



Draft Genome Sequence of *Microaerobacter geothermalis* Nad S1^T, a Microaerophilic Bacterium Isolated from Tenusia Hot Spring

 Kian Mau Goh,^a  Kok Jun Liew,^a  Saleha Shahar,^a  Iffah Izzati Zakaria,^b  Ummirul Mukminin Kahar^b

^aDepartment of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia

^bMalaysia Genome and Vaccine Institute, National Institutes of Biotechnology Malaysia, Kajang, Selangor, Malaysia

Kian Mau Goh and Ummirul Mukminin Kahar contributed equally to this work.

ABSTRACT *Microaerobacter geothermalis* Nad S1^T is a rare *Bacillaceae* thermophile that grows optimally at 55°C and circumneutral pH. Although strain Nad S1^T was discovered >10 years ago, its genome is yet to be described. The release of the Nad S1^T genome sequence serves as reference genetic information for subsequent use.

Microaerobacter *geothermalis* Nad S1^T (= DSM 22679^T = JCM 16213^T) is the only described type strain of the genus *Microaerobacter* (1). The Hammam Sidi Jdidi hot spring (Nabeul, Tunisia), from which the bacterium was isolated, is located near the Mediterranean Sea (1). Strain Nad S1^T is an anaerobic and microaerophilic bacterium that grows optimally in 1.5 to 3.0% (wt/vol) NaCl (1). This sequencing effort aims to address the gap in type strain genome data.

The genomic DNA of *Microaerobacter geothermalis* Nad S1^T was purchased from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH (Braunschweig, Germany). A paired-end library was prepared using the NEBNext Ultra II DNA library preparation kit for Illumina (New England BioLabs, Ipswich, MA, USA), according to the manufacturer's instructions. Sequencing was performed using the NovaSeq 6000 system (Illumina, San Diego, CA, USA) with 150-bp paired-end reads. Sequence adaptors and low-quality reads were filtered using Trimmomatic v0.39 (2). *De novo* assembly was performed using SOAPdenovo v2.40 (3), SPAdes v3.15.3 (4), and ABySS v2.3.4 (5) before integration with Contig Integrator for Sequence Assembly (CISA) v1.3 (6), and the assembly result with the smallest number of scaffolds was selected. Annotation was conducted using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v5.30 (7). The average nucleotide identity (ANI) values for the genome of strain Nad S1^T versus other type species genomes were analyzed using OrthoANI v0.93.1 (8). The Genome Taxonomy Database Toolkit (GTDB-Tk) v1.7.0 was used to classify the genome (9). Default parameters were used for all software tools unless stated otherwise.

The sequencer generated a total of 1.2 Gb in 3.9 million paired-end reads. Upon removal of the low-quality reads, the assembled genome has a size of 3,132,374 bp, contributed by 85 contigs, with up to 350× coverage, an N_{50} value of 57,386 bp, and an average G+C content of 41.35%. A total of 3,159 genes were identified in the genome, including 3,026 protein-encoding genes, 45 pseudogenes, and 88 genes for RNA (76 tRNA, six 5S RNA, one 16S RNA, one 23S RNA, and four noncoding RNA genes). Based on the earlier study, strain Nad S1^T is able to use nitrate and nitrite as electron acceptors under anaerobic conditions, and the cells can reduce nitrate (1). We found multiple gene sequences for nitrite reductase, nitrate reductase subunit alpha, nitrate reductase subunit beta, nitrate reductase molybdenum cofactor assembly chaperone, and respiratory nitrate reductase subunit gamma. Nad S1^T transports nitrate and nitrite via its NarK family nitrate/nitrite major facilitator superfamily (MFS) transporter.

Editor Frank J. Stewart, Montana State University

Copyright © 2022 Goh et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Kian Mau Goh, gohkianmau@utm.my, or Ummirul Mukminin Kahar, ummirul@nibm.my.

The authors declare no conflict of interest.

Received 30 January 2022

Accepted 23 March 2022

Published 7 April 2022

The 16S rRNA gene of *Microaerobacter geothermalis* Nad S1^T shared 90 to 92% sequence identity with those of other species affiliated with the family *Bacillaceae* (1). The genome comparison analyses indicated that the Nad S1^T genome shared 65.4 to 66.8% ANI with other type species genomes, such as those of *Tepidibacillus fermentans*, *Vulcanibacillus modesticaldus*, *Melghiribacillus thermohalophilus*, and *Heyndrickxia oleronia* (10–13). In addition, none of the curated genomes listed in the GTDB (9) was closely related to the Nad S1^T genome. The genome sequence of *Microaerobacter geothermalis* Nad S1^T will serve as reference genetic information for future research.

Data availability. The whole-genome shotgun sequence of *Microaerobacter geothermalis* Nad S1^T has been deposited in NCBI GenBank under BioProject accession number [PRJNA797672](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA797672), BioSample accession number [SAMN25026507](https://www.ncbi.nlm.nih.gov/biosample/SAMN25026507), and GenBank accession number [JAKIHL000000000](https://www.ncbi.nlm.nih.gov/genbank/JAKIHL000000000). The version described in this paper is the first version, [JAKIHL010000000](https://www.ncbi.nlm.nih.gov/genbank/JAKIHL010000000). The raw sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA) with accession number [SRX13800198](https://www.ncbi.nlm.nih.gov/sra/SRX13800198). The 16S rRNA gene sequence of *Microaerobacter geothermalis* Nad S1^T has been deposited in NCBI GenBank with accession number [FN552009.1](https://www.ncbi.nlm.nih.gov/genbank/FN552009.1).

ACKNOWLEDGMENTS

This work was supported by Malaysia Fundamental Research Grant Scheme (FRGS) 5F241 and 5F245 with the respective reference codes FRGS/1/2019/STG03/UTM/02/1 and FRGS/1/2019/STG04/UTM/02/4 awarded to K.M.G. and S.S. U.M.K. and I.I.Z. are grateful for funding received from a Malaysia FRGS grant (grant FRGS/1/2020/WAB11/MESTECC/02/1) and a National Institutes of Biotechnology Malaysia (NIBM) Seed Fund Grant 2021 (grant NIBM/SF/2021-13). K.J.L. appreciates the support from Universiti Teknologi Malaysia.

REFERENCES

1. Khelifi N, Romdhane EB, Hedi A, Postec A, Fardeau M-L, Hamdi M, Tholozan J-L, Ollivier B, Hirschler-Réa A. 2010. Characterization of *Microaerobacter geothermalis* gen. nov., sp. nov., a novel microaerophilic, nitrate- and nitrite-reducing thermophilic bacterium isolated from a terrestrial hot spring in Tunisia. *Extremophiles* 14:297–304. <https://doi.org/10.1007/s00792-010-0308-5>.
2. 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>.
3. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung D, Yiu S, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam T, Wang J. 2015. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 4:30. <https://doi.org/10.1186/s13742-015-0069-2>.
4. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. *Curr Protoc Bioinform* 70:e102. <https://doi.org/10.1002/cpbi.102>.
5. Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, Jahesh G, Khan H, Coombe L, Warren RL, Biro I. 2017. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res* 27:768–777. <https://doi.org/10.1101/gr.214346.116>.
6. Lin S-H, Liao Y-C. 2013. CISA: contig integrator for sequence assembly of bacterial genome. *PLoS One* 8:e60843. <https://doi.org/10.1371/journal.pone.0060843>.
7. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki E, Zaslavsky L, Lomsadze A, Pruitt K, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
8. Lee I, Kim YO, Park S-C, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103. <https://doi.org/10.1099/ijsem.0.000760>.
9. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-TK: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>.
10. Owusu-Darko R, Allam M, Mtshali S, Ismail A, Buys EM. 2017. Draft genome sequence of *Bacillus oleronius* DSM 9356 isolated from the termite *Reticulitermes santonensis*. *Genom Data* 12:76–78. <https://doi.org/10.1016/j.gdata.2017.03.005>.
11. Abin CA, Hollibaugh JT. 2016. Draft genome sequence for the type strain *Vulcanibacillus modesticaldus* BR, a strictly anaerobic, moderately thermophilic, and nitrate-reducing bacterium isolated from deep-sea hydrothermal vents of the mid-Atlantic ridge. *Genome Announc* 4:e01246-16. <https://doi.org/10.1128/genomeA.01246-16>.
12. Addou NA, Schumann P, Spröer C, Hania WB, Hacene H, Fauque G, Cayol J-L, Fardeau M-L. 2015. *Melghiribacillus thermohalophilus* gen. nov., sp. nov., a novel filamentous, endospore-forming, thermophilic and halophilic bacterium. *Int J Syst Evol Microbiol* 65:1172–1179. <https://doi.org/10.1099/ijms.0.000075>.
13. Slobodkina GB, Panteleeva AN, Kostrikina NA, Kopitsyn DS, Bonch-Osmolovskaya EA, Slobodkin AI. 2013. *Tepidibacillus fermentans* gen. nov., sp. nov.: a moderately thermophilic anaerobic and microaerophilic bacterium from an underground gas storage. *Extremophiles* 17:833–839. <https://doi.org/10.1007/s00792-013-0564-2>.