accepted 4 May 2025
Published 27 June 2025

TABLE 1 Condensed table of isolate characteristics[a]

Sample origin	Sample description	SRA accession(s)	NCBI BioSample accession for co-assembly
GW271 groundwater well	0.1 µm filter, filtered onsite at ORNL	SRS18494355, SRS18494357, SRS18494359, SRS18494358	SAMN36877165
GW271 groundwater well	0.2 µm filter, filtered at LBNL	SRS18494376, SRS18494377, SRS18494378, SRS18494379	SAMN36877163
GW271 groundwater well	0.2 µm filter, filtered onsite at ORNL	SRS18494380, SRS18494381, SRS18494382, SRS18494384, SRS18494385	SAMN36877164
FW106 groundwater well	0.1 µm filter, filtered onsite at ORNL	SRS18494386, SRS18494387, SRS18494388	SAMN36877159
FW106 groundwater well	0.2 µm filter, filtered at LBNL	SRS18494390, SRS18494389, SRS18494391, SRS18494409	SAMN36877160
FW106 groundwater well	0.2 µm filter, filtered onsite at ORNL	SRS18494408, SRS18494299, SRS18494302, SRS18494301, SRS18494300, SRS18494303	SAMN36877158
FW115-24 groundwater well	0.1 µm filter, filtered onsite at ORNL	SRS18494305	Same as SRA accession
FW115-24 groundwater well	0.2 µm filter, filtered at LBNL	SRS18494304, SRS18494306, SRS18494308	SAMN36877162
FW115-24 groundwater well	0.2 µm filter, filtered onsite at ORNL	SRS18494307, SRS18494310, SRS18494311	SAMN36877161
Groundwater extraction negative control	Control	SRS18494312	No co-assembly attempted
Groundwater reagent water extraction negative control	Control	SRS18494313	No co-assembly attempted
EB271-02-01 sediment core	Vadose zone, 91.44–128.85 cm bgs	SRS18494297, SRS18494296, SRS18494322, SRS18494333	SAMN36877171
EB271-02-02 sediment core	Vadose zone, 128.85–166.25 cm bgs	SRS18494360, SRS18494371, SRS18494397, SRS18494345	SAMN36877172
EB271-02-03 sediment core	Vadose zone, 166.25–182.88 cm bgs	SRS18494356, SRS18494383, SRS18494298, SRS18494309	SAMN36877173
EB271-03-01 sediment core	Vadose zone, 182.88–211 cm bgs	SRS18494314, SRS18494315, SRS18494316, SRS18494317	SAMN36877174
EB271-03-02 sediment core	Vadose zone, 211.26–239.64 cm bgs	SRS18494318, SRS18494319, SRS18494320, SRS18494321	SAMN36877175
EB271-03-03 sediment core	Vadose zone, 239.64–274.32 cm bgs	SRS18494323, SRS18494324, SRS18494325, SRS18494326	SAMN36877176
EB271-04-01 sediment core	Variably saturated zone, 274.32–297.83 cm bgs	SRS18494327, SRS18494328, SRS18494329, SRS18494330	SAMN36877177
EB271-04-02 sediment core	Variably saturated zone, 297.83–321.35 cm bgs	SRS18494332, SRS18494331, SRS18494334, SRS18494335	SAMN36877178
EB271-04-03 sediment core	Variably saturated zone, 321.35–344.86 cm bgs	SRS18494336, SRS18494337, SRS18494339, SRS18494338	SAMN36877179
EB271-04-04 sediment core	Variably saturated zone, 344.86–365.76 cm bgs	SRS18494340, SRS18494341, SRS18494342, SRS18494343	SAMN36877180
EB271-05-01 sediment core	Saturated zone, 365.76–383.27 cm bgs	SRS18494361, SRS18494364	SAMN36877181
EB271-05-02 sediment core	Saturated zone, 383.27–400.78 cm bgs	SRS18494363, SRS18494362	SAMN36877182
EB271-05-03 sediment core	Saturated zone, 400.78–418.29 cm bgs	SRS18494365, SRS18494366, SRS18494368, SRS18494367	SAMN36877183
EB271-05-04 sediment core	Saturated zone, 418.29–438.72 cm bgs	SRS18494370	Same as SRA accession
EB106-02-01 sediment core	Vadose zone, 91.44–127.22 cm bgs	SRS18494369, SRS18494372, SRS18494373, SRS18494374	SAMN36877166
EB106-02-02 sediment core	Vadose zone, 127.22–163.00 cm bgs	SRS18494375, SRS18494393, SRS18494392, SRS18494394, SRS18494395	SAMN36877167
EB106-02-03 sediment core	Vadose zone, 163.00–182.88 cm bgs	SRS18494396, SRS18494398, SRS18494399, SRS18494400	SAMN36877168
EB106-03-01 sediment core	Vadose zone, 182.88–197.58 cm bgs	SRS18494401, SRS18494402, SRS18494403, SRS18494404	SAMN36877169

(*Continued on next page*)

TABLE 1 Condensed table of isolate characteristics[a] (*Continued*)

Sample origin	Sample description	SRA accession(s)	NCBI BioSample accession for co-assembly
EB106-03-02 sediment core	Vadose zone, 197.58–212.27 cm bgs	SRS18494405, SRS18494406, SRS18494407, SRS18494344	SAMN36877170
EB106-03-03 sediment core	Vadose zone, 212.27–226.97 cm bgs	SRS18494346	Same as SRA accession
EB106-03-04 sediment core	Vadose zone, 226.97–241.66 cm bgs	SRS18494348	Same as SRA accession
EB106-03-05 sediment core	Vadose zone, 241.66–256.36 cm bgs	SRS18494347	Same as SRA accession
EB106-04-01 sediment core	Vadose zone, 274.32–291.47 cm bgs	SRS18494349	Same as SRA accession
EB106-04-04 sediment core	Variably saturated zone, 325.76–342.90 cm bgs	SRS18494351	Same as SRA accession
EB106-05-04 sediment core	Saturated zone, 410.65–425.61 cm bgs	SRS18494350	Same as SRA accession
EB106-05-06 sediment core	Saturated zone 440.57–457.20 cm bgs	SRS18494352	Same as SRA accession
Sediment extraction negative control	Control	SRS18494353	No co-assembly attempted
Sediment extraction negative control	Control	SRS18494354	No co-assembly attempted

[a]See Figshare (https://doi.org/10.6084/m9.figshare.29053628.v1) for a full table that includes SRA accessions, sampling date, more sample details, and more assembly statistics. µm, micron; cm, centimeter; bgs, below ground surface; ORNL, Oak Ridge National Laboratory; LBNL, Lawrence Berkeley National Laboratory.

Metagenomics sequencing libraries were prepared using the Illumina Nextera Flex Kit (now called Illumina DNA Prep Kit), according to the manufacturer's instructions. Library concentration was measured using a Qubit 3 Fluorometer. Samples were sequenced by QB3 Genomics, UC Berkeley, Berkeley, CA (RRID:SCR_022170) using 2 × 150 bp on an Illumina HiSeq4000.

Illumina reads were filtered for phiX174 and trimmed to remove any residual adapter sequences using BBTools Version 38.86 (8) using parameters "bf1 ktrim=r k=23 mink=11" along with a database of Illumina adapters and phiX174 spike-in. The resulting read set was assembled using SPAdes Version 3.15.4 (9, 10) with parameters "—memory 1024 —only-assembler —meta -k 21,33,55,77,99,127." The SPAdes built-in error correction was disabled as it isn't generally necessary and requires more computer memory than we had available. Metagenomes from the same sample were co-assembled, resulting in 26 sediment metagenomes and nine groundwater metagenomes. Co-assemblies improved assembly metrics, generally increasing N50 to 1.7-2X the individual assembly N50s.

ACKNOWLEDGMENTS

We thank the Hazen, Zhou, and Banfield labs for discussions about DNA extraction from sediment. We also thank Dominique Joyner for assistance with the transport of samples from the Oak Ridge National Laboratory.

This research by ENIGMA—Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Science Focus Area Program at the Lawrence Berkeley National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract number DE-AC02-05CH11231.

AUTHOR AFFILIATIONS

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA
[2]Center for Biofilm Engineering, Montana State University, Bozeman, Montana, USA
[3]Department of Bioengineering, University of California, Berkeley, California, USA
[4]Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
[5]University of Tennessee, Knoxville, Tennessee, USA
[6]Department of Microbiology and Immunology, Montana State University, Bozeman, Montana, USA

AUTHOR ORCIDs

Lauren M. Lui http://orcid.org/0000-0001-8720-5268
Torben N. Nielsen http://orcid.org/0000-0002-0987-7189
Heidi J. Smith http://orcid.org/0000-0001-7234-9146
John-Marc Chandonia http://orcid.org/0000-0002-5153-9079
Jennifer V. Kuehl http://orcid.org/0000-0003-2813-2518
Fangchao Song http://orcid.org/0000-0002-9427-9879
Andrew Sczesnak http://orcid.org/0000-0002-0152-9745
Terry C. Hazen http://orcid.org/0000-0002-2536-9993
Matthew W. Fields http://orcid.org/0000-0001-9053-1849
Adam P. Arkin http://orcid.org/0000-0002-4999-2931

FUNDING

Funder	Grant(s)	Author(s)
U.S. Department of Energy	DE-AC02-05CH11231	Lauren M. Lui
		Torben N. Nielsen
		Heidi J. Smith
		John-Marc Chandonia
		Jennifer V. Kuehl
		Andrew Sczesnak
		Andrew Hendrickson
		Terry C. Hazen
		Matthew W. Fields
		Adam P. Arkin

AUTHOR CONTRIBUTIONS

Lauren M. Lui, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Writing – original draft, Writing – review and editing | Torben N. Nielsen, Data curation, Formal analysis, Investigation, Methodology, Software | Heidi J. Smith, Conceptualization, Data curation, Investigation, Writing – original draft, Writing – review and editing | John-Marc Chandonia, Data curation, Funding acquisition, Writing – review and editing | Jennifer V. Kuehl, Investigation, Resources | Andrew Sczesnak, Investigation | Andrew Hendrickson, Investigation | Terry C. Hazen, Funding acquisition, Supervision, Writing – review and editing | Matthew W. Fields, Funding acquisition, Supervision, Writing – review and editing | Adam P. Arkin, Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing.

DATA AVAILABILITY

The sequences and assembled metagenomes were deposited at the National Center for Biotechnology Information (NCBI) database under the accession numbers listed in Table 1 in BioProject PRJNA1001011.

REFERENCES

1. Smith HJ, Zelaya AJ, De León KB, Chakraborty R, Elias DA, Hazen TC, Arkin AP, Cunningham AB, Fields MW. 2018. Impact of hydrologic boundaries on microbial planktonic and biofilm communities in shallow terrestrial subsurface environments. FEMS Microbiol Ecol 94:fiy191. https://doi.org/10.1093/femsec/fiy191
2. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun 7:13219. https://doi.org/10.1038/ncomms13219
3. Lui LM, Majumder EL-W, Smith HJ, Carlson HK, von Netzer F, Fields MW, Stahl DA, Zhou J, Hazen TC, Baliga NS, Adams PD, Arkin AP. 2021. Mechanism across scales: a holistic modeling framework integrating laboratory and field studies for microbial ecology. Front Microbiol 12:642422. https://doi.org/10.3389/fmicb.2021.642422
4. Smith MB, Rocha AM, Smillie CS, Olesen SW, Paradis C, Wu L, Campbell JH, Fortney JL, Mehlhorn TL, Lowe KA, et al. 2015. Natural bacterial communities serve as quantitative geochemical biosensors. MBio 6:e00326-15. https://doi.org/10.1128/mBio.00326-15

5. Green SJ, Prakash O, Jasrotia P, Overholt WA, Cardenas E, Hubbard D, Tiedje JM, Watson DB, Schadt CW, Brooks SC, Kostka JE. 2012. Denitrifying bacteria from the genus *Rhodanobacter* dominate bacterial communities in the highly contaminated subsurface of a nuclear legacy waste site. Appl Environ Microbiol 78:1039–1047. https://doi.org/10.1128/AEM.06435-11
6. Moon J-W, Paradis CJ, Joyner DC, von Netzer F, Majumder EL, Dixon ER, Podar M, Ge X, Walian PJ, Smith HJ, et al. 2020. Characterization of subsurface media from locations up- and down-gradient of a uranium-contaminated aquifer. Chemosphere 255:126951. https://doi.org/10.1016/j.chemosphere.2020.126951
7. Wu X, Gushgari-Doyle S, Lui LM, Hendrickson AJ, Liu Y, Jagadamma S, Nielsen TN, Justice NB, Simmons T, Hess NJ, Joyner DC, Hazen TC, Arkin AP, Chakraborty R. 2023. Distinct depth-discrete profiles of microbial communities and geochemical insights in the subsurface critical zone. Appl Environ Microbiol 89:e0050023. https://doi.org/10.1128/aem.00500-23
8. BBTools. 2016. DOE Joint Genome Institute. Available from: https://jgi.doe.gov/data-and-tools/software-tools/bbtools. Retrieved 25 Jul 2023.
9. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116
10. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes *de novo* assembler. Curr Protoc Bioinformatics 70:e102. https://doi.org/10.1002/cpbi.102