



Whole-Genome Sequence of *Geobacillus thermoleovorans* ARTRW1, Isolated from Armutlu Geothermal Spring, Turkey

 Merve Oztug,^{a,b} Anil Cebeci,^{a,c} Hande Mumcu,^a Muslum Akgoz,^b  Nevin Gul Karaguler^a

^aIstanbul Technical University, Faculty of Science and Letters, Department of Molecular Biology and Genetics, Istanbul, Turkey

^bNational Metrology Institute, TUBITAK UME, Kocaeli, Turkey

^cHalic University, Faculty of Science and Letters, Department of Molecular Biology and Genetics, Istanbul, Turkey

ABSTRACT The thermophilic microorganism *Geobacillus thermoleovorans* ARTRW1 was isolated from water samples collected in the Armutlu hot spring in Turkey. Here, the whole-genome sequence and its annotations are reported.

Geobacillus thermoleovorans is a thermophilic Gram-positive bacterium isolated from various environments, including air (1), hot springs (2, 3), a gold mine (4), and sugar refineries (5). The *Geobacillus thermoleovorans* strain ARTRW1 is a rod-shaped, Gram-positive, and aerobic bacterium isolated from water samples collected from the vent of the Armutlu hot spring (74.9°C, pH 7.1). Briefly, the water samples were collected in 50-ml sterile tubes and transferred to the laboratory within 3 h. The GPS coordinates of the sample collection site are 40.545557°N, 28.841777°E. Next, 1 ml of water sample was used to inoculate 50 ml of Luria-Bertani (LB) liquid medium and was incubated at 60°C for 72 h. The culture was streaked onto LB nutrient broth solidified with Gelzan (0.4% by mass) and incubated at 60°C for 24 h. A single colony was picked and purified by repeated streaking (five times) onto fresh culture plates. The strain was grown on the LB liquid medium at 60°C overnight prior to the DNA isolation. The genomic DNA was isolated using the Promega Wizard genomic DNA purification kit according to the manufacturer's protocol. A TruSeq Nano DNA low-throughput library prep kit (Illumina, USA) was used to construct the sequencing libraries. Quality control for the size distribution and quantity of the libraries created was performed using a 2100 bioanalyzer (Agilent Technologies, USA). Synthesis sequencing (SBS) was performed using a HiSeq rapid SBS kit v2 (Illumina) with single fragment readings (2 × 150-bp paired-end [PE] reads) obtained by splicing the two fragments end to end. The cBot v2 system (Illumina) was used to construct clusters by bridge amplification, and the Illumina HiSeq 2500 platform was used for sequencing. A total of 17,383,190 sequences were generated.

The Geneious Prime v2019.1 program was used to control the quality of the sequences. The paired ends were merged using the BBmerge tool, duplicate reads were removed using the Dedupe tool, and the BBNorm tool was used for error correction and normalization. The Geneious assembler method was used to map the reads to the *Geobacillus thermoleovorans* CCB US3 UF5 reference genome obtained from the NCBI, and a consensus sequence was formed. Reads whose quality was lower than 30 were filtered out during the mapping process using the Geneious assembler method. Default parameters were used for all tools. The consensus assembly generated a single sequence of a circular genome with 3,596,341 bp (50-fold coverage). The average chromosome G+C content was 52.27%. From the annotation of *Geobacillus thermoleovorans* strain ARTRW1 using the software PGAP with GeneMarkS, 3,885 coding sequences (CDS), 4,002 genes, and 88 tRNAs were found (6). In addition, 1,381 unidentified hypothetical protein sequences (hypothetical protein CDS) were found in the

Citation Oztug M, Cebeci A, Mumcu H, Akgoz M, Karaguler NG. 2020. Whole-genome sequence of *Geobacillus thermoleovorans* ARTRW1, isolated from Armutlu geothermal spring, Turkey. *Microbiol Resour Announc* 9:e00269-20. <https://doi.org/10.1128/MRA.00269-20>.

Editor Julia A. Maresca, University of Delaware

Copyright © 2020 Oztug 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 Nevin Gul Karaguler, karaguler@itu.edu.tr.

Received 12 March 2020

Accepted 20 May 2020

Published 11 June 2020

5,835 CDS. The average nucleotide identity (ANI) value was calculated using EzBioCloud TruBAC ID software v1 (7), revealing 98.32% similarity to the *Geobacillus thermoleovorans* group. The ANI value with respect to the *Geobacillus thermoleovorans* CCB US3 UF5 strain was calculated to be 99.90% using EzBioCloud ANI calculator v1 (8). The Comprehensive Antibiotic Resistance Database (9) was searched for the presence of antibiotic resistance genes, and no match for resistance genes was found.

Data availability. The genome sequence of *Geobacillus thermoleovorans* ARTRW1 has been deposited at DDBJ/EMBL/GenBank under the accession number [CP042251](https://ncbi.nlm.nih.gov/nuccore/CP042251). The associated BioProject and BioSample accession numbers are [PRJNA556596](https://ncbi.nlm.nih.gov/bioproject/PRJNA556596) and [SAMN12360852](https://ncbi.nlm.nih.gov/biosample/SAMN12360852), respectively. The raw reads were deposited in the Sequence Read Archive under the SRA accession number [SRX7882959](https://ncbi.nlm.nih.gov/sra/SRX7882959).

REFERENCES

- Gaultier NE, Junqueira ACM, Uchida A, Purbojati RW, Houghton JN, Chénard C, Wong A, Kolundžija S, Clare ME, Kushwaha KK, Panicker D, Putra A, Kee C, Premkrishnan BNV, Heinle CE, Lim SBY, Vettath VK, Drautz-Moses DI, Schuster SC. 2018. Genome sequence of *Geobacillus thermoleovorans* SGAir0734, isolated from Singapore air. *Genome Announc* 6:e00636-18. <https://doi.org/10.1128/genomeA.00636-18>.
- Muhd Sakaff MKL, Abdul Rahman AY, Saito JA, Hou S, Alam M. 2012. Complete genome sequence of the thermophilic bacterium *Geobacillus thermoleovorans* CCB_US3_UF5. *J Bacteriol* 194:1239. <https://doi.org/10.1128/JB.06580-11>.
- Bose S, Mukherjee T, Sen U, Roy C, Jameela Rameez M, Ghosh W, Mukhopadhyay SK. 2016. Genome sequence of the multiple-protease-producing strain *Geobacillus thermoleovorans* N7, a thermophilic bacterium isolated from Panipala hot spring, West Bengal, India. *Genome Announc* 4:e01202-16. <https://doi.org/10.1128/genomeA.01202-16>.
- DeFlaun MF, Fredrickson J, Dong H, Pfiffner SM, Onstott TC, Balkwill DL, Streger SH, Stackebrandt E, Knoessen S, van Heerden E. 2007. Isolation and characterization of a *Geobacillus thermoleovorans* strain from an ultra-deep South African gold mine. *Syst Appl Microbiol* 30:152-164. <https://doi.org/10.1016/j.syapm.2006.04.003>.
- Tai S-K, Lin H-PP, Kuo J, Liu J-K. 2004. Isolation and characterization of a cellulolytic *Geobacillus thermoleovorans* T4 strain from sugar refinery wastewater. *Extremophiles* 8:345-349. <https://doi.org/10.1007/s00792-004-0395-2>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res*. 2016;44:6614-6624. <https://doi.org/10.1093/nar/gkw569>.
- Ha S-M, Kim CK, Roh J, Byun J-H, Yang S-J, Choi S-B, Chun J, Yong D. 2019. Application of the whole genome-based bacterial identification system, TrueBac ID, using clinical isolates that were not identified with three matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) systems. *Ann Lab Med* 39:530-536. <https://doi.org/10.3343/alm.2019.39.6.530>.
- Yoon S-M, Ha J, Lim S, Kwon J, Chun S-H. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281-1286. <https://doi.org/10.1007/s10482-017-0844-4>.
- Alcock BP, Raphenya AR, Lau TT, Tsang KK, Megane B, Edalatmand A, Huynh W, Nguyen A-LV, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran H-K, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FS, Hsiao WW, Domselaar GV, McArthur AG. 2019. CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 48:D517-D525. <https://doi.org/10.1093/nar/gkz935>.