Complete Genomic Sequence of the Thermophilic and Hydrogenotrophic Methanogen *Methanothermobacter* sp. Strain KEPCO-1

Byoung Seung Jeon,[a,b] Hyunjin Kim,[a] Young Wook Go,[a] Young Gook Kim,[c] Ji Sun Joo,[c] Okkyoung Choi,[a] Byoung-In Sang[a]

aDepartment of Chemical Engineering, Hanyang University, Seoul, Republic of Korea
bCenter for Applied Geosciences, University of Tübingen, Tübingen, Germany
cKorea Electric Power Research Institute, Daejeon, Republic of Korea

Byoung Seung Jeon and Hyunjin Kim contributed equally to this work as first authors. Author order was determined in order of increasing seniority.

Okkyoung Choi and Byoung-In Sang contributed equally to this work as corresponding authors.

ABSTRACT Here, we describe the complete genome of *Methanothermobacter* sp. strain KEPCO-1, a thermophilic and hydrogenotrophic methanogen that was isolated from an anaerobic digester in Seoul, Republic of Korea. The genome of KEPCO-1 shares 96.98% of its sequence with *Methanothermobacter marburgensis* strain DSM 2133 and consists of 1,741,029 bp, with 1,822 protein-coding genes, 44 noncoding RNAs, and a GC content of 48.47%. The development of this genome will facilitate future genomic studies of KEPCO-1.

Power-to-gas (P2G) is an electrical storage system (1) that has recently been considered as an approach to addressing load fluctuations associated with renewable energy (2, 3). Using this system, surplus electricity from renewable sources can be converted by water electrolysis into hydrogen, which can then be utilized in methane production with carbon dioxide by using the methanation process (3).

Here, we describe the isolation of *Methanothermobacter* sp. strain KEPCO-1 and the generation of its complete genome. KEPCO-1 grows optimally at 60°C with a pH of 7.0 to 7.5 and produces methane under pressurized conditions (approximately 150 to 200 kPa), using CO_2 and H_2 (4). The isolation process was conducted through sequential subculture using general agar (4, 5). The strain was identified by 16S rRNA gene sequencing using the following primers: 40F (5′-GAT TAA GCC ATG CAA GTC GAA CGA-3′), 450F (5′-CTT CTG GAA TAA GGG CTG GGC A-3′), 765R (5′-CAT CGT TTA CGG CCA GGA CTA C-3′), and 1430R (5′-CTC CTC AAA GAA CCC AGA TTC GAC-3′). Once the strain was identified, whole-genome sequencing was performed.

KEPCO-1 was grown in basal medium at 60°C and 200 kPa for 7 days, and genomic DNA was extracted using the DNeasy UltraClean microbial kit (Qiagen Korea Ltd., Seoul, Republic of Korea) (6). The KEPCO-1 genome was analyzed using the PacBio RS II platform (PacBio single-molecule real-time sequencing) at Macrogen Co. Ltd. (Seoul, Republic of Korea), and the library was prepared using the SMRTbell template preparation kit (7). The raw data were assembled *de novo* by RS HGAP v3 and polished using Quiver v1 in SMRT Portal v2.3.0 (8).

During the preassembly step, filtering and assembly were conducted using preAssembler Filter v1 and preAssembler v2. Default parameters were used for all software unless otherwise specified. After the genome was assembled, gene functions were annotated via the NCBI Prokaryotic Genome Annotation Pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_prok).

The filtered data included a total of 1,099,392,337 bases. There were 81,220 post-filtered reads, encompassing 964,138,339 sequenced bases. The mean subread and N_{50} read lengths were 7,204 bases and 10,409 bases, respectively. The genome is composed of a single chromosome (1,741,029 bp), with an average reference coverage of $421\times$. One contig was generated and the two ends of the contig overlapped, which indicates that the genome is circular. Genome annotations identified a total of 1,866 genes, of which 1,822 were predicted to contain coding sequences, 37 were tRNAs, and 7 were rRNAs. The GC content was 48.75%.

The most closely related strain for KEPCO-1 was *Methanothermobacter marburgensis* DSM 2133 (9, 10), and the average nucleotide identity of KEPCO-1 with respect to the nucleotide sequence of DSM 2133 was 96.98% (11, 12). Furthermore, genes that were involved in energy production and conversion represented the largest proportion of genes in the KEPCO-1 genome (12). The number of genes and the proportion were 169 and 8.9%, respectively. This category is related to genes that are involved in the synthesis of enzymes, coenzymes, and prosthetic groups that are involved in CO_2 and H_2 utilization to produce methane (12).

Data availability. This whole-genome project (in GenBank format) and the corresponding raw files (in fastq format) have been deposited under GenBank accession numbers CP042937 and SRR10298064, respectively.

REFERENCES

1. Clegg S, Mancarella P. 2016. Storing renewables in the gas network: modelling of power-to-gas seasonal storage flexibility in low-carbon power systems. IET Gener Transm Distrib 10:566–575. https://doi.org/10.1049/iet-gtd.2015.0439.
2. Clegg S, Mancarella P. 2015. Integrated modeling and assessment of the operational impact of power-to-gas (P2G) on electrical and gas transmission networks. IEEE Trans Sustain Energy 6:1234–1244. https://doi.org/10.1109/TSTE.2015.2424885.
3. Vo TT, Xia A, Wall DM, Murphy JD. 2017. Use of surplus wind electricity in Ireland to produce compressed renewable gaseous transport fuel through biological power to gas systems. Renew Energy 105:495–504. https://doi.org/10.1016/j.renene.2016.12.084.
4. Choi O, Kim M, Go Y, Hong M-G, Kim B, Shin Y, Lee S, Kim YG, Joo JS, Jeon BS, Sang B-I. 2019. Selective removal of water generated during hydrogenotrophic methanation from culture medium using membrane distillation. Energies 12:4130. https://doi.org/10.3390/en12214130.
5. Angelidaki I, Sanders W. 2004. Assessment of the anaerobic biodegradability of macropollutants. Rev Environ Sci Biotechnol 3:117–129. https://doi.org/10.1007/s11157-004-2502-3.
6. Göttig S, Frank D, Mungo E, Nolte A, Hogardt M, Besier S, Wichelhaus TA. 2019. Emergence of ceftazidime/avibactam resistance in KPC-3-producing *Klebsiella pneumoniae* in vivo. J Antimicrob Chemother 74:3211. https://doi.org/10.1093/jac/dkz330.

7. Rhoads A, Au KF. 2015. PacBio sequencing and its applications. Genomics Proteomics Bioinformatics 13:278–289. https://doi.org/10.1016/j.gpb.2015.08.002.
8. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
9. Stackebrandt E, Goebel BM. 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Evol Microbiol 44:846–849. https://doi.org/10.1099/00207713-44-4-846.
10. Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.
11. Kim M, Oh H-S, Park S-C, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351. https://doi.org/10.1099/ijs.0.059774-0.
12. Liesegang H, Kaster A-K, Wiezer A, Goenrich M, Wollherr A, Seedorf H, Gottschalk G, Thauer RK. 2010. Complete genome sequence of *Methanothermobacter marburgensis*, a methanoarchaeon model organism. J Bacteriol 192:5850–5851. https://doi.org/10.1128/JB.00844-10.