

Draft Genome Sequence of *Bacillus coagulans* NL01, a Wonderful L-Lactic Acid Producer

Zhaojuan Zheng,^{a,b} Ting Jiang,^a Xi Lin,^a Jie Zhou,^a Jia Ouyang^{a,c}

College of Chemical Engineering, Nanjing Forestry University, Nanjing, People's Republic of China^a; Jiangsu Key Lab of Biomass-based Green Fuels and Chemicals, Nanjing, People's Republic of China^b; Key Laboratory of Forest Genetics & Biotechnology of the Ministry of Education, Nanjing, People's Republic of China^c

Here, we report the draft genome sequence of *Bacillus coagulans* NL01, which could produce high optically pure L-lactic acid using xylose as a sole carbon source. The draft genome is 3,505,081 bp, with 144 contigs. About 3,903 protein-coding genes and 92 rRNAs are predicted from this assembly.

Received 13 May 2015 Accepted 18 May 2015 Published 18 June 2015

Citation Zheng Z, Jiang T, Lin X, Zhou J, Ouyang J. 2015. Draft genome sequence of *Bacillus coagulans* NL01, a wonderful L-lactic acid producer. *Genome Announc* 3(3): e00635-15. doi:10.1128/genomeA.00635-15.

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

Address correspondence to Jia Ouyang, hgouyj@njfu.edu.cn.

Bacillus coagulans is a wonderful L-lactic acid producer, which is also thought to be a promising producer for high-value chemicals. However, little is known about this species (1). Recently, there has been an extraordinary accumulation of knowledge on the genomics of these organisms. *B. coagulans* NL01 was isolated from soil using a medium with xylose as the sole carbon source (2). In our previous studies, a maximum lactic acid concentration of 75.0 g/liter was achieved from xylose by *B. coagulans* NL01 in batch fermentation. Moreover, this strain also exhibited high fermentation ability from low-cost lignocellulosic feedstocks (2–4). With glucose-rich lignocellulosic hydrolyzate as a substrate, 75.0 g/liter lactic acid was produced (4). When using xylose-rich corn stove prohydrolysate as a substrate, 23.5 g/liter lactic acid was obtained in 36 h (3). To better understand the fermentation process of *B. coagulans* NL01, especially for xylose fermentation, the genome of *B. coagulans* NL01 was sequenced.

The genomic DNA of strain NL01 was extracted using a bacterial DNA kit (Omega). Its quality and quantity were examined and measured using a NanoDrop2000 spectrophotometer (Thermo Scientific) and the Quant-iT PicoGreen double-stranded DNA kit (Invitrogen). Whole-genome sequencing was performed using Illumina MiSeq and generated 1,578,349 paired-end reads of 300 bp with 200-fold coverage on average. Paired-end reads were merged by FLASH (5), and the long-reads were assembled *de novo* into contigs using Newbler (6). The draft genome yields 144 contigs with a G+C content of 46.23% and a total assembly length of 3,505,081 bp. The annotation of the genome was performed using RAST (<http://rast.nmpdr.org>) (7).

Annotation in RAST revealed 92 RNAs, 3,903 coding sequences, and 406 subsystems. There were 450 genes involved in carbohydrates, which contain all the necessary genes for trichloroacetic acid (TCA) and lactic acid production. Also, three L-lactate dehydrogenases and one lactate permease were predicted by RAST. RAST also identified three genes involved in the subsystems of xylose utilization, including xylose isomerase, xylulose kinase, and a repressor, open reading frame, kinase (ROK)-family xylose-responsive transcription regulator. Little genetic informa-

tion was found by RAST, which may be due to the lack of knowledge about this strain. In addition to the lactic acid production, genes required for the production of high-value chemicals (acetoin and butanediol) were found in the annotation, which implies the potential application as a biocatalyst for high-value chemicals. Like other *B. coagulans* strains (1), we also identified the genes involved in the clusters of regularly interspaced short palindromic repeat (CRISPR)-cas systems, such as *cas1-3*, *cas5*, and *cas6*. The CRISPR-cas system was believed to act as a bacterial immune system to keep the host from phage infection. In-depth comparative analysis among different species is now the focus of our work.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. **LBMQ00000000**. The version described in this paper is version LBMQ01000000.

ACKNOWLEDGMENTS

We are grateful to the National Natural Science Foundation of China (31300487 and 31200443), the Natural Science Foundation of Jiangsu Province of China (BK20130970), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) for financial support.

REFERENCES

- Su F, Xu P. 2014. Genomic analysis of thermophilic *Bacillus coagulans* strains: efficient producers for platform bio-chemicals. *Sci Rep* 4:3926. <http://dx.doi.org/10.1038/srep03926>.
- Ouyang J, Cai C, Chen H, Jiang T, Zheng Z. 2012. Efficient non-sterilized fermentation of biomass-derived xylose to lactic acid by a thermotolerant *Bacillus coagulans* NL01. *Appl Biochem Biotechnol* 168:2387–2397. <http://dx.doi.org/10.1007/s12010-012-9944-9>.
- Zheng Z, Cai C, Jiang T, Zhao M, Ouyang J. 2014. Enhanced L-lactic acid production from biomass-derived xylose by a mutant *Bacillus coagulans*. *Appl Biochem Biotechnol* 173:1896–1906. <http://dx.doi.org/10.1007/s12010-014-0975-2>.
- Ouyang J, Ma R, Zheng Z, Cai C, Zhang M, Jiang T. 2013. Open fermentative production of L-lactic acid by *Bacillus* sp. strain NL01 using lignocellulosic hydrolyzates as low-cost raw material. *Bioresour Technol* 135:475–480. <http://dx.doi.org/10.1016/j.biortech.2012.09.096>.

5. Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27:2957–2963. <http://dx.doi.org/10.1093/bioinformatics/btr507>.
6. Nederbragt AJ. 2014. On the middle ground between open source and commercial software—the case of the Newbler program. *Genome Biol* 15: 113. <http://dx.doi.org/10.1186/gb4173>.
7. 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>.