

Draft Genome Sequence of *Geobacillus subterraneus* Strain K, a Hydrocarbon-Oxidizing Thermophilic Bacterium Isolated from a Petroleum Reservoir in Kazakhstan

Andrey B. Poltarau, ^a Diyana S. Sokolova, ^b Denis S. Grouzdev, ^b Timopheyy M. Ivanov, ^b Sophia G. Malakho, ^c Alena V. Korshunova, ^b Tatiyana P. Tourova, ^b Tamara N. Nazina ^b

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russian Federation^a; Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russian Federation^b; Federal State Budgetary Institution, "Federal Science Center for Physical Culture and Sport," Moscow, Russian Federation^c

The draft genome sequence of *Geobacillus subterraneus* strain K, a thermophilic aerobic oil-oxidizing bacterium isolated from production water of the Uzen high-temperature oil field in Kazakhstan, is presented here. The genome is annotated for elucidation of the genomic and phenotypic diversity of thermophilic alkane-oxidizing bacteria.

Received 9 June 2016 Accepted 10 June 2016 Published 4 August 2016

Citation Poltarau AB, Sokolova DS, Grouzdev DS, Ivanov TM, Malakho SG, Korshunova AV, Tourova TP, Nazina TN. 2016. Draft genome sequence of *Geobacillus subterraneus* strain K, a hydrocarbon-oxidizing thermophilic bacterium isolated from a petroleum reservoir in Kazakhstan. *Genome Announc* 4(4):e00782-16. doi:10.1128/genomeA.00782-16.

Copyright © 2016 Poltarau 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 Tamara N. Nazina, nazinatn@gmail.com.

Geobacillus subterraneus strain K (=VKM B-2225) was isolated from a formation water sample of the Uzen high-temperature oil field (Kazakhstan) (44°44'33.3" N, 52°78'22.2" E) and was described as a reference strain of the species (1). *G. subterraneus* strain K is an aerobic endospore-forming bacterium able to grow at temperatures ranging from 48 to 70°C with an optimum at 60°C and utilizing fatty acids, sugars, amino acids, aromatic compounds, and *n*-alkanes of crude oil. In the genome of strain K growing on C₁₆ and C₂₂ *n*-alkanes, a range of *alkB* gene homologs was revealed by PCR amplification (2). The *alkB* and *ladAB* genes coding the monooxygenases associated with *n*-alkane oxidation have been detected in several thermophilic bacilli (3–5).

Genomic DNA was isolated from the biomass using the DNeasy Blood & Tissue Kit (Qiagen, Germany) according to the manufacturer's instructions. The TruSeq DNA Sample Preparation Kit (Illumina, United States) was used to create the libraries for genome sequencing. Genomic DNA was sequenced using the MiSeq Reagent Kit version 2 (Illumina). The shotgun library was constructed with a 500 bp-span paired-end library.

Approximately 11,888,884 paired-end reads were generated providing 250-fold genome coverage. The resulting reads were assembled with SPAdes version 3.1.0 (6). The remaining 158 contigs were submitted to the ProDeGe website (7) for automatic decontamination. The assembled sequence was submitted to the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP) (8) for annotation.

The draft genome sequence of *G. subterraneus* strain K revealed a genome size of 3,309,142 bp, with an average G+C content of 51.43%, slightly exceeding the 49.7% value determined earlier (1). The genome contained 3,275 genes, of which 3,058 were coding DNA sequences, 18 coded rRNAs (5S, 16S, and 23S), 70 coded tRNAs, and five belonged to noncoding RNAs (ncRNAs). There were no complete genes coding 16S rRNA, but the *gyrB* and *parE* gene sequences revealed in the genome had 99.9% and 99.6%

similarity, respectively, with those determined earlier for the strain (9). The results of phenotypic studies characterizing the strain as an aerobic spore-forming bacterium utilizing a wide range of organic substrates were confirmed by the detection of respective genes. Despite utilization of a range of *n*-alkanes by strain K, neither *alkB* nor *ladA* genes were found in the draft genome. A range of protein-coding sequences involved in the β -oxidation pathway was detected in the genome of the strain K. Among them were alcohol dehydrogenases, aldehyde dehydrogenase, long-chain fatty acid-CoA ligases, acyl-CoA dehydrogenase, enoyl-CoA hydrogenase, hydroxyacyl-CoA dehydrogenase, and thiolase. Further studies of the genes of alkane biodegradation by aerobic thermophilic bacteria are necessary.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. [LWHE000000000](https://www.ncbi.nlm.nih.gov/nuccore/LWHE000000000). The version described in this paper is version LWHE01000000.

ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation (grant 16-14-00028) and the Russian Foundation for Basic Research (grant 15-04-02622). This work, including the efforts of T.N.N., was funded by the Russian Science Foundation (RSF) (16-14-00028); the efforts of A.B.P. were funded by the Russian Foundation for Basic Research (RFBR) (15-04-02622).

FUNDING INFORMATION

This work, including the efforts of Andrey B. Poltarau, was funded by Russian Foundation for Basic Research (RFBR) (15-04-02622). This work, including the efforts of Diyana Sh. Sokolova and Tamara N. Nazina, was funded by Russian Science Foundation (RSF) (16-14-00028).

REFERENCES

- Nazina TN, Tourova TP, Poltarau AB, Novikova EV, Grigoryan AA, Ivanova AE, Lysenko AM, Petrunyaka VV, Osipov GA, Belyaev SS,

- Ivanov MV. 2001. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. thermocatenulatus*, *G. thermoleovorans*, *G. kaustophilus*, *G. thermoglucosidasius* and *G. thermodenitrificans*. *Int J Syst Evol Microbiol* 51:433–446. <http://dx.doi.org/10.1099/00207713-51-2-433>.
2. Korshunova AV, Tourova TP, Shestakova NM, Mikhailova EM, Poltaraus AB, Nazina TN. 2011. Detection and transcription of *n*-alkane biodegradation genes (*alkB*) in the genome of a hydrocarbon-oxidizing bacterium *Geobacillus subterraneus* K. *Microbiology* 80:682–691. <http://dx.doi.org/10.1134/S0026261711050079>.
 3. Tourova TP, Nazina TN, Mikhailova EM, Rodionova TA, Ekimov AN, Mashukova AV, Poltaraus AB. 2008. *alkB* homologs in thermophilic bacteria of the genus *Geobacillus*. *Mol Biol* 42:217–226. <http://dx.doi.org/10.1134/S0026893308020076>.
 4. Liu X, Dong Y, Zhang J, Zhang A, Wang L, Feng L. 2009. Two novel metal-independent long-chain alkyl alcohol dehydrogenases from *Geobacillus thermodenitrificans* NG80-2. *Microbiology* 155:2078–2085. <http://dx.doi.org/10.1016/j.resmic.2009.08.010>.
 5. Boonmak C, Takahashi Y, Morikawa M. 2014. Cloning and expression of three *ladA*-type alkane monooxygenase genes from an extremely thermophilic alkane-degrading bacterium *Geobacillus thermoleovorans* B23. *Extremophiles* 18:515–523. <http://dx.doi.org/10.1007/s00792-014-0636-y>.
 6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
 7. Tennessen K, Andersen E, Clingenpeel S, Rinke C, Lundberg DS, Han J, Danyl JL, Ivanova N, Woyke T, Kyrpides N, Pati A. 2016. ProDeGe: a computational protocol for fully automated decontamination of genomes. *ISME J* 10:269–272. <http://dx.doi.org/10.1038/ismej.2015.100>.
 8. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufu S, Li W. 2013. *Prokaryotic genome annotation pipeline*. In *The NCBI Handbook*, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.
 9. Tourova TP, Korshunova AV, Mikhailova EM, Sokolova DS, Poltaraus AB, Nazina TN. 2010. Application of *gyrB* and *parE* sequence similarity analyses for differentiation of species within the genus *Geobacillus*. *Microbiology* 79:356–369. <http://dx.doi.org/10.1134/S0026261710030124>.