

Complete genome sequence of *Limosilactobacillus fermentum* strain VNUA2025

Thuy Thi Thu Trinh,¹ Son Truong Dinh,¹ Hoan Van Duong,¹ Canh Xuan Nguyen¹

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT Here, we present the complete genome sequence of *Limosilactobacillus fermentum* strain VNUA2025, isolated from diseased fish samples in Vietnam. The genome was obtained by integrating sequencing data generated using Oxford Nanopore's PromethION platform with an R10.4.1 flow cell. The fully assembled circular genome is 2,110,961 bp, with a GC content of 51.51%.

KEYWORDS *Limosilactobacillus fermentum*, VNUA2025, probiotic

Limosilactobacillus fermentum strain VNUA2025 was isolated from the intestines of grass carp collected in Hai Phong Province, Vietnam. Intestinal samples were either stored at 4°C prior to processing or immediately subjected to isolation using the serial dilution method (1).

The VNUA2025 strain was cultured in 20 mL of MRS medium at 30°C, with shaking at 200 rpm for 48 h. The culture was then centrifuged at 10,000g for 10 min to harvest the cells. Total genomic DNA was extracted using a standard phenol–chloroform extraction protocol followed by isopropanol precipitation (2). The concentration, integrity, and purity of the genomic DNA (gDNA) were then evaluated. The genomic DNA was sequenced on the Oxford Nanopore's PromethION platform using an R10.4.1 flow cell.

The DNA library was prepared using the Native Barcoding Kit 24 V14 protocol (SQK-NBD114.24, Oxford Nanopore Technologies, UK) with slight modifications: the end-repair and ligation steps were extended to 30 min, and wide-bore pipette tips with gentle flicking were used to minimize DNA shearing. Raw FAST5 files were base-called with Guppy v6.4.6 in super-accuracy mode, generating a total of 930,898 reads, corresponding to approximately 8.37 Gb of raw data. The reads had a median length of 7.0 kb and a mean length of 9.0 kb, with an N50 of 12.8 kb. Mean and median quality scores were 21.4 and 21.3, respectively, and notably, 100% of reads exceeded Q20.

Base calling was performed using Guppy v6.4.6 in super-accurate mode. Reads shorter than 5,000 bp or with a Q score ≤20 were removed using Chopper v0.5 (3). *De novo* assembly of the filtered reads was performed using the Flye v2.9-b1768 using the parameters --nano-hq and --read-error 0.03 on Nanopore super-accuracy base-called data (4), yielding the complete circular genome of *L. fermentum* strain VNUA2025 of approximately 2.11 Mb with an average coverage of ~3,145×.

The completeness of the genome assembly was assessed using BUSCO v5.4.5 (5) in prok_genome mode, utilizing the bacteria_odb10 lineage data set (released on 8 January 2024). The VNUA2025 assembly showed >98.4% completeness. Default parameters were used for all software unless otherwise noted. These metrics indicate a high-quality, contiguous genome assembly, suitable for downstream genomic analyses.

The VNUA2025 genome showed an ANI score of 98.83% and an OrthoANI score of 98.88% when compared with the genome of *L. fermentum* IMDO 130101 (GCF_900205745.1) (6), thereby confirming its identification as *L. fermentum*. Genome annotation, including the identification of genes, proteins, and functional features,

Editor Karthik Anantharaman, University of Wisconsin-Madison, Madison, Wisconsin, USA

Address correspondence to Canh Xuan Nguyen, nxcanh@vnua.edu.vn.

The authors declare no conflict of interest.

Received 4 December 2025

Accepted 6 March 2026

Published 20 April 2026

Copyright © 2026 Trinh 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/).

was performed using the RAST server (Rapid Annotation using Subsystem Technology; <https://rast.nmpdr.org/>) (7, 8), which identified 57 tRNAs and 15 rRNAs. Additionally, GeneMark (<http://exon.gatech.edu/GeneMark/genemarks2.cgi>) predicted 2,120 protein-coding genes.

These high-quality, contiguous genome assemblies are suitable for downstream genomic analyses.

AUTHOR AFFILIATION

¹Faculty of Biotechnology, Vietnam National University of Agriculture, Hanoi, Vietnam

AUTHOR ORCID*s*

Son Truong Dinh  <http://orcid.org/0000-0002-5538-0173>

Canh Xuan Nguyen  <http://orcid.org/0000-0002-7791-6397>

AUTHOR CONTRIBUTIONS

Thuy Thi Thu Trinh, Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review and editing | Son Truong Dinh, Formal analysis, Methodology, Resources, Software, Writing – original draft | Hoan Van Duong, Data curation, Investigation, Writing – original draft, Writing – review and editing | Canh Xuan Nguyen, Data curation, Funding acquisition, Methodology, Resources, Supervision, Writing – review and editing

DATA AVAILABILITY

The whole-genome sequence and associated data have been deposited at GenBank (accession number [CP197324](https://doi.org/10.1128/mra.01422-25)) and the Sequence Read Archive ([SRR33928577](https://doi.org/10.1128/mra.01422-25)).

REFERENCES

1. Blaize JF, Suter E, Corbo CP. 2023. Serial dilution method and plating in microbial enumeration. Department of Biological Sciences, Wagner College, Staten Island, NY, USA.
2. Nguyen Thu T, Dinh Son T, Nguyen Canh X. 2023. Genome characterization of *Streptomyces* sp. strain VNUA74, a potential biocontrol against pathogenic fungus *Colletotrichum* spp. Microbiol Resour Announc 12:e0087323. <https://doi.org/10.1128/MRA.00873-23>
3. De Coster W, Rademakers R. 2023. NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics 39. <https://doi.org/10.1093/bioinformatics/btad311>
4. 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>
5. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
6. Lee I, Ouk Kim Y, Park SC, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66:1100–1103. <https://doi.org/10.1099/ijsem.0.000760>
7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. <https://doi.org/10.1186/1471-2164-9-75>
8. 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>