



Complete Genome Sequence of *Lacticaseibacillus paracasei* Strain NSMJ15

Ji Young Jung,^a Kook-il Han,^a Hyun Mi Jin,^a Young Ho Nam,^a Hye Kyeong Kang,^a Mi Hwa Lee^a

^aMicrobial Research Department, Nakdonggang National Institute of Biological Resources (NNIBR), Sangju-si, Gyeongsangbuk-do, South Korea

ABSTRACT We report here the complete genome sequence of *Lacticaseibacillus paracasei* NSMJ15, isolated from makgeolli (a traditional Korean fermented liquor) and shown to have potentially probiotic characteristics. The genome consisted of a 2.79-Mbp chromosome contig and four plasmids having a total of 2,947 genes, including 2,690 coding sequences.

Probiotics are generally considered a means to restore the balance of gut microbes and promote beneficial functions of the gut (1, 2). Lactic acid bacteria (LAB) are a major group of probiotic microorganisms in both humans and animals and are being extensively investigated in studies on probiotics (1–4). *Lacticaseibacillus paracasei* (formerly “*Lactobacillus paracasei*”) strain NSMJ15 was isolated from makgeolli, a traditional Korean fermented liquor, and *in vitro* study showed its probiotic potential, in particular, its antimicrobial properties against gut pathogens (5). Here, we report the complete genome sequence of strain NSMJ15, which supports its characteristics as a potential probiotic candidate.

Strain NSMJ15 was first isolated by serial dilutions of a makgeolli suspension and plating onto MRS agar (Difco, USA) (5). For whole-genome sequencing, genomic DNA (gDNA) was extracted from strain NSMJ15 grown in MRS broth (Difco) at 30°C for 48 h using a Maxwell 16 DNA purification kit (Promega, USA). The gDNA was sequenced using the PacBio RS II platform and Illumina HiSeq X ten platform (2 × 151 bp) at Macrogen, Inc. (South Korea). Intact gDNA was sheared to approximately 20 kb using a g-TUBE device (Covaris, Inc., USA) and purified using AMPure PB magnetic beads, and a sequencing library was constructed using the PacBio SMRTbell template prep kit v1.0 (PacBio). For the Illumina sequencing, 100 ng gDNA was sheared using an LE220 focused ultrasonicator (Covaris, Inc.), and a sequencing library with a mean size of 350 bp was prepared using the TruSeq Nano DNA library prep kit (Illumina). A total of 50,575 subreads (0.60 Gbp; coverage, 202.53-fold; mean subread length, 11,845 bp; N_{50} , 16,212 bp) generated using the PacBio RS II platform were used for *de novo* genome assembly using HGAP v3 (6). The overlap of both ends of each contig was checked to determine whether it was circular, and the overlapping ends were trimmed. A total of 5,902,004 quality-filtered Illumina paired-end reads (0.89 Gbp; coverage, 301.27-fold), of which ≥90% bases in each read had a Phred quality score of ≥30, were used for error correction with Pilon v1.21 to obtain the final genome assembly (7). Default parameters were used except where otherwise noted.

The final genome assembly had a mean sequencing depth of 105-fold and a GC content of 46.40%. It consisted of a 2,791,177-bp circular chromosome and four plasmids (pLPN-1, pLPN-2, pLPN-3, and pLPN-4). The genome assembly and annotation statistics are shown in Table 1. Average nucleotide identity (ANI) analysis was performed using OrthoANu (8) to identify strain NSMJ15. It shared 98.34% sequence similarity with *Lacticaseibacillus paracasei* subsp. *paracasei* JCM8130^T (GenBank accession numbers AP012541 to AP012543). This value was higher than the ANI threshold range of 95 to 96% (9), indicating that strain NSMJ15 belongs to the species *Lacticaseibacillus paracasei*.

Editor Julia A. Maresca, University of Delaware

Copyright © 2021 Jung 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 Ji Young Jung, jjjung@nnibr.re.kr.

The authors declare no conflict of interest.

Received 20 October 2021

Accepted 19 November 2021

Published 9 December 2021

TABLE 1 Summary of assembly and annotation statistics for *Lacticaseibacillus paracasei* strain NSMJ15

Genetic element	Genome size (bp)	GC content (%)	No. of coding sequences	No. of rRNAs	No. of tRNAs	Sequencing depth (×)	GenBank accession no.
Chromosome	2,791,177	46.62	2,524	15	60	102	CP049324
pLPN-1	51,668	43.93	60	0	0	264	CP049325
pLPN-2	50,646	43.60	48	0	0	156	CP049326
pLPN-3	46,080	41.26	45	0	0	114	CP049327
pLPN-4	18,602	41.33	13	0	0	88	CP049328
Total	2,958,173	46.40	2,690	15	60	105	

Genome annotation using NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.11 (10) identified 2,690 protein-coding genes, 15 rRNA genes, 60 tRNA genes, 3 noncoding RNAs (ncRNAs), and 179 pseudogenes. Data obtained from the RAST annotation server (11) showed that the genome contained 3,075 coding DNA sequences (CDSs) and 231 subsystems. Genes for metabolism of carbohydrates showed the highest prevalence (17.0%), followed by genes for protein metabolism (13.2%) and amino acid and derivatives (12.1%). Several cell surface proteins (classes A and C sortase), LPXTG-motif cell wall anchor domain proteins, and α -alanyl-lipoteichoic acid biosynthesis proteins (dltABCD) were detected in the strain NSMJ15 genome, which suggests its potential to adhere to intestinal epithelial cells (12). In particular, putative bacteriocin-encoding gene clusters were identified *in silico* using the BAGEL4 software tool (13), which revealed two areas of interest (AOIs) in the chromosome and four open reading frames (ORFs) encoding the core peptides of bacteriocin. Strain NSMJ15 also contained genes encoding ABC transporters and bacteriocin immunity for bacteriocin export and self-immunity to bacteriocin, respectively.

Data availability. The genome sequence and raw sequencing reads for strain NSMJ15 were deposited under GenBank accession numbers CP049324, CP049325, CP049326, CP049327, and CP049328, BioProject accession number PRJNA607656, BioSample accession number SAMN14142699, and SRA accession numbers SRX9113284 and SRX9113285.

ACKNOWLEDGMENT

This work was carried out with the support of a grant (project no. NNIBR202102103) from the Nakdonggang National Institute of Biological Resources, funded by the Ministry of Environment, South Korea.

REFERENCES

- Hemarajata P, Versalovic J. 2013. Effects of probiotics on gut microbiota: mechanisms of intestinal immunomodulation and neuromodulation. *Therap Adv Gastroenterol* 6:39–51. <https://doi.org/10.1177/1756283X12459294>.
- Fong W, Li Q, Yu J. 2020. Gut microbiota modulation: a novel strategy for prevention and treatment of colorectal cancer. *Oncogene* 39:4925–4943. <https://doi.org/10.1038/s41388-020-1341-1>.
- Pandey KR, Naik SR, Vakil BV. 2015. Probiotics, prebiotics and synbiotics—a review. *J Food Sci Technol* 52:7577–7587. <https://doi.org/10.1007/s13197-015-1921-1>.
- Markowiak P, Śliżewska K. 2018. The role of probiotics, prebiotics and synbiotics in animal nutrition. *Gut Pathog* 10:21. <https://doi.org/10.1186/s13099-018-0250-0>.
- Jung JY, Han S-S, Kim Z-H, Kim MH, Kang HK, Jin HM, Lee MH. 2021. In-vitro characterization of growth inhibition against the gut pathogen of potentially probiotic lactic acid bacteria strains isolated from fermented products. *Microorganisms* 9:2141. <https://doi.org/10.3390/microorganisms9102141>.
- Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Walker BJ, Abeel T, Shea T, Priest M, Bouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Yoon SH, Ha SM, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>.
- Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68:461–466. <https://doi.org/10.1099/ijsem.0.002516>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvermin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>.
- Lebeer S, Vanderleyden J, De Keersmaecker SC. 2008. Genes and molecules of lactobacilli supporting probiotic action. *Microbiol Mol Biol Rev* 72:728–764. <https://doi.org/10.1128/MMBR.00017-08>.
- van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. 2018. BAGEL4: a user-friendly Web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res* 46:W278–W281. <https://doi.org/10.1093/nar/gky383>.