




# Genome Sequence of *Geobacillus thermoleovorans* SGAir0734, Isolated from Singapore Air

Nicolas E. Gaultier,<sup>a</sup>  Ana Carolina M. Junqueira,<sup>a\*</sup> Akira Uchida,<sup>a</sup> Rikky W. Purbojati,<sup>a</sup> James N. I. Houghton,<sup>a</sup> Caroline Chénard,<sup>b</sup> Anthony Wong,<sup>a</sup> Sandra Kolundžija,<sup>a</sup> Megan E. Clare,<sup>a</sup> Kavita K. Kushwaha,<sup>a</sup> Deepa Panicker,<sup>a</sup> Alexander Putra,<sup>a</sup> Carmon Kee,<sup>a</sup> Balakrishnan N. V. Premkrishnan,<sup>a</sup> Cassie E. Heinle,<sup>a</sup> Serene B. Y. Lim,<sup>a</sup> Vineeth Kodengil Vettath,<sup>a</sup> Daniela I. Drautz-Moses,<sup>a</sup> Stephan C. Schuster<sup>a</sup>

<sup>a</sup>Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore

<sup>b</sup>Asian School of the Environment, Nanyang Technological University, Singapore

**ABSTRACT** The thermophilic bacterium *Geobacillus thermoleovorans* was isolated from a tropical air sample collected in Singapore. The genome was sequenced on the PacBio RS II platform and consists of one chromosome with 3.6 Mb and one plasmid with 75 kb. The genome comprises 3,509 protein-coding genes, 88 tRNAs, and 27 rRNAs.

*Geobacillus thermoleovorans* is a rod-shaped Gram-positive thermophilic bacterium (1) classified in the phylum *Firmicutes*. It was first reported as strain LEH-1 in a study on heat sensitivity of thermophilic bacteria (2) and later designated *Bacillus thermoleovorans* (3). *Geobacillus thermoleovorans* is found in a variety of environments, including gold mines (4), sugar refineries (5), and hot springs (6). This organism's thermophilic characteristic makes it interesting for potential industrial applications (7).

SGAir0734 was isolated from an air sample collected in Singapore (global position system coordinates 1.345771°N, 103.6801°E) using the Spin Air air sampler (IUL, Spain). The airborne particles were impacted onto Reasoner's 2A agar (Becton, Dickinson). Isolation of colonies was carried out by culturing on Trypticase soy agar (Becton, Dickinson) at 50°C and in Luria broth overnight at 30°C prior to DNA extraction. Genomic DNA was purified using the Wizard genomic DNA purification kit (Promega, USA) according to the manufacturer's protocol. Library preparation was performed with the SMRTbell template prep kit 1.0 (Pacific Biosciences), followed by single-molecule real-time (SMRT) sequencing on the PacBio RS II platform.

Sequencing resulted in a total of 55,892 subreads. The reads were subsequently used for *de novo* assembly with Hierarchical Genome Assembly Process (HGAP) version 3 (8), which is included in the PacBio SMRT Analysis 2.3.0 package. In order to polish and correct the errors of the assembly, Quiver (8) and Pilon version 1.16 (9) were used. The consensus assembly generated two contigs, including one chromosome with 3,615,819 bp (146-fold coverage) and one plasmid, unnamed\_3, with 75,441 bp (259-fold coverage). Both contigs were unable to be circularized using Circlator (10). The average chromosome G+C content was 52.19%.

Using the average nucleotide identity (ANI) method with Microbial Species Identifier (MiSI) (11), taxonomical identification revealed a 99.0% marker similarity with the available reference genome of *Geobacillus thermoleovorans*.

Annotation was completed using NCBI's Prokaryotic Genome Annotation Pipeline (PGAP) version 4.2 (12). A total of 3,850 genes were predicted, including 3,509 protein-coding genes (PCGs), 27 rRNA subunits (9 genes for each of the 5S, 16S, and 23S subunits), 88 tRNAs, 5 noncoding RNAs, and 221 pseudogenes. Functional annotation with Rapid Annotations using Subsystems Technology (RAST) (13–15) showed that 116 genes were associated with dormancy and sporulation, which could indicate the ability to enter long-term survival states in hostile environments. Seventeen genes were found

Received 2 June 2018 Accepted 7 June 2018 Published 5 July 2018

**Citation** 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 Announcements* 6:e00636-18. <https://doi.org/10.1128/genomeA.00636-18>.

**Copyright** © 2018 Gaultier 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 Stephan C. Schuster, [SCSchuster@ntu.edu.sg](mailto:SCSchuster@ntu.edu.sg).

\* Present address: Ana Carolina M. Junqueira, Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

N.E.G. and A.C.M.J. contributed equally to this article.

to be potentially involved in stress response, specifically heat shock, which may contribute to the thermophilicity of this organism.

**Accession number(s).** The genome and plasmid sequences of *Geobacillus thermoleovorans* SGAir0734 are available in DDBJ/EMBL/GenBank under accession numbers CP027303 and CP027305, respectively.

## ACKNOWLEDGMENT

This work was supported by a Singapore Ministry of Education Academic Research Fund tier 3 grant (MOE2013-T3-1-013).

## 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. <https://doi.org/10.1099/00207713-51-2-433>.
- Merkel GJ, Perry JJ. 1977. Effect of growth substrate on thermal death of thermophilic bacteria. *Appl Environ Microbiol* 34:626–629. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC242721/>.
- Zarilla KA, Perry JJ. 1987. *Bacillus thermoleovorans*, sp. nov., a species of obligately thermophilic hydrocarbon utilizing endospore-forming bacteria. *Syst Appl Microbiol* 9:258–264. [https://doi.org/10.1016/S0723-2020\(87\)80031-0](https://doi.org/10.1016/S0723-2020(87)80031-0).
- DeFlaun MF, Fredrickson JK, 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. <https://doi.org/10.1007/s00792-004-0395-2>.
- 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>.
- McMullan G, Christie JM, Rahman TJ, Banat IM, Ternan NG, Marchant R. 2004. Habitat, applications and genomics of the aerobic, thermophilic genus *Geobacillus*. *Biochem Soc Trans* 32:214–217. <https://doi.org/10.1042/bst0320214>.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Hunt M, De Silva N, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <https://doi.org/10.1186/s13059-015-0849-0>.
- Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, Pati A. 2015. Microbial species delineation using whole genome sequences. *Nucleic Acids Res* 43:6761–6771. <https://doi.org/10.1093/nar/gkv657>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- 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, Arrello 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>.
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
- Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, 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>.