Complete Genome Sequence of *Bacillus subtilis* Strain CGMCC 12426, an Efficient Poly-γ-Glutamate Producer

Hui Dong, Jinglin Chang, Xin He, Qinlian Hou, Wei Long

Key Laboratory of Tianjin Radiation and Molecular Nuclear Medicine, Institute of Radiation Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China

ABSTRACT *Bacillus subtilis* CGMCC 12426 is an efficient producer of poly-γ-glutamate with regular stereochemistry. Here, the complete genome sequence of *B. subtilis* CGMCC 12426 is presented, which may facilitate the design of rational strategies for further strain improvements with industrial potential.

Poly-γ-glutamate (γ-PGA) is an anionic, biodegradable, water-soluble biopolymer that is edible and nontoxic to humans and the environment (1). This biopolymer has various functions and has been used in a broad range of industrial fields such as food, cosmetics, pharmaceuticals, and water treatment (2, 3). γ-PGA was first discovered in *Bacillus anthracis* at the start of the 20th century, and γ-D-PGA was later produced by this type of strain (4). Several bacteria (mostly from the genus *Bacillus*) have been shown to secrete γ-PGA into the medium as a product of fermentation. The most intensively studied are *B. subtilis* and *B. licheniformis* (5, 6). Compared with *B. licheniformis*, *B. subtilis* shows a higher productivity, and there has been growing interest in the application of this type of strain (7, 8). Recently, a strain of *B. subtilis* was isolated from soil and named *B. subtilis* KH2, which is an efficient γ-PGA producer. It has been deposited in the China General Microbiological Culture Collection Center (CGMCC no. 12426).

Genomic DNA from *B. subtilis* CGMCC 12426 was extracted using the QIAamp DNA minikit (Qiagen, CA). The quantity and quality of genomic DNA were evaluated on the Agilent 2100 Bioanalyzer (Agilent, USA). Genomic DNA was used to construct a 10-kb insert SMRTbell library, and then sequenced on the single molecule real-time (SMRT) DNA sequencing platform using the Pacific Biosciences (PacBio) RS II sequencer (Pacific Biosciences, CA) (9). A total of 150,292 polymerase reads on one SMRT cell for 3-h movie times led to a total of 1,386,478,854 nucleotide bases. After filtering to remove any reads having low accuracy values less than 0.8, 1,256,996,185 read bases were obtained with 0.866 read quality. All of the filtered sequences were *de novo* assembled using the RS hierarchical genome assembly process (HGAP) assembly protocol 2.0 in SMRT analysis software version 2.3.0 (Pacific Biosciences) (10). The length of the complete circular chromosome is 4,138,265 bp, with a 74,165-bp length plasmid. The annotation was performed by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP), resulting in the prediction of 4,581 genes, including 4,222 coding sequences (CDSs), and 87 tRNA and 30 rRNA (5S rRNA, 16S rRNA, and 23S rRNA) sequences.

The genome sequence of *B. subtilis* CGMCC 12426 could serve as a basis for further elucidation of the genetic background of this promising strain, and provide significant opportunities for investigating the metabolic and regulatory mechanisms underlying the formation of ethanol, organic acids, amino acids, etc. Importantly, all of the genes responsible for PGA biosynthesis and degradation were successfully annotated. This genome sequence may also facilitate the identification of suitable target genes that can

Address correspondence to Hui Dong, donghui@irm-cams.ac.cn, or Wei Long, longway@irm-cams.ac.cn.

assist with the development of superior microbial cell factories with higher concentration, yield, and productivity of PGA by systems metabolic engineering.

Accession number(s). The complete genome information of *B. subtilis* KH2 (CGMCC 12426) was deposited in GenBank under the accession numbers CP018184 and CP018185.

REFERENCES

1. Kimura K, Tran LS, Do TH, Itoh Y. 2009. Expression of the *pgsB* encoding the poly-gamma-DL-glutamate synthetase of *Bacillus subtilis* (*natto*). Biosci Biotechnol Biochem 73:1149–1155. https://doi.org/10.1271/bbb.80913.
2. Ashiuchi M, Shimanouchi K, Nakamura H, Kamei T, Soda K, Park C, Sung MH, Misono H. 2004. Enzymatic synthesis of high-molecular-mass poly-γ-glutamate and regulation of its stereochemistry. Appl Environ Microbiol 70:4249–4255. https://doi.org/10.1128/AEM.70.7.4249-4255.2004.
3. Wu Q, Xu H, Shi N, Yao J, Li S, Ouyang P. 2008. Improvement of poly(gamma-glutamic acid) biosynthesis and redistribution of metabolic flux with the presence of different additives in *Bacillus subtilis* CGMCC 0833. Appl Microbiol Biotechnol 79:527–535. https://doi.org/10.1007/s00253-008-1462-x.
4. Candela T, Fouet A. 2006. Poly-gamma-glutamate in bacteria. Mol Microbiol 60:1091–1098. https://doi.org/10.1111/j.1365-2958.2006.05179.x.
5. Ashiuchi M, Kamei T, Baek DH, Shin SY, Sung MH, Soda K, Yagi T, Misono H. 2001. Isolation of *Bacillus subtilis* (chungkookjang), a poly-γ-glutamate producer with high genetic competence. Appl Microbiol Biotechnol 57:764–769. https://doi.org/10.1007/s00253-001-0848-9.
6. Bajaj IB, Singhal RS. 2010. Effect of aeration and agitation on synthesis of poly (γ-glutamic acid) in batch cultures of *Bacillus licheniformis* NCIM 2324. Biotechnol Bioprocess Eng 15:635–640. https://doi.org/10.1007/s12257-009-0059-2.
7. Jiang Y, Tang B, Xu Z, Liu K, Xu Z, Feng X, Xu H. 2016. Improvement of poly-γ-glutamic acid biosynthesis in a moving bed biofilm reactor by *Bacillus subtilis* NX-2. Bioresour Technol 218:360–366. https://doi.org/10.1016/j.biortech.2016.06.103.
8. Ashiuchi M, Misono H. 2002. Biochemistry and molecular genetics of poly-γ-glutamate synthesis. Appl Microbiol Biotechnol 59:9–14. https://doi.org/10.1007/s00253-002-0984-x.
9. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
10. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-time DNA sequencing from single polymerase molecules. Science 323:133–138. https://doi.org/10.1126/science.1162986.