

# Draft Genome Sequence of *Bacillus subtilis* subsp. *natto* Strain CGMCC 2108, a High Producer of Poly- $\gamma$ -Glutamic Acid

Siyuan Tan, Yonghong Meng, Anping Su, Chen Zhang, Yuanyuan Ren

College of Food Engineering and Nutritional Science, Shaanxi Normal University, Shaanxi, China

Here, we report the 4.1-Mb draft genome sequence of *Bacillus subtilis* subsp. *natto* strain CGMCC 2108, a high producer of poly- $\gamma$ -glutamic acid ( $\gamma$ -PGA). This sequence will provide further help for the biosynthesis of  $\gamma$ -PGA and will greatly facilitate research efforts in metabolic engineering of *B. subtilis* subsp. *natto* strain CGMCC 2108.

Received 8 April 2016 Accepted 13 April 2016 Published 26 May 2016

**Citation** Tan S, Meng Y, Su A, Zhang C, Ren Y. 2016. Draft genome sequence of *Bacillus subtilis* subsp. *natto* strain CGMCC 2108, a high producer of poly- $\gamma$ -glutamic acid. *Genome Announc* 4(3):e00426-16 doi:10.1128/genomeA.00426-16.

**Copyright** © 2016 Tan 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 Yonghong Meng, mengyonghong@hotmail.com.

*Bacillus subtilis* subsp. *natto*, a spore-forming, aerobic, gram-positive bacterium is thought to be a potential host strain used in the production of the traditional Japanese food “natto” made from soybeans and is a potent producer of poly- $\gamma$ -glutamic acid ( $\gamma$ -PGA) (1). As a consequence, there has been growing interest in the biotechnological applications of this host strain due to both established biological information and intriguing physiological characteristics. Recently, efforts have further expanded the genetic toolbox for *B. subtilis* subsp. *natto* and isolated strains that are high-level producers of  $\gamma$ -PGA.

Although type strain BEST195 DNA has been available for some time (1), it is not the highest producer of  $\gamma$ -PGA. *B. subtilis* subsp. *natto* CGMCC 2018 has been found to be a higher producer of  $\gamma$ -PGA (2), and is widely used in the fields of foods, cosmetics, pharmaceuticals, and water treatment due to its outstanding qualities such as biodegradability, film-reforming property, moisture-retention property, and edibility (3–7). We studied the effect of calcium on the activity of  $\alpha$ -ketoglutaric acid dehydrogenase (OGDH) *in vivo* and *in vitro*, as well as expression abundances of OGDH and poly-glutamate synthesizing enzyme (PGSB). As a result, we found that the activities of OGDH and PGSB are regulated by Ca<sup>2+</sup> at the transcriptional level and, in turn, have a significant effect on  $\gamma$ -PGA production (8).

Samples were sent to the Beijing Genomics Institute at the High-Throughput Sequencing Facility (HTSF) for whole-genome shotgun sequencing using an Illumina HiSeq2000 sequencer (Beijing, China). The raw sequence data comprise a total of 90 reads that provide high sampling coverage of the genome (34.39-fold coverage). The reads were assembled using SOAPdenovo version 1.05 (9). This led to a genome assembly containing 35 contigs (each at a length of  $\geq$ 310 bp). The gaps among the scaffolds were filled by a long PCR experiment and three large sequences were finally constructed, including *B. natto* CGMCC 2018 with 4,122,154 bp containing two plasmids, one with 5,820 bp and the other with 65,774 bp.

The draft genome of strain CGMCC 2108 was analyzed and annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) ([http://www.ncbi.nlm.nih.gov/genome/annotation\\_prok](http://www.ncbi.nlm.nih.gov/genome/annotation_prok)). The draft genome sequence of strain CGMCC 2108 has a

G+C content of 43.26%. The genome contains 4,487 predicted open reading frames and 4,365 predicted protein-coding sequences. There are 87 tRNA genes, 10 rRNA genes, and 5 noncoding RNA (ncRNA) genes predicted from this assembly.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited at DDBJ/EMBL/Gen Bank under the accession numbers [CP014471](https://www.ncbi.nlm.nih.gov/nuccore/CP014471) to [CP014473](https://www.ncbi.nlm.nih.gov/nuccore/CP014473).

## ACKNOWLEDGMENTS

This research work was supported by projects from the Overall Planning of Shaanxi Science and Technology (project 2015KTCQ02-07) and the China Agriculture Research System-28 (CA-28).

## FUNDING INFORMATION

This work, including the efforts of Yonghong Meng, was funded by Overall Planning of Shaanxi Science and Technology (2015KTCQ02-07). This work, including the efforts of Yonghong Meng, was funded by China Agriculture Research System-28 (CA-28).

## REFERENCES

- Nishito Y, Osana Y, Hachiya T, Popendorf K, Toyoda A, Fujiyama A, Itaya M, Sakakibara Y. 2010. Whole genome assembly of a natto production strain *Bacillus subtilis* natto from very short read data. *BMC Genomics* 11:243. <http://dx.doi.org/10.1186/1471-2164-11-243>.
- Ren Y, Huang B, Meng Y, Wei L, Chen Z. 2015. Metabolic and phylogenetic analyses based on nitrogen in a new poly- $\gamma$ -glutamic acid-producing strain of *Bacillus subtilis*. *Biotechnol Lett* 37:1221–1226. <http://dx.doi.org/10.1007/s10529-015-1776-6>.
- Choi HJ, Kunioka M. 1995. Preparation conditions and swelling equilibria of hydrogel prepared by  $\gamma$ -irradiation from microbial poly( $\gamma$ -glutamic acid). *Radiat Phys Chem* 46:175–179. [http://dx.doi.org/10.1016/0969-806X\(95\)00009-M](http://dx.doi.org/10.1016/0969-806X(95)00009-M).
- Shih IL, Van YT. 2001. The production of poly ( $\gamma$ -glutamic acid) from microorganism and its various applications. *Bioresour Technol* 79: 207–225. [http://dx.doi.org/10.1016/S0960-8524\(01\)00074-8](http://dx.doi.org/10.1016/S0960-8524(01)00074-8).
- Yokoi H, Arima T, Hirose J, Hayashi S, Takasaki Y. 1996. Flocculation properties of poly ( $\gamma$ -glutamic acid) produced by *Bacillus subtilis*. *J Ferment Bioeng* 82:84–87.
- Mitsuiki M, Mizuno A, Tanimoto H. 1998. Relationship between the antifreeze activities and the chemical structures of alio- and poly(glutamic acids). *J Agric Food Chem* 46:891–895.

7. Shih IL, Van YT, Yeh LC, Lin HG, Chang YN. 2001. Production of a biopolymer flocculant from *Bacillus licheniformis* and its flocculation properties. *Bioresour Technol* 78:267–272. [http://dx.doi.org/10.1016/S0960-8524\(01\)00027-X](http://dx.doi.org/10.1016/S0960-8524(01)00027-X).
8. Huang B, Qin P, Xu Z, Zhu R, Meng Y. 2011. Effects of CaCl<sub>2</sub> on viscosity of culture broth, and on activities of enzymes around the 2-oxoglutarate branch, in *Bacillus subtilis* CGMCC 2108 producing poly-(c-glutamic acid). *Bioresour Technol* 102:3595–3598. <http://dx.doi.org/10.1016/j.biortech.2010.10.073>.
9. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1:18. <http://dx.doi.org/10.1186/2047-217X-1-18>.