Microbiology Resource Announcements

| Environmental Microbiology | Announcement

Draft genome of methanol-oxidizing *Methylobacterium fujisawaense* strain LAC1

Jooho Chung,[1] Jinha Kim,[1] John C. Blazier,[2] Myung Hwangbo,[1,3] Kung-Hui Chu[1]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We report the draft genome of *Methylobacterium fujisawaense* LAC1 isolated from an acidic aquifer in Indian Head, MD, USA. The genome contains 5,883,000 bp and has a GC content of 70% with 5,434 protein-encoding genes with functional assignments. This strain can grow on methanol with lanthanum, a rare earth element.

KEYWORDS *Methylobacterium fujisawaense*, rare earth elements, methanol oxidation, acidophiles

Methylobacterium fujisawaense LAC1 (designated as LAC1 hereafter) is an acidophilic, pink-pigmented methylotrophic bacterium that can grow on methanol in acidic medium (pH 5.5) with lanthanum (La), a rare earth element. LAC1 was isolated from an enrichment culture inoculated with the groundwater of a contaminated acidic aquifer site at the Indian Head Division Naval Surface Warfare Center, Charles County, MD. The enrichment culture was established in calcium- and copper-free ATCC 2157 medium (1) with 0.1% methanol and 30 µM La^{3+} at room temperature. ATCC 2157 medium was prepared without the addition of calcium and copper salts. The enrichment culture was then incubated with shaking at room temperature. Following several respikes of methanol, a loopful of the enrichment culture was used to streak on Gelrite (RPI, IL) plates containing the same medium with 0.5% methanol and 30 µM La^{3+}. After restreaking several times, a pure colony of LAC1 was obtained. The genomic DNA of LAC1 was extracted using a FastDNA SPIN kit for soil (MP biomedical, CA), and the extracted gDNA quantity was determined using the Qubit High Sensitivity (HS) Assay Kit on the Qubit Fluorometer 4.0 (Invitrogen, CA), and the integrity was confirmed by the Genomic DNA ScreenTape Assay Kit (Agilent, CA) on the 4200 TapeStation (Agilent, CA). Library of LAC1 was prepared using Illumina DNA Prep Kit with Illumina IDT for Illumina UD indexes (Plate A; Illumina, CA). The same HS Assay Kit was used for post-library quantification and quality controlled by D1000 ScreenTape Assay Kit (Agilent, CA) on the 4200 TapeStation. Sequencing was conducted by the Illumina MiSeq platform (Illumina, CA) with the paired-end 2 × 250 bp strategy, resulting in a total of 518,277 paired-end reads. The reads were trimmed, adapter sequences were removed, and quality control was performed using FastQC within Trim Galore v0.6.7 (2). Sequence reads were assembled *de novo* in SPAdes v3.15.3 (3) using the "isolate" setting, on the Texas A&M University's Grace computing cluster. BLAST search and BV-BRC's MinHash k-mer-based "Similar Genome Finder" tool were used to identify the isolate as the most closely related to *Methylobacterium radiotolerans* strain JCM 2831. Annotation of the assembly with the RASTtk-based (4) custom annotation pipeline at BV-BRC (5) indicated a high level of contamination (>10%). The SPAdes contigs showed a bimodal distribution of k-mer coverage values, with one group of contigs having k-mer coverage values ~ 8 and another group with k-mer coverage values ~2. Removal of the lower k-mer coverage contigs using a publicly available Bash one liner (https://github.com/ECBSU/oneliners) yielded a high-quality (high completeness and low contamination) assembly when annotated again with the

Editor J. Cameron Thrash, University of Southern California, Los Angeles, California, USA

Address correspondence to Kung-Hui Chu, kchu@civil.tamu.edu.

The authors declare no conflict of interest.

Received 19 April 2023
Accepted 28 July 2023
Published 20 September 2023

Copyright © 2023 Chung et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Downloaded from https://journals.asm.org/journal/mra on 19 December 2023 by 129.113.53.71.

custom annotation pipeline at BV-BRC. The filtered contigs were then submitted and annotated through NCBI PGAP v4.6 (6). Genome coverage was calculated by SAMtools v1.16.1 (7) after mapping the reads to the assembled genome using BWA-MEM2 v2.2.1 (8). Default parameters were used for all software unless otherwise specified.

The assembled draft genome of LAC1 is 5,883,000 bp long with 28.0× genome coverage and 70% GC content. It consists of 72 contigs, with an N_{50} value of 185,086 bp. The genome contains 5,556 total genes including 5,434 total protein-encoding genes, 3 5S rRNA genes, 1 16S rRNA gene, 1 23S rRNA gene, and 49 tRNA genes. The 16S rRNA gene sequence of LAC1 has 92% similarity to *M. fujisawaense* DSM 5686.

ACKNOWLEDGMENTS

We further acknowledge the staff members in the Institute for Genome Sciences and Society at Texas A&M University for their assistance in sequencing and sequence data processing.

AUTHOR AFFILIATIONS

[1]Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, Texas, USA
[2]Texas A&M Institute for Genome Sciences and Society, Texas A&M University, College Station, Texas, USA
[3]School of Earth, Environmental and Marine Sciences, The University of Texas - Rio Grande Valley, Brownsville, Texas, USA

AUTHOR ORCIDs

Jooho Chung http://orcid.org/0009-0009-2258-5402
Jinha Kim http://orcid.org/0000-0002-2380-3688
Myung Hwangbo http://orcid.org/0000-0001-9216-0519
Kung-Hui Chu http://orcid.org/0000-0002-4212-1789

AUTHOR CONTRIBUTIONS

Jooho Chung, Data curation, Writing – original draft, Writing – review and editing, Methodology | Jinha Kim, Data curation, Writing – review and editing | John C. Blazier, Formal analysis, Methodology, Writing – review and editing | Myung Hwangbo, Writing – review and editing | Kung-Hui Chu, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review and editing, Conceptualization, Resources

DATA AVAILABILITY

The draft genome sequence was deposited in GenBank under the accession number JARWLY000000000. The BioProject, BioSample, and SRA accession numbers are PRJNA948434, SAMN33903885, and SRR24019894, respectively.

REFERENCES

1. ATCC. ATCC 2157 medium recipe, https://www.atcc.org/~/media/01d07065549946ac863c486728b33349.ashx
2. Krueger F. 2015. Trim Galore!: a wrapper around cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. Babraham Institute.
3. 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
4. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365
5. Olson RD, Assaf R, Brettin T, Conrad N, Cucinell C, Davis JJ, Dempsey DM, Dickerman A, Dietrich EM, Kenyon RW, Kuscuoglu M, Lefkowitz EJ, Lu J, Machi D, Macken C, Mao C, Niewiadomska A, Nguyen M, Olsen GJ, Overbeek JC, Parrello B, Parrello V, Porter JS, Pusch GD, Shukla M, Singh I, Stewart L, Tan G, Thomas C, VanOeffelen M, Vonstein V, Wallace ZS, Warren AS, Wattam AR, Xia F, Yoo H, Zhang Y, Zmasek CM, Scheuermann RH, Stevens RL. 2023. Introducing the bacterial and viral bioinformatics

Downloaded from https://journals.asm.org/journal/mra on 19 December 2023 by 129.113.53.71.

resource center (BV-BRC): a resource combining PATRIC, IRD and ViPR. Nucleic Acids Res 51:D678–D689. https://doi.org/10.1093/nar/gkac1003
6. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44:6614–6624. https://doi.org/10.1093/nar/gkw569
7. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352
8. Vasimuddin M, Misra S, Li H, Aluru S. Efficient architecture-aware acceleration of BWA-MEM for multicore systems 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS); Rio de Janeiro, Brazil: p 314–324. https://doi.org/10.1109/IPDPS.2019.00041

Downloaded from https://journals.asm.org/journal/mra on 19 December 2023 by 129.113.53.71.