AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

OMICS DATA SETS

Metagenomes and Metagenome-Assembled Genomes from Substrate-Amended Hot Spring Sediment Incubations from Yellowstone National Park

Nicholas J. Reichart,[a] Robert M. Bowers,[b] Tanja Woyke,[b] Roland Hatzenpichler[a,c,d]

[a]Department of Chemistry and Biochemistry, Montana State University, Bozeman, Montana, USA
[b]Department of Energy, Joint Genome Institute, Berkeley, California, USA
[c]Thermal Biology Institute, Montana State University, Bozeman, Montana, USA
[d]Center for Biofilm Engineering, Montana State University, Bozeman, Montana, USA

ABSTRACT Here, we report on eight sediment metagenomes obtained from an alkaline hot spring, with their corresponding metagenome-assembled genomes. Samples had been incubated for 48 h with various substrate amendments in conjunction with the amino acid analog L-homopropargylglycine in a study targeted at identifying anabolicly active uncultured thermophilic archaea and bacteria.

Hot spring ecosystems harbor diverse microbial life, much of which currently lacks cultured representatives (1, 2). The metabolic connectivity of hot spring microbiomes makes it difficult to determine conditions conducive to cultivation. However, incubations of hot spring material with substrate analogs, in conjunction with click-chemistry-mediated fluorescence labeling to identify translationally active members within a microbial community, represent a culture-independent alternative for obtaining information on environmental microbes (3–5). Here, we subjected hot spring material to such incubations, followed by metagenomic sequencing of cells sorted via fluorescence-activated cell sorting (FACS), to accompany previous 16S rRNA gene amplicon data for translationally active cells from this hot spring (6). The data set generated here can serve as a publicly available resource to assess the metabolic potential of the translationally active microbiome from substrate-amended hot spring samples.

A sediment-water microbiome from a hot spring within the Five Sisters hot spring group (74°C, pH 8.4) in Yellowstone National Park (YNP) (44.532619, −110.797272) was incubated with the amino acid analog L-homopropargylglycine (HPG) (50 μM) at 74°C for 48 h in the dark under different conditions (amendment with cellobiose or glycine under oxic conditions or without substrate amendment under anoxic headspace). Seven bulk samples were selected for sequencing based on significant shifts in their active microbial populations, as determined via FACS (6). In addition, one unamended, unincubated sample was sequenced (Table 1). Genomic DNA was extracted using the MP Biomedical FastDNA SPIN kit for soil according to the manufacturer's directions. Metagenomes were sequenced from either 10 ng (taxon 3300038414) or 1 ng (taxons 3300040930, 3300040931, 3300040932, 3300040933, 3300040934, 3300040935, and 3300041472) of starting genomic DNA. Libraries for the 10-ng sample were prepared using the KAPA HyperPrep kit (Kapa Biosystems) with 5 cycles of PCR. Libraries for the 1-ng samples were prepared with Nextera XT tagmentation (Illumina) following 12 cycles of PCR. All samples were multiplexed and sequenced on the Illumina NovaSeq 6000 system using NovaSeq XP v1 reagents with a S4 flow cell (2 × 150-bp indexed run).

The average base-calling score across all metagenomes was 36, with an average percentage above Q30 of 93%. Individual quality scores for the metagenomes are listed in Table 1 at https://doi.org/10.6084/m9.figshare.16905058. Reads were assembled using

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

Address correspondence to Roland Hatzenpichler, roland.hatzenpichler@montana.edu.

The authors declare no conflict of interest.

Received 16 November 2021
Accepted 27 February 2022
Published 30 March 2022

TABLE 1 Sample information

Taxon object identification no.	Sample name	Substrate	Headspace	HPG	Incubation time (h)	No. of MAGs
3300038414	FS5_t0	None	Atmospheric	No		15
3300040930	FS5_Glycine_1	Glycine	Atmospheric	Yes	48	3
3300040931	FS5_Glycine_2	Glycine	Atmospheric	Yes	48	11
3300040932	FS5_Glycine_3	Glycine	Atmospheric	Yes	48	3
3300040933	FS5_Cellobiose_2	Cellobiose	Atmospheric	Yes	48	8
3300040934	FS_N2_1	None	100% N_2	Yes	48	12
3300040935	FS_N2_2	None	100% N_2	Yes	48	10
3300041472	FS5_HPG_2	None	Atmospheric	Yes	48	12

metaSPAdes v3.13.0 or v3.14.1, and genes were annotated using the Joint Genome Institute (JGI) Integrated Microbial Genomes and Microbiomes (IMG/M) database (7) pipeline v5.0.14 or v5.0.19. For all metagenomes, an average of 99% of reads were mapped to contigs. Metagenomes were then binned according to IMG/M methods,

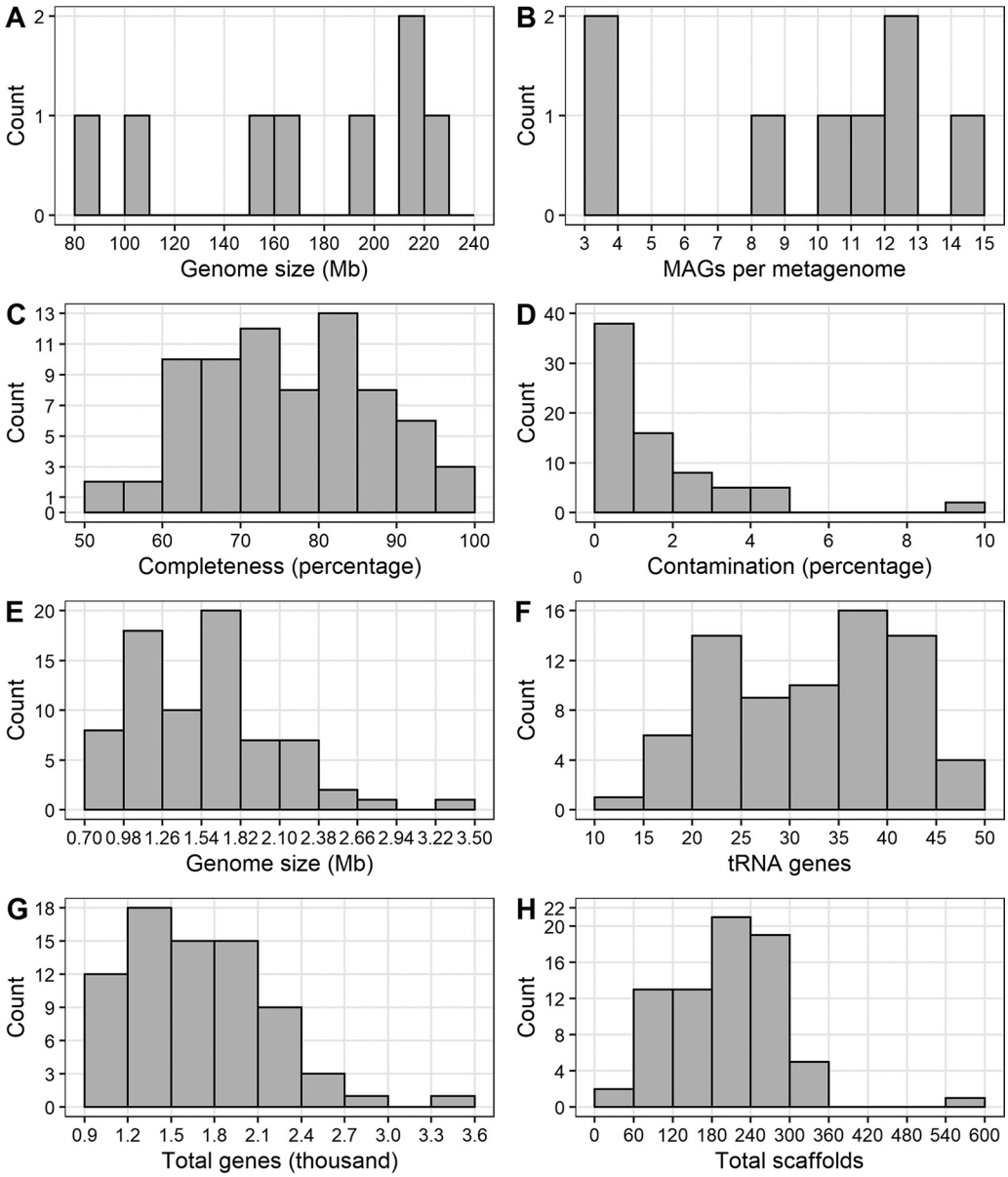

FIG 1 Sequencing statistics for assembled metagenomes (A and B) and medium- and high-quality MAGs (C to H) generated by the IMG/M pipeline.

which used MetaBAT v2.12.1 or v2.2.15 and metagenome-assembled genome (MAG) statistics generated with CheckM v1.0.12 or v1.1.3. Seven high-quality MAGs and 67 medium-quality MAGs, classified according to MIMAG standards (8), were represented in the data set (Fig. 1). Finally, MAGs were phylogenetically classified with GTDB-tk v0.2.2 or v1.3.0, which used GTDB database release 86 or 95, respectively. Metagenome-processing versions and metagenome statistics are available in Table 1 and all IMG/M bin numbers, statistics, and processing methods for MAGs can be found in Table 2 at https://doi.org/10.6084/m9.figshare.16905058. The MAGs represented 12 phyla and 18 classes. The metagenomic data sets and MAGs provide a unique resource to be analyzed with previously generated activity data in efforts to describe microbial function from hot spring communities more thoroughly.

Data availability. The assembled metagenomes were deposited in the IMG/M database according to the taxon object identification numbers listed in Table 1. NCBI SRA accession numbers for the metagenomic samples are listed in Table 1 and details of the MAGs, with corresponding IMG/M bin identification numbers, are in Table 2 at https://doi.org/10.6084/m9.figshare.16905058.

ACKNOWLEDGMENTS

This work was supported through NSF RII Track-2 FEC award DBI-1736255 and NASA Exobiology award 80NSSC19K0449 (to R.H.). A portion of this research was performed under the Facilities Integrating Collaborations for User Science (FICUS) initiative (proposal 10.46936/fics.proj.2017.49972/6000002 to R.H.) and used resources at the DOE JGI, which is a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research and operated under contract DE-AC02-05CH11231. This material is based on work supported by the U.S. DOE, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education (ORISE) for the DOE. ORISE is managed by Oak Ridge Associated Universities (ORAU) under contract DE-SC0014664.

We appreciate the U.S. National Park Service, particularly Annie Carlson at the Yellowstone Center for Resources, for permitting work in YNP under permit YELL-SCI-8010.

All opinions expressed in this paper are the authors' and do not necessarily reflect the policies and views of the DOE, ORAU, or ORISE.

REFERENCES

1. Eloe-Fadrosh EA, Paez-Espino D, Jarett J, Dunfield PF, Hedlund BP, Dekas AE, Grasby SE, Brady AL, Dong H, Briggs BR, Li WJ, Goudeau D, Malmstrom R, Pati A, Pett-Ridge J, Rubin EM, Woyke T, Kyrpides NC, Ivanova NN. 2016. Global metagenomic survey reveals a new bacterial candidate phylum in geothermal springs. Nat Commun 7:10476. https://doi.org/10.1038/ncomms10476.
2. Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L. 2018. Phylogenetically novel uncultured microbial cells dominate earth microbiomes. mSystems 3:e00055-18. https://doi.org/10.1128/mSystems.00055-18.
3. Hatzenpichler R, Scheller S, Tavormina PL, Babin BM, Tirrell DA, Orphan VJ. 2014. In situ visualization of newly synthesized proteins in environmental microbes using amino acid tagging and click chemistry. Environ Microbiol 16:2568–2590. https://doi.org/10.1111/1462-2920.12436.
4. Hatzenpichler R, Connon SA, Goudeau D, Malmstrom RR, Woyke T, Orphan VJ. 2016. Visualizing in situ translational activity for identifying and sorting slow-growing archaeal-bacterial consortia. Proc Natl Acad Sci U S A 113:E4069–E4078.
5. Hatzenpichler R, Krukenberg V, Spietz RL, Jay ZJ. 2020. Next-generation physiology approaches to study microbiome function at single cell level. Nat Rev Microbiol 18:241–256. https://doi.org/10.1038/s41579-020-0323-1.
6. Reichart NJ, Jay ZJ, Krukenberg V, Parker AE, Spietz RL, Hatzenpichler R. 2020. Activity-based cell sorting reveals responses of uncultured archaea and bacteria to substrate amendment. ISME J 14:2851–2861. https://doi.org/10.1038/s41396-020-00749-1.
7. Chen I-MA, Chu K, Palaniappan K, Ratner A, Huang J, Huntemann M, Hajek P, Ritter S, Varghese N, Seshadri R, Roux S, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpides NC. 2021. The IMG/M data management and analysis system v.6.0: new tools and advanced capabilities. Nucleic Acids Res 49:D751–D763. https://doi.org/10.1093/nar/gkaa939.
8. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Genome Standards Consortium, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, Schriml L, Banfield JF, Hugenholtz P, Woyke T. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 35:725–731. https://doi.org/10.1038/nbt.3893.