



Complete Genome Sequence of the Poly- γ -Glutamate-Synthesizing Bacterium *Bacillus subtilis* Bs-115

Fengqing Wang,^a Lijuan Gong,^a Lihong Zhou,^a Jinzhong Liang^b

^aDepartment of Bioscience Engineering, Sichuan University of Science & Engineering, Zigong, People's Republic of China

^bKey Laboratory of Food Science and Engineering, College of Food Engineering, Harbin University of Commerce, Harbin, People's Republic of China

ABSTRACT *Bacillus subtilis* Bs-115 was isolated from the soil of a corn field in Yutai County, Jinan City, Shandong Province, People's Republic of China, and is characterized by the efficient synthesis of poly- γ -glutamate (γ -PGA), with corn saccharification liquid as the sole energy and carbon source during the process of γ -PGA formation. Here, we report the complete genome sequence of *Bacillus subtilis* Bs-115 and the genes associated with poly- γ -glutamate synthesis.

Bacillus subtilis Bs-115 was isolated from the soil of a corn field in Yutai County, Jinan City, Shandong Province, People's Republic of China. Strain Bs-115 can use glucose and other common carbon sources for growth. Currently, there are many strains of this genus that can be retrieved in NCBI that have been reported to have poly- γ -glutamate (γ -PGA) synthesis ability. Strain Bs-115 was the only one that could use corn saccharification liquid as the only carbon and energy source for γ -PGA synthesis. γ -PGA is an extremely important substance with unique properties that have been exploited for a wide array of useful applications, such as in food, cosmetics, medicine, agriculture, oil recovery, and water treatment (1). The number of potential applications for γ -PGA is still increasing.

The genome sequence of *Bacillus subtilis* Bs-115 was determined by Nextomics Biosciences Co., Ltd. (Wuhan, People's Republic of China), using PacBio single-molecule real-time sequencing technology. A total of 1,687,494,897 bp of original data was obtained. The depth of sequencing was about 367.31 \times , and the average length was 15,134 bp. Celera assembler and minimus2 software were used to assemble reads and connect contigs (2), respectively. The whole information map of the *Bacillus subtilis* Bs-115 genome was drawn using Circos (3) to obtain gene structure prediction results.

The genome of *Bacillus subtilis* Bs-115 consists of a single circular chromosome of 4,142,593 bp and has an average GC content of 43.35%. There are a total of 4,661 putative open reading frames (with an average size of 792.91 bp) according to Glimmer (4). Eighty-eight tRNA genes for all 20 amino acids and 10 16S-23S-5S rRNA operons were identified by tRNAscan-SE (5) and RNAmmer (6), respectively.

Sixteen simple sequence repeats (SSR) and 86 tandem repeats were identified by *misa* and *TRF* (7), respectively. DNA base modification (such as m4C and m6A type base methylation modification) plays an important role in biological processes, including growth and aging, immunity, pathogenicity of bacteria, and the development of disease. Surface modification analysis (8) showed that there were 388,984 m4C and 11,504 m6C base modifications. The consistencies of the restriction modification system of the genome with putative type I specificity, a putative type I methyltransferase system, a putative type I restriction enzyme, and a putative type IV methyl-directed restriction enzyme were 100%, 94.75%, 95.04%, and 99.23%, respectively.

Received 20 February 2018 Accepted 8 March 2018 Published 19 April 2018

Citation Wang F, Gong L, Zhou L, Liang J. 2018. Complete genome sequence of the poly- γ -glutamate-synthesizing bacterium *Bacillus subtilis* Bs-115. *Genome Announc* 6:e00197-18. <https://doi.org/10.1128/genomeA.00197-18>.

Copyright © 2018 Wang 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 Fengqing Wang, wangfengqing-100@163.com.

The genome of *Bacillus subtilis* Bs-115 contains the *pgsBCAE* gene, which is related to γ -PGA synthase enzymes, and the *pgdS* and *ggt* genes, which are related to γ -PGA degradation enzymes. Much work has been done to improve γ -PGA production based on the heterologous expression of synthase genes. However, little research has been performed to investigate ways to improve the yield and molecular weight of γ -PGA by reducing the activity of degradation enzyme gene expression. In particular, research on *in situ* editing of the target gene (the γ -PGA degradation enzyme gene) to improve the production and molecular weight of γ -PGA using clustered regularly interspaced short palindromic repeat (CRISPR)-Cas genome editing technology has not yet been reported.

Accession number(s). The nucleotide sequence comprising the *Bacillus subtilis* Bs-115 genome was deposited in GenBank with the accession number [CP020722](https://www.ncbi.nlm.nih.gov/nuccore/CP020722).

ACKNOWLEDGMENTS

This work was financially supported by the National Science and Technology Planning Program of China (2012BAD32B08) and the Talent Introduction Project of Sichuan University of Science & Engineering (2017RCL73, 2015RC48, and 2015RC46).

REFERENCES

- Adetoro O, Aditya B, Victor UI, David H, Craig W, Iza R. 2015. Poly- γ -glutamic acid: production, properties and applications. *Microbiology* 161: 1–17. <https://doi.org/10.1099/mic.0.081448-0>.
- Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G. 2008. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* 24:2818–2824. <https://doi.org/10.1093/bioinformatics/btn548>.
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res* 19:1639–1645. <https://doi.org/10.1101/gr.092759.109>.
- 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>.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25: 955–964.
- Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <https://doi.org/10.1093/nar/gkm160>.
- Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573. <https://doi.org/10.1093/nar/27.2.573>.
- Davis BM, Chao MC, Waldor MK. 2013. Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Curr Opin Microbiol* 16:192–198. <https://doi.org/10.1016/j.mib.2013.01.011>.