Draft Genome Sequence from a Putative New Genus and Species in the Family Methanoregulaceae Isolated from the Anoxic Basin of Lake Untersee in East Antarctica

Nicole Y. Wagner,a Aria S. Hahn,b,c Dale Andersen,d Mary Beth Wilhelm,e Connor Morgan-Lang,h,c Mia Vanderwilt,f Sarah Stewart Johnsona,f

aDepartment of Biology, Georgetown University, Washington, DC, USA
bDepartment of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada
cKoonkie Cloud Services, Inc., Menlo Park, California, USA
dCarl Sagan Center, SETI Institute, Mountain View, California, USA
eSpace Science and Astrobiology Division, NASA Ames Research Center, Moffett Field, California, USA
fScience, Technology, and International Affairs Program, Georgetown University, Washington, DC, USA

ABSTRACT Here, we report the draft genome sequence for a new putative genus and species in the Methanoregulaceae family, whose members are generally slow-growing rod-shaped or coccoid methanogenic archaea. The information on this sediment-dwelling organism sheds light on the prokaryotes inhabiting isolated, deep, and extremely cold methane-rich environments.

With methane concentrations reaching as high as 21.8 ± 1.4 mmol liter⁻¹, the anoxic basin of the ice-covered Antarctic Lake Untersee is one of the most methane-rich naturally occurring aquatic ecosystems on Earth (1). An Untersee environmental sample was collected in 2016 (coordinates, 71°21.12’S, 13°26.08’E) from benthic sediments below 100 m of 4°C water using an Ekman dredge. After collection, the samples were frozen and then shipped on dry ice and stored at −80°C.

The sequence was isolated from metagenomic data. DNA was extracted using a Qiagen AllPrep kit. A paired-end library was constructed using a Kapa HyperPlus kit and sequenced on an Illumina MiSeq platform on a V3/600-cycle flow cell, generating 24,799,516 paired-end 300-bp reads for the entire metagenome. The reads were trimmed to remove adapters and bases below a quality score of 30 using Trimmomatic 0.36 (2), assembled using MEGAHIT version 1.1.2 (3), and binned using MetaBAT 0.32.5 (4). Reads were mapped back to the bin using the default parameters in Bowtie 1.2.2 (5) and SAMtools 1.3.1 (6) and reassembled using ABySS version 2.1.5 (7) to improve the metagenome-assembled genome (MAG) quality. This MAG contains 45 contigs and 1,202,444 reads (4.9% of the total metagenomic reads).

The N₅₀ value for this genome is 69,553 bp. CheckM version 1.0.11 (8) calculated completion to be 95.07% and contamination to be 0.0%. The draft genome is 2.367 Mb, with 2,318 open reading frames (ORFs), 1 small subunit (SSU) 16S rRNA gene, and 60 tRNA genes (Fig. 1). The genome was annotated using MetaPathways 2.5.1 (9) and KAAS-KEGG 2.1 (10). Default settings were used for all programs.

The 16S rRNA of this organism is 95% similar to that of Methanosphaerula palustris (11), as calculated using online BLAST-based (12) homology. This genetic distance suggests that the MAG may be classified as a previously unobserved genus in the Methanoregulaceae family (13).

Of the predicted ORFs in the MAG, 17.24% code for translation, ribosomal structure, and biogenesis, 10.49% belong to metabolism, and 10.53% code for amino acid metabolism genes. This MAG has a GC content of 66.4%. An average coverage of...
was calculated using Bowtie 1.2.2 (5) and SAMtools 1.3.1 (6), with default parameters.

The sequence describes a methanogen with the complete mcrABG operon identified on a single contig. It contains the oppA, mppA, and qseF genes for quorum sensing and cheA, cheW, cheR, and cheYB genes associated with chemotaxis. It has ORFs for amino acid transport and synthesis, a potential response to starvation, and the regulation of nitrate uptake in low-nitrogen environments (glnL and glnA genes) (14).

The sequence also includes the ruvABC resolvasome genes, which are absent in Methanosphaerula palustris. These ruvABC genes catalyze the resolution of the Holliday junction during recombination and DNA repair (15); their presence may lead to recombination rates higher than those for Methanosphaerula palustris, as higher re-

FIG 1 Shown from the outer to inner ring of this metagenome-assembled genome (MAG) circos plot is the following: first, the proportion of the total predicted open reading frames (ORFs) (any sequence with a start and stop codon, excluding rRNA genes) with a corresponding annotation in the COG database (17); then, tRNA locations, contigs numbered based on their order in the MAG, and all predicted ORFs. Finally, in the innermost ring, the long subunit (LSU)/SSU/5S location is shown.
combination rates have been observed to lead to higher GC content in prokaryotes (16).

**Data availability.** This whole-genome shotgun project was deposited at DDBJ/ENA/GenBank under accession number SISS00000000, and the reads were deposited under accession number PRJNA521775. This study is version SISS01000000.

**REFERENCES**

1. Wand U, Samarkin VA, Nitzsche H-M, Hubberten H-W. 2006. Biogeochemistry of methane in the permanently ice-covered Lake Untersee, central Dronning Maud Land, East Antarctica. Limnol Oceanogr 51:1180–1194. https://doi.org/10.4319/lo.2006.51.2.1180.
2. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
3. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033.
4. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3:e1165. https://doi.org/10.7717/peerj.1165.
5. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
6. Li H, Handsaker B, Wyssker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map (SAM) format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
7. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. Genome Res 19:1117–1123. https://doi.org/10.1101/gr.089532.108.
8. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
9. Konwar KM, Hanson NW, Pagé AP, Hallam SJ. 2013. MetaPathways: a modular pipeline for constructing pathway/genome databases from environmental sequence information. BMC Bioinformatics 14:202. https://doi.org/10.1186/1471-2105-14-202.
10. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:W182–W185. https://doi.org/10.1093/nar/gkm321.
11. Cadillo-Quiroz H, Browne P, Kyrpides N, Woyke T, Goodwin L, Detter C, Yavitt JB, Zinder SH. 2015. Complete genome sequence of Methanosphaerula palustris E1-9C, a hydrogenotrophic methanogen isolated from a minerotrophic fen peatland. Genome Announc 3:e01280-15. https://doi.org/10.1128/genomeA.01280-15.
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
13. Claridge JE, III. 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin Microbiol Rev 17:840–862. https://doi.org/10.1128/CMR.17.4.840-862.2004.
14. Miller AJ, Fan X, Shen Q, Smith SJ. 2008. Amino acids and nitrate as signals for the regulation of nitrogen acquisition. J Exp Bot 59:111–119. https://doi.org/10.1093/jxb/erm208.
15. Dickman MJ, Ingleston SM, Sedelnikova SE, Rafferty JB, Lloyd RG, Grasby JA, Hornby DP. 2002. The RuvABC resolvasome: quantitative analysis of RuvA and RuvC assembly on junction DNA. Eur J Biochem 269:5482–5501. https://doi.org/10.1046/j.1432-1033.2002.03250.x.
16. Reichenberger ER, Rosen G, Hershberg U, Hershberg R. 2015. Prokaryotic nucleotide composition is shaped by both phylogeny and the environment. Genome Biol Evol 7:1380–1389. https://doi.org/10.1093/gbe/evv063.
17. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36. https://doi.org/10.1093/nar/28.1.33.