

Draft Genome Sequence of *Bacillus litoralis* C44, Isolated from Chinese Scholar Tree (*Sophora japonica*) Forest Soil

Weili Lyu,^{a,b} Yunlong Wu,^{a,b} Ya Liu,^{a,b} Zhitang Lyu^{a,b,c}

College of Life Sciences, Hebei University, Baoding, People's Republic of China^a; Key Laboratory of Microbial Diversity Research and Application of Hebei Province, Baoding, People's Republic of China^b; Key Laboratory of Medicinal Chemistry and Molecular Diagnosis (Hebei University), Ministry of Education, Baoding, People's Republic of China^c

***Bacillus litoralis* C44 can hydrolyze rutin to produce isoquercetin by the enzyme α -L-rhamnosidase. We report here the genome sequence and annotation result of strain C44. The genomic information will serve as references to the physiology, genetics, and evolution of this species and further genetic engineering research in this species.**

Received 6 August 2016 Accepted 9 August 2016 Published 29 September 2016

Citation Lyu W, Wu Y, Liu Y, Lyu Z. 2016. Draft genome sequence of *Bacillus litoralis* C44, isolated from Chinese scholar tree (*Sophora japonica*) forest soil. *Genome Announc* 4(5):e01059-16. doi:10.1128/genomeA.01059-16.

Copyright © 2016 Lyu 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 Zhitang Lyu, lzt325@126.com.

Bacillus litoralis is a Gram-positive, motile, endospore-forming, rod-shaped bacterium. This species was first isolated and described by Jung-Hoon Yoon in 2005 (1). However, there are no more reports about whole-genome information of this species until today. *Bacillus litoralis* C44 was isolated by Yating Zhang in 2010 (38°52'0"N 115°30'00" E). This strain can hydrolyze rutin to isoquercetin efficiently and specifically by coding α -L-rhamnosidase (2). To find the genes of α -L-rhamnosidase and probe the mechanism of enzyme specificity, a whole-genome shotgun sequencing strategy was carried out.

The whole-genome DNA was extracted from bacterial overnight cultures by a phenol-chloroform method (3) and then was sequenced by a 2 × 100-bp paired-end approach using the Illumina HiSeq TM2000 platform at Novogene (Tianjin, China). *De novo* genome assembly was performed using SOAPdenovo2 version 2.04 (4), which generated 46 contigs with an N_{50} value of 248,449 bp and a largest contig size of 771,368 bp. Then, gene prediction was performed with GeneMarkS version 3.25 (5), while tRNA was predicted with tRNAscan-SE version 1.3.1 (6), rRNA was predicted with RNAmmer version 1.2 (7), small RNA (sRNA) was predicted with Rfam version 12.0 (8), and short tandem repeats (STR) were predicted with Repeat Masker version 4.05. Finally, the gene functions were annotated into the COG (9), KEGG (10), NR databases with BLASTx, respectively, and the GO database with InterProScan 5 (11).

The final assembled genome is 5,616,296 bp and has a G+C content of 35.54%. The organism possesses 5,403 coding sequences (CDSs), which yielded a coding capacity of 4,696,614 nucleotides, 115 predicted tRNA, nine predicted sRNA, and 208 predicted STR. The coding density of the draft genome is 83.62%, with an average gene length of 869 bp. According to the annotation of the whole-genome sequence of the four databases mentioned above, there are four α -L-rhamnosidase genes predicted, which were designated ORF0334, ORF0338, ORF2809, and ORF2811, respectively. ORF0334 and ORF0338 are located on contig 1 (accession no. LWSG01000001), while both

ORF2809 and ORF2811 are located on contig 8 (accession no. LWSG01000045). Here, we announce the first genome sequence of a *Bacillus litoralis* strain.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. [LWSG00000000](https://www.ncbi.nlm.nih.gov/nuclseq/LWSG00000000/). The version described in this paper is version LWSG01000000.

ACKNOWLEDGMENTS

We are grateful to the Hebei Provincial Key Discipline of Bioengineering (1050-5030023), Biology Discipline of Hebei Provincial Level with Distinctive Characteristic (1050-5030004), and the Key Laboratory of Microbial Diversity Research and Application of Hebei Province for providing scientific research platform.

FUNDING INFORMATION

This work, including the efforts of Zhitang Lyu, was funded by Natural Science Foundation of Hebei Province (Hebei Provincial Natural Science Foundation) (C2015201155).

REFERENCES

1. Yoon JH, Oh TK. 2005. *Bacillus litoralis* sp. nov., isolated from a tidal flat of the Yellow Sea in Korea. *Int J Syst Evol Microbiol* 55:1945–1948. [http://dx.doi.org/10.1099/ijs.0.63332-0](https://doi.org/10.1099/ijs.0.63332-0).
2. Lu Z, Wang J, Lin S, Zhang Y. 2012. Degradation of rutin into isoquercitrin by *Bacillus litoralis* strain C44. *IOSR J Eng* 2:1154–1161. [http://dx.doi.org/10.9790/3021-020511541161](https://doi.org/10.9790/3021-020511541161).
3. Marmur J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J Mol Biol* 3:208–218. [http://dx.doi.org/10.1016/S0022-2836\(61\)80047-8](https://doi.org/10.1016/S0022-2836(61)80047-8).
4. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* 1:18. [http://dx.doi.org/10.1186/2047-217X-1-18](https://doi.org/10.1186/2047-217X-1-18).
5. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29:2607–2618. [http://dx.doi.org/10.1093/nar/29.12.2607](https://doi.org/10.1093/nar/29.12.2607).

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. <http://dx.doi.org/10.1093/nar/25.5.0955>.
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. <http://dx.doi.org/10.1093/nar/gkm160>.
8. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J, Finn RD. 2015. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res* 43:D130–D137. <http://dx.doi.org/10.1093/nar/gku1063>.
9. Powell S, Szklarczyk D, Trachana K, Roth A, Kuhn M, Muller J, Arnold R, Rattei T, Letunic I, Doerks T, Jensen LJ, von Mering C, Bork P. 2012. eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Res* 40:D284–D289. <http://dx.doi.org/10.1093/nar/gkr1060>.
10. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28:27–30. <http://dx.doi.org/10.1093/nar/28.1.27>.
11. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240. <http://dx.doi.org/10.1093/bioinformatics/btu031>.