Complete and Draft Genome Sequences of Aerobic Methanotrophs Isolated from a Riparian Wetland

Ohana Yonara de Assis Costa,[a] Marion Meima-Franke,[a] Paul L. E. Bodelier[a]

[a]Department of Microbial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands

ABSTRACT Wetlands are important sources of methane emissions, and the impacts of these emissions can be mitigated by methanotrophic bacteria. The genomes of methanotrophs *Methylomonas* sp. strain LL1 and *Methylosinus* sp. strain H3A, as well as *Methylocystis* sp. strains H4A, H15, H62, and L43, were sequenced and are reported here.

W etlands are among the largest source of the greenhouse gas methane (1). Methanotrophic bacteria are the only biological filter, mitigating emission to the atmosphere (1). Here, we report the complete genome sequence of *Methylomonas* sp. strain LL1 and the draft genome sequences of *Methylosinus* sp. strain H3A and *Methylocystis* sp. strains H4A, H15, H62, and L43. These bacteria were isolated as previously described (2) from the same riparian wetland where composition and functioning of the methanotrophic community were described (3–5).

Strains were grown as described previously (2) in a nitrate mineral salts (NMS) medium, at 25°C in an atmosphere of 30% CH_4 in air, and were harvested in mid-exponential phase. Bacterial genomic DNA was extracted using either the GNOME DNA isolation kit (MP Biomedicals, USA) (all *Methylocystis* sp. strains) or the Genomic-tip 100/G kit (Qiagen Benelux BV, The Netherlands) (*Methylomonas* sp. strain LL1 and *Methylosinus* sp. strain H3A), according to the manufacturers' instructions. The sequencing libraries of the genomes of *Methylomonas* sp. strain LL1 and *Methylosinus* sp. strain H3A were prepared using the SMRTbell template prep kit 1.0 and sequenced in one single-molecule real-time (SMRT) cell using the PacBio RS II (Pacific Biosciences, Inc., USA) sequencing platform, which was executed by the Genomics Facility of the School of Medicine of the University of Maryland (Baltimore, USA). Quality control, raw read filtering, and genome assembly were performed with the help of SMRT analysis software v2.3.0 (Pacific Biosciences, Inc.) featuring Hierarchical Genome Assembly Process algorithm v3 (HGAP3) (6). The HGAP3 data processing pipeline comprised PreAssembler v1 for filtering, Celera assembler v8.1 for assembly (7), BLASR v1 (8) for mapping, and Quiver v1 (6) for consensus polishing using only unambiguously mapped reads. HGAP3 defaults settings were applied, except for the genome size estimate parameter, set to 5.0 Mbp. The completeness of the *Methylomonas* sp. strain LL1 genome was assessed using the benchmarking universal single-copy orthologs (BUSCO; v4.14) software (https://busco.ezlab.org) (9) with the gammaproteobacteria_odb10 database (quality score, 98.9%; 366 total BUSCO groups searched). The genomes of all *Methylocystis* sp. strains were sequenced using an Illumina HiSeq 2000 instrument. Library preparation (10), sequencing, and sequence quality control and trimming using internally developed software SOAPnuke v1.4.0 and parameters -l 15 -q 0.2 -n 0.05 (11) were performed at BGI Tech Solutions (Hong Kong, China). The sequencing library was prepared according to BGI protocols. Briefly, 1 µg genomic DNA was randomly fragmented by a Covaris sonicator. DNA fragments were end repaired and A tailed. Next, Illumina adapters were ligated to the 3'-adenylated DNA, which was purified using AxyPrep Mag PCR clean up kit. The resulting fragments

TABLE 1 General genome information

Strain	Sequencing platform	Total size of assembly (Mb)	Assembler	No. of raw reads	Raw read N$_{50}$ (bp)	Coverage (×)	Completion	G+C content (%)	No of CDS[a]	No of scaffolds	Scaffold N$_{50}$ (bp)	Core metabolic pathways[b]	GenBank (SRA) accession no.[c]
Methylomonas sp. LL1	PacBio RS II	4,923,893	Celera v8.1	78,774	14,377	16	Yes	50.9	4,426	2	4,798,577	pMMO, Mox, PQQ, sMMO, SC, PPP, RuMP, EDD, EMP, TCA	CP064653.1, CP064652.1 (SRR13259185)
Methylosinus sp. H3A	PacBio RS II	5,384,582	Celera v8.1	63,638	15,418	16	No	64.2	4,974	11	4,391,449	pMMO, Mox, PQQ, sMMO, SC, EMP, TCA	JADNQW000000000 (SRR13259182)
Methylocystis sp. H15	Illumina HiSeq 2000	4,053,493	Ray 2.0.1	6,891,600		49.8	No	62.3	3,913	39	217,286	pMMO, Mox, PQQ, sMMO, SC, EMP, TCA	JADNQX000000000 (SRR13259184)
Methylocystis sp. H4A	Illumina HiSeq 2000	4,288,666	Ray 2.0.1	6,990,370		49.8	No	62.3	4,214	38	238,602	pMMO, Mox, PQQ, sMMO, SC, EDD, EMP, TCA	JADNQY000000000 (SRR13259180)
Methylocystis sp. H62	Illumina HiSeq 2000	4,540,521	Ray 2.0.1	6,914,734		43.9	No	62.1	4,408	63	216,123	pMMO, Mox, PQQ, sMMO, SC, EMP, TCA	JADNQZ000000000 (SRR13259181)
Methylocystis sp. L43	Illumina HiSeq 2000	4,061,225	Ray 2.0.1	6,899,554		48.6	No	62.4	3,919	38	145,473	pMMO, Mox, PQQ, sMMO, SC, EMP, TCA	JADNRA000000000 (SRR13259183)

[a] CDS, coding DNA sequences.
[b] pMMO, membrane-bound methane monooxygenase; Mox, PQQ-linked methanol and formaldehyde dehydrogenases; PQQ, pyrroloquinoline quinone biosynthesis; sMMO, soluble methane monooxygenase; SC, serine cycle; PPP, pentose phosphate pathway; RuMP, assimilatory ribulose monophosphate pathway; EDD, Entner-Doudoroff pathway; EMP, Embden-Meyerhof-Parnas pathway; TCA, tricarboxylic acid cycle.
[c] The BioProject number of the study is PRJNA674997.

were size selected by agarose gel electrophoresis. The library quality and quantity were assessed using the Agilent Technologies 2100 bioanalyzer and ABI StepOnePlus real-time PCR system. The quality of all reads was checked by FastQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Clean reads were assembled using Ray v2.0.1 (12). Automatic gene prediction and annotation were performed by using NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (13). Default parameters were used for all software unless otherwise specified. The *Methylomonas* sp. LL1 genome includes a 0.12-Mb plasmid. General genome statistics and predicted metabolic pathways are detailed in Table 1.

Data availability. GenBank and SRA (raw data) accession numbers are listed in Table 1.

ACKNOWLEDGMENTS

REFERENCES

1. Bodelier PLE, Pérez G, Veraart AJ, Krause SMB. 2019. Methanotroph ecology, environmental distribution and functioning, p 1–38. *In* Lee EY (ed), Methanotrophs. Springer International Publishing, Cham, Switzerland.
2. Bodelier PLE, Meima-Franke M, Zwart G, Laanbroek HJ. 2005. New DGGE strategies for the analyses of methanotrophic microbial communities using different combinations of existing 16S rRNA-based primers. FEMS Microbiol Ecol 52:163–174. https://doi.org/10.1016/j.femsec.2004.11.004.
3. Bodelier PLE, Bär-Gilissen M-J, Meima-Franke M, Hordijk K. 2012. Structural and functional response of methane-consuming microbial communities to different flooding regimes in riparian soils. Ecol Evol 2:106–127. https://doi.org/10.1002/ece3.34.
4. Bodelier PLE, Meima-Franke M, Hordijk CA, Steenbergh AK, Hefting MM, Bodrossy L, von Bergen M, Seifert J. 2013. Microbial minorities modulate methane consumption through niche partitioning. ISME J 7:2214–2228. https://doi.org/10.1038/ismej.2013.99.
5. Krause S, Meima-Franke M, Hefting MM, Bodelier PLE. 2013. Spatial patterns of methanotrophic communities along a hydrological gradient in a riparian wetland. FEMS Microbiol Ecol 86:59–70. https://doi.org/10.1111/1574-6941.12091.
6. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
7. Denisov G, Walenz B, Halpern AL, Miller J, Axelrod N, Levy S, Sutton G. 2008. Consensus generation and variant detection by Celera assembler. Bioinformatics 24:1035–1040. https://doi.org/10.1093/bioinformatics/btn074.
8. Chaisson MJ, Tesler G. 2012. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. BMC Bioinformatics 13:238. https://doi.org/10.1186/1471-2105-13-238.
9. Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. Methods Mol Biol 1962:227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.
10. Ulanova D, Uenaka Y, Sakama M, Sakurai T. 2020. Draft genome sequence of *Salinispora* sp. strain H7-4, isolated from deep-sea sediments of the Shikoku Basin. Microbiol Resour Announc 9:e00834-20. https://doi.org/10.1128/MRA.00834-20.
11. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, Zhang X, Wang J, Yang H, Fang L, Chen Q. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. GigaScience 7:gix120. https://doi.org/10.1093/gigascience/gix120.
12. Boisvert S, Laviolette F, Corbeil J. 2010. Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. J Comput Biol 17:1519–1533. https://doi.org/10.1089/cmb.2009.0238.
13. 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.