

# Draft Genome Sequence of *Bacillus megaterium* BHG1.1, a Strain Isolated from Bar-Headed Goose (*Anser indicus*) Feces on the Qinghai-Tibet Plateau

Wen Wang,<sup>a,b</sup> Si-Si Zheng,<sup>a,c</sup> Hao Sun,<sup>a</sup> Jian Cao,<sup>a,c</sup> Fang Yang,<sup>a</sup> Xue-Lian Wang,<sup>a</sup> Lai-Xing Li<sup>a</sup>

Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xi'ning, China<sup>a</sup>; Center of Growth, Metabolism and Aging, College of Life Sciences, Sichuan University, Chengdu, China<sup>b</sup>; University of the Chinese Academy of Sciences, Beijing, China<sup>c</sup>

***Bacillus megaterium* is a soil-inhabiting Gram-positive bacterium that is routinely used in industrial applications for recombinant protein production and bioremediation. Studies involving *Bacillus megaterium* isolated from waterfowl are scarce. Here, we report a 6.26-Mbp draft genome sequence of *Bacillus megaterium* BHG1.1, which was isolated from feces of a bar-headed goose.**

Received 7 March 2016 Accepted 22 March 2016 Published 12 May 2016

**Citation** Wang W, Zheng S-S, Sun H, Cao J, Yang F, Wang X-L, Li L-X. 2016. Draft genome sequence of *Bacillus megaterium* BHG1.1, a strain isolated from bar-headed goose (*Anser indicus*) feces on the Qinghai-Tibet plateau. *Genome Announc* 4(3):e00317-16. doi:10.1128/genomeA.00317-16.

**Copyright** © 2016 Wang 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 Lai-Xing Li, lxli@nwipb.cas.cn.

*Bacillus megaterium* is a soil-inhabiting Gram-positive, spore-forming, and rod-shaped bacterium. This microbe was first described by Anton De Bary in 1884 (1). Due to its large cell size, *B. megaterium* is well suited for research on cell morphology, sporulation, and protein localization (2). It has also exhibited many advantages in the industrial applications for recombinant protein and vitamin production (3, 4), as well as for bioremediation (5). Beyond these traditional applications, *B. megaterium* is now gaining particular interest as probiotic supplements for use in animal feeds because of its heat stability and ability to survive the gastric barrier (6, 7). Our previous comparative study showed that the relative abundance of genus *Bacillus* was significantly higher in the gut microbiota of wild bar-headed geese compared to farmed ones (8, 9). Therefore, there is a great need to isolate some candidate probiotic strains of *Bacillus megaterium* from a wild bar-headed goose. In this direction, the work here presents the draft genome sequence of *Bacillus megaterium* BHG1.1, which was isolated from a pooled fecal sample collected from a wild bar-headed goose.

The whole-genome DNA was extracted and then sequenced using the Illumina HiSeq 4000 platform (one-lane, paired-end run [2 × 150 bp]). A total of 9,771,088 paired-end reads with lengths of 150 nucleotides of 1,465,663,200 bp were generated, which, after quality control, resulted in 9,270,858 clean reads of 1,366,096,694 bp. *De novo* assembly of these clean reads was carried out using the newly developed SPAdes algorithm version 3.6.2 (10) followed by error correction using the Bayes Hammer program (11). The reads were assembled into 92 contigs, with an  $N_{50}$  value of 1.2 Mb and the largest contig size of 2.1 Mb. Then, gene prediction was performed with Prodigal version 2.6.2 (12), while tRNA was predicted with tRNAscan-SE version 1.3.1 (13), and rRNA was predicted with RNAmmer version 1.2 (14). Subsequently, the gene functions were annotated into the KEGG (15) and COG (16) databases using BLASTx.

The genome is 6,263,758 bp long, which appears to be larger than other finished genomes of *B. megaterium*, with a G+C con-

tent of 37.20%. The most similar strain to strain BHG1.1 in databases is *B. megaterium* QM B1551 (GenBank accession no. NC\_014019.1) with an average nucleotide identity (17) value of 96.76%. This microbe possesses 6,682 genes, 6,426 coding sequences (CDSs), 112 tRNAs, and 5 rRNAs. Approximately 60.29% ( $n = 3,874$ ) of the CDSs were assigned to functional COG categories and 38.80% ( $n = 2,493$ ) were assigned a KEGG orthology (KO) number. Based on KEGG pathway analysis, most genes encode proteins involved in metabolism and environmental information processing.

Further studies are under way to determine the probiotic potential of *B. megaterium* BHG1.1, including its ability to suppress the growth of pathogenic microbes and to productively influence the immune response.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited in GenBank under the accession number [LUCO00000000](https://www.ncbi.nlm.nih.gov/nuccore/LUCO00000000). The version described in this paper is the first version, LUCO01000000.

## ACKNOWLEDGMENTS

This work was supported by the State Forestry Administration Program (no. Y31I351B01) and the National Key Basic Research Program of China (973 Program, no. 2010CB530301).

## FUNDING INFORMATION

This work, including the efforts of Lai-Xing Li, was funded by State Forestry Administration Program (Y31I351B01). This work, including the efforts of Lai-Xing Li, was funded by National Key Basic Research Program of China (2010CB530301).

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## REFERENCES

- De Bary A. 1884. Vergleichende Morphologie und Biologie der Pilze, Mycetozenen und bacterien. Wilhelm Engelmann, Leipzig, Germany.
- Christie G, Götzke H, Lowe CR. 2010. Identification of a receptor sub-

- unit and putative ligand-binding residues involved in the *Bacillus megaterium* QM B1551 spore germination response to glucose. *J Bacteriol* 192: 4317–4326. <http://dx.doi.org/10.1128/JB.00335-10>.
3. Vary PS, Biedendieck R, Fuerch T, Meinhardt F, Rohde M, Deckwer W-D, Jahn D. 2007. *Bacillus megaterium*—from simple soil bacterium to industrial protein production host. *Appl Microbiol Biotechnol* 76: 957–967. <http://dx.doi.org/10.1007/s00253-007-1089-3>.
  4. Korneli C, David F, Biedendieck R, Jahn D, Wittmann C. 2013. Getting the big beast to work—systems biotechnology of *Bacillus megaterium* for novel high-value proteins. *J Biotechnol* 163:87–96. <http://dx.doi.org/10.1016/j.jbiotec.2012.06.018>.
  5. Liu M, Luo K, Wang Y, Zeng A, Zhou X, Luo F, Bai L. 2014. Isolation, identification and characteristics of an endophytic quinclorac degrading bacterium *Bacillus megaterium* Q3. *PLoS One* 9:e108012. <http://dx.doi.org/10.1371/journal.pone.0108012>.
  6. Barbosa TM, Serra CR, La Ragione RM, Woodward MJ, Henriques AO. 2005. Screening for *Bacillus* isolates in the broiler gastrointestinal tract. *Appl Environ Microbiol* 71:968–978. <http://dx.doi.org/10.1128/AEM.71.2.968-978.2005>.
  7. Cutting SM. 2011. *Bacillus* probiotics. *Food Microbiol* 28:214–220. <http://dx.doi.org/10.1016/j.fm.2010.03.007>.
  8. Wang W, Cao J, Li JR, Yang F, Li Z, Li LX. 2016. Comparative analysis of the gastrointestinal microbial communities of bar-headed goose (*Anser indicus*) in different breeding patterns by high-throughput sequencing. *Microbiol Res* 182:59–67. <http://dx.doi.org/10.1016/j.micres.2015.10.003>.
  9. Wang W, Cao J, Yang F, Wang XL, Zheng SS, Sharshov K, Li LX. 2016. High-throughput sequencing reveals the core gut microbiome of Bar-headed goose (*Anser indicus*) in different wintering areas in Tibet. *MicrobiologyOpen* 5:287–295. <http://dx.doi.org/10.1002/mbo3.327>.
  10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
  11. Medvedev P, Scott E, Kakaradov B, Pevzner P. 2011. Error correction of high-throughput sequencing datasets with non-uniform coverage. *Bioinformatics* 27:i137–i141. <http://dx.doi.org/10.1093/bioinformatics/btr208>.
  12. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <http://dx.doi.org/10.1186/1471-2105-11-119>.
  13. 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>.
  14. 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/gkml60>.
  15. 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>.
  16. 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>.
  17. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57: 81–91. <http://dx.doi.org/10.1099/ijs.0.64483-0>.