JOURNAL OF BACTERIOLOGY, Sept. 2011, p. 4563–4564 Vol. 193, No. 17
0021-9193/11/$12.00 doi:10.1128/JB.05378-11

Genome Sequence of the Thermophilic Strain *Bacillus coagulans* 2-6, an Efficient Producer of High-Optical-Purity L-Lactic Acid▽

Fei Su,[1] Bo Yu,[2] Jibin Sun,[3] Hong-Yu Ou,[1] Bo Zhao,[2] Limin Wang,[2] Jiayang Qin,[3] Hongzhi Tang,[1] Fei Tao,[1] Michael Jarek,[5] Maren Scharfe,[5] Cuiqing Ma,[4] Yanhe Ma,[2] and Ping Xu[1]*

State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, People's Republic of China[1]; Institute of Microbiology, Chinese Academy of Sciences, Beijing, People's Republic of China[2]; Tianjin Industrial Biotechnology R&D Center, Chinese Academy of Sciences, Tianjin, People's Republic of China[3]; State Key Laboratory of Microbial Technology, Shandong University, Jinan, People's Republic of China[4]; and Helmholtz Center for Infection Research, Inhoffenstrasse 7, Braunschweig, Germany[5]

Received 23 May 2011/Accepted 13 June 2011

***Bacillus coagulans* 2-6 is an efficient producer of lactic acid. The genome of *B. coagulans* 2-6 has the smallest genome among the members of the genus *Bacillus* known to date. The frameshift mutation at the start of the D-lactate dehydrogenase sequence might be responsible for the production of high-optical-purity L-lactic acid.**

Bacillus coagulans, from spoiled canned milk, was first described in 1915 by Hammer (8). Because of stable high performance in the utilization of renewable resources and nonsterilization fermentation at high temperature, the thermophilic *B. coagulans* strains have been suggested to be superior producers of lactic acid (11, 13). In addition to the production of lactic acid, *B. coagulans* has also been found to be a source of many other commercially valuable products, such as thermostable enzymes and coagulin, an antimicrobial peptide (6). Compared with other probiotic bacteria such as *Lactobacillus* species, some strains of *B. coagulans* are able to survive in the environment of extremes of heat, acidity of the stomach, and bile acids (3). However, little genetic information is known. Here, we present the genome sequence of *B. coagulans* strain 2-6, which is an efficient producer of high-optical-purity L-lactic acid with the advantages of high carbon efficiency, less by-product formation, and thermotolerance (13).

The whole genome of *B. coagulans* 2-6 was sequenced using the Illumina GA system performed by the Helmholtz Center for Infection Research in Germany with a combination of paired-end library and mate pair. Reads were assembled with Velvet (14). According to the draft sequence of *B. coagulans* 36D1 and contigs from different assembly softwares (Edena [5], Euler-SR [2], and SOAPdenovo [7]), the complete genome sequence of strain 2-6 was completed. Closure of the gaps was finished by Bubble PCR primer walking using the routine Sanger method and edited in the Phred/Phrap/Consed (4) package. Finally, Illumina data were used to correct potential base errors and increase consensus quality by mapping the reads to the genome. The genome sequence of *B. coagulans* 2-6 was annotated with the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (12) and functional annotation using Clusters of Orthologous Genes and KEGG (9).

The genome of *B. coagulans* 2-6, which is the smallest of the known *Bacillus* genomes, is composed of a 3,073,079-bp single circular chromosome with a mean GC content of 47.3% and a 9,910-bp plasmid whose mean GC content is 38.0%. We identified 2,975 protein-coding sequences (CDS) in the chromosome and 10 CDS in the plasmid. No putative biological functions were predicted for the plasmid. The CDS in the chromosome constitute 79.9% of the genome. Putative biological functions were assigned to 2,332 (78.4%) predicted proteins based on BLAST (1) results. The frameshift mutation at the start of the D-lactate dehydrogenase sequence might be responsible for the production of high-optical-purity L-lactic acid (optical purity, >99%) by strain 2-6 (13). Only a fragment of pyrophosphokinase in the phosphoketolase pathway was predicted, which suggested the pentose mainly lost in the transaldolase/transketolase pathway. Compared with the phosphoketolase pathway, the transaldolase/transketolase pathway could produce 1.67 mol of lactic acid per mol pentose, whereas the phosphoketolase pathway produces only 1 mol of lactic acid in addition to 1 mol acetate (10).

Nucleotide sequence accession number. The complete genome sequence of *B. coagulans* 2-6 has been submitted to GenBank under accession number CP002472.

This work was supported by grants from the National Basic Research Program of China (2007CB707803), the Chinese National Programs for High Technology Research and Development (2006AA020102 and 2007AA10Z360), the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-YW-G-005), and the National Natural Science Foundation of China (30900022).

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Chaisson, M. J., and P. A. Pevzner.** 2008. Short read fragment assembly of bacterial genomes. Genome Res. **18:**324–330.
3. **Endres, J. R., et al.** 2009. Safety assessment of a proprietary preparation of a novel probiotic, Bacillus coagulans, as a food ingredient. Food Chem. Toxicol. **47:**1231–1238.

* Corresponding author. Mailing address: School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China. Phone: 86 21 34206647. Fax: 86 21 34206723. E-mail: pingxu@sjtu.edu.cn.

▽ Published ahead of print on 24 June 2011.

4. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
5. **Hernandez, D., P. Francois, L. Farinelli, M. Osteras, and J. Schrenzel.** 2008. *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. Genome Res. **18:**802–809.
6. **Hyronimus, B., C. Le Marrec, and M. C. Urdaci.** 1998. Coagulin, a bacteriocin-like inhibitory substance produced by *Bacillus coagulans* I4. J. Appl. Microbiol. **85:**42–50.
7. **Li, R., et al.** 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. Genome Res. **20:**265–272.
8. **Nakamura, L., I. Blumenstock, and D. Claus.** 1988. Taxonomic study of *Bacillus coagulans* Hammer 1915 with a proposal for *Bacillus smithii* sp. nov. Int. J. Syst. Evol. Microbiol. **38:**63.
9. **Ogata, H., S. Goto, W. Fujibuchi, and M. Kanehisa.** 1998. Computation with the KEGG pathway database. Biosystems **47:**119–128.
10. **Patel, M., M. Ou, L. Ingram, and K. Shanmugam.** 2004. Fermentation of sugar cane bagasse hemicellulose hydrolysate to L(+)-lactic acid by a thermotolerant acidophilic *Bacillus* sp. Biotechnol. Lett. **26:**865–868.
11. **Patel, M. A., et al.** 2006. Isolation and characterization of acid-tolerant, thermophilic bacteria for effective fermentation of biomass-derived sugars to lactic acid. Appl. Environ. Microbiol. **72:**3228–3235.
12. **Pruitt, K. D., T. Tatusova, W. Klimke, and D. R. Maglott.** 2009. NCBI Reference Sequences: current status, policy and new initiatives. Nucleic Acids Res. **37:**D32–D36.
13. **Qin, J., et al.** 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.
14. **Zerbino, D. R., and E. Birney.** 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. **18:**821–829.