Data in Brief 65 (2026) 112574

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Draft genome sequence of *Lacticaseibacillus paracasei* isolated from fermented banana pseudostem: genomic insights into probiotic and anti-diabetic potential

Prajwal Bhanu R B[a], Lakshmi Jayaram[b], Gagan Deep V S[b], Chandrashekar Srinivasa[a,*], Ling Shing Wong[c], Ramith Ramu[b,*]

[a] *Department of studies in Biotechnology, Davangere University, Shivagangotri, Davangere-577007, Karnataka, India*
[b] *Department of Biotechnology and Bioinformatics, School of Life Sciences, JSS Academy of Higher Education & Research, Mysuru 570015, Karnataka, India*
[c] *Faculty of Health and Life Sciences, INTI International University, Nilai, Malaysia*

ARTICLE INFO

Article history:
Received 3 October 2025
Revised 17 December 2025
Accepted 5 February 2026
Available online 13 February 2026

Keywords:
Lacticaseibacillus paracasei
Fermented banana pseudostem
Whole-genome sequencing
Probiotic potential
Anti-diabetic activity

ABSTRACT

Lacticaseibacillus paracasei RAMULABWGS01 was isolated from fermented banana pseudostem and sequenced to explore its functional, probiotic, and anti-diabetic potential. Genomic DNA was isolated & whole-genome sequence (WGS) using the Illumina NovaSeq platform. The assembled draft genome comprises 3101,562 base pairs (bp) with N50 of 407,841 bp & GC content of 46.4%. Average Nucleotide Identity (ANI) via JSpeciesWS showed 98.27% similarity to the *L. paracasei* reference genome, confirming its identity. This genomic resource provides valuable insights into the probiotic capabilities of *L. paracasei* and supports comparative genomics within the genus. The genome is publicly available at NCBI under BioProject PRJNA1230373 and BioSample SAMN47177163.

* Corresponding authors.
E-mail addresses: chandru.s@davangereuniversity.ac.in (C. Srinivasa), ramith.gowda@gmail.com (R. Ramu).

https://doi.org/10.1016/j.dib.2026.112574

Specifications Table

Subject	Biology
Specific subject area	This dataset presents the draft genome sequence of *Lacticaseibacillus paracasei* RAMULABWGS01, a probiotic strain derived from fermented banana pseudostem. The focus is on microbial genomics and probiotic research, highlighting functional genomics, probiotic features, and potential anti-diabetic properties.
Type of data	Raw data and analyzed
Data collection	Using the phenol-chloroform method genomic DNA of RAMULABWGS01 was isolated from a pure culture. Whole-genome sequencing was done using the Illumina NovaSeq platform, which produced paired-end reads. FastQC was used to check read quality and remove adapters. Genome assembly was performed using **SPAdes v3.15.5.** Phylogenomic analysis was done through the Type Strain Genome Server (TYGS). A graphical circular genome map was created with Proksee. Functional genome annotation was completed using the Rapid Annotation Using Subsystem Technology (RAST). Secondary metabolite biosynthetic gene clusters identified in antiSMASH v8.0.
Data source location	The strain *Lacticaseibacillus paracasei* RAMULABWGS01 was isolated from fermented banana pseudostem collected in Nanjangud, Karnataka, India.
Data Accessibility	The draft genome sequence data of *Lacticaseibacillus paracasei* have been deposited in the NCBI GenBank repository. The data are accessible under BioProject accession number PRJNA1230373, with the associated BioSample accession number SAMN47177163.
Related research article	None

1. Value of the Data

The draft genome sequence data has been submitted to NCBI GenBank repository with BioProject ID PRJNA1230373 & BioSample ID SAMN47177163.

Data Significance

- The genome sequence of *Lacticaseibacillus paracasei* RAMULABWGS01 provides a reliable reference for genomic characterization and taxonomic classification within the *Lacticaseibacillus* genus using ANI and digital DNA–DNA hybridization.
- This dataset enables in-depth comparative genomics to identify biosynthetic gene clusters (BGCs), functional genes, and evolutionary signatures that contribute to host–microbe interactions, metabolic potential.
- The genome provides valuable insights into probiotic functions and bioactive metabolic pathways, facilitating future research on health-promoting traits and the development of probiotic functional food applications.

2. Background

Lacticaseibacillus paracasei RAMULABWGS01 is a gram-positive, anaerobic bacterium isolated from fermented banana pseudostem [1]. Members of the *Lacticaseibacillus* genus are extensively studied for their probiotic properties [2], particularly their ability to modulate gut microbiota and promote metabolic health. Although the complete spectrum of bioactive compounds in RAMULABWGS01 is not fully elucidated, lactic acid bacteria are known to produce metabolites such as short-chain fatty acids, bacteriocins, and exopolysaccharides, which contribute to host health and disease prevention [3]. Notably, several strains of *L. paracasei* have demonstrated beneficial effects in improving glucose metabolism, reducing systemic inflammation, and supporting

glycemic control factors essential for managing type II diabetes. Given these attributes, RAMULABWGS01 holds promise as a candidate for developing functional foods or probiotic-based interventions targeting metabolic disorders [4]. However, further genomic and functional studies are required to clarify its specific bioactive pathways and therapeutic mechanisms.

3. Data Description

This work reports the draft genome sequence of *Lacticaseibacillus paracasei* (RAMULABWGS01), shown in Fig. 1, a strain isolated from fermented pseudostem. The genome annotation includes predicted gene functions, metabolic pathways, and potential probiotic traits. The assembled genome is represented by a single scaffold of 3101,562 bp, with an N50 of 407,841 bp, a GC content of 46.4 %, and sequencing coverage of 96.93 % (Table 1). Although represented as a single scaffold, the assembly contains internal Ns regions; therefore, it reflects a scaffold-level draft genome than a completely closed chromosome. Genome quality assessment using the FastQC tool revealed 99.12 % completeness, indicating a high-quality assembly. Average nucleotide identity (ANI) analysis using JSpeciesWS showed RAMULABWGS01 shares 98.27 % identity with the reference genome strain in the *Lacticaseibacillus* genus Fig 2. A diagrammatic representation using a phylogenomic tree illustration showing similar strains, done using the TYGS, as shown in Fig 3. Functional annotation of the genome identified several genes linked to probiotic functions, including bile salt hydrolase (bsh), which plays a key role in bile acid deconjugation and is associated with regulation of lipid and glucose metabolism, and glutamate decarboxylase (gad), responsible for GABA production, a pathway related to improved insulin sensitivity and glucose regulation. These gene functions support host gut adaptation and provide genomic

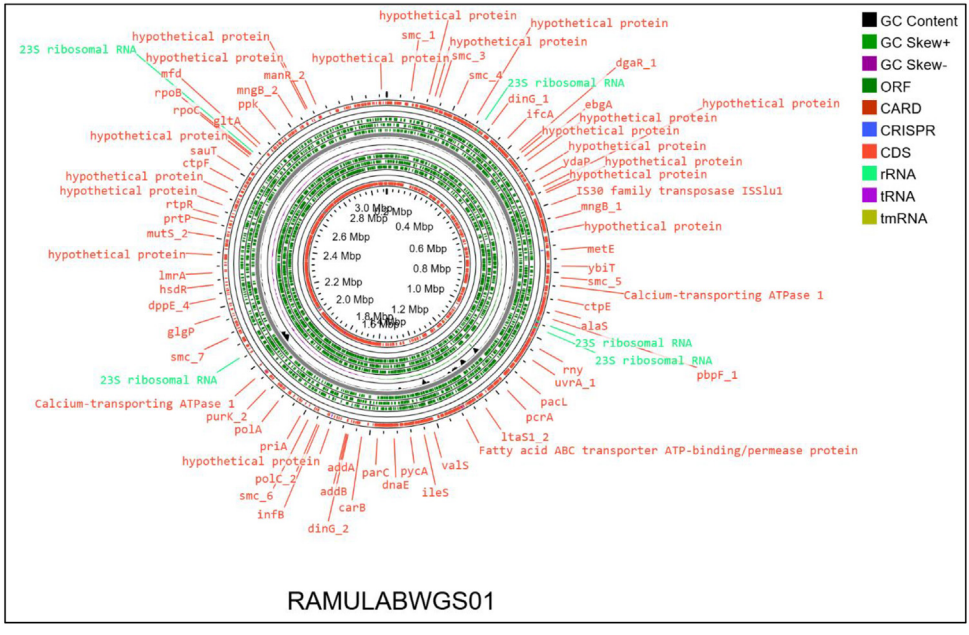

Fig. 1. Circular genome representation of RAMULABWGS01, highlighting genomic features related to its probiotic and therapeutic potential against diabetes. Outer rings display annotated coding sequences (CDS) and hypothetical proteins; colored labels mark functionally important genes, including those potentially linked to host metabolic regulation. Functional elements such as rRNAs (green), CARD antibiotic resistance genes (blue), CRISPR (orange), tRNAs (violet), and tmRNAs (olive) are marked. Inner rings depict GC content (black) and GC skew (green for positive, purple for negative).

Table 1
Genome features and assembly statistics of *Lacticaseibacillus paracasei*.

Feature	Value
Genome	*Lacticaseibacillus paracasei*
Domain	Bacteria
Taxonomy ID	1597
Genome size (bp)	3101,562
GC content	46.4
N50	407,841
L50	3
Number of scaffolds	1
Number of Coding Sequences	2917
Number of RNAs	74
CRISPR element	1

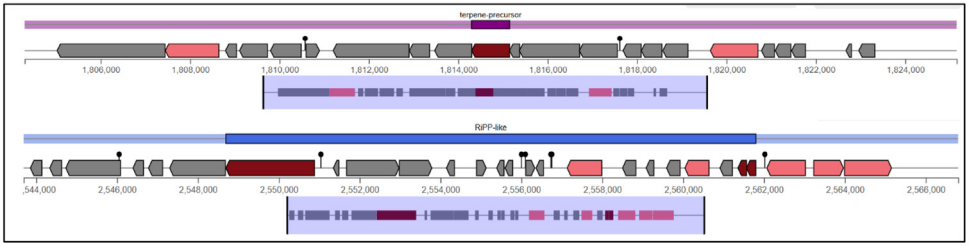

Fig. 2. Predicted biosynthetic gene clusters (BGCs) in RAMULABWGS01. The upper cluster (purple) encodes a terpene biosynthetic pathway, while the lower cluster (blue) represents a RiPP-like gene cluster. Arrows indicate open reading frames (ORFs), with functional predictions denoted by color. The presence of these clusters suggests the strain's ability to produce bioactive compounds potentially beneficial for metabolic regulation and antidiabetic effects.

evidence for the strain's potential anti-diabetic activity. Additional significant genes include multiple exopolysaccharide biosynthesis genes (fig|1597.2221.peg.284–286), D-alanylation pathway genes (fig|1597.2221.peg.2890, fig|1597.2221.peg.2897), and Clp stress-response proteases (fig|1597.2221.peg.2822, fig|1597.2221.peg.2827). The draft genome sequence of *L. paracasei* RAMULABWGS01 provides a valuable reference for comparative genomics within the *Lacticaseibacillus* spp. It may facilitate the discovery of functional genes and bioactive pathways relevant to probiotic applications, especially in the context of Type II diabetes management and functional food development.

4. Experimental Design, Materials and Methods

4.1. Bacterial isolation and genomic DNA extraction

Lactic acid bacteria (LAB) strains were isolated from fermented banana pseudostem of *Musa* sp. cultivar Nanjangud rasa bale collected in Nanjangud, Karnataka, India. The banana pseudostem was washed thoroughly, cut into small cubes, salted with one tablespoon, and allowed to ferment in a glass jar for one week, during which the pH decreased to 3.4 from 6.7±0.2, supporting the selective enrichment of lactic acid bacteria. Following fermentation, the sample was diluted & spread on MRS agar plates, then incubated under anaerobic conditions at 37°C for 24–48 hours. Colonies displaying distinct morphologies were selected and Colonies were tested for Gram stain and catalase. Only Gram-positive, catalase-negative colonies were grown in MRS broth. A total of seven LAB strains were isolated and maintained as 40 % glycerol stocks at -80°C. Preliminary identification followed Bergey's Manual. DNA was extracted from the isolated LAB strains using optimized protocols to ensure a high-quality yield. The integrity and concen-

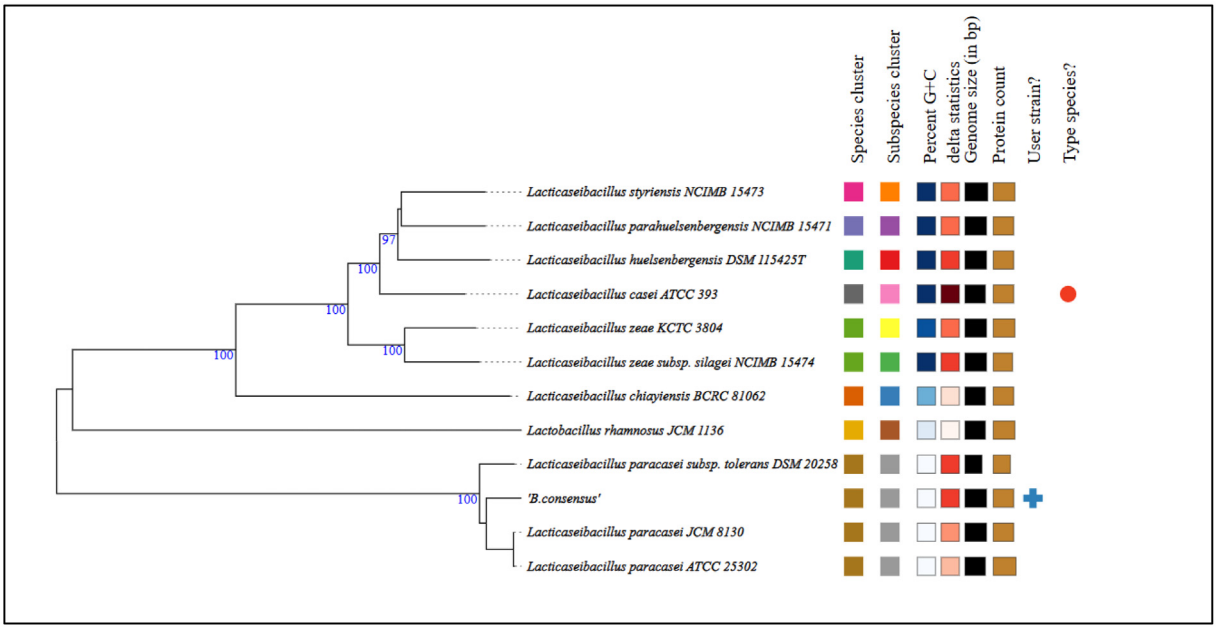

Fig. 3. Phylogenomic tree showing the relationship of *Lacticaseibacillus paracasei* RAMULABWGS01 (indicated by a blue plus symbol under "User strain?") with closely related *Lacticaseibacillus* type strains, constructed using WGS data via TYGS platform. Branch support values are based on Genome Blast Distance Phylogeny (GBDP) with 100 pseudo-bootstrap replicates, with only values ≥70 % shown. The mean branch support is across 75.0 %. Metadata columns provide species and subspecies clusters, genome G + C content, genome size, number of predicted proteins, & strain type designation.

tration of the DNA were evaluated using a Qubit fluorometer, a Nanodrop spectrophotometer, and agarose gel electrophoresis, followed by PCR using universal primers 27F and 1492R.

4.2. Genome sequence assembly and annotation

WGS was performed on the Illumina NovaSeq 6000 platform using 150 bp reads. Raw data were quality controlled using FastQC (v0.12.1), trimmed with Trimmomatic (v0.39), and Genome assembly was performed using SPAdes v3.15.5, aligned to reference genomes [5]. Genome annotation was done with RAST [6] and performed using the RASTtk pipeline. As described by Overbeek et al. (2014), RASTtk uses Glimmer3 for ORF prediction, SEED subsystem rules for functional assignment, tRNAscan-SE for tRNA identification, and standard rRNA prediction tools integrated within the RASTtk framework. and biosynthetic gene clusters were predicted using antiSMASH 8.0.[7]. Phylogenomic analysis was conducted with TYGS [8].

4.3. Genomic similarity analysis using ANI and digital DNA-DNA hybridization (dDDH)

ANI between the query and reference genomes was determined using the JSpeciesWS tool. The assembled genome was uploaded, and pairwise ANI values were calculated with a BLAST-based approach. dDDH was carried out using the Genome-to-Genome Distance Calculator (GGDC) v4.0 with the BLAST method [9].

Limitations

Not applicable

Ethics Statement

This study complies with the ethical requirements of *Data in Brief* and does not include human participants, animal testing, or social media data

Credit Author Statement

Prajwal Bhanu: Software, Formal analysis, Investigation; **Lakshmi J:** Software, Formal analysis, Investigation, Data curation, Visualization, Writing-review & editing, Writing – original draft; **Gagan Deep V S:** Software, Formal analysis, Data curation; **Chandrashekar Srinivasa:** Super- vision; Writing-review & editing**. Ling Shing Wong:** Supervision, Project administration, Funding acquisition, Writing-review & editing. **Ramith Ramu**: Conceptualization, Project administration, Supervision, Software, Formal analysis, Data curation, Validation, Writing-review & editing.

Data Availability Statement

The draft genome sequence data of *Lacticaseibacillus paracasei* are publicly available in the NCBI GenBank repository under BioProject accession number PRJNA1230373 and BioSample accession number SAMN47177163 (https://www.ncbi.nlm.nih.gov/sra/SRR32539447).

Acknowledgments

All the authors thank JSS Academy of Higher Education and Research (Mysore, Karnataka, India) for their kind support and encouragement.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] V.B.C. Kumari, S.S. Huligere, R. Ramu, S. Naik Bajpe, M.Y. Sreenivasa, E. Silina, V. Stupin, R.R. Achar, Evaluation of probiotic and antidiabetic attributes of Lactobacillus strains isolated from fermented beetroot, Front. Microbiol. 13 (2022) 911243, doi:10.3389/fmicb.2022.911243.
[2] L. Jayaram, R. Ramu, D. Puttegowda, Khang Wen Goh, C. Kumari VB, A protocol for assessing the adhesion potential of lactic acid bacteria on various cell types, STAR Protocols 6 (1) (2025) 103619, doi:10.1016/j.xpro.2025.103619.
[3] E.E. Imade, S.E. Omonigho, O.O. Babalola, B.J. Enagbonma, Lactic acid bacterial bacteriocins and their bioactive properties against food-associated antibiotic-resistant bacteria, Ann. Microbiol. 71 (1) (2021) 44, doi:10.3390/ijerph120910254.
[4] J. Yang, J. Huang, Z. Huang, Y. Xu, W. Li, S. Zhu, Y. Zhao, et al., Cardiometabolic benefits of lacticaseibacillus paracasei 8700: 2: A randomized double-blind placebo-controlled trial, Clin. Nutrit. 42 (9) (2023) 1637–1646, doi:10.1016/j.clnu.2023.07.017.
[5] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (15) (2014) 2114–2120, doi:10.1093/bioinformatics/btu170.
[6] R. Overbeek, R. Olson, G.D. Pusch, G.J. Olsen, J.J. Davis, T. Disz, R.A. Edwards, et al., The SEED and the rapid annotation of microbial genomes using Subsystems Technology (RAST), Nucleic Acids Res. 42 (D1) (2014) D206–D214, doi:10.1093/nar/gkt1226.
[7] K. Blin, S. Shaw, H.E. Augustijn, Z.L. Reitz, F. Biermann, M. Alanjary, T. Weber, antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation, Nucl. Acids Res. 51 (2023) W46–W50, doi:10.1093/nar/gkad344.
[8] J.P. Meier-Kolthoff, M. Göker, TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy, Nat. Commun. 10 (2019) 2182, doi:10.1038/s41467-019-10210-3.
[9] M. Richter, R. Rosselló-Móra, F. Oliver Glöckner, J. Peplies, JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison, Bioinformatics 32 (2016) 929–931, doi:10.1093/bioinformatics/btv681.