AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Geotoga petraea* Strain HO-Geo1, Isolated from a Petroleum Reservoir in Russia

Denis S. Grouzdev,[a] Ekaterina M. Semenova,[b] Diyana S. Sokolova,[b] Tatiyana P. Tourova,[b] Andrey B. Poltaraus,[c] Tamara N. Nazina[b]

[a]Institute of Bioengineering, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, Russian Federation
[b]Winogradsky Institute of Microbiology, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, Russian Federation
[c]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russian Federation

ABSTRACT The draft genome sequence of *Geotoga petraea* strain HO-Geo1, a bacterium isolated from production water of the Vostochno-Anzirskoe petroleum reservoir in Russia, is presented. The genome of strain HO-Geo1 is annotated for elucidation of the metabolic potential and its possible function in the subsurface microbial community and biotechnological application.

At the time this manuscript was prepared, only two *Geotoga* species were known, *G. petraea* and *G. subterranea* (1, 2). Members of the genus *Geotoga* were isolated from oil deposits and are anaerobes fermenting carbohydrates and proteinaceous compounds, are capable of reducing elemental sulfur to hydrogen sulfide, and cause pipeline corrosion and oil souring (1). The genome of the *G. petraea* type strain is presently unavailable, and no data exist on the intraspecies phenotypic diversity, which made it necessary to investigate the new *G. petraea* HO-Geo1 strain and to sequence its genome.

Strain HO-Geo1 (VKM B-3300) was isolated from the Vostochno-Anzirskoe petroleum reservoir (Russia) (3) by sequential transfers from the highest dilutions on RM liquid medium containing (in grams per liter) NaCl, 20.0; $MgCl_2 \cdot 2H_2O$, 18.0; Na_2SO_4, 3.0; $CaCl_2 \cdot 6H_2O$, 1.5; KCl, 0.5; NH_4Cl, 0.33; $NaHCO_3$, 0.2; yeast extract, 1.0; peptone, 5.0; ferric citrate, 0.02; and $Na_2S \cdot 9H_2O$, 0.2 (pH 6.8 to 7.0) at 48°C. When examined on an Axio Imager D1 epifluorescence microscope (Carl Zeiss, Germany), the cells of strain HO-Geo1 were motile non-spore-forming rods with a sheath-like outer structure specific to members of the order *Thermotogales*. The strain was grown in the RM medium for 7 days at 48°C under anaerobic conditions. DNA was purified from the cell biomass using the cetyltrimethylammonium bromide (CTAB) method (4). The 16S rRNA gene was amplified with 27F and 1492R primers (5), and purified PCR products were sequenced with an ABI Prism 3730 DNA analyzer (Applied Biosystems, USA). The 16S rRNA sequence analysis using a BLASTn (6) search against the NCBI database revealed that HO-Geo1 shares 99.2% similarity with *Geotoga petraea* T5^T (GenBank accession no. NR_104910).

DNA libraries were prepared with the NEBNext DNA library prep kit for Illumina (New England BioLabs). Next-generation shotgun sequencing of the genomic DNA was carried out using the HiSeq 1500 platform (Illumina, Inc., USA). A total of 1,946,596 250-bp single-end reads were obtained from strain HO-Geo1. Low-quality reads were trimmed using Trimmomatic version 0.36 (7) with the default settings for single-end reads. Subsequently, the quality-filtered reads were *de novo* assembled with SPAdes version 3.12.0 using the default settings (8). The final draft genome assembly of HO-Geo1 contained 22 contigs, covering a total of 2,150,220 bp, with an N_{50} value of 331,226 bp, L_{50} of 2, G+C content of 29.38%, and average sequence coverage of 140×.

Citation Grouzdev DS, Semenova EM, Sokolova DS, Tourova TP, Poltaraus AB, Nazina TN. 2019. Draft genome sequence of *Geotoga petraea* strain HO-Geo1, isolated from a petroleum reservoir in Russia. Microbiol Resour Announc 8:e00706-19. https://doi.org/10.1128/MRA.00706-19.

Editor J. Cameron Thrash, University of Southern California

Address correspondence to Tamara N. Nazina, nazina@inmi.ru.

Received 12 June 2019
Accepted 28 June 2019
Published 18 July 2019

Identification of the protein-coding sequences and primary annotation were performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP; version 4.7) (9), which identified 2,058 genes, 1,997 protein-coding sequences, 9 pseudogenes, and 52 RNA genes.

Functional analysis performed with the Rapid Annotations using Subsystems Technology (RAST; version 2.0) server (10) via the RAST*tk* pipeline with the default settings (11) showed that 149 genes were associated with protein metabolism, 114 genes were associated with the metabolism of amino acids and their derivatives, 100 genes were associated with carbohydrate metabolism, and 49 genes were associated with the metabolism of cofactors, vitamins, prosthetic groups, and pigments.

The genome sequence provided here is expected to broaden our knowledge regarding the genetic and functional characteristics of the genus *Geotoga*.

Data availability. The GenBank/EMBL/DDBJ accession number for the 16S rRNA gene sequence of strain HO-Geo1 is MK984240. This whole-genome shotgun project was deposited at DDBJ/ENA/GenBank under the accession no. SRME00000000. The version described in this paper is version SRME01000000. The associated BioProject, BioSample, and SRA accession numbers are PRJNA530105, SAMN11295181, and SRR8846986, respectively.

ACKNOWLEDGMENTS

Work on genome sequencing of the strain was supported by the Russian Science Foundation (RSF) through grant 16-14-00028. Physiological and taxonomic studies of the strain were supported by the Ministry of Science and Higher Education of the Russian Federation. The funders had no role in the study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. Davey ME, Wood WA, Key R, Nakamura K, Stahl D. 1993. Isolation of three species of *Geotoga* and *Petrotoga*: two new genera, representing a new lineage in the bacterial line of descent distantly related to the "*Thermotogales*." Syst Appl Microbiol 16:191–200. https://doi.org/10.1016/S0723-2020(11)80467-4.
2. Davey ME, Wood WA, Key R, Nakamura K, Stahl D. 1993. Validation of the publication of new names and new combinations previously effectively published outside of the IJSB, list no. 47. Int J Syst Bacteriol 43:864–865.
3. Nazina TN, Sokolova DSh, Babich TL, Semenova EM, Ershov AP, Bidzhieva SKh, Borzenkov IA, Poltaraus AB, Khisametdinov MR, Tourova TP. 2017. Microorganisms of low-temperature heavy oil reservoirs (Russia) and their possible application for enhanced oil recovery. Microbiology 86:773–785. https://doi.org/10.1134/S0026261717060121.
4. Wilson K. 2001. Preparation of genomic DNA from bacteria. Curr Protoc Mol Biol Chapter 2:Unit 2.4. https://doi.org/10.1002/0471142727.mb0204s56.
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. 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.
7. 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.
8. 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.
9. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W. 2013. Prokaryotic Genome Annotation Pipeline. The NCBI handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.
10. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
11. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.