1 **JAST (Journal of Animal Science and Technology) TITLE PAGE**

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Genome Announcement
Article Title (within 20 words without abbreviations)	Complete genome sequence of candidate probiotic *Limosilactobacillus fermentuum* KUFM407
Running Title (within 10 words)	Genome of potential probiotic *Limosilactobacillus fermentuum* KUFM407
Author	Bogun Kim[1], Ji yu Heo[1], Xiaoyue Xu[1], Hyunju Lee[1], Duleepa Pathiraja[1], Jae-Young Kim[1, 2], Yi Hyun Choi[1], In-Geol Choi[1], Sae Hun Kim[1, 2]
Affiliation	[1] College of Life Sciences and Biotechnology, Korea University, Seoul 02841, Republic of Korea [2] Institute of Life Science and Natural Resources, Korea University, Seoul 02841, Republic of Korea
ORCID (for more information, please visit https://orcid.org)	Bogun Kim (https://orcid.org/0000-0002-5493-1674) Ji yu Heo (https://orcid.org/0009-0009-5334-0125) Xiaoyue Xu (https://orcid.org/0000-0001-9654-8676) Hyunju Lee (https://orcid.org/0009-0000-5423-4812) Duleepa Pathiraja (https://orcid.org/0000-0001-6239-5958) Jae-Young Kim (https://orcid.org/0000-0003-1937-9535) Yi Hyun Choi (https://orcid.org/0009-0006-0513-1503) In-Geol Choi (https://orcid.org/0000-0001-7403-6274) Sae Hun Kim (https://orcid.org/0000-0002-0990-2268)
Competing interests	No potential conflict of interest to report.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through High Value-added Food Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA)(321034053HD020) and supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry, and Fisheries, funded by the Ministry of Agriculture, Food, and Rural Affairs (MAFRA)(32136051SB010 & 1545027002).
Acknowledgements	Not applicable.
Availability of data and material	The complete genome sequence has been deposited in the NCBI GenBank under the accession number GCA_030290995.1. The BioProject accession number is PRJNA981335. The BioSample accession number is SAMN35673550.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Sae Hun Kim, In-Geol Choi Data curation: Bogun Kim, Ji yu Heo, Duleepa Pathiraja, Xiaoyue Xu, Sae Hun Kim, In-Geol Choi Formal analysis: Bogun Kim, Ji yu Heo, Duleepa Pathiraja, Xiaoyue Xu Methodology: Bogun Kim, Ji yu Heo, Xiaoyue Xu, Duleepa Pathiraja Software: Bogun Kim, Duleepa Pathiraja, Xiaoyue Xu Validation: Bogun Kim, Ji yu Heo, Xiaoyue Xu, Hyunju Lee, Duleepa Pathiraja, Jae-Young Kim, Yi Hyun Choi Investigation: Ji yu Heo, Xiaoyue Xu, Hyunju Lee, Duleepa Pathiraja, Jae-Young Kim, Yi Hyun Choi Writing - original draft: Ji yu Heo, Bogun Kim, Yi Hyun Choi, Jae-Young Kim

	Writing - review & editing: Bogun Kim, Ji yu Heo, Xiaoyue Xu, Hyunju Lee, Duleepa Pathiraja, Jae-Young Kim, Yi Hyun Choi, Sae Hun Kim, In-Geol Choi
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

2

3 **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	Sae Hun Kim
Email address – this is where your proofs will be sent	saehkim@korea.ac.kr
Secondary Email address	saehkim@gmail.com
Address	College of Life Sciences and Biotechnology, Korea University, Seoul, 02841, Korea
Cell phone number	010-9071-3055
Office phone number	02-3290-3055
Fax number	3290-3499
First name, middle initial, last name	In-Geol Choi
Email address – this is where your proofs will be sent	igchoi@korea.ac.kr
Secondary Email address	igchoi@gmail.com
Address	Department of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul, 02841, Korea
Cell phone number	010-7448-3289
Office phone number	02-3290-3152
Fax number	

4

Genome Announcement

Complete genome sequence of candidate probiotic *Limosilactobacillus fermentuum* KUFM407

Title page: Genome of potential probiotic *Limosilactobacillus fermentuum* KUFM407

Bogun Kim[a], Ji yu Heo[a], Xiaoyue Xu[a], Hyunju Lee[a], Duleepa Pathiraja[a], Jae-Young Kim[a, b], Yi Hyun Choi[a], In-Geol Choi[a], Sae Hun Kim[a, b]

[a] College of Life Sciences and Biotechnology, Korea University, Seoul 02841, Republic of Korea

[b] Institute of Life Science and Natural Resources, Korea University, Seoul 02841, Republic of Korea

Abstract

It has been reported that the administration of *Limosilactobacillus fermentum* alleviates diseases such as osteoporosis and colitis. In this study, we report the complete genome sequence of *Limosilactobacillus fermentum* KUFM407, a probiotic strain of LAB isolated from Korean traditional fermented food, Kimchi. Whole genome sequencing of *L. fermentum* KUFM407 was performed on the Illumina MiSeq and Oxford Nanopore MinION platform. The genome consisted of one circular chromosome (2,077,616 bp) with a GC content of 51.5% and one circular plasmid sequence (13,931 bp). Genome annotation identified 1,932 protein-coding genes, 15 rRNAs, and 58 tRNAs in the assembly. The function annotation of the predicted proteins revealed genes involved in the biosynthesis of bacteriocin and fatty acids. The complete genome of *L. fermentum* KUFM407 could provide valuable information for the development of new probiotic food and health supplements.

Keywords: *Limosilactobacillus fermentum*; KUFM407; complete genome sequence; probiotics;

Limosilactobacillus fermentum has been widely used in the fermentation of various foods and is considered a strain with high probiotic potential. Probiotics are "live microorganisms that, when administered in adequate amounts, confer a health benefit on the host" [1]. Strains of *L. fermentum* have high survival rates in the gastrointestinal tract. They strongly attach to enterocytes and produce antimicrobial compounds. In addition, *L. fermentum* has been shown to benefit host and human health by regulating immune responses and improving intestinal health.

For lactic acid bacteria (LAB) to act as functional probiotic strains, properties such as the ability to adhere to mucosal surfaces and resistance to low pH and high bile concentrations are required [2]. For acid tolerance confirmation, 0.1 mL aliquots of each active culture were inoculated in 10 mL De Man–Rogosa–Sharpe (MRS) broth (BD Co., Franklin Lakes, USA) broth acidified to pH 2.5 and supplemented with 1000 U mL−1 of porcine pepsin (Sigma-Aldrich, St. Louis, MO, USA). The samples were then incubated at 37 °C for 3h. To determine bile salt tolerance, 0.1 mL aliquots of each active culture were inoculated in 10 mL MRS broth containing 0.3% oxgall bile salt (Sigma-Aldrich, St. Louis, MO, USA) and incubated at 37 °C for 24h. Following incubation, cell suspensions were spread on MRS agar plates, and viable cell counts were determined through plate counting methods. *L. fermentum* KUFM407 (KUFM407) showed high stability against acid and bile salts (Table 1).

Table 1. Acid and bile tolerance of *Limosilactobacillus fermentum* KUFM407 (log CFU/ml)

	Initial mean counts	Resistant to gastric juice (3h)	Bile tolerance (24h)
KUFM407	7.71±0.10	7.45±0.06*	8.52±0.08*
Strain A	7.30±0.08	7.24±0.14	4.55±0.04*
L. rhamnosus GG	7.02±0.06	6.97±0.07	8.26± 0.23*

Each value represents mean ± standard deviation (SD) from three trials (log CFU/ml); Viable counts (log CFU/ml) of each strain at 3h, 24h were compared with initial counts (0h); *p<0.05 (Student's t-test, two tailed).

KUFM407 obtained from the Food Microbiology Laboratory, Division of Food Bioscience and Technology, Korea University (Seoul, South Korea) was cultivated in MRS broth for 24h at 37°C and sub-cultured three times before the extraction of genomic DNA (gDNA).

Subcultured strains were washed three times with PBS buffer, and 1 mL aliquots of the washed strains were adjusted to the OD600 range of 1.0 to 2.0. Exgene™ Cell SV (Geneall, Seoul, Korea) was used to extract gDNA after gram-positive bacteria-specific pretreatment. The presence of a single strain of KUFM407 was confirmed by gel electrophoresis and 16S rRNA sequencing.

The extracted gDNA was prepared for short-read sequencing using an Illumina® DNA Prep Kit (Illumina, San Diego, CA, USA). Short-read sequencing was performed on an Illumina MiSeq sequencer using the Illumina MiSeq® Reagent Kit v3 (Illumina), resulting in paired-end reads of 300 base pairs (bp) in length. A long-read sequencing library was prepared using an Oxford Nanopore Ligation Sequencing Kit (Oxford Nanopore, Oxford, UK). Long-read sequencing was performed on a MinION sequencing device (Oxford Nanopore) using an R9.4.1 flow cell (Oxford Nanopore). Illumina short-read sequencing yielded 1,699,990 paired-end reads (419,571,925 bp), and Oxford Nanopore long-read sequencing produced 53,365 reads totaling 298,111,808 bp.

The draft genome sequence was constructed from the long reads using Flye assembler (v. 2.9.2) [3] after two polishing iterations. Adapter sequences were removed, and short reads were quality controlled using TrimGalore (v. 0.6.7) [4] in paired-end mode. The quality of the draft genome assembly was improved by error correction with PolyPolish (v. 0.5.0) [5] using quality controlled short reads. Genome and functional annotations of predicted genes were performed using the Prokaryotic Genome Annotation Pipeline (v. 6.4) [6]. Genome completeness was assessed with BUSCO (v. 5.4.6) [7] using the Lactobacillales _odb10 dataset. Default parameters were used for all software unless otherwise noted.

The complete genome sequence of KUFM407 consisted of a circular chromosome (2,077,616 bp) with a G+C ratio of 51.5% and a circular plasmid sequence of 13,931 bp (Table 2). The genome was 99.7% complete. A total of 2,143 genes, including 1,932 protein-coding, 15 rRNA, and 58 tRNA genes, and 135 pseudogenes were predicted in the genome sequence (Fig. 1). Biological functions were assigned to 1,729 (89.5%) of the protein-coding genes. The most assigned proteins were associated with replication, recombination and repair; amino acid transport and metabolism; translation, ribosomal structure and biogenesis; transcription; and carbohydrate transport and metabolism (207, 170, 155, 137, 123 genes, respectively).

78 Table 2. Genome features of *Limosilactobacillus fermentum* KUFM407

	Length (bp)	GC (%)	Depth	CDSs	tRNA	rRNA
Chromosome	2,077,616	51.5	122.0	1,920	58	15
Plasmid	13,931	40.5	36.0	12	0	0
Total	2,091,547	51.4	121.4	1,932	58	15

79 bp, base pair; G, guanine; C, cytosine; tRNA, transfer RNA; rRNA, ribosomal RNA

80 In the plasmid sequence of KUFM407, four genes (*garQ, garI, garC, garD*) known to be involved in the
81 production of the garvicin Q family class II bacteriocin were found [8]. Also, fatty acid biosynthetic gene cluster
82 was identified in the chromosome and short-chain fatty acids such as acetate, propionate, and butyrate produced
83 by gut microbes are known to have anti-inflammatory effects [9] (Table 3). The genomic information of *L.*
84 *fermentum* KUFM407 could provide insights for future research on the characteristics of this strain as a
85 functional food and health supplement.

86

87 Table 3. Genes with biosynthetic functions in *Limosilactobacillus fermentum* KUFM407

Predicted function	Gene name	Functions	Gene position	Length (aa)
Bacteriocin production				
	garQ	Garvicin Q family class II bacteriocin	p(7,514 -7,720)	68
	garI	Bacteriocin immunity protein	p(7,720 – 8,031)	103
	garC	Peptide cleavage/export ABC transporter	p(8,421 – 10,580)	719
	garD	Bacteriocin secretion accessory protein	p(10,591 – 11,967)	458
Fatty acid biosynthetic gene cluster				
	fabZ1	3-hydroxyacyl-ACP dehydratase	183,265 - 183,711	148
	marR	MarR family transcriptional regulator	183,796 - 184,236	146
	fabH	Ketoacyl-ACP synthase III	184,260 - 185,219	319
	accP	Acyl carrier protein	185,248 - 185,496	82
	fabD	ACP S-malonyltransferase	185,496 - 186,443	315
	fabG	3-oxoacyl-ACP reductase FabG	186,427 - 187,158	243
	fabF	Beta-ketoacyl-ACP synthase II	187,171 - 188,409	412
	accB	Acetyl-CoA carboxylase biotin carboxyl carrier protein	188,412 - 188,858	148
	fabZ2	3-hydroxyacyl-ACP dehydratase	188,861 - 189,295	144
	accC	Acetyl-CoA carboxylase biotin carboxylase subunit	189,316 - 190,713	465
	accD	Acetyl-CoA carboxylase carboxyltransferase subunit beta	190,682 - 191,530	282
	accA	Acetyl-CoA carboxylase carboxyltransferase subunit alpha	191,523 - 192,296	257
	fabI	Enoyl-ACP reductase	192,314 - 193,078	254

88 aa, amino acid; p, plasmid
89

Data Availability

The complete genome sequence has been deposited in NCBI GenBank under accession number GCA_030290995.1. The BioProject accession number is PRJNA981335 and the BioSample accession number is SAMN35673550.

Acknowledgements

Not applicable.

FUNDING

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through High Value-added Food Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA)(321034053HD020) and supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry, and Fisheries, funded by the Ministry of Agriculture, Food, and Rural Affairs (MAFRA)(32136-05-1-SB010, 122030-3).

References

1. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, et al. The International Scientific Association for probiotics and prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol. 2014;11:506–14. https://doi.org/10.1038/nrgastro.2014.66

2. Bao Y, Zhang Y, Zhang Y, Liu Y, Wang S, Dong X, et al. Screening of potential probiotic properties of Lactobacillus fermentum isolated from traditional dairy products. Food Control. 2010;21:695–701. https://doi.org/10.1016/j.foodcont.2009.10.010

3. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37:540–6. https://doi.org/10.1038/s41587-019-0072-8

4. Krueger F, James F, Ewels P, Afyounian E, Schuster-Boeckler B. FelixKrueger/TrimGalore. v.0.6.7-DOI via Zenodo. Zenodo; 2021

5. Wick RR, Holt KE. Polypolish: Short-read polishing of long-read bacterial genome assemblies. PLOS Comp Biol. 2022;18:e1009802. https://doi.org/10.1371/journal.pcbi.1009802

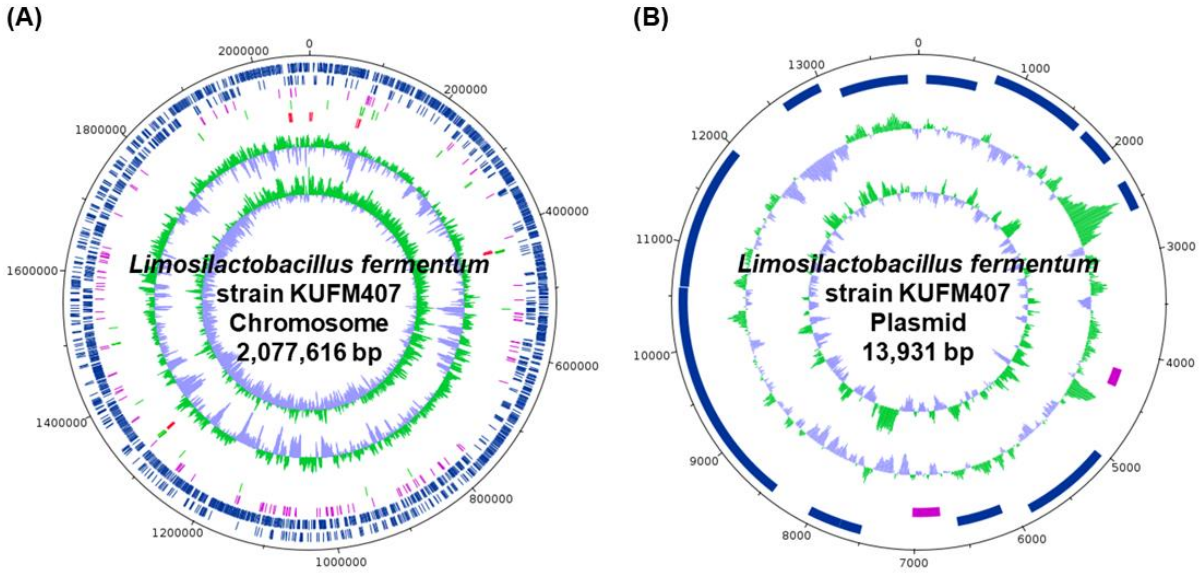
6. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016;44:6614–24. https://doi.org/10.1093/nar/gkw569

7. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021;38:4647–54. https://doi.org/10.1093/molbev/msab199

8. Desiderato CK, Hasenauer KM, Reich SJ, Goldbeck O, Holivololona L, Ovchinnikov KV, et al. Garvicin Q: characterization of biosynthesis and mode of action. Microb Cell Fact. 2022;21:1-8. https://doi.org/10.1186/s12934-022-01952-9

9. Rivière A, Selak M, Lantin D, Leroy F, De Vuyst L. Bifidobacteria and butyrate-producing colon bacteria: importance and strategies for their stimulation in the human gut. Front Microbio. 2016;7:979. https://doi.org/10.3389/fmicb.2016.00979

129 Figure 1



130

131 Figure 1. Circular chromosome and plasmid maps of *Limosilactobacillus fermentum* KUFM407. Marked
132 features are shown from the periphery to the center; protein-coding genes (forward strand), protein-coding genes
133 (reverse strand), pseudogenes, tRNA, rRNA, GC content, and GC skew. (A) Chromosome, (B) Plasmid. bp:
134 base pair.

135

136