AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Metagenome-Assembled Genome of USCα AHI, a Potential High-Affinity Methanotroph from Axel Heiberg Island, Canadian High Arctic

Calvin Rusley,[a] Tullis C. Onstott,[a] Tatiana A. Vishnivetskaya,[b] Alice Layton,[b] Archana Chauhan,[b]* Susan M. Pfiffner,[b] Lyle G. Whyte,[c] Maggie C. Y. Lau[a]*

[a]Department of Geosciences, Princeton University, Princeton, New Jersey, USA
[b]Center for Environmental Biotechnology, University of Tennessee, Knoxville, Tennessee, USA
[c]Department of Natural Resource Sciences, McGill University, Montreal, Québec, Canada

ABSTRACT Metagenomic sequencing of active-layer cryosols from the Canadian High Arctic has yielded a nearly complete genome for an atmospheric CH_4-oxidizing bacterium belonging to upland soil cluster α (USCα). This genome contains genes involved in CH_4 metabolism, H_2 metabolism, and multiple carbon assimilation pathways.

Recent studies have shown that mineral cryosols from the Canadian High Arctic Axel Heiberg Island (AHI) act as CH_4 sinks during the summer (1), drawing CH_4 from both the atmosphere and underlying hypoxic cryosols (2, 3), and harbor metabolically active upland soil cluster α (USCα) proteobacteria (1). Twenty-one metagenomic data sets of active-layer cryosols (4) from long-term core incubation experiments were used to construct the draft genome of this USCα. Sequencing and sample collection methods were published by Chauhan et al. (4).

Raw reads were filtered using the Princeton University Galaxy server using "filter by quality" to keep reads having 90% of the bases with a Phred score of >30. Nextera transposase adaptor sequences and the last five bases at the 3′ end were removed using Trim Galore. IDBA-UD v1.1.1 (with the settings mink = 20, maxk = 100, and step = 20) was used to create 21 individual assemblies and 1 coassembly from reads longer than 50 nucleotides (nt) (5). Bins were created using MetaBAT v0.32.4 (6) (–very sensitive option), evaluated using CheckM v1.0.6 (7), and annotated using PROKKA v1.12-beta (8) and BLAST v2.2.29+ (9). Default parameters were used for all software unless otherwise specified. The coassembly yielded a 90.56% complete genome with 0.31% contamination, containing a USCα-like particulate methane monooxygenase β-subunit (*pmoA*) gene. CheckM assigned this genome as an unknown species within the *Beijerinckiaceae*.

As CheckM analysis indicated that 4 of the 21 individual assemblies had unknown *Beijerinckiaceae* bins (6.43 to 36.49% complete), we extracted *Beijerinckiaceae* reads from these 4 metagenomes (SRA accession numbers SRR1586250, SRR1586265, SRR1586287, and SRR1586310). We then mapped the quality-filtered reads onto the USCα bin and four *Beijerinckiaceae* genomes having different phylogenetic distances from USCα (10), namely, *Methylocapsa acidiphila* B2 (NZ_ATYA01000001), *Methylocella silvestris* BL2 (NC_011666), *Methylocystis* sp. strain SC2 (NC_018485), and *Methylosinus trichosporium* OB3b (NZ_ADVE02000003), using Bowtie2 v2.3.2 (11). All mapped reads were pooled and reassembled using SPAdes v3.10.1 (12). Binning using MetaBAT v0.32.4 (–very sensitive option) yielded a single bin. Evaluated by CheckM v1.0.6, this final genome had slightly improved completeness and less contamination (Table 1). This genome was annotated using PROKKA v1.12-beta (8), BLAST v2.2.29+ (9) against

Citation Rusley C, Onstott TC, Vishnivetskaya TA, Layton A, Chauhan A, Pfiffner SM, Whyte LG, Lau MCY. 2019. Metagenome-assembled genome of USCα AHI, a potential high-affinity methanotroph from Axel Heiberg Island, Canadian High Arctic. Microbiol Resour Announc 8:e01178-19. https://doi.org/10.1128/MRA.01178-19.

Editor J. Cameron Thrash, University of Southern California

Address correspondence to Maggie C. Y. Lau, maglau@idsse.ac.cn.

* Present address: Archana Chauhan, Department of Zoology, Panjab University, Chandigarh, India; Maggie C. Y. Lau, Laboratory of Extraterrestrial Ocean Systems, Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences, Sanya, Hainan, China.

Received 20 September 2019
Accepted 11 October 2019
Published 14 November 2019

TABLE 1 Statistics summary of the coassembled and reassembled USCα genomes[a]

CheckM output	Beijerinckiaceae bin from coassembly	USCα AHI genome from reassembly
Marker lineage	o__Rhizobiales (UID3654)	o__Rhizobiales (UID3654)
No. of genomes	92	92
No. of markers	481	481
No. of marker sets	319	319
0 copies (missing)	**36**	**32**
1 copy	**444**	**449**
2 copies	**1**	**0**
3 copies	0	0
4 copies	0	0
≥5 copies	0	0
Completeness (%)	**90.56**	**91.64**
Contamination (%)	**0.31**	**0.00**
Strain heterogeneity (%)	0.00	0.00
No. of unique markers (of 43)	42	42
No. of multicopy markers	0	0
Insertion branch UID	UID3666	UID3666
Taxonomy (contained)	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria; o__Rhizobiales;f__Beijerinckiaceae	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria; o__Rhizobiales;f__Beijerinckiaceae
Taxonomy (sister)	Unresolved	Unresolved
GC content (%)	**59.1**	**59**
Genome size (Mbp)	**3.03**	**3.26**
Gene count	**3,388**	**3,928**
Coding density (fraction)	**0.82**	**0.81**
Translation table	11	11
No. of descendant genomes	3	3
Lineage		
GC content (%)		
Mean	60.6	60.6
SD	2.6	2.6
Genome size (Mbp)		
Mean	4.28	4.28
SD	0.13	0.13
Gene count		
Mean	3,861	3,861
SD	86	86

[a] Values that are different between the two draft genomes are marked in bold font.

the SILVA SSU v128 and NCBI databases, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) automatic annotation server v2.1 (13). A phylogenetic tree using single-copy genes (14) was created using Anvi'o v5.2 (15) phylogenomic analysis for *Beijerinckiaceae* genomes selected by referencing Tveit et al. (10). Average nucleotide identity (ANI) and average amino acid identity (AAI) values were calculated using the scripts ani.rb (with the options –win, 1,000; –step, 200; –len, 700; –id, 70) and aai.rb (with the options –len-fraction, 0.8; –id, 20), respectively, from the enveomics package v1.4.4 (16).

The USCα AHI genome belongs within the *Beijerinckiaceae* (Fig. 1) and possesses a 416-nt-long 16S rRNA gene that is 98.1 to 98.6% similar to published USCα 16S rRNA genes (10, 17). Its *pmoA* and *pmoB* genes match 99.7 to 100% with DNA and RNA sequences previously reported from AHI that were phylogenetically determined as the high-affinity form for CH_4 oxidation (1). USCα AHI is able to assimilate C from CH_4 and from CO_2 via the serine cycle, the reductive glycine pathway, and the Calvin-Benson-Bassham cycle. USCα AHI can utilize various carbon sources via the pentose phosphate and Entner-Doudoroff pathways, including acetate in its tricarboxylic acid (TCA) cycle, although the acetate transporter gene (*actP*) is absent. The [NiFe] group 1h hydrogenase for H_2 metabolism is also present.

Data availability. The draft genome sequence of USCα AHI has been deposited at NCBI GenBank under the accession number VDMG00000000 (BioSample number SAMN11877018 and BioProject number PRJNA545288). The version described in this

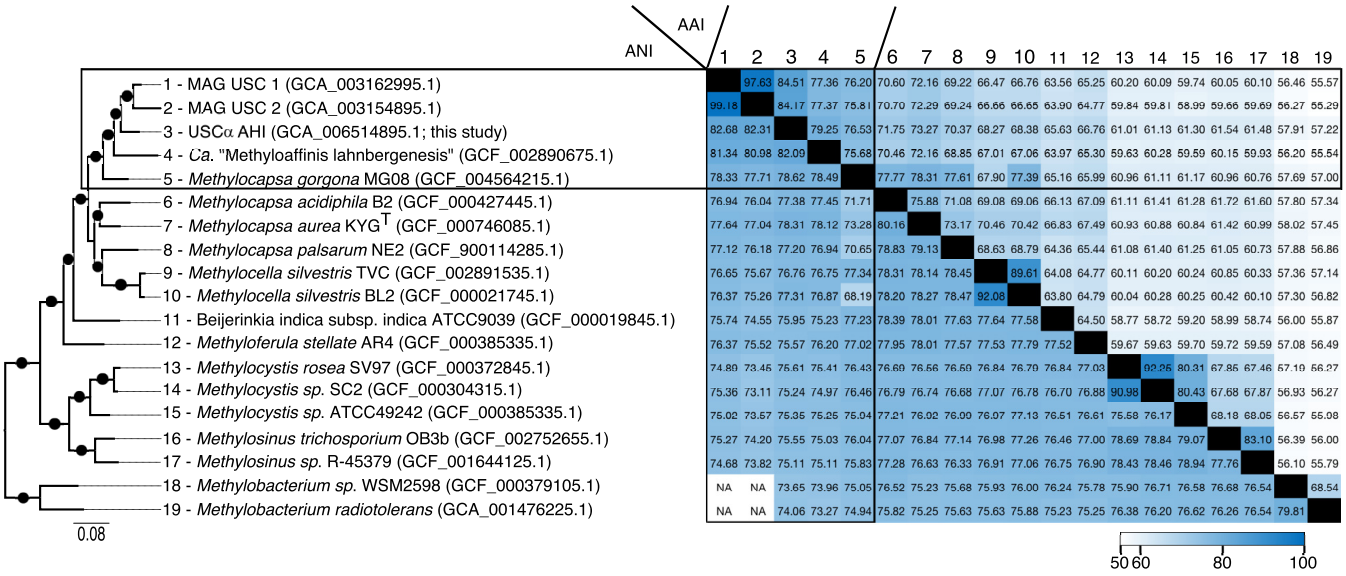

FIG 1 Genomic comparison between USCα AHI and genomes of methanotrophs within the *Beijerinckiaceae*. (Left) Phylogenomic tree constructed from 86 concatenated single-copy genes. The scale bar indicates the probability of substitution in amino acid residues. Filled circles indicate local support of 0.99 calculated using CAT approximation in FastTree v2.1.10 (included in Anvi'o v5.2). (Right) Matrix of pairwise ANI and AAI values ordered as indicated for the left panel. Black rectangles mark ANI and AAI values of USCα genomes. Color intensity indicates values between 55 and 100. NA, not available because fewer than 100 fragments (700 nt) shared an identity of >70%.

paper is VDMG01000000. The raw reads of 21 metagenomes have been deposited at the NCBI Sequence Read Archive under the accession number SRP047512 (4).

ACKNOWLEDGMENTS

The sequencing was supported by U.S. Department of Energy grant DE-SC0004902 to S.M.P.

We thank the Princeton University Research Computing Office of Information Technology staff for their support with the computational analyses and the GEO523 2016 class for filtering the metagenomic reads on the Galaxy platform.

M.C.Y.L. conceived the analysis. T.A.V. performed the total DNA extraction and submitted it to A.L. and A.C. for sequencing. A.C. and A.L. performed the initial quality filtering. C.R. and M.C.Y.L. assembled the sequenced reads. C.R. performed the mapping, binning, reassembly, gene prediction, and annotation with consultation from M.C.Y.L., and C.R., M.C.Y.L., and T.C.O. contributed to the interpretation of the data and production of the manuscript.

We declare no conflict of interest.

REFERENCES

1. Lau MCY, Stackhouse BT, Layton AC, Chauhan A, Vishnivetskaya TA, Chourey K, Ronholm J, Mykytczuk NCS, Bennett PC, Lamarche-Gagnon G, Burton N, Pollard WH, Omelon CR, Medvigy DM, Hettich RL, Pfiffner SM, Whyte LG, Onstott TC. 2015. An active atmospheric methane sink in high Arctic mineral cryosols. ISME J 9:1880–1891. https://doi.org/10.1038/ismej.2015.13.
2. Stackhouse BT, Vishnivetskaya TA, Layton A, Chauhan A, Pfiffner S, Mykytczuk NC, Sanders R, Whyte LG, Hedin L, Saad N, Myneni S, Onstott TC. 2015. Effects of simulated spring thaw of permafrost from mineral cryosol on CO_2 emissions and atmospheric CH4 uptake. J Geophys Res Biogeosci 120:1764–1784. https://doi.org/10.1002/2015JG003004.
3. Stackhouse BT, Lau MCY, Vishnivetskaya TA, Burton N, Wang R, Southworth A, Whyte LG, Onstott TC. 2017. Atmospheric CH_4 oxidation by Arctic permafrost and mineral cryosols as a function of water saturation and temperature. Geobiology 15:94–111. https://doi.org/10.1111/gbi.12193.
4. Chauhan A, Layton AC, Vishnivetskaya TA, Williams D, Pfiffner SM, Rekepalli B, Stackhouse B, Lau MCY, Phelps TJ, Mykytczuk N, Ronholm J, Whyte LG, Onstott TC, Sayler GS. 2014. Metagenomes from thawing low-soil-organic-carbon mineral cryosols and permafrost of the Canadian high Arctic. Genome Announc 2:e01217-14. https://doi.org/10.1128/genomeA.01217-14.
5. Peng Y, Leung HCM, Yiu SM, Chin F. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428. https://doi.org/10.1093/bioinformatics/bts174.
6. 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.
7. 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.
8. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation.

Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
9. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.
10. Tveit AT, Grethe A, Robinson SL, Schintlmeister A, Dedysh SN. 2019. Widespread soil bacterium that oxidizes atmospheric methane. Proc Natl Acad Sci U S A 116:8515–8524. https://doi.org/10.1073/pnas.1817812116.
11. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
13. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:182–185.
14. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, Darling AE, Malfatti S, Swan BK, Gies EA, Dodsworth J. a, Hedlund BP, Tsiamis G, Sievert SM, Liu W-T, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature 499:431–437. https://doi.org/10.1038/nature12352.
15. Eren A, Esen Ö, Quince C, Vineis J, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319. https://doi.org/10.7717/peerj.1319.
16. Rodriguez-R L, Konstantinidis K. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. PeerJ Prepr 4:e1900v1. https://peerj.com/preprints/1900/.
17. Pratscher J, Vollmers J, Wiegand S, Dumont MG, Kaster A-K. 2018. Unravelling the identity, metabolic potential and global biogeography of the atmospheric methane-oxidizing upland soil cluster α. Environ Microbiol 20:1016–1029. https://doi.org/10.1111/1462-2920.14036.