

# Genome Sequence of *Sporolactobacillus laevolacticus* DSM442, an Efficient Polymer-Grade D-Lactate Producer from Agricultural Waste Cottonseed as a Nitrogen Source

Hongyan Wang,<sup>a,c</sup> Limin Wang,<sup>a</sup> Jiansong Ju,<sup>c</sup> Bo Yu,<sup>a</sup> Yanhe Ma<sup>b</sup>

CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China<sup>a</sup>; Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China<sup>b</sup>; College of Life Science, Hebei Normal University, Shijiazhuang, China<sup>c</sup>

***Sporolactobacillus laevolacticus* DSM442 is an efficient polymer-grade D-lactic acid producer from low-cost agricultural waste cottonseed powder as the sole nitrogen source. Here we present a 3.59-Mb assembly of its genome sequence, which might provide useful information to further improve the strain for higher production titers.**

Received 19 November 2013 Accepted 26 November 2013 Published 26 December 2013

**Citation** Wang H, Wang L, Ju J, Yu B, Ma Y. 2013. Genome sequence of *Sporolactobacillus laevolacticus* DSM442, an efficient polymer-grade D-lactate-producer from agricultural waste cottonseed as a nitrogen source. *Genome Announc.* 1(6):e01100-13. doi:10.1128/genomeA.01100-13.

**Copyright** © 2013 Wang 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.

The microorganism *Sporolactobacillus laevolacticus* was first isolated and described as *Bacillus laevolacticus* by Nakayama and Yanoshi in 1967 (1). Andersch et al. reclassified *Bacillus laevolacticus* as *Sporolactobacillus laevolacticus* according to phylogenetic analyses, as well as chemotaxonomic and physiological characterizations in 1994 (2). *S. laevolacticus* is a Gram-positive and motile microorganism and also shows microaerophilic and catalase-negative properties (3, 4). Previous studies of *S. laevolacticus* suggested that it is a homofermentative microorganism that completely converts glucose to D-lactic acid (3, 5). Recent study has shown that *S. laevolacticus* DSM442 can use cottonseed powder as the sole nitrogen source to produce D-lactic acid of high optical purity. A high concentration (144.4 g/liter) was obtained during fed-batch fermentation, with an average productivity of 4.13 g/(liter·h) and a high yield of 0.96 g/g glucose, indicating the predominant potential of *S. laevolacticus* DSM442 (6).

Here, we present the draft genome sequence of *S. laevolacticus* DSM442 obtained by using the Illumina HiSeq 2000 system (300-bp paired-end sequences), which was performed by the Chinese National Human Genome Center at Shanghai, China. With the Velvet program (7), we obtained 8,499,698 high-quality-read base pairs, which provided 474-fold coverage. Through data assembly, we obtained 61 contigs with a contig  $N_{50}$  of 112,365 bp. The average length of each contig is 58,855 bp and the largest one is 306,117 bp, with a total length of 3.59 Mb.

Gene prediction was performed in line with the predicted result of Glimmer 3.02, Genemark, and the Z-Curve program software, and 3,636 genes were predicted. The GC content of the predicted gene is 43.7%, and the average length of coding sequences (CDS) is 830 bp. The maximum CDS length is 7,092 and the minimum is 114 bp. There were 60 tRNA sequences found by tRNAscan and 9 rRNA sequences predicted via RNAmmer (8). Gene function annotations were made through searching the NCBI nucleotide and KEGG protein databases as well as the SEED protein databases of NCBI (9). The clusters of orthologous groups (COG) classification was performed using

the CDD databases. The metabolic pathways of *S. laevolacticus* DSM442 were built through the KEGG databases. The annotation results showed that 2,688 proteins have clear biological functions, of which 1,737 proteins have KEGG orthologs and 2,683 proteins have COG classifications. All the matched homologous proteins are derived from 412 species, including *S. inulinus* CASD, which has the highest percentage, reaching 53.2%. *S. laevolacticus* DSM442 is predicted to possess complete metabolic pathways, including those for glycolysis, the tricarboxylic acid cycle, and the pentose phosphate pathway. Additionally, there are 107 genes annotated for inorganic ion transport and metabolism, including Na/Pi cotransporter II-like protein and calcium-transporting ATPase (EC 3.6.3.8), which might provide useful information to further increase the bacterium's capacity to withstand salt duress and thus enhance lactate production.

**Nucleotide sequence accession numbers.** This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AWTC000000000](https://www.ncbi.nlm.nih.gov/nuccore/AWTC000000000). The version described in this paper is version AWTC01000000.

## ACKNOWLEDGMENTS

The work was partly supported by grants from the National Basic Research Program of China (2011CBA00800), the Chinese National Programs for High Technology Research & Development (2011AA02A202), and the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EW-G-2-3).

## REFERENCES

1. Claus D, Berkeley RCW. 1986. Genus *Bacillus* Cohn 1872, p 174<sup>AL</sup>. In Bergey DH, Holt JG, Kreig NR (ed), *Bergey's manual of systematic bacteriology*, vol. 2. Williams & Wilkins, Baltimore, MD.
2. Hatayama K, Shoun H, Ueda Y, Nakamura A. 2006. *Tuberibacillus calidus* gen. nov., sp. nov., isolated from a compost pile and reclassification of *Bacillus naganensis* Tomimura et al. 1990 as *Pullulanibacillus naganensis* gen. nov., comb. nov. and *Bacillus laevolacticus* Andersch et al. 1994 as *Sporolactobacillus laevolacticus* comb. nov. *Int. J. Syst. Evol. Microbiol.* 56: 2545–2551.

3. Andersch I, Pianka S, Fritze D, Claus D. 1994. Description of *Bacillus laevolacticus* (ex Nakayama and Yanoshi 1967) sp. nov., nom. rev. Int. J. Syst. Evol. Microbiol. 44:659–664.
4. Claus D, Fritze D, Kocur M. 2006. Genera related to the genus *Bacillus*—*Sporolactobacillus*, *Sporosarcina*, *Planococcus*, *Filibacter* and *Caryophanon*. Prokaryotes 4:631–653.
5. de Boer JP, Cronenberg CC, de Beer D, van den Heuvel JC, de Mattos MJ, Neijssel OM. 1993. pH and glucose profiles in aggregates of *Bacillus laevolacticus*. Appl. Environ. Microbiol. 59:2474–2478.
6. Li Y, Wang L, Ju J, Yu B, Ma Y. 2013. Efficient production of polymer-grade D-lactate by *Sporolactobacillus laevolacticus* DSM442 with agricultural waste cottonseed as the sole nitrogen source. Bioresour. Technol. 142: 186–191.
7. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. 18:821–829.
8. 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.
9. Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28:27–30.