

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	Complete genome sequence of probiotic <i>Lactobacillus johnsonii</i> 7409N31 isolated from a healthy Hanwoo calf
Running Title (within 10 words)	Genome sequence of <i>Lactobacillus johnsonii</i> 7409N31 from a Hanwoo calf
Author	Young Joon Oh ¹ , Jieun Lee ² , Seul Ki Lim ¹ , Min-Sung Kwon ¹ , Sulhee Lee ¹ , Sang-Pil Choi ¹ , Dohyeon Yu ³ , Yeon-su Oh ⁴ , Jinho Park ^{5*} , and Hak-Jong Choi ^{1*}
Affiliation	¹ <i>Technology Innovation Research Division, World Institute of Kimchi, Gwangju 61755, Korea</i> ² <i>Kimchi Industry Promotion Division, World Institute of Kimchi, Gwangju 61755, Korea</i> ³ <i>College of Veterinary Medicine, Gyeongsang National University, Jinju 52828, Korea</i> ⁴ <i>College of Veterinary Medicine and Institute of Veterinary Science, Kangwon National University, Chuncheon 24341, Korea</i> ⁵ <i>Veterinary Internal Medicine, College of Veterinary Medicine, Jeonbuk National University, Iksan 54596, Korea</i>
ORCID (for more information, please visit https://orcid.org)	Young Joon Oh (https://orcid.org/0000-0002-0422-4525) Jieun Lee (https://orcid.org/0000-0002-1584-9520) Seul Ki Lim (https://orcid.org/0000-0001-7371-2402) Min-Sung Kwon (https://orcid.org/0000-0001-8673-0255) Sulhee Lee (https://orcid.org/0000-0002-9481-128X) Sang-Pil Choi (http://orcid.org/0000-0001-8797-8332)

	<p>Dohyeon Yu (https://orcid.org/0000-0001-7645-6926)</p> <p>Yeon-su Oh (https://orcid.org/0000-0001-5743-5396)</p> <p>Jinho Park (https://orcid.org/0000-0001-5235-5717)</p> <p>Hak-Jong Choi (https://orcid.org/0000-0003-1185-0919)</p>
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This research was supported by Technology Development Program (Project No. 1116043-1) for Bio-industry, Ministry for Agriculture, Food and Rural Affairs, Korea and partially supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C100517111), by a research grant from the World Institute of Kimchi funded by the Ministry of Science and ICT, Korea (KE2201-1), and by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry, and Fisheries (IPET) through Agriculture Microbiome R&D Program, which is funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (918018-04-4-SB010), Korea.
Acknowledgements	Not applicable.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	<p>Conceptualization: Oh YJ, Lee J, Park J, Choi HJ.</p> <p>Data curation: Oh YJ, Lee J.</p> <p>Formal analysis: Oh YJ, Lim SK, Kwon MS, Lee S, Choi SP.</p> <p>Methodology: Oh YJ, Lee J, Lim SK.</p> <p>Software: Oh YJ, Lim SK.</p> <p>Validation: Park J, Choi HJ.</p> <p>Investigation: Oh YJ, Lee J, Lim SK.</p>

	Writing - original draft: Oh YJ, Yu D, Oh Y. Writing - review & editing: Oh YJ, Park J, Choi HJ.
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

4

5 CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Jinho Park and Hak-Jong Choi
Email address – this is where your proofs will be sent	Jinho Park: jpark@jbnu.ac.kr Hak-Jong Choi: hjchoi@wikim.re.kr
Secondary Email address	
Address	Jinho Park: Veterinary Internal Medicine, College of Veterinary Medicine, Jeonbuk National University, Iksan 54596, Republic of Korea. Hak-Jong Choi: Technology Innovation Research Division, World Institute of Kimchi, Gwangju 61755, Republic of Korea.
Cell phone number	Jinho Park, Tel: +82-10-2670-3046 Hak-Jong Choi, Tel: +82-10-3362-1984
Office phone number	Jinho Park, Tel: +82-63-850-0949 Hak-Jong Choi, Tel: +82-62-610-1729
Fax number	Jinho Park, Fax: +82-63-850-0910 Hak-Jong Choi, Fax: +82-62-610-1853

6

7

8 **Abstract**

9 *Lactobacillus johnsonii* 7409N31 was isolated from the feces of a healthy 11-day-old Hanwoo calf from a farm in
10 Geochang-gun, Gyeongsangnam-do, Korea. The genome of the strain was completely sequenced using the PacBio
11 RSII sequencing system, and it was confirmed that it was composed of one circular chromosome. The size of the
12 entire genome was 2,198,442 bp, and it had 35.01 mol% G + C content and 2,222 protein-coding sequences, 24
13 rRNA, 3 ncRNA, and 112 tRNA genes. Strain 7409N31 possessed genes encoding enzymes involved in the
14 hydrolysis of both fibrous and non-fibrous carbohydrates. These data provide a comprehensive theoretical
15 understanding for developing industrial probiotic feed additives that improve nutrient digestibility.

16
17 **Keywords:** *Lactobacillus johnsonii*, Complete genome sequence, Hanwoo calf, Probiotics

18
19
ACCEPTED

Lactic acid bacteria (LAB) have been associated with the fermentation and preservation of food since ancient times and are one of the most important groups of industrial microorganisms with a multi-billion dollar market. They are naturally found in fermented foods as well as in human and animal cavities, including the gastrointestinal tract [1-2]. LAB colonizes the intestines of animals as a part of the normal intestinal flora, inhibit the growth of harmful bacteria, prevent diarrhea, and inhibit the absorption of toxic substances [3]. Ruminants utilize intestinal microbes to break down cellulose into glucose so that it can be used as an energy source, and some LAB are known to use cellulose as a carbon source [4]. Ruminant roughage contains both fibrous (cell wall material) and non-fibrous (cell content) carbohydrates, but most of them are composed of fibers such as cellulose, hemicellulose, and pectin [5].

In the present study, *Lactobacillus johnsonii* 7409N31 (KCCM 13026P) was isolated from the feces of a healthy 11-day-old Hanwoo calf. Strain 7409N31 was anaerobically grown in deMan, Rogosa, and Sharpe (MRS, Difco, Franklin Lakes, NJ, USA) medium at 35 °C for 24 h. Genomic DNA of 7409N31 strain was extracted as described previously [6]. The complete genome of *L. johnsonii* 7409N31 was sequenced by using Pacific Biosciences (PacBio) reagents (DNA Link Inc., Seoul, Korea). A total number of 735,552 reads with a mean subread length of 5,638 bases (N50, 6,459 bases) were obtained with PacBio sequencing. These sequences were assembled *de novo* using the Hierarchical Genome Assembly Process (HGAP, version 3.0) workflow [7]. The genome of the *L. johnsonii* 7409N31 was annotated using the NCBI Prokaryotic Genome Annotation Pipeline and the Pathosystems Resource Integratin Center (PATRIC, version 3.6.12) genome data base [8]. Genome assembly completeness was evaluated using Benchmarking Universal Single-Copy Orthologous suite (BUSCO, version 5.1.3) [9], and evolutionary genealogy of genes: Non-supervised Orthologous Group-mapper (EggNOG-mapper, version 2.0, <http://eggno-mapper.embl.de>) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (www.genome.jp/kegg) were used for functional annotation. The complete genome of strain 7409N31 has one circular chromosome (2,198,442 bp), with a G + C content of 35.01 mol%, 2,222 protein-coding sequences, 24 rRNA, 3 ncRNA, and 112 tRNA genes (Fig. 1 and Table 1). Further analysis showed that about 80% of the protein-coding genes (1,778 proteins) could be clustered in 21 functional categories of cluster of orthologous groups (COG) (Fig. 2). Most of the proteins of strain 7409N31 were classified into categories related to 'core functions' such as replication, recombination and repair (category L: 10.2%), transcription (category K: 8.9%), translation, ribosomal structure and biogenesis (category J: 8.7%), and carbohydrate transport and metabolism (category G: 7.5%). This suggests that strain 7409N31 has a relatively strong carbohydrate metabolism ability similar to other strains in the genus *Lactobacillus* [10].

The genome of *L. johnsonii* 7409N31 possessed the *bcsZ* gene that encodes an endoglucanase involved in the hydrolysis of the D-glycosidic bond of cellulose. In addition, the presence of genes such as beta-fructofuranosidase (*sacA*), cellobiose PTS components (*celB* and *chbC*), and oligo-1,6-glucosidase (IMA, *malL*) involved in the metabolism of non-fibrous carbohydrates such as starch was confirmed. These results suggest that *L. johnsonii* strain 7409N31 has the potential to be developed as an industrial probiotic feed additive because it can facilitate digestion of both fibrous and non-fibrous carbohydrates.

Nucleotide sequence accession number

The GenBank accession number for the genome of *L. johnsonii* strain 7409N31 is CP084221.

ACCEPTED

- 61 1. Tannock GW. Analysis of the intestinal microflora, a renaissance. *Antonie van Leeuwenhoek*.
62 1999;76:265-78.
- 63 2. Vaughan EE, de Vries MC, Zoetendal EG, Ben-Amor K, Akkermans AD, de Vos WM. The intestinal
64 LABs. *Antonie van Leeuwenhoek*. 2002;82:341-352.
- 65 3. Shahani, KM, Ayebo AD. Role of dietary lactobacilli in gastrointestinal microecology. *Am J Clin Nutr*.
66 1980;33:2448-57. <https://doi.org/10.1093/ajcn/33.11.2448>.
- 67 4. Matthews C, Crispie F, Lewis E, Reid M, O'Toole PW, Cotter PD. The rumen microbiome: a crucial
68 consideration when optimising milk and meat production and nitrogen utilisation efficiency. *Gut Microbes*.
69 2019;10:115-132. <https://doi.org/10.1080/19490976.2018.1505176>.
- 70 5. Fontaine AS, Bout S, Barriere Y, Vermerris W. Variation in cell wall composition among forage maize
71 (*Zea mays* L.) inbred lines and its impact on digestibility: Analysis of neutral detergent fibre composition
72 by pyrolysis-gas chromatography-mass spectrometry. *J Agric Food Chem*. 2003;51:8080-7.
73 <https://doi.org/10.1021/jf034321g>.
- 74 6. Oh YJ, Kim JY, Lee J, Lim SK, Yu D, Oh YS, Park J, Choi HJ. Complete genome sequence of
75 *Lactobacillus amylovorus* 1394N20, a potential probiotic strain, isolated from a Hanwoo calf. *J Anim Sci*
76 *Technol*. 2021;63:1207-10. <https://doi.org/10.5187/jast.2021>.
- 77 7. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial
78 genome assemblies from long-read SMRT sequencing data. *Nat Methods*. 2013;10:563-9.
79 <https://doi.org/10.1038/nmeth.2474>.
- 80 8. Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, et al. The PATRIC Bioinformatics
81 Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res*. 2020;48:D606–12.
82 <https://doi.org/10.1093/nar/gkz943>.
- 83 9. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome
84 assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31:3210-3212.
85 <https://doi.org/10.1093/bioinformatics/btv351>.
- 86 10. Boekhorst J, Siezen RJ, Zwahlen MC, Vilanova D, Pridmore RD, Mercenier A, et al. The complete
87 genomes of *Lactobacillus plantarum* and *Lactobacillus johnsonii* reveal extensive differences in
88 chromosome organization and gene content. *Microbiology*. 2004;150:3601-3611.
89 <https://doi.org/10.1099/mic.0.27392-0>.

Table and Figure

Table 1. Genomic features of *Lactobacillus johnsonii* strain 7409N31

Properties	Value
BioProject	PRJNA766157
BioSample	SAMN21619988
Accession No.	CP084221
Sequencing method	PacBio RSII
Assembly method	HGAP version 3.0
Genome size (bp)	2,198,442
Contig	1
Total CDSs	2,222
rRNA genes	24
tRNA genes	112
G + C content (mol%)	35.01

HGAP, hierarchical genome assembly process; CDSs, coding sequences; G + C, guanine + cytosine

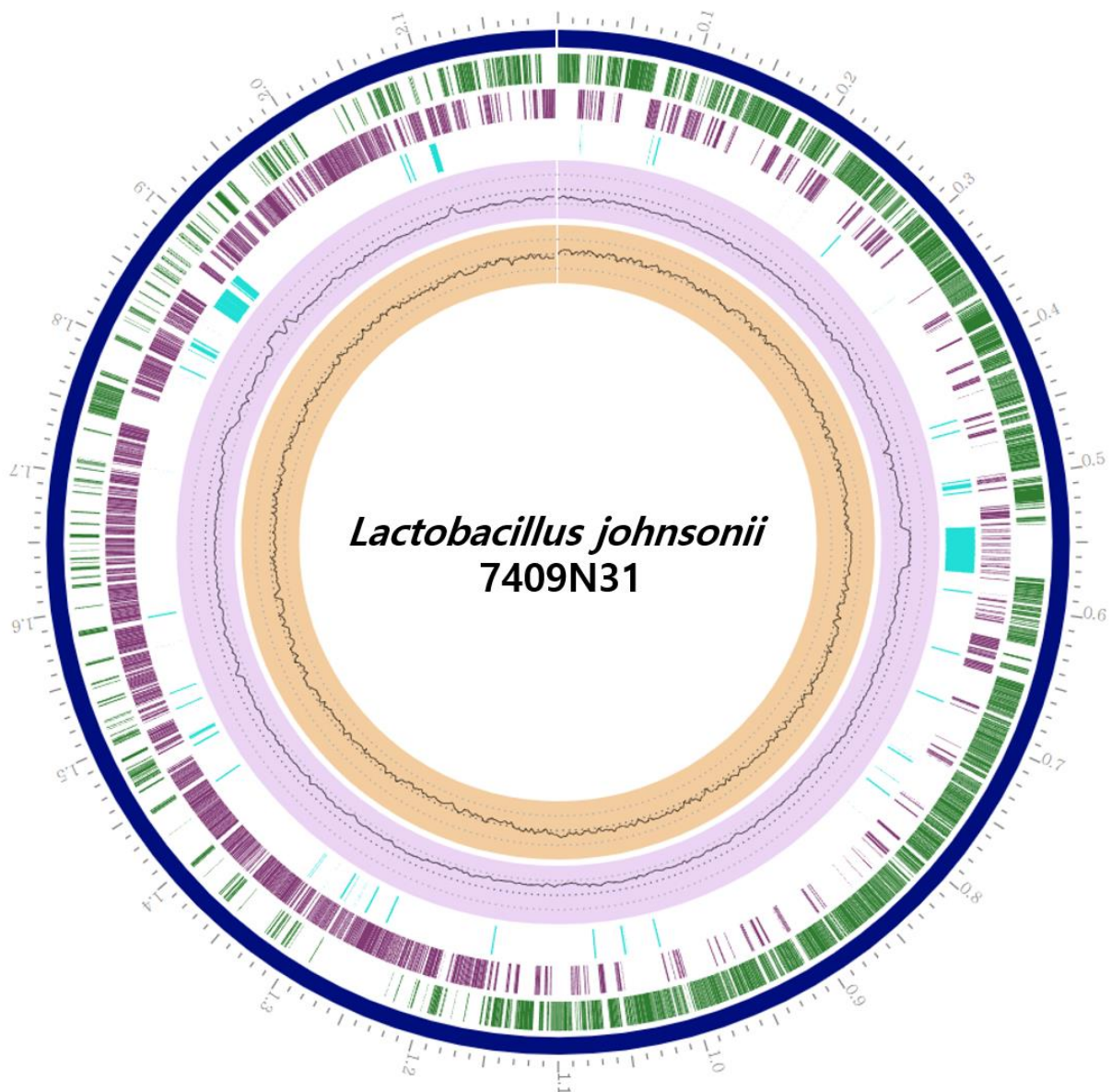


Figure 1. Circular view of the genome of *Lactobacillus johnsonii* 7409N31 showing the physical map of its significant features generated using PATRIC. From outside to inside: contigs (blue), annotated reference genes (particularly, coding sequences [CDSs]) on the forward strand (green), and annotated reference genes on the reverse strand (purple). The fourth circle shows non-CDSs featured in the genome (light blue). The next circle indicates guanine–cytosine (GC) content (lavender/light purple), while the innermost circle indicates the GC skew (peach).

COG functional category

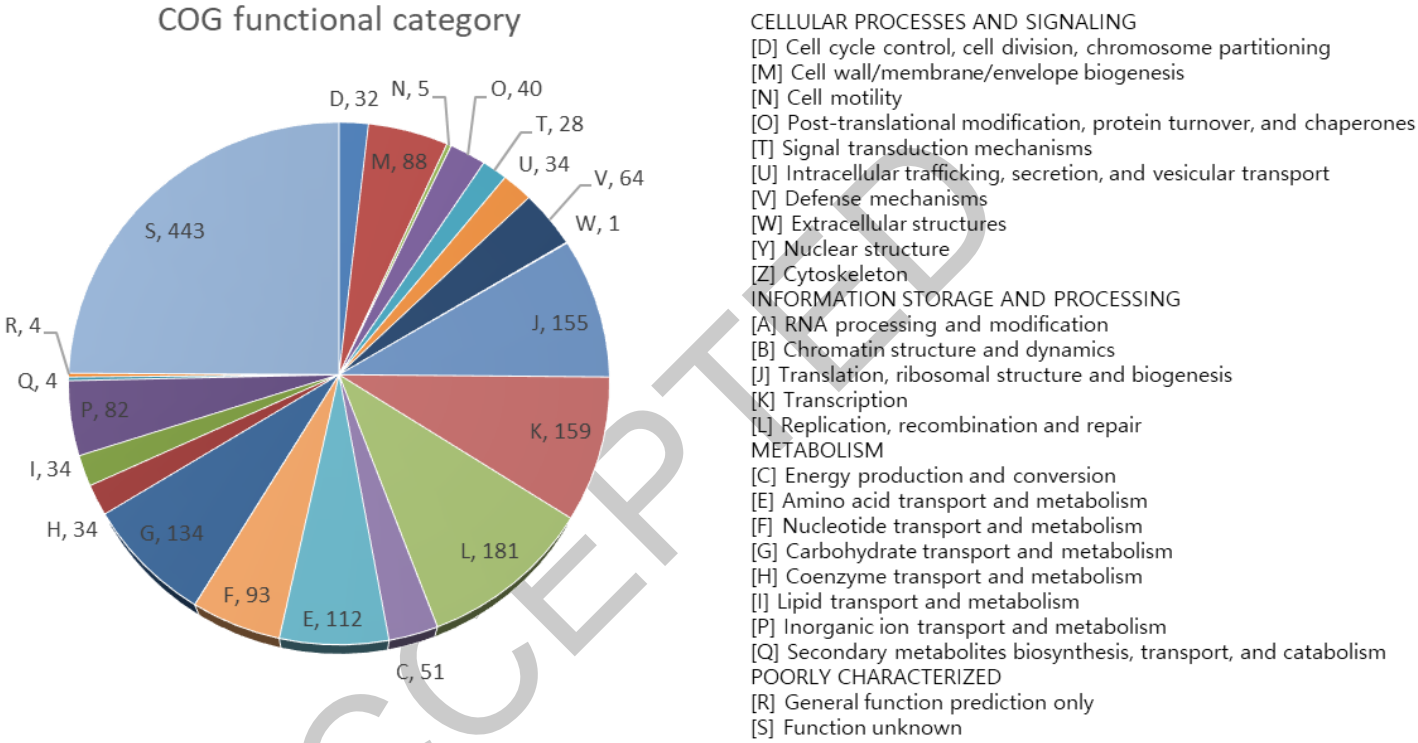


Figure 2. COG functional annotation of *Lactobacillus johnsonii* 7409N31.