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 Chu1

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 La3+ 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 La3+. 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). 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/loneliers) yielded a high-quality (high completeness and low contamination) assembly when annotated again with the
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 N50 value of 185,086 bp. The genome contains 5,556 total genes including 5,434 total protein-encoding genes, 355 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

1Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, Texas, USA
2Texas A&M Institute for Genome Sciences and Society, Texas A&M University, College Station, Texas, USA
3School 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 PRJNA948434, SAMN33903885, and SRR24019894, respectively.

REFERENCES

1. ATCC. ATCC 2157 medium recipe, https://www.atcc.org/~/media/01d07065549946ac863c486728b33349.ashx

