



# Complete Genome Sequence of *Geobacillus* sp. Strain E55-1, Isolated from Mine Geyser in Japan

 Kentaro Miyazaki,<sup>a,b</sup> Erina Hase,<sup>a</sup> Natsuko Tokito<sup>a,b</sup>

<sup>a</sup>Department of Life Science and Biotechnology, Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan

<sup>b</sup>Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan

**ABSTRACT** We report here the complete genome sequence of *Geobacillus* sp. strain E55-1, isolated from the hot sediments of Mine Geyser in Japan. This strain exhibited ~85% average nucleotide identity and ~98.5% 16S rRNA sequence identity with the most closely related *Geobacillus* species.

The genus *Geobacillus* comprises a group of Gram-positive thermophilic bacterial species (1) that were originally classified as group 5 within the genus *Bacillus* but were later reclassified as a new genus, *Geobacillus* (2). Researchers are interested in *Geobacillus* from a biotechnological point of view, particularly as a source of thermophilic enzymes (3).

We collected hot sediment samples at Mine Geyser in Japan (4, 5), which were spread over Lennox-LB agar (1.5% [wt/vol]) plates. After overnight incubation at 55°C, several single colonies were isolated. Colony PCR was conducted to amplify the nearly full-length 16S rRNA gene using a set of primers, Bac8f(C) and UN1541r(U) (6). DNA sequencing analysis suggested that many of them were affiliated with one of the following thermophilic bacilli: *Bacillus*, *Geobacillus*, or *Anoxybacillus*. One of the strains, designated E55-1, showed 98.5% identity with the 16S rRNA genes of a *Geobacillus* species; this strain was subjected to whole-genome analysis by combining Oxford Nanopore Technologies (ONT) and Illumina sequencing technologies. All software analyses were implemented with default settings.

Cells were grown in Lennox-LB broth at 55°C for 18 h. Genomic DNA was extracted from pelleted cells using lysozyme, proteinase K, and achromopeptidase for enhanced cell lysis efficiency (7). For long-read sequencing, genomic DNA was treated with the Short Read Eliminator XS kit (Circulomics). The resulting DNA was used to construct a library using a ligation sequencing kit (SQK-LSK109; ONT). Sequencing was performed using a GridION X5 system with a FLO-MIN106 R9.41 flow cell (ONT) for 6 h. Base calling was performed using Guppy v.3.0.3 (ONT) to generate 68,238 reads corresponding to 781 Mb of genome with an average length of 11,447 bases. The raw reads were filtered (quality [Q],  $\geq 10$ ; read length,  $\geq 1,000$  bases) using NanoFilt v.2.3.0 (8). The longest read length was 198,777 bases.

For short-read sequencing, a library with ~350-bp inserts was generated using the Nextera DNA Flex library prep kit (Illumina), which was subjected to 156-bp paired-end sequencing on the Illumina MiSeq platform. Adapter sequences and low-quality data were trimmed (Q,  $\geq 30$ ; read length,  $\geq 10$  bases) using fastp v.0.20.0 (9), yielding 1.03 million paired-end reads, spanning 308 Mb with an average length of 149 bases.

The trimmed long- and short-read data were assembled using Unicycler v.0.4.8 (10) and polished with Pilon v.1.23 (11), generating a single circular chromosome (3,754,053 bp, 51.0 mol% G+C content) and two circular putative plasmids, pGspE55-1 (126,981 bp, 45.6 mol% G+C content) and pGspE55-2 (43,187 bp, 43.9 mol% G+C content). Coverages

**Citation** Miyazaki K, Hase E, Tokito N. 2020. Complete genome sequence of *Geobacillus* sp. strain E55-1, isolated from Mine Geyser in Japan. *Microbiol Resour Announc* 9:e00339-20. <https://doi.org/10.1128/MRA.00339-20>.

**Editor** John J. Dennehy, Queens College

**Copyright** © 2020 Miyazaki 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 Kentaro Miyazaki, [miyazaki-kentaro@aist.go.jp](mailto:miyazaki-kentaro@aist.go.jp).

**Received** 1 April 2020

**Accepted** 24 April 2020

**Published** 14 May 2020

of short reads to the chromosome (72.9×) and plasmid (155.7× for pGspE55-1 and 274.4× for pGspE55-2) sequences suggested a copy number ratio of 1:2:4 in this case.

Automated annotation was performed using DFAST v.1.2.4 (12). The chromosome contained 3,697 coding sequences, 89 tRNA genes, and 26 rRNA genes, while pGspE55-1 and pGspE55-2 contained 134 and 65 coding sequences, respectively. A Web-based JSpeciesWS search (13) revealed that E55-1 showed the highest ~85% average nucleotide identities with *G. subterraneus* KCTC 3922 (GenBank accession number [NZ\\_CP014342](#); 14) and *G. kaustophilus* HTA426 ([NC\\_006510](#); 15), values that were below the cutoff (95%) for novel species (16).

**Data availability.** The complete genome sequences are available from DDBJ/EMBL/GenBank under the accession numbers [AP022557](#) (chromosome), [AP022558](#) (pGspE55-1), and [AP022559](#) (pGspE55-2). Raw sequencing data were deposited in the DDBJ/SRA database under the accession numbers [DRX197760](#) (Illumina MiSeq) and [DRX197759](#) (GridION) (BioProject accession number, [PRJDB9279](#); BioSample accession number, [SAM00204259](#)).

## ACKNOWLEDGMENT

This work was partly supported by two grants awarded to K.M. from the Japan Society for the Promotion of Science (JSPS), a Grant-in-Aid for Scientific Research (A; 19H00936) and a Grant-in-Aid for Challenging Research (Pioneering; 19H05538).

## REFERENCES

- Hussein AH, Lisowska BK, Leak DJ. 2015. The genus *Geobacillus* and their biotechnological potential. *Adv Appl Microbiol* 92:1–48. <https://doi.org/10.1016/bs.aambs.2015.03.001>.
- 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 uzensus* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenuatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. thermocatenuatus*, *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>.
- Kananaviciute R, Citavicius D. 2015. Genetic engineering of *Geobacillus* spp. *J Microbiol Methods* 111:31–39. <https://doi.org/10.1016/j.mimet.2015.02.002>.
- Oshima T, Imahori K. 1974. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa. *Int J Syst Evol Microbiol* 24:102–112. <https://doi.org/10.1099/00207713-24-1-102>.
- Miyazaki K. 2019. Complete genome sequencing of *Thermus thermophilus* strain HC11, isolated from Mine Geyser in Japan. *Microbiol Resour Announc* 8:e00873-19. <https://doi.org/10.1128/MRA.00873-19>.
- Miyazaki K, Sato M, Tsukuda M. 2017. PCR primer design for 16S rRNAs for experimental horizontal gene transfer test in *Escherichia coli*. *Front Bioeng Biotechnol* 5:14. <https://doi.org/10.3389/fbioe.2017.00014>.
- Morita H, Kuwahara T, Ohshima K, Sasamoto H, Itoh K, Hattori M, Hayashi T, Takami H. 2007. An improved DNA isolation method for metagenomic analysis of the microbial flora of the human intestine. *Microb Environ* 22:214–222. <https://doi.org/10.1264/jsm.22.214>.
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- 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>.
- Tanizawa Y, Fujisawa T, Kaminuma E, Nakamura Y, Arita M. 2016. DFAST and DAGA: Web-based integrated genome annotation tools and resources. *Biosci Microbiota Food Health* 35:173–184. <https://doi.org/10.12938/bmfh.16-003>.
- Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a Web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32:929–931. <https://doi.org/10.1093/bioinformatics/btv681>.
- Lee Y-J, Park M-K, Park G-S, Lee S-J, Lee SJ, Kim BS, Shin J-H, Lee D-W. 2017. Complete genome sequence of the thermophilic bacterium *Geobacillus subterraneus* KCTC 3922<sup>T</sup> as a potential denitrifier. *J Biotechnol* 251:141–144. <https://doi.org/10.1016/j.jbiotec.2017.04.022>.
- Takami H, Nishi S, Lu J, Shimamura S, Takaki Y. 2004. Genomic characterization of thermophilic *Geobacillus* species isolated from the deepest sea mud of the Mariana Trench. *Extremophiles* 8:351–356. <https://doi.org/10.1007/s00792-004-0394-3>.
- Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 106:19126–19131. <https://doi.org/10.1073/pnas.0906412106>.