Genome Sequences of Two Morphologically Distinct and Thermophilic *Bacillus coagulans* Strains, H-1 and XZL9

Ke Xu, Fei Su, Fei Tao, Chao Li, Jun Ni, Ping Xu

State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

K.X. and F.S. contributed equally to this work.

Two thermophilic *Bacillus coagulans* strains, H-1 and XZL9, both of which were isolated from soils, have different morphological properties. Strain XZL9 but not H-1 is an efficient pentose-utilizing producer of important platform compounds, such as L-lactic acid and 2,3-butanediol. Here we announce the 2.86- and 3.43-Mb sequences of their genomes.

Address correspondence to Ping Xu, pingxu@sjtu.edu.cn.

Many strains present morphological, cytological, and physiological alterations under different environments. Morphological features of microbial cells may play a role in fermentations (1–3). In previous studies, *Bacillus coagulans* was proposed to be a potential good producer for platform chemicals, such as L-lactic acid, 2,3-butanediol, and ethanol (4–9). Both of the *B. coagulans* strains H-1 (CCTCC M 2013105) and XZL9 (DSM 23184) are good L-lactic acid producers. These two strains were isolated by our group from soils and were found to share many good characteristics, such as high optical purity, nonsterilization fermentation under high temperatures, and high productivity. However, a morphological change of strain H-1, which is unable to utilize pentose, transforming it from a filamentous to a rod shape, is accompanied by reduced production rates and even growth limits. Meanwhile, the stable morphology of strain XZL9 can produce high L-lactic acid (130 g/liter) from xylose, with a high optical purity of 99% and a high yield of 98% at a temperature of 50°C (data not shown). To discriminate the relationship between physicochemical and metabolic properties, we determined the whole-genome sequences of strains H-1 and XZL9.

We obtained the raw data sets of strains H-1 and XZL9 using the Illumina HiSeq 2000 system. A total of 9,915,274 filtered reads for strain H-1 were assembled into 337 contigs, and a total of 15,532,424 filtered reads for strain XZL9 were assembled into 45 contigs using VELVET (10). The genome annotations were performed by the RAST server (11). The functional descriptions were determined using Clusters of Orthologous Genes (12) and KEGG (13). The genes encoding tRNAs and rRNAs were identified by tRNAscan-SE (14) and RNAmmer (15), respectively.

The draft genome sequences of strains H-1 and XZL9 consist of 2,862,880 bases and 3,426,041 bases, with GC contents of 47.3% and 46.5%, respectively. There are 104 and 101 predicted RNAs in strains H-1 and XZL9, respectively. A total of 329 subsystems for strain H-1 and 531 subsystems for strain XZL9 were identified by RAST. In the results of RAST, we did not find significant differences in the genes involved in cell division of the two strains. Both of them possess the key structure of a cytokinetic ring, including *ftsAKLIWZ* and missing the essential gene *zipA*, which connects the ring with the cell membrane. This implies that *B. coagulans* has a common cell division process similar to that of typical rod-shaped cells, such as *Escherichia coli* and *Bacillus subtilis*. Besides these, in the subsystem of the cell wall and capsule, there are 74 genes and 108 genes for strains H-1 and XZL9, respectively. We have carefully compared the two strains, finding that the systems of rhamnose biosynthesis and sialic acid metabolism were completely absent in strain H-1. This might lead to the unstable morphology of strain H-1, which needs to be determined by further experiments.

Nucleotide sequence accession numbers. The whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession numbers ANAQ00000000 for strain H-1 and ANAP00000000 for strain XZL9. The versions described in this paper are the first version, ANAQ01000000 and ANAP01000000.

REFERENCES

1. **Saito II, Honda H, Kawabe T, Mukumoto F, Shimizu M, Kobayashi T.** 2000. Comparison of biotin production by recombinant *Sphingomonas* sp. under various agitation conditions. Biochem. Eng. J. **5**:129–136.
2. **Zhang A, Yang ST.** 2009. Engineering *Propionibacterium acidipropionici* for enhanced propionic acid tolerance and fermentation. Biotechnol. Bioeng. **104**:766–773.
3. **Jiang L, Wang J, Liang S, Cai J, Xu Z, Cen P, Yang S, Li S.** 2011. Enhanced butyric acid tolerance and bioproduction by *Clostridium tyrobutyricum* immobilized in a fibrous bed bioreactor. Biotechnol. Bioeng. **108**:31–40.
4. **Su F, Yu B, Sun J, Ou HY, Zhao B, Wang L, Qin J, Tang H, Tao F, Jarek M, Scharfe M, Ma C, Ma Y, Xu P.** 2011. Genome sequence of the thermophilic strain *Bacillus coagulans* 2-6, an efficient producer of high-optical-purity L-lactic acid. J. Bacteriol. **193**:4563–4564.

5. **Su F, Xu K, Zhao B, Tai C, Tao F, Tang H, Xu P.** 2011. Genome sequence of the thermophilic strain *Bacillus coagulans* XZL4, an efficient pentose-utilizing producer of chemicals. J. Bacteriol. **193:** 6398–6399.

6. **Qin J, Zhao B, Wang X, Wang L, Yu B, Ma Y, Ma C, Tang H, Sun J, Xu P.** 2009. Non-sterilized fermentative production of polymer-grade L-lactic acid by a newly isolated thermophilic strain *Bacillus* sp. 2-6. PLoS One **4:**e4359.

7. **Wang L, Zhao B, Liu B, Yu B, Ma C, Su F, Hua D, Li Q, Ma Y, Xu P.** 2010. Efficient production of L-lactic acid from corncob molasses, a waste by-product in xylitol production, by a newly isolated xylose utilizing *Bacillus* sp. strain. Bioresour. Technol. **101:**7908–7915.

8. **Wang L, Xue Z, Zhao B, Yu B, Xu P, Ma Y.** 2013. Jerusalem artichoke powder: a useful material in producing high-optical-purity L-lactate using an efficient sugar-utilizing thermophilic *Bacillus coagulans* strain. Bioresour. Technol. **130:**174–180.

9. **Ou MS, Mohammed N, Ingram LO, Shanmugam KT.** 2009. Thermophilic *Bacillus coagulans* requires less cellulases for simultaneous saccharification and fermentation of cellulose to products than mesophilic microbial Biocatalysts. Appl. Biochem. Biotechnol. **155:** 379–385.

10. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18:**821–829.

11. **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.

12. **Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA.** 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics **4:**41.

13. **Ogata H, Goto S, Fujibuchi W, Kanehisa M.** 1998. Computation with the KEGG pathway database. J. Biol. Syst. **47:**119–128.

14. **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.

15. **Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108.