Draft Genome Sequence of *Bacillus subtilis* GXA-28, a Thermophilic Strain with High Productivity of Poly-γ-Glutamic Acid

Wei Zeng,[a,b,c] Guiguang Chen,[c] Zhen Tang,[a,b,c] Hao Wu,[a,b,c] Lin Shu,[c] Zhiqun Liang[a,b,c]

State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Guangxi University, Guangxi, China[a]; Key Laboratory of Ministry of Education for Microbial and Plant Genetic Engineering, Guangxi University, Guangxi, China[b]; College of Life Science and Technology, Guangxi University, Guangxi, China[c]

W.Z. and G.C. contributed equally to this work.

***Bacillus subtilis* GXA-28 is a thermophilic strain that can produce high yield and high molecular weight of poly-γ-glutamic acid under high temperature. Here, we report the draft genome sequence of this strain, which may provide the genomic basis for the high productivity of poly-γ-glutamic acid.**

Address correspondence to Zhiqun Liang, zqliang@gxu.edu.cn.

*B*acillus subtilis strain GXA-28 (CCTCC M2012347) is a Gram-positive, thermophilic, spore-forming bacterium with capability of high productivity of poly-γ-glutamic acid under high temperature, which was isolated from marine sands of Beihai, Guangxi, China (1). Poly-γ-glutamic acid or γ-PGA is a promising biomaterial with wide application in industry, agriculture, medicine, food, cosmetics, and wastewater treatment. Compared with the laboratory strain of *B. subtilis* Marburg 168 (2), the strain of GXA-28 can produce larger amounts of γ-PGA. The yield of γ-PGA, which was obtained at 45°C after 22 h, increased by 2.27 and 10 times over that obtained at 37°C and 28°C, respectively. Furthermore, the molecular weight reached 3.03×10^6, which belongs to the range of ultra-high molecular weight (3). In order to explore the biosynthesis mechanism of high yield and high molecular weight of γ-PGA under high temperature, the genome of GXA-28 was sequenced and released.

The genome of GXA-28 was sequenced by using a shotgun strategy combining the Illumina HiSeq 2000 and Illumina MiSeq platforms, which produced paired reads totaling 1,260 Mb with 300-fold coverage. All of the sequence data were processed and *de novo* assembled into 13 contigs with an N_{50} of 1,100,410 bp, an N_{90} of 292,933 bp, and a maximum contig size of 1,156,606 bp using SPAdes version 3.0 (4). Annotation was conducted by RAST (5), RNAmmer (6), tRANScan (7), and BLAST against the RefSeq database. The genome of GXA-28 is 4,261,421 bp with a G + C content of 43.6%, containing 4,468 protein coding genes (CDSs), 71 tRNA genes, and 20 rRNA operons.

BLAST analysis of the genome sequence of GXA-28 against the RefSeq database revealed that it provided a complete set of genes related to γ-PGA biosynthesis, including the glutamate racemase genes *yrpC* and *racE*; the glutamate symport protein genes *gltT* and *gltP*; the synthetase genes *pgsBCAE*; the depolymerase genes *pgdS*, *ywrD*, and *ggt*; and the regulator genes *comPA*, *degSU*, *degQ*, and *swrA*. It has been reported that the two-component system (TCS) genes *comPA* and *degSU* were the key regulatory factors in the γ-PGA biosynthesis (8, 9). Interestingly, two temperature-responsive TCS genes, *vicKR* and *desKR*, which are involved in the regulation of the components of the cell wall/membrane and the desaturation of membrane phospholipids, respectively, are found in the genome of GXA-28. These may be associated with the high productivity of γ-PGA in GXA-28 under high temperature and provided novel information on the γ-PGA biosynthesis in *Bacillus* species, although much evidence is needed to verify this.

The availability of the genome sequence of a thermophilic strain GXA-28 provides us the opportunity to further understand the genetic differences between Marburg 168 and GXA-28 that affect γ-PGA biosynthesis, to explain the genetic reasons for high productivity of γ-PGA in GXA-28 under high temperature, and to get more genes that regulate or participate in the process of γ-PGA production.

Nucleotide sequence accession numbers. The whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number JPNZ00000000. The version described in this paper is version JPNZ01000000.

REFERENCES

1. **Zeng W, Lin Y, Qi Z, He Y, Wang D, Chen G, Liang Z.** 2013. An integrated high-throughput strategy for rapid screening of poly(γ-glutamic acid)-producing bacteria. Appl. Microbiol. Biotechnol. **97:**2163–2172. http://dx.doi.org/10.1007/s00253-013-4717-0.
2. **Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessières P, Bolotin A, Borchert S, Borriss R, Boursier L, Brans A, Braun M, Brignell SC, Bron S, Brouillet S, Bruschi CV, Caldwell B, Capuano V, Carter NM, Choi SK, Cordani JJ, Connerton IF, Cummings NJ, Daniel RA, Denziot F, Devine KM, Düsterhöft A, Ehrlich SD, Emmerson PT, Entian KD, Errington J, Fabret C, Ferrari E, Foulger D, Fritz C, Fujita M, Fujita Y, Fuma S, Galizzi A, Galleron N, Ghim SY,**

Glaser P, Goffeau A, Golightly EJ, Grandi G, Guiseppi G, Guy BJ, Haga K, Haiech J, Harwood CR, Hènaut A, Hilbert H, Holsappel S, Hosono S, Hullo MF, Itaya M, Jones L, Joris B, Karamata D, Kasahara Y, Klaerr-Blanchard M, Klein C, Kobayashi Y, Koetter P, Koningstein G, Krogh S, Kumano M, Kurita K, Lapidus A, Lardinois S, Lauber J, Lazarevic V, Lee SM, Levine A, Liu H, Masuda S, Mauël C, Médigue C, Medina N, Mellado RP, Mizuno M, Moestl D, Nakai S, Noback M, Noone D, O'Reilly M, Ogawa K, Ogiwara A, Oudega B, Park SH, Parro V, Pohl TM, Portelle D, Porwollik S, Prescott AM, Presecan E, Pujic P, Purnelle B, Rapoport G, Rey M, Reynolds S, Rieger M, Rivolta C, Rocha E, Roche B, Rose M, Sadaie Y, Sato T, Scanlan E, Schleich S, Schroeter R, Scoffone F, Sekiguchi J, Sekowska A, Seror SJ, Serror P, Shin BS, Soldo B, Sorokin A, Tacconi E, Takagi T, Takahashi H, Takemaru K, Takeuchi M, Tamakoshi A, Tanaka T, Terpstra P, Togoni A, Tosato V, Uchiyama S, Vandebol M, Vannier F, Vassarotti A, Viari A, Wambutt R, Wedler H, Weitzenegger T, Winters P, Wipat A, Yamamoto H, Yamane K, Yasumoto K, Yata K, Yoshida K, Yoshikawa HF, Zumstein E, Yoshikawa H, Danchin A. 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. Nature **390**:249–256. http://dx.doi.org/10.1038/36786.

3. **Zeng W, Chen G, Wang Q, Zheng S, Shu L, Liang Z.** 2014. Metabolic studies of temperature control strategy on poly(γ-glutamic acid) production in a thermophilic strain *Bacillus subtilis* GXA-28. Bioresour. Technol. **155**:104–110. http://dx.doi.org/10.1016/j.biortech.2013.12.086.

4. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS,** Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. **19**:455–477. http://dx.doi.org/10.1089/cmb.2012.0021.

5. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75. http://dx.doi.org/10.1186/1471-2164-9-75.

6. **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. http://dx.doi.org/10.1093/nar/gkm160.

7. **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. http://dx.doi.org/10.1093/nar/25.5.0955.

8. **Tran LS, Nagai T, Itoh Y.** 2000. Divergent structure of the ComQXPA quorum-sensing components: molecular basis of strain-specific communication mechanism in *Bacillus subtilis*. Mol. Microbiol. **37**:1159–1171. http://dx.doi.org/10.1046/j.1365-2958.2000.02069.x.

9. **Stanley NR, Lazazzera BA.** 2005. Defining the genetic differences between wild and domestic strains of *Bacillus subtilis* that affect poly-γ-DL-glutamic acid production and biofilm formation. Mol. Microbiol. **57**:1143–1158. http://dx.doi.org/10.1111/j.1365-2958.2005.04746.x.