

Complete genome sequencing and probiotic trait analysis of *Lactocaseibacillus rhamnosus* LR110, a human stool isolate from the NORDBIOTIC collection

A. Szczepankowska,¹ B. Cukrowska,² T. Aleksandrak-Piekarczyk¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We present the complete genome of *Lactocaseibacillus rhamnosus* strain LR110, a human stool isolate from the NORDBIOTIC strain collection. The genome consists of a 2,867,184-bp chromosome with 46.8% GC content. Genomic analysis revealed genes related to thiamine salvage, lactose metabolism, and putrescine biosynthesis, providing insight into the strain's potential probiotic properties.

KEYWORDS probiotics, *Lactocaseibacillus rhamnosus*, lactobacilli, vitamin B1, lactose metabolism, bacteriocins, gut homeostasis, ornithine decarboxylase, putrescine biosynthesis, anti-obesity

Lactocaseibacillus rhamnosus is a rod-shaped, gram-positive lactic acid bacterium renowned for its probiotic properties (1). *L. rhamnosus* LR110 from the NORDBIOTIC collection originates from the feces of a healthy individual and has been deposited in the DSMZ Collection (DSMZ33794). LR110, along with other NORDBIOTIC strains, has been shown to alleviate symptoms of viral respiratory infections (2).

For this study, LR110 was taken directly from the NORDBIOTIC collection and cultured overnight from a single colony in MRS liquid medium (Oxoid) at 37°C under aerobic conditions. Cell pellets were pre-treated with lysozyme and mutanolysin prior to DNA extraction (3). DNA was isolated using a cetyltrimethylammonium bromide/lysozyme extraction protocol (4). The obtained DNA sample was sequenced using hybrid technology, and the resulting data are shown in Table 1.

Short-read sequencing was performed using the Illumina MiSeq platform (2 × 300 bp) with the NEB Ultra II FS library kit (New England Biolabs). Quality control and filtering of raw sequencing data were done using FASTQC (version 0.12.0) (5) and fastp (version 0.23.2) (6), resulting in 535,134 reads and 250,676,561 nt of final sequencing data.

For long-read sequencing, the GridION platform was used with the SQK-LSK109 native barcoding expansion kit (EXP-NBD103) and R9.4.1 flow cell (Oxford Nanopore Technologies). Basecalling was done using Guppy (version 6.1.3), adaptor removal with Porechop (version 0.2.4) (<https://github.com/rrwick/Porechop>), and quality control using NanoFilt (version 2.8.0) and NanoPlot (version 1.41.6) (7). This yielded 32,659 reads, with 295,055,194 bp of data and an N_{50} value of 13,393 bp.

Assembly of the nanopore reads, circularization, and contig rotation were performed using the Tricycler pipeline (version 0.5.3) (8) with multiple assemblers: Flye (version 2.9) (9), Unicycler (version 0.4.8) (10), Raven (version 1.8.1) (11), and Miniasm (version 0.3-r179) (12). Consensus sequences were corrected using Racon (version 1.5.0) (13) and Medaka (version 1.7.2) (14). Illumina short reads were then aligned to the assembled long-read contigs using Polypolish (version 0.5.0) (15) and POLCA (version 4.0.5) (16). All software was used with default parameters.

Editor David Rasko, University of Maryland School of Medicine, Baltimore, Maryland, USA

Address correspondence to T. Aleksandrak-Piekarczyk, tamara@ibb.waw.pl.

The authors declare no conflict of interest.

Received 11 October 2024

Accepted 17 January 2025

Published 24 February 2025

Copyright © 2025 Szczepankowska 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/).

TABLE 1 *L. rhamnosus* LR110 sequencing data

Description	Sequence ID LR110
GenBank accession no.	CP147912
Size (bp)	2,867,184
%GC	46.8
Coding sequences	2,599
Illumina raw data	
Number of sequences	557,312
Sequence length (bp)	294,330,379
Illumina trimmed data	
Illumina reads	535,134
Illumina data (nt)	250,676,561
Illumina coverage	56
ONT raw data	
Number of sequences	44,324
Sequence length (bp)	300,375,248
ONT trimmed data	
Number of sequences	32,659
Sequence length (bp)	295,055,194
<i>N</i> ₅₀	13,393
Average length	9,034.4
ONT coverage	76

The resulting circular genome assembly was 2,867,184 bp in length, with a GC content of 46.8% and 188.0× coverage. Genome annotation, performed using the NCBI Prokaryotic Genome Annotation Pipeline (version 6.6) (17), identified 2,676 genes, including 2,599 coding sequences and 77 RNA genes. Notably, LR110 contains genes involved in the thiamine (vitamin B1) salvage pathway, important for short-chain fatty acid synthesis and gut microbiota homeostasis, as well as class IIb and IIc bacteriocins, which together may contribute to modulating microbial dysbiosis and managing conditions, like inflammatory diseases, obesity, and allergy (2, 18). Additionally, components of the lactose-specific phosphotransferase system (EIIA and EIIBC) and complete tagatose pathway were identified, indicating the strain's potential for efficient lactic acid production from lactose. Moreover, a gene encoding an ornithine decarboxylase (EC:4.1.1.17) involved in putrescine biosynthesis was detected, which has been linked to beneficial effects in patients with obesity and type 2 diabetes (19).

ACKNOWLEDGMENTS

Genome sequencing was carried out at the DNA Sequencing and Synthesis Facility, Institute of Biochemistry and Biophysics, Polish Academy of Sciences.

AUTHOR AFFILIATIONS

¹Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland

²Department of Pathomorphology, The Children Memorial Health Institute, Warsaw, Poland

AUTHOR ORCIDs

A. Szczepankowska  <http://orcid.org/0000-0002-8733-3283>

T. Aleksandrak-Piekarczyk  <http://orcid.org/0000-0002-4725-760X>

AUTHOR CONTRIBUTIONS

A. Szczepankowska, Conceptualization, Investigation, Validation, Visualization, Writing – original draft, Writing – review and editing | B. Cukrowska, Validation, Writing –

review and editing | T. Aleksandrak-Piekarczyk, Investigation, Methodology, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

Complete sequencing data have been deposited in GenBank under BioProject [PRJNA1071652](https://ncbi.nlm.nih.gov/bioproject/PRJNA1071652) and BioSample [SAMN39639086](https://ncbi.nlm.nih.gov/biosample/SAMN39639086). The whole genome sequence is available under accession number [CP147912](https://ncbi.nlm.nih.gov/assembly/GCA_021116240.1). Illumina SRA reads are available under accession number [SRX23536627](https://ncbi.nlm.nih.gov/sra/SRX23536627), and Oxford Nanopore SRA reads under accession number [SRX23536626](https://ncbi.nlm.nih.gov/sra/SRX23536626).

REFERENCES

1. Mathipa-Mdakane MG, Thantsha MS. 2022. *Lactocaseibacillus rhamnosus*: a suitable candidate for the construction of novel bioengineered probiotic strains for targeted pathogen control. *Foods* 11:785. <https://doi.org/10.3390/foods11060785>
2. Skrzydło-Radomańska B, Prozorow-Król B, Cichoż-Lach H, Majsiak E, Biera JB, Kanarek E, Sowińska A, Cukrowska B. 2021. The effectiveness and safety of multi-strain probiotic preparation in patients with diarrhea-predominant irritable bowel syndrome: a randomized controlled study. *Nutrients* 13:756. <https://doi.org/10.3390/nu13030756>
3. Szczepankowska AK, Cukrowska B, Aleksandrak-Piekarczyk T. 2024. Whole-genome sequencing and characterization of human fecal isolate *Lactocaseibacillus casei* LC130 from NORDBIOTIC collection. *Microbiol Resour Announc* 13:e0050724. <https://doi.org/10.1128/mra.00507-24>
4. Wilson K. 2001. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol* Chapter 2:Unit. <https://doi.org/10.1002/0471142727.mb0204s56>
5. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
6. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
7. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
8. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 22:266. <https://doi.org/10.1186/s13059-021-02483-z>
9. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
10. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
11. Vaser R, Šikić M. 2021. Time- and memory-efficient genome assembly with Raven. *Nat Comput Sci* 1:332–336. <https://doi.org/10.1038/s43588-021-00073-4>
12. Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 32:2103–2110. <https://doi.org/10.1093/bioinformatics/btw152>
13. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746. <https://doi.org/10.1101/gr.214270.116>
14. Wright C, Wykes M. 2023. Medaka: sequence correction provided by ONT research. <https://github.com/nanoporetech/medaka>
15. Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 18:e1009802. <https://doi.org/10.1371/journal.pcbi.1009802>
16. Zimin AV, Salzberg SL. 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput Biol* 16:e1007981. <https://doi.org/10.1371/journal.pcbi.1007981>
17. Tatusova T, DiCuccio M, Badredin A, Chetvernin 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>
18. Wan Z, Zheng J, Zhu Z, Sang L, Zhu J, Luo S, Zhao Y, Wang R, Zhang Y, Hao K, Chen L, Du J, Kan J, He H. 2022. Intermediate role of gut microbiota in vitamin B nutrition and its influences on human health. *Front Nutr* 9:1031502. <https://doi.org/10.3389/fnut.2022.1031502>
19. Bui TI, Britt EA, Muthukrishnan G, Gill SR. 2022. Probiotic induced synthesis of microbiota polyamine as a nutraceutical for metabolic syndrome and obesity-related type 2 diabetes. *Front Endocrinol (Lausanne)* 13:1094258. <https://doi.org/10.3389/fendo.2022.1094258>