





# Draft Genome Sequence of an Epiphytic Strain, *Bacillus* sp. Strain WL1, Isolated from the Surface of *Nostoc flagelliforme* Colonies in Yinchuan, Ningxia, China

 Xiang Gao (高翔)<sup>a,b</sup> Tao Zheng,<sup>a</sup> Xiaolong Yuan,<sup>a</sup> Zhengke Li,<sup>a</sup>  Yongjun Wei (魏勇军)<sup>c</sup>

<sup>a</sup>School of Food and Biological Engineering, Shaanxi University of Science and Technology, Xi'an, Shaanxi Province, China

<sup>b</sup>School of Life Sciences, Central China Normal University, Wuhan, Hubei Province, China

<sup>c</sup>Laboratory of Synthetic Biology, Key Laboratory of Advanced Drug Preparation Technologies, Ministry of Education, School of Pharmaceutical Sciences, Zhengzhou University, Zhengzhou, Henan Province, China

**ABSTRACT** *Bacillus* species are Gram-positive, aerobic, spore-forming bacteria that are widely spread in soil, dust, and water. One strain, *Bacillus* sp. strain WL1, was isolated from the surface of the cyanobacterium *Nostoc flagelliforme* in Yinchuan, Ningxia, China. The draft genome sequence of this strain was 5.62 Mbp.

*Nostoc flagelliforme* is a terrestrial blue-green alga that is distributed throughout the arid and semiarid steppes of the western and northwestern parts of China (1). It exhibits a predominantly filamentous colony shape. It was previously reported that the epiphytic bacteria of *N. flagelliforme* were mainly *Bacillus* spp., including *Bacillus subtilis*, *B. licheniformis*, *B. megaterium*, *B. cereus*, and *B. circulans*. In this study, we attempted to isolate and identify those epiphytic bacteria on the surface of *N. flagelliforme* collected on the eastern side of Helan Mountain in Yinchuan, Ningxia, China (2). The epiphytic bacteria were washed from the *N. flagelliforme* filaments with sterilized water and cultivated with Luria-Bertani (LB) solid medium at 37°C. One dominant strain with a uniform shape and size was isolated. The 16S rRNA gene was amplified with the primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTGTTACGACTT-3') as previously described (3). This strain was identified as a *Bacillus* sp. based on 16S rRNA gene sequencing and named *Bacillus* sp. strain WL1.

*Bacillus* sp. strain WL1 was stored in glycerol solution (20%, vol/vol) at -80°C. The strain was streaked onto solid LB medium and cultivated at 37°C. One single *Bacillus* sp. strain WL1 colony was inoculated into liquid LB medium for shake cultivation overnight at 30°C, and the cell culture was collected by centrifugation for DNA extraction using a DNeasy PowerMax soil kit (Qiagen, Germany) (4). The DNA was used to build a 2 × 150-bp sequencing library and sequenced using the Illumina HiSeq 2500 platform at Genesky Bio-Tech Co., Ltd. (Shanghai, China); the raw reads of *Bacillus* sp. strain WL1 were filtered using Trim Galore v0.4.4 ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)). The total number of reads from the paired-end sequencing (2 × 150 bp) was 4,372,767, and a total of 658.84 Mbp of data was obtained. The sequence coverage of this genome was 117×. The draft genome sequences were assembled using SOAPdenovo v2.04 with default parameters, and the resulting assembly was used for further genomic analysis (5, 6). The gaps in the draft genome sequence were corrected using GapCloser v1.12 software (6). RNAmmer v1.2 and tRNAscan-SE v2.0 were used for rRNA and tRNA prediction, respectively (7, 8). Glimmer3 v3.02 (<http://ccb.jhu.edu/software/glimmer/index.shtml>) was used for gene prediction. Default parameters were used for all software unless otherwise specified.

A total of 420 contigs comprised the draft genome assembly of *Bacillus* sp. strain WL1. The total length of the draft genome sequence is 5,624,893 bp, with a GC content of 35.2%. The length of the largest contig is 226,802 bp, and the  $N_{50}$  contig length is

**Citation** Gao X, Zheng T, Yuan X, Li Z, Wei Y. 2021. Draft genome sequence of an epiphytic strain, *Bacillus* sp. strain WL1, isolated from the surface of *Nostoc flagelliforme* colonies in Yinchuan, Ningxia, China. *Microbiol Resour Announc* 10:e00382-21. <https://doi.org/10.1128/MRA.00382-21>.

**Editor** David Rasko, University of Maryland School of Medicine

**Copyright** © 2021 Gao 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/).

Address correspondence to Xiang Gao, [xianggao@sust.edu.cn](mailto:xianggao@sust.edu.cn), or Yongjun Wei, [yongjunwei@zzu.edu.cn](mailto:yongjunwei@zzu.edu.cn).

**Received** 13 April 2021

**Accepted** 23 June 2021

**Published** 29 July 2021

38,057 bp. A total of 5,697 genes are predicted in the draft genome sequence of *Bacillus* sp. strain WL1, and the average gene length is 804.9 bp. The gene length composed 81.1% of the whole draft genome sequence. The public version of the draft genome was annotated using PGAP v5.1 through NCBI (9), and the annotated genome sequence was deposited in the GenBank database (9). Genome-based phylogenetic analysis suggested that this strain was most similar to *B. cereus* ATCC 14579.

**Data availability.** The draft genome sequence of *Bacillus* sp. strain WL1 was deposited in GenBank under accession number [JAGKST0000000000](https://ncbi.nlm.nih.gov/nucl/JAGKST0000000000), BioProject accession number [PRJNA733760](https://ncbi.nlm.nih.gov/bioproject/PRJNA733760), BioSample accession number [SAMN19458293](https://ncbi.nlm.nih.gov/biosample/SAMN19458293), and SRA accession number [SRR14099160](https://ncbi.nlm.nih.gov/sra/SRR14099160).

## ACKNOWLEDGMENTS

This work has received support from the self-determined research funds of CCNU from the Colleges' Basic Research and Operation fund of the MOE (grant number CCNU16GD012) and the Key Project of Natural Science of Shaanxi Province, China (grant number 2020JZ-51).

## REFERENCES

1. Gao K. 1998. Chinese studies on the edible blue-green alga, *Nostoc flagelliforme*: a review. *J Appl Phycol* 10:37–49. <https://doi.org/10.1023/A:1008014424247>.
2. Gao X, Xu H, Ye S, Liang W. 2016. A proposal on the restoration of *Nostoc flagelliforme* for sustainable improvement in the ecology of arid steppes in China. *Environments* 3:14. <https://doi.org/10.3390/environments3020014>.
3. Li Y, Fu L, Li X, Wang Y, Wei Y, Tang J, Liu H. 2021. Novel strains with superior degrading efficiency for lincomycin manufacturing biowaste. *Ecotoxicol Environ Saf* 209:111802. <https://doi.org/10.1016/j.ecoenv.2020.111802>.
4. Wei Y, Wu Y, Zhang L, Zhou Z, Zhou H, Yan X. 2020. Genome recovery and metatranscriptomic confirmation of functional acetate-oxidizing bacteria from enriched anaerobic biogas digesters. *Environ Pollut* 265:114843. <https://doi.org/10.1016/j.envpol.2020.114843>.
5. Wei Y, Gao X. 2021. Complete genome sequence of *Curtobacterium* sp. strain YC1, isolated from the surface of *Nostoc flagelliforme* colonies in Yinchuan, Ningxia, China. *Microbiol Resour Announc* 10:e01467-20. <https://doi.org/10.1128/MRA.01467-20>.
6. Hurgobin B. 2016. Short read alignment using SOAP2. *Methods Mol Biol* 1374:241–252. [https://doi.org/10.1007/978-1-4939-3167-5\\_13](https://doi.org/10.1007/978-1-4939-3167-5_13).
7. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <https://doi.org/10.1093/nar/gkm160>.
8. Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences, p 1–14. *In* Kollmar M (ed), *Gene prediction: methods and protocols*. Humana Press, New York, NY.
9. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.