Draft Genome Sequences of Two Thermophilic *Moorella* sp. Strains, Isolated from an Acidic Hot Spring in Japan

Yuto Fukuyama,[a] Ayumi Tanimura,[a] Masao Inoue,[a] Kimiho Omae,[a] Takashi Yoshida,[a] Yoshihiko Sako[a]

[a]Laboratory of Marine Microbiology, Graduate School of Agriculture, Kyoto University, Kitashirakawa Oiwake, Sakyo, Kyoto, Japan

ABSTRACT The thermophilic *Moorella* sp. strains E308F and E306M were isolated from an acidic hot spring in Japan. Here, we report the draft genome sequences of E308F (3.06 Mbp; G+C content, 54.0%) and E306M (2.99 Mbp; G+C content, 54.4%), to advance the genomic information available on the genus *Moorella*.

The genus *Moorella* is often represented by anaerobic, thermophilic, and CO-utilizing bacteria that use the Wood-Ljungdahl pathway for CO or CO_2 fixation and acetogenesis (1). Recently, some strains and species of *Moorella* were reported as hydrogenogenic CO-oxidizing bacteria (2, 3). Here, we successfully isolated two novel *Moorella* sp. strains, E308F and E306M, under 100% CO cultivation, and we report their draft genome sequences.

Samples were collected from Unagi-onsen (31°13′39″N, 130°36′46″E), an acidic hot spring in Japan. For isolation, the cells were grown at 65°C on modified hypotonic artificial seawater (hASW) medium (4) under 100% CO. Modifications of this medium were as follows: the concentration of yeast extract was increased to 1.0 g/liter, and sodium pyruvate (1.0 g/liter) was supplemented. To identify the isolates (strains E308F and E306M), we extracted their genomic DNA using a DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA). Their partial 16S rRNA genes were amplified using the universal primers B27f (5′-AGAGTTTGATCCTGGCTCAG-3′) and U533r (5′-TTACCGCGGC KGCTGRCAC-3′) (5, 6) for Sanger sequencing. The sequence data showed the highest sequence identity with the 16S rRNA genes of *Moorella humiferrea* (96.2% and 97.4% for E308F and E306M, respectively) by BLASTN search (7), suggesting the assignment of the two strains to the genus *Moorella*.

For genome sequencing, the isolates were grown on modified hASW medium. Mate pair libraries were prepared from DNA using a Nextera mate pair library preparation kit (Illumina, Inc., San Diego, CA, USA), followed by sequencing with the Illumina MiSeq platform using a v3 reagent kit (2×300-bp mate pair reads). Totals of 4,147,082 and 3,134,908 reads were generated for E308F and E306M, respectively. The generated reads were filtered and trimmed using Trimmomatic v0.36 (8). The Nextera mate pair junction adaptor sequences were trimmed using NxTrim v0.4.1 (9), and only proper mate pair reads (3,796,450 and 2,817,228, respectively) were assembled into scaffolds using SPAdes v3.13.0 (10). After the correction of large-scale misassemblies by examination of the distribution of Nextera mate pair insert sizes using NxRepair v0.13 (11), the assembled scaffolds were analyzed using DFAST server v1.2.0 (12) to predict open reading frames (ORFs) with annotation. The average nucleotide identity (ANI) was calculated by JSpeciesWS (13). All analyses were performed with default parameter settings.

The draft genomes of E308F and E306M were assembled into 10 and 6 scaffolds with 373× and 283× average genome coverages, N_{50} sizes of 1,211,500 bp and 2,985,197 bp, total lengths of 3,057,355 bp and 2,985,197 bp, and average G+C contents of 54.0% and 54.4%, containing 3,106 and 3,051 predicted ORFs, respectively. The

ANI percentage between the two strains was 98.9%. In contrast, the ANI percentages between strain E308F and other published *Moorella* genomes ranged from 78.5% (*Moorella thermoacetica* ATCC 39073; GCF_000013105) to 90.1% (*Moorella mulderi* DSM 14980; GCF_001594015), which were below the proposed 95% cutoff for the genome definition of a species (14).

Both strains possessed two CO dehydrogenase (CODH) genes, and these amino acid sequence alignments showed their conserved active site residues (15, 16). Of these, one CODH gene was flanked by an energy-converting hydrogenase, implying the capability of hydrogenogenic carboxydotrophy (17).

Data availability. The genome sequences of *Moorella* sp. strains E308F and E306M have been deposited in the DNA Data Bank of Japan (DDBJ) under accession numbers BJKN01000000 and BJKO01000000, respectively. Sequence data have been deposited in the DDBJ Sequence Read Archive under the accession numbers DRX170528 and DRX170529, respectively.

REFERENCES

1. Drake HL, Gössner AS, Daniel SL. 2008. Old acetogens, new light. Ann N Y Acad Sci 1125:100–128. https://doi.org/10.1196/annals.1419.016.
2. Jiang B, Henstra AM, Paulo PL, Balk M, van Doesburg W, Stams AJM. 2009. Atypical one-carbon metabolism of an acetogenic and hydrogenogenic *Moorella thermoacetica* strain. Arch Microbiol 191:123–131. https://doi.org/10.1007/s00203-008-0435-x.
3. Alves JI, van Gelder AH, Alves MM, Sousa DZ, Plugge CM. 2013. *Moorella stamsii* sp. nov., a new anaerobic thermophilic hydrogenogenic carboxydotroph isolated from digester sludge. Int J Syst Evol Microbiol 63:4072–4076. https://doi.org/10.1099/ijs.0.050369-0.
4. Yoneda Y, Yoshida T, Yasuda H, Imada C, Sako Y. 2013. A thermophilic, hydrogenogenic and carboxydotrophic bacterium, *Calderihabitans maritimus* gen. nov., sp. nov., from a marine sediment core of an undersea caldera. Int J Syst Evol Microbiol 63:3602–3608. https://doi.org/10.1099/ijs.0.050468-0.
5. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc Natl Acad Sci U S A 82:6955–6959. https://doi.org/10.1073/pnas.82.20.6955.
6. Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC. 1989. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. Nucleic Acids Res 17:7843–7853. https://doi.org/10.1093/nar/17.19.7843.
7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
8. 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.
9. O'Connell J, Schulz-Trieglaff O, Carlson E, Hims MM, Gormley NA, Cox AJ. 2015. NxTrim: optimized trimming of Illumina mate pair reads. Bioinformatics 31:2035–2037. https://doi.org/10.1093/bioinformatics/btv057.
10. 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.
11. Murphy RR, O'Connell J, Cox AJ, Schulz-Trieglaff O. 2015. NxRepair: error correction in de novo sequence assembly using Nextera mate pairs. PeerJ 3:e996. https://doi.org/10.7717/peerj.996.
12. 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.
13. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a Web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931. https://doi.org/10.1093/bioinformatics/btv681.
14. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol 68:461–466. https://doi.org/10.1099/ijsem.0.002516.
15. Dobbek H, Svetlitchnyi V, Gremer L, Huber R, Meyer O. 2001. Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster. Science 293:1281–1285. https://doi.org/10.1126/science.1061500.
16. Svetlitchnyi V, Dobbek H, Meyer-Klaucke W, Meins T, Thiele B, Römer P, Huber R, Meyer O. 2004. A functional Ni-Ni-[4Fe-4S] cluster in the monomeric acetyl-CoA synthase from *Carboxydothermus hydrogenoformans*. Proc Natl Acad Sci U S A 101:446–451. https://doi.org/10.1073/pnas.0304262101.
17. Inoue M, Nakamoto I, Omae K, Oguro T, Ogata H, Yoshida T, Sako Y. 2019. Structural and phylogenetic diversity of anaerobic carbon-monoxide dehydrogenases. Front Microbiol 9:3353. https://doi.org/10.3389/fmicb.2018.03353.