

Draft genome sequence of *Limosilactobacillus fermentum* PJG11 isolated from wild yam, *Dioscorea luzonensis* in Ilocos Norte, Philippines

Ma. Joy Theresa Agcaoili,^{1,2} Lemuel Ray Balloyan,¹ Jimmbeth Zenila Fabia,^{1,2} Marvielyn Olivar,^{1,2} Patrick James Macugay,³ Peter James Icalia-Gann^{1,4}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT This report presents the draft genome sequence of a lactic acid bacterium, *Limosilactobacillus fermentum* PJG11, isolated from wild yam, *Dioscorea luzonensis*, an endemic root crop in Ilocos Norte, Philippines. The discovered genomic characteristics of the strain shall have implications in the study of its beneficial properties and toward comparative genomic analyses involving strains of the same species.

KEYWORDS whole genome sequencing, probiotics

Lactic acid bacteria (LAB) have been established as key sources of probiotic strains exhibiting genomic and metabolic characteristics that offer health benefits toward their hosts (1–3), with the potential to treat certain cancers as one of them (4, 5). Studies have highlighted *Limosilactobacillus fermentum* strains as one of many probiotic candidates demonstrating significant anticancer treatment properties *in vitro* and *in vivo* (6–10). Here, the genome of a locally isolated *L. fermentum* strain designated as PJG11 was sequenced and assembled hoping to contribute an in-depth understanding and aid in the investigation of potential locally-isolated anticancer probiotic treatments.

Strain PJG11 was isolated from the indigenous *Dioscorea luzonensis* harvested in Ilocos Norte, Philippines (18.0564° N, 120.5650° E). Plant tubers were initially peeled, cut into smaller cubes, and then washed with sterile distilled water. The cut tissues were then submerged in de Man, Rogosa, and Sharpe (MRS) broth and incubated for 4 days. The resulting solution was then directly streaked into MRS agar and incubated at 37°C for 24–48 h. Colonies displaying typical LAB morphology were picked and subjected to successive re-streaking with the same medium and growth conditions. Pure cultures were preserved with 50% (vol/vol) glycerol and stored at –80°C for long-term storage.

For whole genome sequencing, the PJG11 isolate was revived and cultured in 20 mL of MRS broth and incubated with shaking for 24 h at 37°C. Cells were collected through centrifugation of approximately 1.5 mL of broth culture for DNA extraction using the GF-1 Bacterial DNA Extraction Kit (Vivantis Technologies, Malaysia) following the manufacturer's protocol. Quality and concentration of the extracted DNA was assessed by agarose gel electrophoresis and ScanDrop2 spectrophotometer (Analytik Jena, Germany). DNA extracts were submitted to the DNA Sequencing Core Facility of the Philippine Genome Center for library preparation and sequencing. DNA libraries were constructed using the Nextera XT DNA Library Preparation Kit (Illumina Inc., USA) and assessed using the Agilent TapeStation 2200 system (Agilent Technologies, USA). The whole genome was sequenced using the Illumina MiSeq platform (Illumina Inc., USA) with a paired-end read of 150 bp.

Read qualities of raw sequencing data were assessed using FastQC v0.12.0 (11), in which read pairs totaled 4,241,473, with the number of bases amounting to 638.4

Editor Zhenjiang Zech Xu, Nanchang University, Nanchang, Jiangxi, China

Address correspondence to Peter James Icalia-Gann, pjicalia@gmail.com.

The authors declare no conflict of interest.

Received 17 March 2026

Accepted 2 April 2026

Published 23 April 2026

Copyright © 2026 Agcaoili 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 Summary of the assembly and annotation results of *L. fermentum* PJG11

Parameter	Value
No. of Contigs	173
Genome Size	2,028,528 bp
GC Content	51.68%
Ave. Coverage Depth	618x
N50	40,877
L50	14
Coding Density	85.4%
CheckM2 Completeness	99.98%
Contamination	0.11%
BUSCO Completeness (lactobacillaceae_odb12)	98.8%
Protein-coding Genes	1964
tRNA	59
ncRNA	5
ncRNA-regions	25

Mbp. Reads with a Phred score below Q20 and <36 bp in length were trimmed using Trimmomatic v0.40 (12). *De novo* genome assembly was conducted using SPAdes v4.2.0 (13). The quality of the refined assembly was assessed using QUAST v5.3.0, CheckM2 v1.1.0, and BUSCO v6.0.0 using the lactobacillaceae_odb12 parameter (14–16). The genome was annotated using Bakta v.1.11.3 (17) in compliant mode. Default parameters of all software were applied unless stated otherwise. Taxonomic identification and ANI calculation using the MiGA webserver reveal the isolate's closest relative as *Limosilactobacillus fermentum* LMG5902 with an ANI of 97.86% (18). Assembly and annotation are summarized in Table 1.

ACKNOWLEDGMENTS

This work was financially supported by the Indigenous Food Plants (IFP) Probiotics Program of the Philippines' Commission on Higher Education (CHED). Additionally, the authors wholly acknowledge the technical support provided by the Philippine Genome Center in this study.

AUTHOR AFFILIATIONS

¹Center for Cellular and Molecular Medical Research, College of Medicine, Mariano Marcos State University, Batac, Ilocos Norte, Philippines

²Research Directorate, Mariano Marcos State University, Batac, Ilocos Norte, Philippines

³Department of Biological Sciences, College of Arts and Sciences, Mariano Marcos State University, Batac, Ilocos Norte, Philippines

⁴Department of Biochemistry and Molecular Biochemistry, College of Medicine, Mariano Marcos State University, Batac, Ilocos Norte, Philippines

AUTHOR ORCIDs

Lemuel Ray Balloyan  <http://orcid.org/0009-0002-1661-2418>

Peter James Icalia-Gann  <http://orcid.org/0009-0003-5342-4464>

AUTHOR CONTRIBUTIONS

Ma. Joy Theresa Agcaoili, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Lemuel Ray Balloyan, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Jimmbeth Zenila Fabia, Investigation, Project administration, Resources | Marvielyn Olivar, Data curation, Resources, Writing – original draft | Patrick James Macugay, Data curation, Formal analysis, Methodology, Writing – original draft | Peter James Icalia-Gann,

Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

The BioProject accession in which the genome data is a part of is [PRJNA1293312](https://ncbi.nlm.nih.gov/bioproject/PRJNA1293312). Additionally, raw reads can be accessed with the SRA accession [SRR35991899](https://ncbi.nlm.nih.gov/sra/SRR35991899), whole genome shotgun data via the GenBank WGS master accession [JBSNBV000000000](https://ncbi.nlm.nih.gov/genbank/JBSNBV000000000), and bacterial sample information with the BioSample accession [SAMN53118818](https://ncbi.nlm.nih.gov/biosample/SAMN53118818).

REFERENCES

1. Akpogheli PO, Edo GI, Ali ABM, Yousif E, Zainulabdeen K, Owheruo JO, Isoje EF, Igbuku UA, Essaghah AEA, Makia RS, Ahmed DS, Umar H, Alamiery AA. 2025. Lactic acid bacteria: nature, characterization, mode of action, products and applications. *Process Biochem* 152:1–28. <https://doi.org/10.1016/j.procbio.2025.02.010>
2. Krishnan AO, Shrivastava R, Jamal MA, Pundhir D, Prakash T. 2025. A genomic investigation into the probiotic attributes among members of the Lactobacillaceae family. *Curr Microbiol* 83:81. <https://doi.org/10.1007/s00284-025-04667-3>
3. Abdul Hakim BN, Xuan NJ, Oslan SNH. 2023. A comprehensive review of Bioactive compounds from lactic acid bacteria: potential functions as functional food in dietetics and the food industry. *Foods* 12:2850. <https://doi.org/10.3390/foods12152850>
4. Ahmad MF, Ahmad FA, Alsayegh AA, Zeyaulah M, Babalghith AO, Faidah H, Ahmed F, Khanam A, Mozaffar B, Kambal N, Bantun F. 2025. Probiotics and cancer: mechanistic insights and organ-specific impact. *Biomolecules* 15:879. <https://doi.org/10.3390/biom15060879>
5. Legesse Bedada T, Feto TK, Awoke KS, Garedew AD, Yifat FT, Birri DJ. 2020. Probiotics for cancer alternative prevention and treatment. *Biomed Pharmacother* 129:110409. <https://doi.org/10.1016/j.biopha.2020.110409>
6. Asoudeh-Fard A, Asoudeh-Fard M, Parsaei A. 2025. Probiotic-enhanced chemotherapy: *Lactobacillus fermentum* synergizes with vincristine to induce apoptosis via dual pathway activation in human cancer cells. *Cancer Cell Int* 25:389. <https://doi.org/10.1186/s12935-025-04032-1>
7. Khalil MA, Sonbol FI, Al-Madboly LA, Aboshady TA, Alqurashi AS, Ali SS. 2022. Exploring the therapeutic potentials of exopolysaccharides derived from lactic acid bacteria and *Bifidobacteria*: antioxidant, antitumor, and periodontal regeneration. *Front Microbiol* 13:803688. <https://doi.org/10.3389/fmicb.2022.803688>
8. Wei Y, Li F, Li L, Huang L, Li Q. 2019. Genetic and biochemical characterization of an exopolysaccharide with *in vitro* antitumoral activity produced by *Lactobacillus fermentum* YL-11. *Front Microbiol* 10:2898. <https://doi.org/10.3389/fmicb.2019.02898>
9. Liu B, Zhang J, Yi R, Zhou X, Long X, Pan Y, Zhao X. 2019. Preventive effect of *Lactobacillus fermentum* CQPC08 on 4-nitroquinoline-1-oxide induced tongue cancer in C57BL/6 mice. *Foods* 8:93. <https://doi.org/10.3390/foods8030093>
10. Baranda SM, Fabia JZP, Jalani GSP, Racho AJJ, Olivar MP, Balloyan LR, Agcaoili MUTT, Gann PJI. 2026. Natural amylase expression enhances starch utilization in *Limosilactobacillus fermentum* PJG11 from *Dioscorea luzonensis*. *J Gen Appl Microbiol*. <https://doi.org/10.2323/jgam.2026.02.01>
11. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
12. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
13. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes *de novo* assembler. *Curr Protoc Bioinformatics* 70:e102. <https://doi.org/10.1002/cpbi.102>
14. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
15. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2024. Author correction: CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 21:735–735. <https://doi.org/10.1038/s41592-024-02248-z>
16. Tegenfeldt F, Kuznetsov D, Manni M, Berkeley M, Zdobnov EM, Kriventseva EV. 2025. OrthoDB and BUSCO update: annotation of orthologs with wider sampling of genomes. *Nucleic Acids Res* 53:D516–D522. <https://doi.org/10.1093/nar/gkaf987>
17. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* 7:11. <https://doi.org/10.1099/mgen.0.000685>
18. Rodríguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT. 2018. The microbial genomes atlas (MiGA) webserver: Taxonomic and gene diversity analysis of Archaea and bacteria at the whole genome level. *Nucleic Acids Res* 46:W282–W288. <https://doi.org/10.1093/nar/gky467>