

Complete genome sequence of *Bacillus licheniformis* HB022, an efficient aflatoxin-degrading bacterial strain

Shubo Li,¹ Si Chen,¹ Shencai Hu¹

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT The genome of *Bacillus licheniformis* HB022, which was isolated from Daqu, a Chinese liquor starter, is 4,369,776 bp long with 4,501 protein-coding genes and a GC ratio of 46.2%. Keyword-based filtering from the functional annotations led to the identification of 359 potential aflatoxin-degrading genes.

KEYWORDS aflatoxin, fermented food, *Bacillus licheniformis*, nanopore

Aflatoxins are primarily a group of toxic metabolites produced by mold strains (1, 2). Aflatoxins pose a severe threat to the safety of the food and result in substantial economic losses annually (3, 4). Microbial degradation of aflatoxins is mainly through the disruption of the toxic structural sites of aflatoxin by enzymes, including laccase, horseradish peroxidase, manganese peroxidase, etc. (5, 6), though the detailed mechanisms still remain unclarified.

To screen aflatoxin-degrading strains, different food fermentation raw materials or fermented products were sampled. Finally, a *Bacillus licheniformis* HB022 with high aflatoxin-degrading capability was screened out. The strain originated from a Daqu sample, which was a type of Chinese liquor starter at Yellow Crane Tower Distillery Co., Ltd. (30.55°N, 114.27°E).

DNA was extracted using the QIAGEN Genomic-tip Kit and randomly fragmented with G-TUBE (Covaris, USA) after quality control. The DNA fragments were then subjected to damage repair (NEBNext F PE DNA Repair Mix), end repair, and the addition of polyA at 3'-end (NEBNext Ultra End Repair/DA-Tailing Module). Addition of barcode sequences and adapters was done using Nanopore EXP-NBD104 Kit. Ligation was performed with Nanopore SQK-LSK109 Ligation Sequencing Kit 1D (PM). Then, samples were purified and quantified using Qubit fluorometer from Thermo Fisher Scientific Incorporation. The concentration of the library was greater than 8 ng/μL. Libraries were then sequenced on the PromethION platform (Oxford Nanopore Technologies) using a single R9.4 FLO-MIN106 flow cell for 48 h. Data were acquired in real-time via MinkNOW software and stored as unprocessed fast5 files. Basecalling was conducted using Guppy software (v3.2.6). We got a total of 134,095 raw reads with N50 of 24,066 bp. Adapters, short reads (<2 kb), and low-quality reads (quality score < 6) were removed, leading to a total of 94,473 reads with N50 of 24,855 bp. Reads were assembled using Canu (v1.5) and wtdbg2 (v2.2) (7), followed by CheckM (v1.1.6) analysis to evaluate the genome completeness. Circlator (v1.5.5) was used to circularize the genome. Coding genes were predicted using Prodigal (v2.6.3) (8). Pseudogenes and putative candidates were identified using GeneWise (v2.2.0) (9). tRNA genes were annotated with tRNAscan-SE (v2.0) (10), while rRNA and other non-coding RNA genes were predicted using Infernal (v1.1.3) (11). Default values of the parameters were used for the software mentioned above. The predicted proteins were subjected to BLASTp searches (*e*-value < 1e-5) against the National Center for Biotechnology Information (NCBI) Nr, Kyoto Encyclopedia of Genes and Genomes (KEGG), Pfam, and COG/eggNOG databases and Blast2GO against

Editor Silvia T. Cardona, University of Manitoba, Winnipeg, Canada

Address correspondence to Shencai Hu, hushencai@whpu.edu.cn.

The authors declare no conflict of interest.

Received 6 March 2026

Accepted 17 April 2026

Published 18 May 2026

Copyright © 2026 Li et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

TABLE 1 Genome assembly and annotation of *Bacillus licheniformis* HB022

Parameter	Value
Assembly	
Genome length (bp)	4,369,776
Raw reads	134,095
Reads	94,473
Contig	1
GC content (%)	46.2
Annotation	
Protein-coding genes	4,501
rRNAs	24
tRNAs	84
ncRNAs	101
Repeat elements	105
CRISPRs	5
Pseudogenes	3
Genomic islands	11

GO for functional annotation. Pathogenicity and antimicrobial resistance determinants were annotated by BLASTp against the databases of CAZy, TCDB, CARD, and VFDB (12–15).

With a coverage of 300×, the reads were assembled into a contig of 4,369,776 bp, which is a complete genome based on CheckM results. The genome was able to be circularized by Circlator (v1.5.5) with the parameter “minimus2 --no_pre_merge.” The start position of the genome is the gene nearest to the center of the contig. We summarized the results of assembly and annotation in Table 1.

AUTHOR AFFILIATION

¹School of Life Science and Technology, Wuhan Polytechnic University, Wuhan, China

AUTHOR ORCID*s*

Shencai Hu  <http://orcid.org/0000-0003-3106-2071>

DATA AVAILABILITY

The complete chromosome sequence of *Bacillus licheniformis* HB022 has been submitted to GenBank with accession number [JBVQOM000000000](https://www.ncbi.nlm.nih.gov/nuclseq/JBVQOM000000000) under BioProject number [PRJNA1429879](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1429879) (BioSample accession number [SAMN56256788](https://www.ncbi.nlm.nih.gov/biosample/SAMN56256788) and SRA accession number [SRR37482908](https://www.ncbi.nlm.nih.gov/sra/SRR37482908)). The supplementary data, including the annotations of genes and the list of the potential aflatoxin-degrading genes, have been deposited in Figshare with DOI <https://doi.org/10.6084/m9.figshare.31557130>.

REFERENCES

- Ellis WO, Smith JP, Simpson BK, Oldham JH. 1991. Aflatoxins in food: occurrence, biosynthesis, effects on organisms, detection, and methods of control. *Crit Rev Food Sci Nutr* 30:403–439. <https://doi.org/10.1080/10408399109527551>
- Nazhand A, Durazzo A, Lucarini M, Souto EB, Santini A. 2020. Characteristics, occurrence, detection and detoxification of aflatoxins in foods and feeds. *Foods* 9:644. <https://doi.org/10.3390/foods9050644>
- Jallow A, Xie H, Tang X, Qi Z, Li P. 2021. Worldwide aflatoxin contamination of agricultural products and foods: from occurrence to control. *Compr Rev Food Sci Food Saf* 20:2332–2381. <https://doi.org/10.1111/1541-4337.12734>
- Hove M, Van Poucke C, Njumbé-Ediagbe E, Nyanga LK, De Saeger S. 2016. Review on the natural co-occurrence of AFB1 and FB1 in maize and the combined toxicity of AFB1 and FB1. *Food Control* 59:675–682. <https://doi.org/10.1016/j.foodcont.2015.06.053>
- Guan Y, Chen J, Nepovimova E, Long M, Wu W, Kuca K. 2021. Aflatoxin detoxification using microorganisms and enzymes. *Toxins (Basel)* 13:46. <https://doi.org/10.3390/toxins13010046>
- Wang J, Ogata M, Hirai H, Kawagishi H. 2011. Detoxification of aflatoxin B1 by manganese peroxidase from the white-rot fungus *Phanerochaete sordida* YK-624. *FEMS Microbiol Lett* 314:164–169. <https://doi.org/10.1111/j.1574-6968.2010.02158.x>
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>

8. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>
9. Birney E, Clamp M, Durbin R. 2004. Genewise and genomewise. *Genome Res* 14:988–995. <https://doi.org/10.1101/gr.1865504>
10. Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol Biol* 1962:1–14. https://doi.org/10.1007/978-1-4939-9173-0_1
11. Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29:2933–2935. <https://doi.org/10.1093/bioinformatics/btt509>
12. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res* 37:D233–D238. <https://doi.org/10.1093/nar/gkn663>
13. Saier MH, Tran CV, Barabote RD. 2006. TCDB: the transporter classification database for membrane transport protein analyses and information. *Nucleic Acids Res* 34:D181–D186. <https://doi.org/10.1093/nar/gkj001>
14. Jia B, Raphenya AR, Alcock B, Wagglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FSL, Wright GD, McArthur AG. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* 45:D566–D573. <https://doi.org/10.1093/nar/gkw1004>
15. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. 2005. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* 33:D325–D328. <https://doi.org/10.1093/nar/gki008>