

Genome Sequence of *Bacillus licheniformis* CGMCC3963, a Stress-Resistant Strain Isolated in a Chinese Traditional Solid-State Liquor-Making Process

Qun Wu,^a Suqin Peng,^a Yao Yu,^b Yixue Li,^b Yan Xu^a

State Key Laboratory of Food Science and Technology, Key Laboratory of Industrial Biotechnology of Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China^a; Key Laboratory of Systems Biology/Key Laboratory of Synthetic Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China^b

***Bacillus licheniformis* CGMCC3963 is an important mao-tai flavor-producing strain. It was isolated from the starter (*Daqu*) of a Chinese mao-tai-flavor liquor fermentation process with solid-state fermentation. We report its genome of 4,525,096 bp here. Many potential insertion genes that are responsible for the unique properties of *B. licheniformis* CGMCC3963 in mao-tai-flavor liquor production were identified.**

Received 21 October 2012 Accepted 5 December 2012 Published 7 February 2013

Citation Wu Q, Peng S, Yu Y, Li Y, Xu Y. 2013. Genome sequence of *Bacillus licheniformis* CGMCC3963, a stress-resistant strain isolated in a Chinese traditional solid-state liquor-making process. *Genome Announc.* 1(1):e00060-12. doi:10.1128/genomeA.00060-12.

Copyright © 2013 Wu 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 Yan Xu, yxu@jiangnan.edu.cn.

Chinese mao-tai-flavor liquor is the symbolic drink of China, just as whisky is of Scotland and brandy is of France (1). It is produced through a complicated spontaneous and solid-state fermentation process, which is subjected to an extremely severe environment, such as high temperatures with acidic and ethanol stresses. It accumulates a distinctive microbial community with specific physiological properties and performance. *Bacillus licheniformis* plays an important role in liquor making, and it also represents one of the major populations of the microbial community (2).

A strain of *B. licheniformis* CGMCC3963 was isolated in this process. It was able to survive under severe conditions, including exposure to acids and ethanol, high temperature, and low water activity. It could also produce specific flavor compounds, including 2,3-butanediol, 3-hydroxy-2-butanone, 2-methylpropionic acid, 3-methylbutanoic acid, furaneol, maltol, and pyrazine (2). Such specific properties are distinct from those of the soil-dwelling type strains *B. licheniformis* 14580 and *B. licheniformis* DSM13 (3, 4).

In order to investigate its unique adaptations to various unfavorable conditions and its capacity to generate specific flavors, the genome of *B. licheniformis* CGMCC3963 was sequenced by a shotgun strategy using Illumina HiSeq 2000. This strategy produced 888 Mb of paired-end data with about 200-fold coverage of the genome. A total of 4,160,727 filtered paired-end reads were assembled into 205 contigs using Velvet v.1.0.14. Open reading frames (ORFs) were identified by Glimmer 3 (5) and were annotated using public databases, including the NCBI nonredundant (NR) database, Swiss-Prot, and Tremble. tRNA and rRNA genes were identified by tRNAscan (6) and RNAmmer (7). The alignment of the assembled genome and reference genomes was performed using Mummer (<http://mummer.sourceforge.net/manual/>).

The sequence of *B. licheniformis* CGMCC3963 is composed of a 4,525,096-bp circular chromosome, which is approximately

302,760 bp larger than that of *B. licheniformis* ATCC 14580. The G+C content is 45.21%. The chromosome contains 4,466 ORFs, 67 tRNA genes, and 10 rRNA operons, which together constitute about 91.35% of the genome. There are 411 unique genes present in this strain compared to *B. licheniformis* ATCC 14580. These unique sequences are important for the ability of this strain to survive under severe conditions and to produce specific flavor compounds.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [AMWQ000000000](https://www.ncbi.nlm.nih.gov/nuccore/AMWQ000000000). The version described in this article is the first version, [AMWQ010000000](https://www.ncbi.nlm.nih.gov/nuccore/AMWQ010000000).

ACKNOWLEDGMENTS

This work was supported by the National High Technology Research and Development Program of China (2012AA021301), the National Natural Science Foundation of China (31000806), the Cooperation Project of Jiangsu Province among Industries, Universities, and Institutes (BY2010116), the 169 Plan of Chinese liquor, and the 111 Project (no. 111-2-06).

REFERENCES

- Xu Y, Ji K. 2011. *Moutai (Maotai)*: production and sensory properties, p 315–330. In Piggott J (ed), *Alcoholic beverages: sensory evaluation and consumer research*. Woodhead Publishing, Cambridge, United Kingdom.
- Wu Q, Xu Y. 2012. Transcriptome profiling of heat-resistant strain *Bacillus licheniformis* CGMCC3962 producing *Maotai* flavor. *J. Agric. Food Chem.* 60:2033–2038.
- Rey MW, Ramaiya P, Nelson BA, Brody-Karpin SD, Zaretsky EJ, Tang M, Lopez de Leon A, Xiang H, Gusti V, Clausen IG, Olsen PB, Rasmussen MD, Andersen JT, Jørgensen PL, Larsen TS, Sorokin A, Bolotin A, Lapidus A, Galleron N, Ehrlich SD, Berka RM. 2004. Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. *Genome Biol.* 5:R77.
- Veith B, Herzberg C, Steckel S, Feesche J, Maurer KH, Ehrenreich P, Bäumer S, Henne A, Liesegang H, Merkl R, Ehrenreich A, Gottschalk G.

2004. The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. *J. Mol. Microbiol. Biotechnol.* 7:204–211.
5. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 27: 4636–4641.
 6. 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.
 7. Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent annotation of rRNA genes in genomic sequences. *Nucleic Acids Res.* 35:3100–3108.