Draft Genome Sequence of *Methanohalophilus mahii* Strain DAL1 Reconstructed from a Hydraulic Fracturing-Produced Water Metagenome

Daniel Lipus,[a,c] Amit Vikram,[a] Daniel E. Ross,[c,d] Kyle Bibby[a,b]

Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania, USA[a]; Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA[b]; National Energy Technology Laboratory (NETL), Pittsburgh, Pennsylvania, USA[c]; AECOM, Pittsburgh, Pennsylvania, USA[d]

We report here the 1,882,100-bp draft genome sequence of *Methanohalophilus mahii* strain DAL1, recovered from Marcellus Shale hydraulic fracturing-produced water using metagenomic contig binning. Genome annotation revealed several key methanogenesis genes and provides valuable information on archaeal activity associated with hydraulic fracturing-produced water environments.

The Marcellus Shale is the largest natural gas reservoir in the United States and is a leading producer of methane through high-volume hydraulic fracturing (1, 2). While the methane in the Marcellus Shale is assumed to be of thermogenic origin, recent observations suggest biogenic methanogenesis potentially contributes to methane production in the Marcellus Shale during the fracturing process (3–5). Several studies have investigated overall microbial populations associated with Marcellus Shale or similar hydraulic fracturing operations (6–8), with *Archaea* being detected at low quantities (6, 9, 10). At this point, little is known about the activity of methanogens in hydraulic fracturing environments.

Here, we present the draft genome sequence of *Methanohalophilus mahii* strain DAL1, recovered from the metagenome of Marcellus Shale hydraulic fracturing-produced water. Sequencing libraries were prepared using Illumina Nextera XT and sequenced using Illumina MiSeq technology (Illumina, San Diego, CA). Sequencing reads were quality trimmed (Q30) and *de novo* assembled into contigs using CLC Genomics Workbench version 8.5.1 (CLC Bio, Aarhus, Denmark) and SPAdes version 3.5.1 (11). Assembled contigs were grouped into genome bins with MaxBin (12) and Vizbin (13) and taxonomy assessed with PhyloPythia (14). Metagenomic reads were mapped against binned contigs and reassembled using SPAdes.

The final draft genome contained 58 contigs of 5,000 bp to 131,826 bp in length and an N_{50} of 52,299 bp. The total genome size was 1,894,170 bp, with a mean G+C content of 42.4% and an average of 90-fold coverage. Draft genome completeness and contamination were estimated using CheckM (15). The final draft genome was found to be 93.9% complete and contain 1.1% contamination.

The draft genome was annotated by Rapid Annotations using Subsystems Technology (RAST) (16, 17), revealing 2,066 gene-coding sequences (CDSs) and 49 RNA sequences (46 tRNA, 16S, 23S, and 5S rRNA). Extraction and phylogenetic analysis of the 16S rRNA gene sequence using RDP (18) and CLC Bio suggested *Methanohalophilus mahii* strain DAL1 to be closely related to *Methanohalophilus mahii* strain DSM 5219 (99% BLASTn [19] nucleotide identity). RAST (16, 17) and KEGG (20) annotation allowed the discovery of all core methanogenesis enzymes necessary for the conversion of CO_2 to methane, including the heterodisulfide reductase (*hdr*), methyl coenzyme M reductase (*mcr*), and methenyl-tetrahydromethanopterin cyclohydrolase (*mch*). Moreover, the coenzyme M methyltransferase (*mtbA*) and the trimethylamine methyltransferase (*mttB*) genes were identified. These observations suggest *Methanohalophilus mahii* strain DAL1 to have the genetic potential for hydrogenotrophic (CO_2 to methane) and methyl reduction (trimethylamine to methane) methanogenesis (21).

Annotation also revealed the potassium uptake proteins TrkA and TrkH and the osmoprotectant transporter ProP. These findings agree with previous observations that members of the genus *Methanohalophilus* accumulate potassium ions and uptake glycine and betaine in response to osmotic stress (22–24).

The analysis of the *Methanohalophilus mahii* sp. DAL1 draft genome indicates the potential for methanogenic activity associated with Marcellus Shale hydraulic fracturing operations.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. MAFX00000000. The version described in this paper is version MAFX01000000.

REFERENCES

1. **Kargbo DM, Wilhelm RG, Campbell DJ**. 2010. Natural gas plays in the Marcellus Shale: challenges and potential opportunities. Environ Sci Technol **44**:5679–5684. http://dx.doi.org/10.1021/es903811p.
2. **Gregory KB, Vidic RD, Dzombak DA**. 2011. Water management challenges associated with the production of shale gas by hydraulic fracturing. Elements **7**:181–186. http://dx.doi.org/10.2113/gselements.7.3.181.
3. **Sharma S, Mulder ML, Sack A, Schroeder K, Hammack R**. 2014. Isotope approach to assess hydrologic connections during Marcellus Shale drilling. Ground Water **52**:424–433. http://dx.doi.org/10.1111/gwat.12083.
4. **Tucker YT, Kotcon J, Mroz T**. 2015. Methanogenic *Archaea* in Marcellus Shale: a possible mechanism for enhanced gas recovery in unconventional shale resources. Environ Sci Technol **49**:7048–7055. http://dx.doi.org/10.1021/acs.est.5b00765.
5. **Vikram A, Lipus D, Bibby KJ**. 25 July 2016. Metatranscriptome analysis of active microbial communities in produced water samples from the Marcellus Shale. Microb Ecol [Epub ahead of print.] http://dx.doi.org/10.1007/s00248-016-0811-z.
6. **Murali A, Hartsock A, Bibby KJ, Hammack RW, Vidic RD, Gregory KB**. 2013. Microbial community changes in hydraulic fracturing fluids and produced water from shale gas extraction. Environ Sci Technol **47**:13141–13150. http://dx.doi.org/10.1021/es402928b.
7. **Mohan A, Hartsock A, Hammack RW, Vidic RD, Gregory KB**. 2013. Microbial communities in flowback water impoundments from hydraulic fracturing for recovery of shale gas. FEMS Microbiol Ecol **86**:567–580. http://dx.doi.org/10.1111/1574-6941.12183.
8. **Cluff MA, Hartsock A, MacRae JD, Carter K, Mouser PJ**. 2014. Temporal changes in microbial ecology and geochemistry in produced water from hydraulically fractured Marcellus Shale gas wells. Environ Sci Technol **48**:6508–6517. http://dx.doi.org/10.1021/es501173p.
9. **Akob DM, Cozzarelli IM, Dunlap DS, Rowan EL, Lorah MM**. 2015. Organic and inorganic composition and microbiology of produced waters from Pennsylvania shale gas wells. Appl Geochem **60**:116–125. http://dx.doi.org/10.1016/j.apgeochem.2015.04.011.
10. **Wuchter C, Banning E, Mincer TJ, Drenzek NJ, Coolen MJ**. 2013. Microbial diversity and methanogenic activity of Antrim Shale formation waters from recently fractured wells. Front Microbiol **4**:367. http://dx.doi.org/10.3389/fmicb.2013.00367.
11. **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. http://dx.doi.org/10.1089/cmb.2012.0021.
12. **Wu YW, Tang Y-H, Tringe SG, Simmons BA, Singer SW, Baker R, Morowitz MJ, Banfield JF**. 2014. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome **2**:1. http://dx.doi.org/10.1186/2049-2618-2-1.
13. **Laczny CC, Sternal T, Plugaru V, Gawron P, Atashpendar A, Margossian HH, Coronado S, van der Maaten LV, Vlassis N, Wilmes P**. 2015. VizBin-an application for reference-independent visualization and human-augmented binning of metagenomic data. Microbiome **3**:1. http://dx.doi.org/10.1186/s40168-014-0066-1.
14. **McHardy AC, Martín HG, Tsirigos A, Hugenholtz P, Rigoutsos I**. 2007. Accurate phylogenetic classification of variable-length DNA fragments. Nat Methods **4**:63–72. http://dx.doi.org/10.1038/nmeth976.
15. **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. http://dx.doi.org/10.1101/gr.186072.114.
16. **Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A**. 2008. The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics **9**:386.
17. **Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R**. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res **42**:D206–D214. http://dx.doi.org/10.1093/nar/gkt1226.
18. **Maidak BL, Olsen GJ, Larsen N, Overbeek R, McCaughey MJ, Woese CR**. 1997. The RDP (ribosomal database project). Nucleic Acids Res **25**:109–110. http://dx.doi.org/10.1093/nar/25.1.109.
19. **Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL**. 2008. NCBI blast: a better Web interface. Nucleic Acids Res **36**:W5–W9. http://dx.doi.org/10.1093/nar/gkn201.
20. **Kanehisa M, Goto S**. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res **28**:27–30. http://dx.doi.org/10.1093/nar/28.1.27.
21. **Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R**. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. Nat Rev Microbiol **6**:579–591. http://dx.doi.org/10.1038/nrmicro1931.
22. **Lai MC, Gunsalus RP**. 1992. Glycine betaine and potassium ion are the major compatible solutes in the extremely halophilic methanogen *Methanohalophilus* strain Z7302. J Bacteriol **174**:7474–7477.
23. **Robertson DE, Lai MC, Gunsalus RP, Roberts MF**. 1992. Composition, variation, and dynamics of major osmotic solutes in *Methanohalophilus* strain FDF1. Appl Environ Microbiol **58**:2438–2443.
24. **Lai MC, Hong TY, Gunsalus RP**. 2000. Glycine betaine transport in the obligate halophilic archaeon *Methanohalophilus portucalensis*. J Bacteriol **182**:5020–5024. http://dx.doi.org/10.1128/JB.182.17.5020-5024.2000.