Identification of a Metagenome-Assembled Genome of an Uncultured *Methyloceanibacter* sp. Strain Acquired from an Activated Sludge System Used for Landfill Leachate Treatment

Shohei Yasuda,a Toshikazu Suenaga,b Laura Orschler,c Shelesh Agrawal,c Susanne Lackner,c Akihiko Teradaa,b

aDepartment of Chemical Engineering, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan
bGlobal Innovation Research Institute, Tokyo University of Agriculture and Technology, Fuchu, Tokyo, Japan
cDepartment of Civil and Environmental Engineering Science, Institute IWAR, Chair of Wastewater Engineering, Technical University of Darmstadt, Darmstadt, Germany

ABSTRACT Using metagenome sequencing, a nearly complete genome sequence was retrieved for the uncultured *Methyloceanibacter* sp. strain A49, recovered from an activated sludge system used for landfill leachate treatment at a closed landfill site. The total size and encoded sequences are 3,407,434 bp and 3,280 genes, respectively.

Until now, methylotrophic bacterial species of the genus *Methyloceanibacter*, belonging to the order *Rhizobiales* (class *Alphaproteobacteria*), have been discovered only in marine environments (1–3). Here, we report for the first time the metagenome-assembled genome (MAG) sequence of an uncultured *Methyloceanibacter* species retrieved from the activated sludge tank of a landfill site that was closed for nitrogen removal from the leachate.

An activated sludge sample was taken from an intermittently aerated sequencing batch reactor removing nitrogen from landfill leachate at a closed landfill site (Tokyo, Japan). The reactor received a periodic methanol supply as an external electron donor, and the leachate had a temperature, pH, total nitrogen concentration, and chloride ion concentration of 28.3°C, 7.65, 159 mg N/liter, and 7,693 mg/liter, respectively. DNA was extracted using the Fast DNA spin kit for soil (MP Biomedicals, Santa Ana, CA, USA) according to the manufacturer's protocol. The quality of the DNA was checked using gel electrophoresis, and the concentration was measured using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). Enzymatic shearing was performed with the Ion Xpress Plus fragment library kit (Thermo Fisher Scientific) to prepare fragmented libraries from genomic DNA for sequencing. Size selection of the library for 600-bp reads was performed with E-Gel Size Select II agarose gel, and libraries were tagged using Ion Xpress barcode adapters (Thermo Fisher Scientific). Template preparation was performed on an Ion Chef system with the Ion 520 and Ion 530 ExT kit. Sequencing was performed on the Ion Torrent Ion S5 system using the 530 chip. Torrent Suite v. 4.4.2 (Thermo Fisher Scientific GmbH, Dreieich, Germany) was utilized for base calling with default parameters according to the manufacturer's protocol, which resulted in 20,244,817 single-end reads with a mean length of 334 bp. Quality filtering of the metagenomic reads was performed using Trimmomatic v. 0.36 (4). Trimming was applied to both sides of the reads to ensure a Q20 quality score, while maintaining a minimum read length of 250 bp. After trimming, 13,005,150 total reads were assembled using MEGAHIT v. 1.1.3 (5) at the minimum contig length of 1,000 bp, producing 44,392 contigs (total, 104,856,336 bp; N_{50}, 8,139 bp; and L_{50}, 2,909 bp). Bowtie 2 v. 2.3.4.2 (6) and SAMtools v. 1.9 (7) were used for mapping the reads. After converting the sequence alignment map (SAM) files to binary alignment map (BAM) files with SAMtools v. 1.9 (7), single-copy bacterial and archaeal genes were identified using HMMER v. 3.2 (8). Default parameters were used for all software unless otherwise specified.

Unsupervised binning by CONCOCT v. 0.4.0 (9) and subsequent manual curation by Anvi'o v. 5 (10) were used to obtain MAGs (with more than 70% completeness and less than 10% redundancy) from the metagenomic assembly. The taxonomic classification of the MAGs was verified using reference genomes from NCBI RefSeq with MiGA (http://microbial-genomes.org) (11). One of the MAGs retrieved by MiGA is proximal to *Methyloceanibacter* sp. strain wino2 (GenBank accession number NZ_CP028960), and the average nucleotide identity (ANI) is 81.38%, suggesting that the MAG may represent a novel *Methyloceanibacter* species. The completeness and contamination are 91.39% and 1.68%, respectively, using CheckM v. 1.0.7 with taxonomic-specific workflow (phylum *Proteobacteria*) (12).

The genome of A49 was annotated using the DDBJ Fast Annotation and Submission Tool (DFAST) pipeline (13). The genome length is 3,407,434 bp, it consists of 447 contigs, and it has a GC content of 62.73%, 26× average coverage, an N_{50} value of 10,298 bp, 1 rRNA, and 42 tRNAs. Functional annotation with KofamKOALA (14) revealed that the MAG harbors the *xoxF* gene, which encodes a pyrroloquinoline quinone-dependent methanol dehydrogenase (15), suggesting that the MAG has a possible methanol oxidation function.

Data availability. The MAG sequence of *Methyloceanibacter* sp. strain A49 has been deposited in DDBJ/EMBL/GenBank under the accession numbers BLYM01000001 through BLYM01000447. The BioSample accession number is SAMD00231436. The read data were also deposited in DDBJ Sequence Read Archive (SRA) under the SRA experiment accession number DRX196218 (run number DRR205819).

REFERENCES

1. Takeuchi M, Katayama T, Yamagishi T, Hanada S, Tamaki H, Kamagata Y, Oshima K, Hattori M, Marumo K, Nedachi M, Maeda H, Suwa Y, Sakata S. 2014. *Methyloceanibacter caenitepidi* gen. nov., sp. nov., a facultatively methylotrophic bacterium isolated from marine sediments near a hydrothermal vent. Int J Syst Evol Microbiol 64:462–468. https://doi.org/10.1099/ijs.0.053397-0.
2. Vekeman B, Kerckhof F-M, Cremers G, de Vos P, Vandamme P, Boon N, Op den Camp HJM, Heylen K. 2016. New *Methyloceanibacter* diversity from North Sea sediments includes methanotroph containing solely the soluble methane monooxygenase. Environ Microbiol 18:4523–4536. https://doi.org/10.1111/1462-2920.13485.
3. Yu W-J, Lee J-W, Nguyen N-L, Rhee S-K, Park S-J. 2018. The characteristics and comparative analysis of methanotrophs reveal genomic insights into *Methylomicrobium* sp. enriched from marine sediments. Syst Appl Microbiol 41:415–426. https://doi.org/10.1016/j.syapm.2018.05.004.
4. 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.
5. 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.
6. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
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. Finn RD, Clements J, Eddy SR. 2011. HMMER Web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37. https://doi.org/10.1093/nar/gkr367.
9. Alneberg J, Bjarnason BS, De Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. Nat Methods 11:1144–1146. https://doi.org/10.1038/nmeth.3103.
10. Eren AM, Esen OC, Quince C, Vineis JH, 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.
11. Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT. 2018. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucleic Acids Res 46:W282–W288. https://doi.org/10.1093/nar/gky467.
12. 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.
13. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037–1039. https://doi.org/10.1093/bioinformatics/btx713.
14. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2020. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36:2251–2252. https://doi.org/10.1093/bioinformatics/btz859.
15. Nakagawa T, Mitsui R, Tani A, Sasa K, Tashiro S, Iwama T, Hayakawa T, Kawai K. 2012. A catalytic role of *XoxF1* as La³⁺-dependent methanol dehydrogenase in *Methylobacterium extorquens* strain AM1. PLoS One 7:e50480. https://doi.org/10.1371/journal.pone.0050480.