Complete Genome Sequence of *Bifidobacterium longum* subsp. *longum* JCM7052

Isamu Yamamoto,a Yumiko Ueno,a Miho Geshi,a Yoshishige Inagaki,b Toshitaka Odamaki,c Kiyotaka Fujitad

aFaculty of Home Economics, Kobe Women’s University, Kobe, Japan  
bFaculty of Literature, Kobe Women’s University, Kobe, Japan  
cNext Generation Science Institute, Morinaga Milk Industry Co. Ltd., Zama, Kanagawa, Japan  
dFaculty of Agriculture, Kagoshima University, Kagoshima, Japan

**ABSTRACT** We report the complete genome sequence of *Bifidobacterium longum* subsp. *longum* JCM7052, isolated from human feces in Japan. This strain has the capability of growing on and utilizing gum arabic as an energy source. The complete genome is 2,273,627 bp long, with 1,929 protein-coding genes and 59.9 mol% G+C content.

*Bifidobacteria* are Gram-positive, non-spore-forming, multimorphogenic rod-shaped, and strictly anaerobic bacteria. Some species of the genus *Bifidobacterium* are used as probiotics, which are defined as live microorganisms that, when administered in adequate amounts, confer a health benefit on the host (1). Among these species, *Bifidobacterium longum* subsp. *longum* is a stable and dominant population in the human gastrointestinal tract, and the genes in its genome vary depending on the habitat, such as the types and quantity of carbohydrates ingested by the host (2). In the genus *Bifidobacterium*, some strains of *B. longum* subsp. *longum* and *B. adolescentis* have been shown to ferment gum arabic as the carbohydrate source (3). We observed that *B. longum* subsp. *longum* JCM7052, which was isolated from the feces of an adult human by Mitsuoka (4), can grow on gum arabic (5).

*B. longum* subsp. *longum* JCM7052 was grown under anaerobic conditions in *Bifidobacterium* medium (5), and DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (6). A sequence library was prepared with the TruSeq DNA sample prep kit (Illumina; FC-121-1001) as described in the manufacturer’s instructions. Sequencing was conducted using a next-generation sequencer, the Illumina HiSeq 2000 (San Diego, CA), and generated 100-bp paired-end reads. A total of 116,426,984 reads and 11,759 Mb of sequence data (~2,500× coverage) was obtained after a quality check of the sequence reads, performed using CASAVA version 1.8.1, a software provided by Illumina. A total of 36 contigs were assembled using Velvet version 1.2.10 (7), and GapCloser version 1.10 was used to close the gaps that emerged during the scaffolding process, conducted using SOAPdenovo (8). The contigs were placed in order compared with the genome sequence of *B. longum* subsp. *longum* JCM1217 (GenBank accession number AP010888) using Mauve (9, 10). Since each contig contained a part of the IS21 insertion sequence (IS) family or rRNA operon in its terminal regions, the primers for PCR were designated referring to the sequences adjacent to the IS and rRNA operon (Table 1), and PCR amplification was performed (Bio-Rad T100 thermal cycler). The amplicons obtained were subjected to sequencing by means of primer walking on an Applied Biosystems 3730 DNA sequencer using a BigDye Terminator version 3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA). The amplicon sequences were assembled into ordered contigs by visual examination. Finally, a circular sequence was obtained. This sequence was annotated using the DDBJ Fast Annotation and Submission Tool (DFAST) (11), and some products were identified using a BLAST (12) search of DDBJ service analytical tools.

Citation Yamamoto I, Ueno Y, Geshi M, Inagaki Y, Odamaki T, Fujita K. 2021. Complete genome sequence of *Bifidobacterium longum* subsp. *longum* JCM7052. Microbiol Resour Announc 10:e01193-20. https://doi.org/10.1128/MRA.01193-20.  
Editor David Rasko, University of Maryland School of Medicine  
Copyright © 2021 Yamamoto et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.  
Address correspondence to Kiyotaka Fujita, k4022897@kadai.jp.  
Received 20 October 2020  
Accepted 11 March 2021  
Published 8 April 2021
The complete genome of *B. longum* subsp. *longum* JCM7052 is 2,273,627 bp long and has a G+C content of 59.9 mol%. The genome contains 1,929 putative proteins and carries 4 RNA operons, 56 tRNAs, and 1 transfer-messenger (tmRNA). Eight gene clusters found in the genome of strain JCM7052 were not found in *B. longum* subsp. *longum* JCM1217, which is a strain that cannot assimilate gum arabic. These genes may provide useful insights into the physiological properties of strain JCM7052.

**Data availability.** This whole-genome project has been deposited in DDBJ/ENA/GenBank under accession number AP022379. The BioProject accession number is PRJD89197. The raw sequence reads have been deposited in the DRA under accession number DRA011367.

**ACKNOWLEDGMENT**

This study was supported by the education fund of Yukiyoshi Gakuen.

**REFERENCES**

1. FAO/WHO. 2006. Probiotics in food: health and nutritional properties and guidelines for evaluation. FAO Food and Nutrition Paper 85. FAO/WHO, Rome, Italy. [http://www.fao.org/3/a-a0512e.pdf](http://www.fao.org/3/a-a0512e.pdf).
2. Odamaki T, Bottacini F, Kato K, Mitsuyama E, Yoshida K, Horigome A, Xiao J-Z, van Sinderen D. 2018. Genomic diversity and distribution of *Biﬁdobacterium* longum subsp. *longum* across the human lifespan. Sci Rep 8:85.
3. Crociani F, Alessandrini A, Mucci MM, Biavati B. 1994. Degradation of complex carbohydrates by *Biﬁdobacterium* spp. Int J Food Microbiol 24:199–210. [https://doi.org/10.1016/0168-1605(94)90119-8](https://doi.org/10.1016/0168-1605(94)90119-8).
4. Matsuda 1969. Vergleichende Untersuchungen über die Biﬁdobakterien aus dem Verdauungstrakt von Menschen und Tieren (Zugleich die Beschreibung von *B. thermophilum* nov. spec. und *B. pseudolongum* nov. spec.). Zentralbl Bakteriol Orig A210:52–64.
5. Sai Shin, Yamamoto I. 2009. α-Galactosidase purified from *Biﬁdobacterium longum* JCM 7052 grown on gum arabic. J Biol Macromol 9:71–80. [http://www.jsb.jp/jbm/2009/0902_4.pdf](http://www.jsb.jp/jbm/2009/0902_4.pdf).
6. Wilson K. 2001. Preparation of genomic DNA from bacteria. Curr Protoc Mol Biol 56:2.4.1–2.4.5. [https://doi.org/10.1002/0471142727.mb0204s56](https://doi.org/10.1002/0471142727.mb0204s56).
7. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. [https://doi.org/10.1101/gr.074492.107](https://doi.org/10.1101/gr.074492.107).
8. 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 S-M, Peng S, Xiaoqian Z, Liu G, Liao Y, Li Y, Yang H, Wang J, Tam W, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascale 1:18. [https://doi.org/10.1186/2047-217X-1-18](https://doi.org/10.1186/2047-217X-1-18).
9. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403. [https://doi.org/10.1101/gr.2289704](https://doi.org/10.1101/gr.2289704).
10. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. [https://doi.org/10.1371/journal.pone.0011147](https://doi.org/10.1371/journal.pone.0011147).
11. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037–1039. [https://doi.org/10.1093/bioinformatics/btx713](https://doi.org/10.1093/bioinformatics/btx713).
12. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, Yu Y-K. 2005. Protein database searches using compositionally adjusted substitution matrices. FEBS Lett 576:210–214. [https://doi.org/10.1016/j.febslet.2005.01.048](https://doi.org/10.1016/j.febslet.2005.01.048).