

# Genome Sequence of *Bacillus coagulans* P38, an Efficient Polymer-Grade L-Lactate Producer from Cellulosic Substrates

Lili Peng,<sup>a</sup> Lifu Song,<sup>b</sup> Lifan Sun,<sup>a</sup> Yumeng Cai,<sup>a</sup> Limin Wang,<sup>a</sup> Bo Yu<sup>a</sup>

CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China<sup>a</sup>; Institute of Bioprocess and Biosystems Engineering, Hamburg University of Technology, Hamburg, Germany<sup>b</sup>

***Bacillus coagulans* P38 is an efficient polymer-grade L-lactic acid producer from a cellulosic carbon source. Here, the draft 3.37-Mb genome sequence of this potential strain may provide useful information to further improve the strain performance for higher titers and, importantly, to understand the mechanism of its high tolerance for 2-furfural.**

Received 10 April 2015 Accepted 16 April 2015 Published 21 May 2015

**Citation** Peng L, Song L, Sun L, Cai Y, Wang L, Yu B. 2015. Genome sequence of *Bacillus coagulans* P38, an efficient polymer-grade L-lactate producer from cellulosic substrates. *Genome Announc* 3(3):e00495-15. doi:10.1128/genomeA.00495-15.

**Copyright** © 2015 Peng et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Bo Yu, yub@im.ac.cn.

Lactic acid is a valuable chemical that finds extensive use in the polymerization of lactic acid to polylactic acid. Many microorganisms, such as fungi, *Lactobacillus* species, and *Bacillus coagulans*, can produce lactic acid (1–3). Compared to the most frequently used lactic acid producers, such as *Lactococcus lactis* and *Lactococcus rhamnosus*, *B. coagulans* can grow optimally at 50 to 55°C, which is expected to minimize contamination in industrial-scale fermentations (4). Therefore, in recent years, there has been an interest in studies on the optical purity of L-lactic acid produced by this species (5, 6). *B. coagulans* strain P38 is an efficient producer of L-lactic acid, with a predominant capacity to tolerate up to 10 g/liter 2-furfural for lactic acid production. Its highly efficient production capability, combined with high inhibitor tolerance, indicates that *B. coagulans* P38 is a promising polymer-grade L-lactic acid producer from cellulosic biomass (6, 7).

Here, we present the draft genome sequence of *B. coagulans* P38 obtained by using the Illumina HiSeq 2000 system, which was performed by the Chinese National Human Genome Center at Shanghai, China. We obtained 9,253,431 high-quality-read base pairs with the Velvet program (8). Through the data assembly, we obtained 137 contigs, and the contig  $N_{50}$  is 57,692 bp. The average length of each contig is 25,385 bp, and the largest one is 272,938 bp, with a total length of 3.37 Mb.

The gene prediction was performed in line with the predicted results of Glimmer 3.02, GeneMark, and the Z-curve program software, with 3,720 genes were predicted. The G+C content of the predicted genes is 47.4%, and the average length of a coding sequence (CDS) is 765 bp. The maximum CDS length is 4,563 bp, and the minimum is 114 bp. Sixty tRNA sequences were found by tRNAscan, and 4 rRNA sequences were predicted via RNAmmer (9). Gene function annotations were made through searching the nucleotide collection, the KEGG protein database, and the SEED protein database of NCBI (10). The Clusters of Orthologous Groups (COG) classifications were performed using the Conserved Domains Database (CDD). The annotation result showed that 2,477 proteins have clear biological functions, of which 1,728 proteins have an ortholog of KEEG, and 2,434 proteins have a

COG classification. All the matched homologous proteins are derived from 146 species, including *B. coagulans* 36D1 (11), with the highest percentage of 72.1%, followed by *B. coagulans* 2-6 (12), with a percentage of 20.1%.

*B. coagulans* P38 was predicted to possess complete metabolic pathways from the genome sequence analysis, including those for glycolysis, the tricarboxylic acid cycle, and the pentose phosphate pathway. One L-lactate dehydrogenase gene and one D-lactate dehydrogenase gene were identified from the genome. A possible lactate/malate dehydrogenase gene was also annotated. Several short-chain dehydrogenase and alcohol dehydrogenase genes were annotated from the genome sequence, which might provide useful information to investigate the 2-furfural tolerance mechanism (13).

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [JSVI00000000](https://www.ncbi.nlm.nih.gov/nuccore/JSVI00000000). The first version (JSVI01000000) is described in this paper.

## ACKNOWLEDGMENTS

This work was partially supported by grants from the National Natural Science Foundation of China (grant 31270108), the Chinese National Programs for High Technology Research & Development (grant 2011AA02A202), and the Key Deployment Project of Chinese Academy of Sciences (grant KSZD-EW-Z-016-3).

## REFERENCES

- Ma K, Maeda T, You H, Shirai Y. 2014. Open fermentative production of L-lactic acid with high optical purity by thermophilic *Bacillus coagulans* using excess sludge as nutrient. *Bioresour Technol* 151:28–35. <http://dx.doi.org/10.1016/j.biortech.2013.10.022>.
- Ye L, Zhou X, Hudari MS, Li Z, Wu JC. 2013. Highly efficient production of L-lactic acid from xylose by newly isolated *Bacillus coagulans* C106. *Bioresour Technol* 132:38–44. <http://dx.doi.org/10.1016/j.biortech.2013.01.011>.
- Okano K, Tanaka T, Ogino C, Fukuda H, Kondo A. 2010. Biotechnological production of enantiomeric pure lactic acid from renewable resources: recent achievements, perspectives, and limits. *Appl Microbiol Biotechnol* 85:413–423. <http://dx.doi.org/10.1007/s00253-009-2280-5>.

4. 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. <http://dx.doi.org/10.1016/j.biortech.2010.05.031>.
5. 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. <http://dx.doi.org/10.1016/j.biortech.2012.11.144>.
6. Peng L, Wang L, Che C, Yang G, Yu B, Ma Y. 2013. *Bacillus* sp. strain p38: an efficient producer of L-lactate from cellulosic hydrolysate, with high tolerance for 2-furfural. *Bioresour Technol* 149:169–176. <http://dx.doi.org/10.1016/j.biortech.2013.09.047>.
7. Peng L, Xie N, Guo L, Wang L, Yu B, Ma Y. 2014. Efficient open fermentative production of polymer-grade L-lactate from sugarcane bagasse hydrolysate by thermotolerant *Bacillus* sp. strain p38. *PLoS One* 9:e107143. <http://dx.doi.org/10.1371/journal.pone.0107143>.
8. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
9. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS Web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 33:W686–W689. <http://dx.doi.org/10.1093/nar/gki366>.
10. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28:27–30. <http://dx.doi.org/10.1093/nar/28.1.27>.
11. Rhee MS, Moritz BE, Xie G, Glavina del Rio T, Dalin E, Tice H, Bruce D, Goodwin L, Chertkov O, Brettin T, Han C, Detter C, Pitluck S, Land ML, Patel M, Ou M, Harbrucker R, Ingram LO, Shanmugam KT. 2011. Complete genome sequence of a thermotolerant sporogenic lactic acid bacterium, *Bacillus coagulans* strain 36D1. *Stand Genomic Sci* 5:331–340. <http://dx.doi.org/10.4056/sigs.2365342>.
12. 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. <http://dx.doi.org/10.1128/JB.05378-11>.
13. Wang X, Yomano LP, Lee JY, York SW, Zheng H, Mullinnix MT, Shanmugam KT, Ingram LO. 2013. Engineering furfural tolerance in *Escherichia coli* improves the fermentation of lignocellulosic sugars into renewable chemicals. *Proc Natl Acad Sci USA* 110:4021–4026. <http://dx.doi.org/10.1073/pnas.1217958110>.