AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Applied and Industrial Microbiology | Announcement

Complete genome sequence of *Bacillus subtilis* GL-4

Shaomin Qin,[1] Wenting Zeng,[1,2] Jue Wei,[1] Fenglian Chen,[1] Ling Ma,[1] Shuying Qin,[1] Jun Lin,[1] Lishi Xu,[1] Peng Zhu,[3] Jinfeng Liu[1]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT A cellulase-producing *Bacillus* was previously isolated from the intestinal tract of domestic bamboo rat and identified as *Bacillus subtilis* GL-4. In this study, we present the complete genome sequence of *B. subtilis* GL-4. The genome is 4,271,214 bp long, with a guanine-cytosine content of 43.45%.

KEYWORDS applied microbiology

Previously, cellulase-producing *Bacillus subtilis* GL-4, isolated on de Man-Rogosa-Sharpe (MRS) medium (Beijing Luqiao Microbial Technology, Beijing, China) at 37°C for 24 h from the intestinal tract of domestic bamboo rat, was preliminarily screened by Congo red plate and identified as *B. subtilis* via physiological and biochemical tests and 16S rDNA PCR method (1). The endoglucanase and filter paper enzyme activities of this strain were 49.32 and 25.13 U/mL, respectively (1). In our previous study, we reported that *B. subtilis* GL-4 exhibits good tolerance to low pH (pH 2.5 and 3.5) and 0.50% bile salts and that it is a potential probiotic candidate strain (2). In the present study, we revealed the complete genome sequence of *B. subtilis* GL-4. For this, the *B. subtilis* GL-4 strain was cultured in MRS medium at 37°C for 24 h. TaKaRa MiniBEST Bacteria Genomic DNA Extraction Kit (v.3.0) was used according to the manufacturer's instructions to extract the total genomic DNA of *B. subtilis* GL-4 from the collected bacterial culture fluid.

The PacBio RS II and Illumina Hiseq 4000 sequencing platforms were utilized to sequence the complete genome of *B. subtilis* GL-4 at Beijing Genomics Institute (Shenzhen, China). The libraries for PacBio sequencing and Illumina sequencing were prepared using the SMRTbell Express Template Prep Kit (v.2.0; PacBio, USA) and Nextera XT DNA Library Preparation Kit (Illumina, USA), respectively, according to the manufacturer's instructions. For the PacBio platform, genomic DNA was sheared using Covaris g-TUBE; sheared DNA was purified using 0.45× AMPure PB beads; four SMRT cells with zero-mode waveguide sequencing arrays were used to generate the subread set. Through PacBio sequencing, 697,148 subreads (8,065,230,617 bases, with 1,888× coverage and a mean subread length of 11,568 bp) were obtained. In contrast, 8,764,210 reads (average insert size 350 bp, with 303× coverage) were obtained from Illumina sequencing. PacBio subreads (length <1 kb) were removed. Quality control and adapter trimming were performed using FastQC (v.0.73) and Trimmomatic (v.0.38.1) (3). Self-correction was performed using Falcon (v.0.3.0) and Proovread (v.2.12) (4). The Celera Assembler (v.8.3) (5) was used to assemble draft genomic unitigs, which are uncontested fragment groups, against a high-quality corrected circular consensus sequence subread set. To improve the accuracy of the genome sequences, single-base corrections were made using GATK (v.1.6–13) (5). To identify plasmid presence, SOAP (v.2) (6) was used to map the filtered reads to a bacterial plasmid database (http://www.ebi.ac.uk/genomes/plasmid.html). Gene prediction and genome annotation were performed using Glimmer (v.3.02) (7) and Diamond (v.3.12) (8), respectively. Default parameters were used except where otherwise noted.

The complete genome sequence of *B. subtilis* GL-4 was noted to be 4,271,214 bp long, with a total guanine-cytosine content of 43.45%. Plasmid was not detected in

Editor Irene L. G. Newton, Indiana University Bloomington, Bloomington, Indiana, USA

Address correspondence to Jinfeng Liu, fjinliu@163.com.

Shaomin Qin and Wenting Zeng contributed equally to this article. Author order was determined both alphabetically and in order of decreasing seniority.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 5 June 2024
Accepted 20 December 2024
Published 15 January 2025

the genome. In total, 4,253 protein-coding sequences, 30 rRNAs, and 87 tRNAs were identified on the chromosome, which were annotated. Genome sequence analysis revealed that *B. subtilis* GL-4 comprises cellulose degradation-associated genes, including three endo-β-1,4-glucanase (EC 3.2.1.4) and five β-glucosidase (EC 3.2.1.21) genes. Furthermore, the genome comprises genes associated with the degradation of other nondigestible carbohydrates, including mannan and xylan.

ACKNOWLEDGMENTS

This study was supported by grants from the Guangxi Natural Science Foundation (2022GXNSFAA035521), Guangxi Science and Technology Major Project (Guike AA 22068099), Special Fund for Basic Scientific Research of Guangxi (Guike 22–7), and Independent Research Project Foundation of Guangxi Key Laboratory of Veterinary Biotechnology (23–035-32-A-02)

AUTHOR AFFILIATIONS

[1]Guangxi Key Laboratory of Veterinary Biotechnology, Key Laboratory of China (Guangxi)-ASEAN Cross-border Animal Disease Prevention and Control, Ministry of Agriculture and Rural Affairs of China, Guangxi Veterinary Research Institute, Nanning, China
[2]Anzhou District Agriculture and Rural Bureau, Mianyang, China
[3]Guangxi Key Laboratory of Beibu Gulf Marine Biodiversity Conservation, Beibu Gulf University, Qinzhou, China

AUTHOR ORCIDs

Shaomin Qin http://orcid.org/0000-0002-1565-0191

FUNDING

Funder	Grant(s)	Author(s)
Guangxi Natural Science Foundation	2022GXNSFAA035521	Shaomin Qin
Guangxi Science and Technology Major Project	Guike AA 22068099	Peng Zhu
Special Fund for Basic Scientific Research of Guangxi	Guike 22-7	Shaomin Qin

DATA AVAILABILITY

The complete sequence of *Bacillus subtilis* GL-4 has been submitted to GenBank and deposited under accession number CP104097. The BioProject and BioSample accession numbers are PRJNA877089 and SAMN30680087, respectively. The raw sequence reads are available under Sequence Read Archive accession numbers SRR29245815 (Illumina) and SRR29261860 (PacBio).

REFERENCES

1. Zeng WT, Qin SM, Wu JM, Liang XL, Bai AB, Liu JF, Chen FL, Qin SY, Ma L. 2021. Isolation, identification and enzyme production characteristics of cellulolytic bacterium from intestine of bamboo rats. Chin J Anim Nutr 33:4142–4152. https://doi.org/10.3969/j.issn.1006-267x.2021.07.055
2. Chentong JY, Qin SM, Wei J, Liu JF, Chen FL, Xu LS, Liang J, Lan GQ, Wu JM. 2024. Study on biological characteristics and preliminary safety evaluation of *Bacillus subtilis* strains GL-4, GL-5 and GL-8 producing cellulases, isolated from bamboo rats. Chin J Anim Sci 60:201–206. https://doi.org/10.19556/j.0258-7033.20230315-02
3. 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
4. Faino L, Seidl MF, Datema E, van den Berg GCM, Janssen A, Wittenberg AHJ, Thomma B. 2015. Single-molecule real-time sequencing combined with optical mapping yields completely finished fungal genome. MBio 6:e00936-15. https://doi.org/10.1128/mBio.00936-15
5. Tsuji M, Kudoh S, Hoshino T. 2015. Draft genome sequence of cryophilic basidiomycetous yeast *Mrakia blollopis* SK-4, isolated from an algal mat of naga-ike lake in the skarvsnes ice-free area, East Antarctica. Genome Announc 3:e01454-14. https://doi.org/10.1128/genomeA.01454-14
6. Badouin H, Hood ME, Gouzy J, Aguileta G, Siguenza S, Perlin MH, Cuomo CA, Fairhead C, Branca A, Giraud T. 2015. Chaos of rearrangements in the mating-type chromosomes of the anther-smut fungus *Microbotryum lychnidis-dioicae*. Genetics 200:1275–1284. https://doi.org/10.1534/genetics.115.177709
7. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673–679. https://doi.org/10.1093/bioinformatics/btm009

8. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29. https://doi.org/10.1038/75556