FEMS Microbes, 2026, **7**, xtag032
https://doi.org/10.1093/femsmc/xtag032
Published: 10 June 2026
Research article – Microbes and Society

Microbial community structure, functional potential, probiotic signatures, and MAG reconstruction of fermented bamboo shoots from Northeast India

Rohit Das [1], **Rakshak Kumar** [2], **Buddhiman Tamang** [1,*]

[1]Department of Microbiology, Sikkim University, 6th Mile, Samdur, Tadong, Gangtok, Sikkim 737102, India
[2]Department of Molecular Biology & Bioinformatics, Tripura University, Suryamaninagar, Tripura 799022, India
*Corresponding author. Department of Microbiology, School of Life Sciences, Sikkim University, 6th Mile, Samdur, Tadong, Gangtok, Sikkim 799022, India.
E-mail: bmtamang3@gmail.com
Editor: Marciane Magnani

Abstract

Fermented bamboo shoot (FBS) products are widely consumed traditional foods across the Northeast region (NER) of India, yet their microbiome structure, functional capacity, biosynthetic potential, and safety attributes remain insufficiently explored. Here, comparative shotgun metagenomics of ten traditional FBS products from six NER states was used to address these gaps integrating previously generated metagenomic data from Tripura with newly generated datasets from Manipur, Meghalaya, Arunachal Pradesh, Nagaland, and Sikkim thereby bringing the total number of samples to 24. Taxonomic profiling revealed a predominance of lactic acid bacteria, primarily members of *Lactiplantibacillus*, *Levilactobacillus*, *Lactobacillus*, *Lactococcus*, and *Pediococcus*, with pronounced product- and region-specific community signatures. Functional annotation demonstrated predominance of genes involved in carbohydrate metabolism, stress response, quorum sensing, ABC transporters, vitamin biosynthesis, and energy metabolism, supporting strong probiotic-associated functional potential across FBS types. AntiSMASH analysis enabled the identification of diverse biosynthetic gene clusters (BGCs) responsible for the production of various secondary metabolites, including bacteriocins, non-ribosomal peptides, terpenes, and siderophores, with higher biosynthetic diversity observed in *Mesu* (Sikkim), *Tuaithar* (Manipur), *Lung-Seij* (Meghalaya), and *Bastenga* (Nagaland). Antimicrobial resistance (AMR) profiling revealed a generally low resistome burden, dominated by intrinsic resistance determinants, with FBS Sikkim and Tripura exhibiting the lowest AMR prevalence among all products. High-quality metagenome-assembled genomes affiliated with *Lactiplantibacillus plantarum*, *Lactobacillus acetotolerans*, and *Pediococcus pentosaceus* exhibited conserved probiotic traits, carbohydrate-active enzymes, biosynthetic pathways, and a limited presence of mobile genetic elements. Overall, the microbiome-based comparative analysis provides a framework for understanding the microbial community structure and functional potential across the NER, demonstrating broad probiotic potential and biosynthetic richness, with *mesu* samples from Sikkim showed a comparatively consistent distribution of functional pathways, biosynthetic gene clusters, and AMR-related features relative to the other FBS samples analysed.

Keywords LAB, secondary metabolite biosynthesis, AMR, metagenome-assembled genomes, pangenome analysis, bioactive compounds

Introduction

Northeast India comprises eight states; Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, Arunachal Pradesh, and Tripura, which are rich in cultural, linguistic, and ethnic diversity. There are about 200 ethnic communities in the region with vibrant cultural and traditional practices (https://indianculture.gov.in/Northeastarchive/history-northeast). The region spreads over almost 18.4 million hectares and is a biodiversity hub of bamboo, contributing more than 66% of India's species of bamboo (Sundriyal et al. 2002), which provides a strong foundation for the development of traditional practices of fermenting bamboo shoot. The traditional knowledge of bamboo shoot fermentation typical to each community in the region, is passed to daughters from mother or grandmother (Das et al. 2023). This non-industrialized knowledge not only transforms the available substrates for later consumption but also transforms them into nutritious and flavoursome product besides detoxifying the anti-nutritional components in the bamboo shoots (Tamang et al. 2009).

Bamboo shoot undergoes spontaneous natural fermentation by the lactic acid bacteria (LAB) when kept in a closed container producing sour product (Tamang et al. 2008, Tamang et al. 2009, Das et al. 2023). However, each of the fermented bamboo shoot (FBS) producing community employs distinct implements, methodology, and duration of fermentation (2 days–1 year) that

Received: 8 January 2026. **Revised:** 25 May 2026. **Accepted:** 25 May 2026

determine the unique organoleptic quality (Tamang et al. 2009, Kikon 2021). The biochemical composition of bamboo shoots used in traditional fermented bamboo foods varies among species, particularly in lignocellulosic components (cellulose ∼35%–46%, hemicellulose ∼21%–30%, and lignin ∼18%–26%), soluble extractives (∼5%–12%), and ash (∼1%–3%), which influence substrate degradability and nutrient availability during fermentation. Phenolic compounds, typically ranging from ∼40 to 125 mg GAE/100 g fresh weight, further contribute to antimicrobial activity and microbial selection during spontaneous fermentation (Satya et al. 2010, Nirmala et al. 2014, Naik et al. 2021, Das et al. 2023, Goyal et al. 2024). Such compositional variability among bamboo species utilized in different regions of Northeast India likely drives differences in fermentation kinetics, metabolite formation, and microbial succession. Therefore, a comparative summary of fermented bamboo shoot varieties, associated bamboo species, dominant LAB, and their structural and biochemical composition relevant to fermentation ecology is provided in Supplementary File 1: ST1.

The major LABs documented from FBS of NER are summarized in Supplementary File 1: ST 1, compiling evidence from previous culture-dependent studies and highlighting the dominant taxa reported from each product. While culture-based analyses have identified several LABs as dominant fermenters contributing to flavour, taste, and preservation, they offer only a partial view of the community and cannot capture the broader spectrum of probiotic attributes that explore the health-promoting potential of these foods (Latif et al. 2023). Traits such as stress tolerance, carbohydrate-active enzyme profiles, antimicrobial compound biosynthesis, and vitamin-forming pathways remain largely inaccessible through culture-dependent approaches. Fermentation of bamboo shoots is a spontaneous fermentation driven by indigenous microbiota and proceeds through sequential microbial stages that transform lignocellulosic substrates into characteristic metabolites. Initial fermentation (0–7 days) is dominated by heterofermentative LAB such as *Leuconostoc* and *Weissella*, along with members of *Enterobacteriaceae*, producing lactic and acetic acids that rapidly reduce the pH from an initial value of ∼6.1–7.0 to around 4.0–4.5 (Badwaik et al. 2014, Chi et al. 2020). Similar observations of acidic conditions (pH 4.84 ± 0.22) during early fermentation stages were also reported by (Singhal et al. 2021). Such acidification creates favorable conditions for LAB succession and stabilization of fermented bamboo shoot microbiota (Li P et al. 2025). Mid-stage fermentation (7–20 days) involves *Lactiplantibacillus*, *Limosilactobacillus*, and *Enterococcus*, which degrade complex polysaccharides to generate mannitol, amino acids, and volatile compounds (Guo et al. 2023). Late-stage fermentation (20–30+ days) is characterized by acid-tolerant LAB such as *Pediococcus*, enhancing amino acid metabolism and phenolic transformations that yield ester-based flavour compounds (Morais et al. 2025) (Supplementary File 1: ST1). Compared with other spontaneous vegetable fermentations such as kimchi and sauerkraut, FBS contains higher lignocellulosic biomass and cyanogenic glycosides requiring specialized detoxification, resulting in a distinctive crisp-fibrous texture, earthy volatiles, minimal effervescence, and high antioxidant and probiotic potential (Hu et al. 2021a, Wu JW et al. 2025, Oh et al. 2026).

Advances in high-throughput metagenomics now allow deeper exploration of traditional fermentations by providing genome-level insights into microbial functions, stress adaptation, and secondary metabolite production. However, the use of next-generation sequencing for FBS from the NER has been sparse, leaving significant gaps in understanding the microbial blueprint. Only a few studies have applied molecular tools to these fermentations, including amplicon metagenomic analysis of *tuaithar* (Das and Deka 2012) and *soidon* (Romi et al. 2015), along with earlier shotgun-based work on fermented bamboo shoots from the region (Tamang et al. 2009). Our group has also contributed through fungal amplicon sequencing of *hirring* and *ekung* (Das et al. 2024) and shotgun-based investigation of FBS from Tripura (Das et al. 2023). While these studies provide valuable initial insights, they primarily focus on individual fermented bamboo shoot products and do not provide a comparative view of microbial communities across different bamboo substrates or their association with probiotic characteristics.

The current literature on FBS lacks comprehensive metagenomic characterization across different types, particularly regarding carbohydrate-active enzymes, metagenome-assembled genomes (MAGs), and antimicrobial resistance genes. We hypothesize that spontaneous fermentation in Northeast India harbours a metabolically active microbiome enriched with probiotic LAB, possessing genes linked to stress tolerance, antimicrobial activity, and essential metabolite biosynthesis. Shotgun metagenomics, coupled with MAG reconstruction, enables comprehensive taxonomic and functional profiling, providing genome-resolved insights into key microbial taxa and their metabolic adaptations.

To comprehensively characterize the taxonomic and functional landscape of FBS across Northeast India, the present study was designed with the following objectives; (i) to predict and annotate functional genes through KEGG pathway and CAZyme analyses for a comparative understanding of metabolic potential across all FBS types; (ii) to reconstruct high-quality MAGs of dominant taxa from these combined datasets to uncover probiotic and metabolic attributes, and (iii) to identify genes and biosynthetic pathways associated with probiotic functionality, bioactive metabolite synthesis, and safety-related traits across the regional FBS microbiome. To achieve these objectives, 24 fermented bamboo shoot samples representing 10 traditional products *bastenga, ekung, hirring, lung-siej, mesu, mileye amileye, midukeye, moiya koshak, moiya pansung,* and *tuaithar* were analysed using shotgun metagenomics to elucidate their microbial composition and functional potential. Taxonomic profiles were integrated with metabolic pathways, probiotic-associated genes, and bioactive biosynthetic clusters to understand their roles in fermentation and nutritional enhancement. Genome-resolved analysis enabled the recovery of dominant LAB and evaluation of their metabolic and probiotic traits. Collectively, this comparative approach reveals how different bamboo substrates shape fermentation microbiomes and identifies functionally enriched, probiotic-rich FBS types, providing a foundation for future targeted isolation and functional validation of key microbial taxa.

Methodology

Sample collection

A total of 24 fermented bamboo shoot samples were analysed in this study, including 18 newly collected samples and 6 previously published samples from Tripura (BioProject: PRJNA912902) (Das et al. 2023) from six states of Northeast India: Nagaland,

Meghalaya, Sikkim, Arunachal Pradesh, Manipur, and Tripura. All samples were obtained at a comparable fermentation stage (∼1 week) to ensure consistency. Households were selected based on prior knowledge of traditional fermenters within the region, using a purposive sampling approach. Local informants and elders from the community assisted in identifying households with a longstanding tradition of preparing FBS. Marketplaces were chosen based on their prominence and accessibility within each region, ensuring they served as major hubs for the local sale of fermented foods. Preference was given to markets known for selling indigenous and homemade fermented products. All samples were collected in pre-sterilized containers and transported to the laboratory in an icebox at 4°C for analysis. The GPS coordinates were recorded using eTrex® 10 (Garmin, USA), mapping was carried out using QGIS 3.34 LTR. Detailed coordinates of sample collection and pictorial representation of sampling location, product types, and the applied methodology are provided in Fig. 1 and Supplementary File 1: ST2.

Metagenomic DNA extraction and library preparation

Ten grams of each sample were homogenized with 90 ml of sterile 0.1 M phosphate buffer saline, pH 6.4, in a stomacher (400 Circulator, Seward, UK). The homogenized mixture was filtered and the filtrate used to extract genomic DNA by use of the XpressDNA Plant Kit (MagGenome, cat no.: MG20PIL-250) following the manufacturer's protocol. DNA samples were analysed for quality and potential degradation on a 0.8% agarose gel at 100 V for 40 min, with concentration verified via Nanodrop to assess suitability for library preparation. According to the manufacturer's instructions, total DNA was subsequently used for DNA library construction and amplification with the TruSeqTM DNA Sample Prep Kit (Illumina, USA) and the cBOT TruSeq PE Cluster Kit v3 (Illumina, USA) reagents. After that, the DNA libraries were sequenced on the TruSeq SBS Kit v3-HS (Illumina, USA) and the Illumina Novaseq 6000 technology following the standard Illumina methods. The DNA libraries with inserted fragments were then paired end sequenced (2 × 150 bp) (Das et al. 2023).

Bioinformatic analysis

Microbial profiling

Raw Illumina NovaSeq sequence data were first checked for quality with the FastQC software (Liu et al. 2021). Then, assemblies were done using MEGAHIT v1.0 (Li et al. 2015), and taxonomic classification of each contig was performed via the Kaiju v1.9.0 pipeline against the database derived from GenBank, using nr + euk to allow for a full match in classification (Menzel et al. 2016). Taxonomic profiling of microbiome community was carried out using MetaPhlAn (v4), which relies on unique clade-specific marker genes for accurate microbial identification. Sequence reads were aligned to the MetaPhlAn marker gene database using Bowtie2 with default settings, and taxonomic assignments were filtered based on high-confidence marker matches to exclude low-complexity regions, chimeric reads, and erroneous alignments (Menzel et al. 2016). For the analysis of amino acid substitution, amino acid sequences were obtained by translating ORFs into amino acid sequence using the BLOcks SUbstitution Matrix (BLOSUM62). The reverse search is then performed by running the algorithm of BWT to facilitate database mapping to identify taxonomic classifications and generate detailed output (Pokrzywa and Polanski 2010).

Predictive functionality and probiotic gene annotation

Protein-coding genes (ORFs) were predicted using Prodigal v2.6.3 (Hyatt et al. 2010), with ORFs shorter than 180 nucleotides excluded from further analysis by default. tRNA genes were identified using the ARAGORN v 1.2.38 program (Laslett and Canback 2004), while ribosomal RNA genes (5S, 5.8S, 16S, 18S, 23S, 28S) were detected and classified using rRNAFinder (Dong and Strous 2019). All ORFs were then searched against the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa and Goto 2000) for KEGG ID annotation. Functional attributes were obtained at three levels: Level 1, Level 2, and Level 3, corresponding to the category, super pathway, and pathways, respectively (Das and Tamang 2021). Probiotic gene annotation was done using Prokka (Seemann 2014). To improve biological relevance and avoid overinterpretation, KEGG pathways associated with human diseases (e.g. cancer-related pathways) were interpreted cautiously. Pathway assignments were not considered indicative of functional activity unless supported by multiple constituent genes. Pathways represented by only a small fraction of their total gene components were considered incomplete and were not emphasized in downstream interpretation. Where appropriate, pathway completeness and gene representation were considered to avoid false-positive functional inference from sparse gene matches.

Secondary metabolite analysis

Secondary metabolite biosynthetic potential was assessed using antiSMASH v8.0 on assembled metagenomic contigs. High-quality contigs (>1 kb) were analysed with default parameters to identify and annotate biosynthetic gene clusters (BGCs) based on conserved domain architecture, gene synteny, and homology to known clusters. Predicted BGCs were classified into major metabolite classes, including non-ribosomal peptide synthetase (NRPS), polyketide synthase (PKS), and ribosomally synthesized and post-translationally modified peptides (RiPPs), terpenes, siderophores, and betalactones, and further compared against the MIBiG database to infer putative metabolite functions. The abundance and distribution of BGC types across samples were summarized to evaluate secondary metabolite diversity and biosynthetic potential (Blin et al. 2025).

Antimicrobial resistance analysis

Antimicrobial resistance (AMR) genes were identified using ABRicate v1.2.0 on assembled metagenomic contigs. High-quality contigs (>1 kb) were screened against curated resistance gene databases (including CARD/ResFinder) using default thresholds for sequence identity and coverage. Detected AMR genes were annotated based on homology to known resistance determinants and classified according to antibiotic classes and resistance mechanisms. The prevalence and diversity of AMR genes across samples were quantified to assess the resistome composition and potential dissemination of antimicrobial resistance (Seemann 2016).

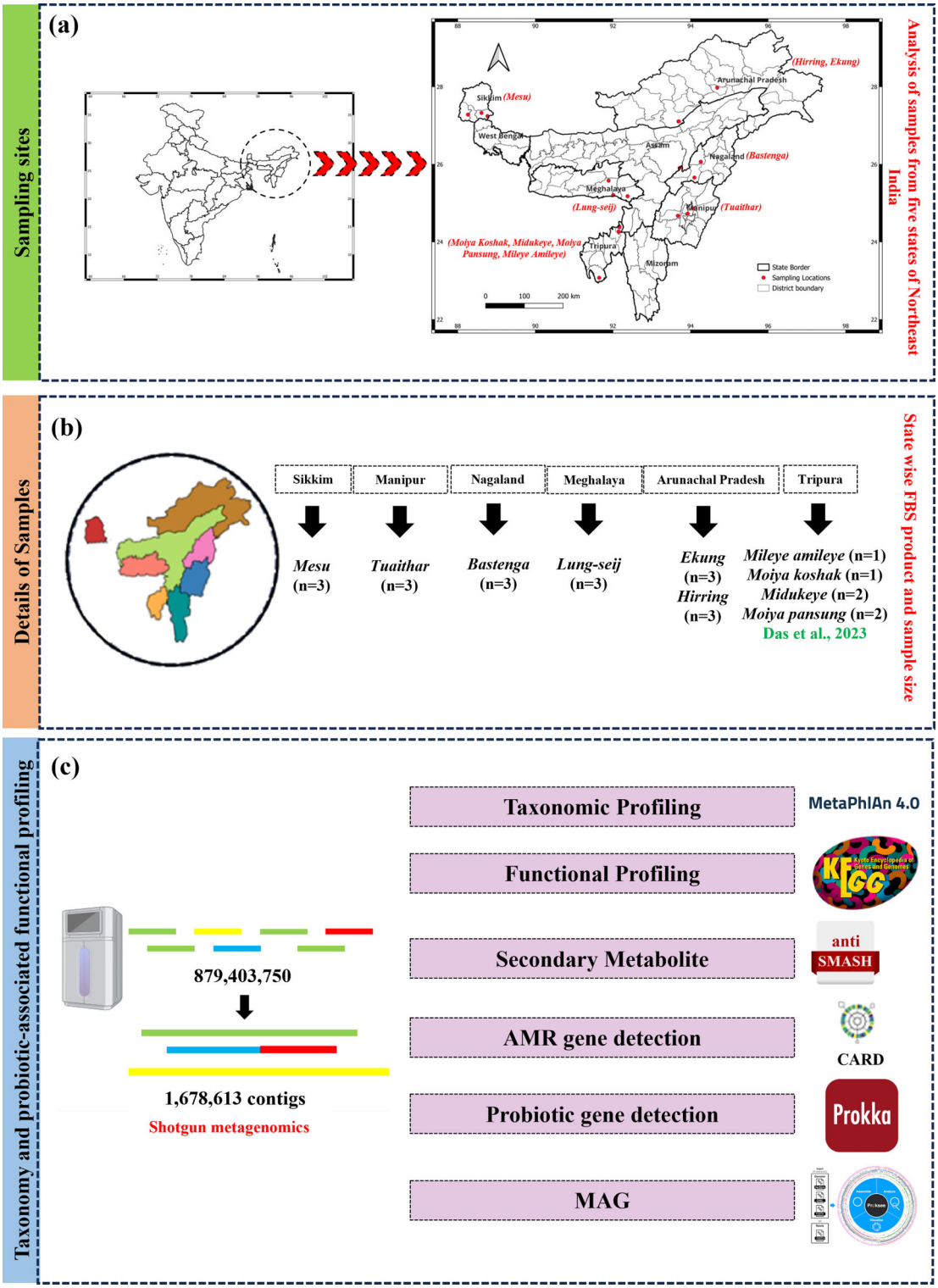

Figure 1 Sampling sites across the northest region (NER) of India, diversity of traditional FBS products, and experimental workflow based on shotgun metagenomic (a) geographic distribution of sampling sites across six states of Northeast India (NER) Sikkim, Manipur, Nagaland, Meghalaya, Arunachal Pradesh, and Tripura highlighting culturally distinct regions from which traditional fermented bamboo shoot products were collected. (b) The mapped locations indicate traditional FBS products included in the study *Mesu* (Sikkim), *Tuaithar* (Manipur), *Bastenga* (Nagaland), *Lung-seij* (Meghalaya), *Ekung* and *Hirring* (Arunachal Pradesh), and *Mileye amileye*, *Moiya koshak*, *Midukeye*, and *Moiya pansung* (Tripura-previously studied) (c) Overview of the experimental workflow in which shotgun DNA from ten FBS products was subjected to high-throughput sequencing and contig assembly. Taxonomic profiling was performed using MetaPhlAn v4.0, functional annotation via KEGG, secondary metabolite gene clusters using antiSMASH, AMR genes using Comprehensive Antibiotic Resistance Database (CARD), and probiotic-associated genes using Prokka. Genome-resolved metagenomics enabled reconstruction of high-quality MAGs, providing an integrated assessment of microbial diversity, functional potential, safety, and probiotic relevance of FBS products.

Metagenomic-assembled genomes

Taxonomy and predictive functionality of MAG

Single-assembly genomes were constructed using MaxBin v2.0 (Wu et al. 2016). To assess the quality of the MAGs, we employed CheckM v1.0.12 to determine their completeness and contamination levels (Parks et al. a). Bins meeting $\geq$ 90% completeness and $\leq$10% contamination were considered for further analysis. The average nucleotide identity (ANI) for all high-quality MAGs were calculated using FastANI (Jain C et al. 2018). Identification of MAGs were performed using the Genome Taxonomy Database-toolkit, GTDB-Tk v2.1.0 (Chaumeil et al. 2022) accessed in 2025 with release 10 (R10-RS226), which spans 715 230 bacterial genome. Secondary metabolites were screened using antiSMASH (https://antismash.secondarymetabolites.org/#!/start) (Blin et al. 2023) accessed in 2025. CAZymes were annotated using dbCAN3 server (https://bcb.unl.edu/dbCAN2/index.php) (Zheng Jinfang et al. 2023) accessed in 2025. and the resulting outputs were then mapped to CAZy database through the webserver (http://www.cazy.org/) (Drula et al. 2022). Gene prediction for the MAGs was done using Prodigal v2.6.1 (Hyatt et al. 2010), and the amino acid sequences from the genes were used for functional prediction with GhostKOALA (Kanehisa et al. 2016) and eggNOG-mapper v2 (Cantalapiedra et al. 2021). Functional attributes of genes were then mapped to the database of KEGG (Kanehisa and Goto 2000) and eggNOG v5.0.2 (Huerta-Cepas et al. 2019). Proksee was used to visualize a circular genome (Grant et al. 2023). For each genome, we conducted predictions and annotations of carbohydrate-active enzymes (CAZymes) by utilizing Carbohydrate-Active Enzymes Database version 3 (dbCAN3) (Zhang et al. 2018).

In silico-Safety assessment of high-quality MAGs

MAGs were analysed to characterize antimicrobial resistance, defense systems, mobile genetic elements, and viral associations using an integrated bioinformatic framework. Predicted open reading frames from MAGs were screened against the CARD (Alcock et al. 2020) to identify antimicrobial resistance genes. CRISPR–Cas systems were detected and classified using dedicated CRISPRCas Finder (Couvin et al. 2018) annotation tools to infer adaptive immune potential. Horizontally acquired genomic regions were identified using AlienHunter (Vernikos and Parkhill 2006), enabling the detection of atypical sequence composition indicative of foreign DNA. Mobile genetic elements were characterized using MobileOG-db (Brown et al. 2022), while prophage regions were predicted using PHASTEST (Wishart et al. 2023) and Phigaro (Starikova et al. 2020), providing complementary prophage identification and boundary resolution. Additionally, viral contigs and virus-associated regions were detected using VirSorter (Roux et al. 2015), allowing classification of bacteriophages and putative viral elements. Presence of plasmid was assessed using PlasmidFinder v2.1.6 (Carattoli et al. 2014). Antiviral defense systems encoded within the MAGs were identified using DefenseFinder, which systematically scans genomic contigs for known bacterial and archaeal defense modules based on curated Hidden Markov Model (HMM) profiles and reference defense system architectures (Tesson et al. 2024).

Pangenome analysis of MAGs

Coding sequence (CDS) prediction and gene annotation were performed using Prokka (Grant et al. 2023). The pan-genome of *Lactobacillus acetotolerans* (DSJ, KB, LM, LT, NH, and WB), *Lactiplantibacillus plantarum* (DB, GEK, GM, LSJ, SSJ, and TT) and *Pediococcus pentosaceus* (BH and GH) was analysed using Roary v3.13.0 (Page et al. 2015), in comparison with other *Lactobacillus acetotolerans, Lactiplantibacillus plantarum,* and *Pediococcus pentosaceus* RefSeq strains, The 16S rRNA gene sequences within the MAGs were evaluated using ContEst16S (Lee et al. 2017) to assess genome completeness, potential contamination, and preliminary taxonomic affiliation. The recovered 16S rRNA sequences were compared against the curated database of EzBioCloud (Chalita et al. 2024) to identify the closest phylogenetic relatives. Based on the highest sequence similarity, representative strains corresponding to the identified taxa were selected. The whole genome sequences of these strains were subsequently retrieved from the EzBioCloud Whole Genome Public DB (Yoon et al. 2017). These reference genomes were used along with the MAGs for downstream comparative and pangenome analyses. Orthologous genes across the strains were identified using OrthoFinder v2.5.4 (Emms and Kelly 2019). Additionally, genome quality assessments for all strains were conducted using CheckM (Gurevich et al. 2013, Parks et al.). The phylogenetic tree analysis was done using FastTree 2.1 (Price et al. 2009) with bootstrap value of 500 repetitions and was edited using iTOL (Letunic and Bork 2021).

Results and discussion

Taxonomic profiling and diversity indices

A total of 879 403 750 reads (R1 and R2) were obtained from eighteen samples (*bastenga, ekung, hirring, mesu, lung-seij,* and *tuaithar,*) with an average of 48 855 763 reads per sample, the lowest read pairs 213 774 (DT) and the highest read pairs 18 945 377 (DBMP). The total assembly length recovered from the samples ranged from 0.2 to 31.19 Mbp (Supplementary File 1: ST2). Shotgun metagenomic sequencing of all fermented bamboo shoot (FBS) samples revealed the presence of four major domains, bacteria, archaea, and eukaryota. Relative abundance analysis across all samples showed that bacteria were overwhelmingly dominant (99.83% of total community abundance), followed by Eukaryota (0.0057%) and Archaea (0.057%) (Fig. 2a). Overall, 7 phyla, 26 families, 112 genera, and 567 species were detected. Across all samples, the abundance of phylum Bacillota (78.3% of total bacterial relative abundance, here and below) was observed (Fig. 2b). At the family level, Lactobacillaceae was the most abundant in all samples (55.5% of total bacterial abundance) (Fig. 2c). At the genus level, notable variations in genus dominance were observed among individual fermented bamboo shoot products. In Arunachal Pradesh, *ekung* and *hirring* were predominantly enriched with *Lactococcus* and *Pediococcus*, exhibiting average relative abundances of 30.07% and 42.3% (of total bacterial abundance), respectively. In Tripura, *mileye amileye*, *moiya pangsung*, *moiya koshak*, and *midukeye* showed differing patterns, with *Lactiplantibacillus* dominating *mileye amileye* (21%), *moiya pangsung* (36.9%), *moiya koshak* (28.6%), while *midukeye* was enriched with *Lactococcus* (33.43%). In Meghalaya, *lung-seij* was dominated by *Lactiplantibacillus* (22.49%), whereas in Nagaland, *bastenga*

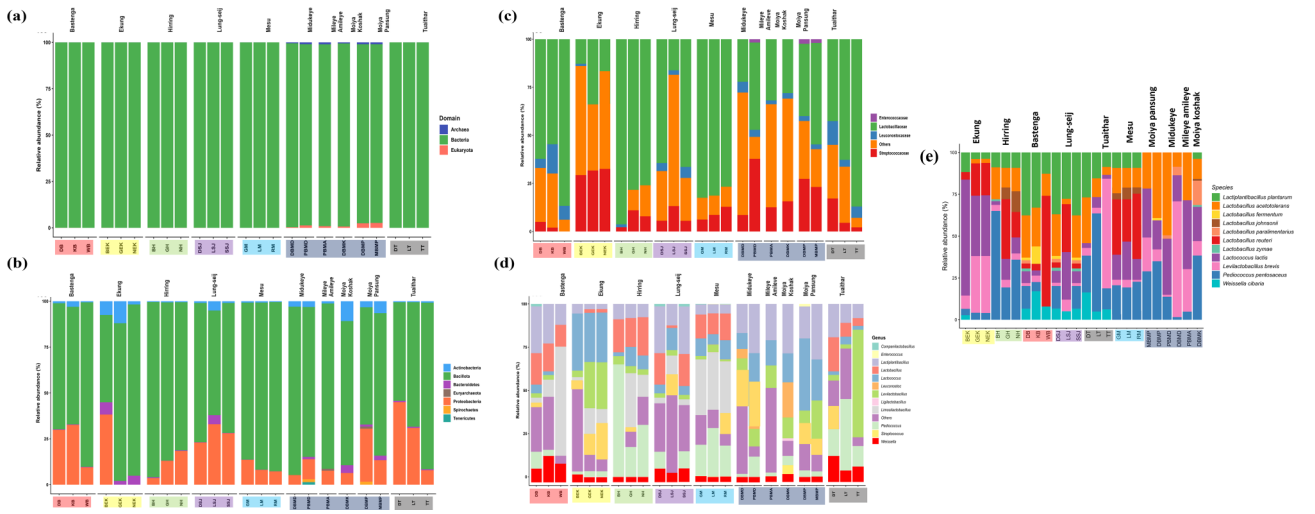

Figure 2 Relative abundance of microbial taxa from domain to genus level in fermented bamboo shoot samples (a) Domain-level taxonomic composition of the fermented bamboo shoot samples showing relative abundances of *Bacteria*, *Eukaryota*, and *Archaea*, based on shotgun metagenomic profiling. (b) Phylum-level distribution of microbial communities across samples, highlighting the dominance and variation of major bacterial phyla, including *Bacillota*, *Proteobacteria*, *Actinobacteriota*, *Bacteroidota*, *Spirochaetota*, *Tenericutes*, and *Euryarchaeota*. (c) Family-level taxonomic composition illustrating the relative abundance of predominant bacterial families, with notable representation of *Lactobacillaceae*, *Streptococcaceae*, *Enterobacteriaceae*, *Leuconostocaceae*, and other minor families across different products. (d) Genus-level microbial profiles of the fermented bamboo shoot samples showing sample-specific dominance and diversity of key LAB and associated genera, including *Lactiplantibacillus*, *Lactococcus*, *Levilactobacillus*, *Pediococcus*, *Weissella*, *Leuconostoc*, and other low-abundance taxa. (e) Species-level distribution of members belonging to the *Lactobacillaceae* family, highlighting the relative abundance of LAB across the samples. Only taxa within the *Lactobacillaceae* family are shown in this panel to emphasize the diversity and distribution of LAB species during fermentation. Colour-coded sample labels represent fermented bamboo shoot samples collected from different states.

exhibited predominance of *Lactiplantibacillus* (37.6%). In Sikkim, *mesu* was primarily enriched with *Limosilactobacillus* (31.8%), and in Manipur, *tuaithar* showed a high abundance of *Lactobacillus* (41.15%). These findings highlight product and region-specific variations in microbial composition at the genus level among fermented bamboo shoot products. Such differences may reflect variations in local fermentation practices; however, they could also be influenced by other factors, including differences in the chemical composition of raw bamboo shoots, environmental conditions, and variations in the fermentation stage at which samples were collected (Fig. 2d). The species-level analysis of the *Lactobacillaceae* family revealed a heterogeneous distribution of LAB across the samples. *Lactobacillus reuteri*, *Lactiplantibacillus plantarum*, and *Levilactobacillus brevis* were among the most dominant species in several fermentation samples, while other species such as *Lactobacillus acetotolerans* and *Lactobacillus paralimentarius* showed variable but notable presence. This pattern highlights the dynamic composition of LAB communities during fermentation (Fig. 2e). Fungal taxa constituted <1% of the total microbial community across all samples, indicating their minimal contribution to the fermentation ecosystem. Therefore, fungal diversity was not emphasized in subsequent analyses, as fermented bamboo shoot is predominantly a bacteria-driven fermentation system dominated by LAB that play the principal role in substrate acidification, preservation, and development of characteristic sensory attributes (Supplementary File 1: ST3).

Across the fermented bamboo shoot products, LAB genera consistently dominated over non-LAB counterparts, though with clear product-specific patterns. *Mesu* showed the strongest LAB predominance, driven by *Limosilactobacillus* (35.7%), with minimal non-LAB presence (*Alcaligenes*, 3.64%). *hirring* and *tuaithar* were characterized by high *Pediococcus* abundance (53.2% and 23.10%, respectively), while *ekung, midukeye*, and *Moiya Pansung* were dominated by *Lactococcus* (49.1%). *Lactiplantibacillus* prevailed in *lung-seij, mileye amileye,* and *moiya koshak* (21.00%–28.58%). In contrast, non-LAB genera such as *Enterobacter*, *Raoultella*, *Comamonas*, *Alcaligenes*, *Aeromonas*, and *Acinetobacter* occurred at lower to moderate levels, with relatively higher representation in *midukeye* and *mileye amileye* (Supplementary File 1: SF1).

No archaea or eukaryotes were found at species level. When the fresh bamboo shoots (*Dendrocalamus latiflorus* Munro) were picked from Dehong (Yunnan, China), it was found that Firmicutes and Proteobacteria were the dominant phyla, while *Lactococcus* and *Lactobacillus* were the dominant genera in sour bamboo shoots (Chen et al. 2022). Metagenomic analysis of *tuaithar* a fermented bamboo shoot of Mizoram revealed *Lactobacillus* as the abundant genus (Deka et al. 2021). Based on the culture-dependent approach, *Lactiplantibacillus plantarum* was predominant in *Soijin, Soidon, Soibum,* and *Gajtenga*. *Soidon* uniquely harbored *Brevibacillus agri*, while *Gajtenga* showed the presence of *Lactobacillus paraplantarum* (Yang et al. 2017). *Melye-amiley*, a distinct fermented bamboo shoot variety associated with the Mog and Chakma tribes of Tripura, exhibited predominance of *Levilactobacillus brevis*, *Lactococcus lactis*, and *L. plantarum* (Barge et al. 2024). Few LAB species such as *Lactobacillus paralimentarius, Lactobacillus zymae, Lactobacillus amylolyticus, Lactobacillus spicheri, Lactobacillus gallinarum, Lactobacillus pentosus, Lactobacillus rossiae, Lactobacillus farciminis,*

Lactobacillus panis, Lactobacillus secaliphilus, Paucilactobacillus vaccinostercus, Lactococcus garvieae, and Weissella paramesenteroides were also found in all FBS of Northeast India that were not reported in culture- dependent method in earlier studies (Das et al. 2025). Interestingly, while certain opportunistic bacteria such as *Stenotrophomonas maltophilia* (0.23%), *Comamonas testosteroni* (0.67%), *Ochrobactrum anthropi* (0.75%), *Acinetobacter* sp., *Aeromonas veronii* (0.43%), *Serratia marcescens* (0.51%), and *Staphylococcus* sp. were detected at >1% relative abundance in some samples. Quality of the water used during the processing of raw bamboo shoot and the hygienic condition may be the reason for the presence of these microorganisms in some of the samples (Denton and Kerr 1998, Tiwari and Nanda 2019, Anjana et al. 2023). Additionally, some of these genera have previously been reported as a common reagent or laboratory contaminants in microbiome sequencing studies and may originate from DNA extraction kits, polymerase chain reaction (PCR) reagents, or laboratory environments, particularly when analyzing low-biomass samples. Previous studies have identified genera such as *Acinetobacter, Bacillus, Bradyrhizobium, Burkholderia, Corynebacterium, Microbacterium, Propionibacterium, Pseudomonas, Ralstonia,* and *Sphingomonas* as frequent contaminants detected in negative controls and reagent kits. These contaminant taxa can increase substantially in relative abundance and may even dominate sequencing profiles in low-biomass samples. Therefore, these taxa were interpreted cautiously in the present study in accordance with previously reported contaminant lists and recommendations for microbiome analyses (Salter et al. 2014).

Alpha diversity analysis revealed notable variations in microbial diversity among the fermented bamboo shoot samples. Among all samples, *Lung-seij* exhibited the highest diversity, as indicated by the greatest Simpson's diversity index (0.873) and Shannon index (2.333), along with the highest species richness estimated by Chao-1 (14.67). In contrast, *Hirring* showed the lowest diversity, with the minimum Simpson's index (0.722), Shannon index (1.619), and Chao-1 value (7.67), indicating comparatively reduced microbial richness. Evenness values ranged from 0.585 to 0.803, with *Mesu* demonstrating the highest evenness (0.803), suggesting a more uniform distribution of microbial taxa within that sample. Overall, the results indicate that traditional fermented bamboo shoot products harbor variable microbial diversity and richness, reflecting differences in fermentation practices, raw materials, and environmental conditions across preparation methods. Detailed values of Simpson's index, Shannon index, evenness, and Chao-1 richness for all samples are provided in (Supplementary File 1: ST4). Higher diversity may result from longer or less controlled fermentations that allow multiple microbes to coexist, whereas lower diversity likely reflects strong selective pressure favouring a few well-adapted fermentative bacteria (Paganin et al. 2021). In a study by Deka et al. (2021), elevated Chao1 and Shannon indices indicate greater species richness and overall community complexity in bamboo shoots (*Tuaithar*) compared with other substrates (*Bekang* and *Sa-um*).

Bray–Curtis beta diversity analysis using PCoA revealed distinct clustering patterns among the fermented bamboo shoot products, indicating differences in microbial community composition. *Lung-seij, tuaithar, bastenga,* and *hirring* grouped closely, suggesting relatively similar microbial profiles. *Midukeye, mileye amileye,* and *moiya koshak* formed another cluster, representing a distinct microbial assemblage. *Ekung* and *moiya pansung* appeared separated from the other groups, indicating greater compositional divergence. Notably, *mesu* formed an independent clade, suggesting a unique microbial community structure compared with the other fermented products (Fig. 3a). Jaccard diversity indices showed distinct clustering. *Moiya pangsung* and *Moiya koshak* grouped together along the negative axis of PCoA, forming a well-defined cluster indicative of similar taxonomic composition. *Midukeye* and *mileye amileye* formed the same cluster, suggesting shared species presence distinct from the other products. On the positive side of PCoA, *bastenga, ekung, tuaithar,* and *lung-seij* clustered closely, reflecting considerable overlap in community membership. *Mesu* and *hirring* grouped together but remained clearly separated from the *bastenga–ekung–lung-seij-tuaithar* cluster, highlighting a distinct set of shared taxa (Fig. 3b). Beta diversity analysis showed clear grouping of fermented bamboo shoot products based on their microbial composition. Some products clustered closely, indicating similar microbial communities, while others formed separate clusters, suggesting distinct microbiome structures. In a study by Wang et al. (2025) the beta-diversity of sour bamboo samples from China indicated significant differences in microbial community composition among regions. In a study by Wu et al. (2025) from Guangxi, China unconstrained PCoA analysis of Bray–Curtis distances revealed that fermentation technique was a major driver of microbial community variation in sour bamboo shoots, as modern and traditional methods generally formed distinct clusters. Samples BEK, NEK, GEK, MBMP, DBMP, and TT clustered closely together, indicating a high degree of similarity in their microbial community composition. A second distinct cluster comprising PBMD, PBMA, DBMD, and DBMK also exhibited strong internal similarity. In contrast, BH, NH, DT, LT, DB, KB, DSJ, and SSJ formed a separate cluster, reflecting considerable variation from the other groups. Additionally, RM, GM, GH, LM, WB, and LSJ grouped into another distinct cluster, suggesting a unique community structure within these samples (Fig. 3c) using Bray-Curtis algorithm. PCoA by use of Jaccard showed the presence of four clear clusters suggesting their presence-absence patterns of the microbial communities across the samples. The samples of PBMD, PBMA, DBMP, DBMK, and BH were clustered and indicated that this group has a common composition of taxa. Another group of MBMP, LM, DBMD, and RM was highly divergent implying a special group of taxa. Additionally, another cluster comprised BEK, GEK, NEK, LSJ, and GM, whereas TT, GH, NH, LT, DT, KB, SSJ, DB, DSJ, and WB significantly differed, which is characteristic of community affiliation (Fig. 3d).

Overall, the taxonomic composition and diversity indices suggest that FBS products from the NER harbour distinct, yet consistently LAB-dominated microbiomes shaped by region-specific fermentation practices. Among the products analysed, *lung-seij* from Meghalaya showed the highest overall microbial as well as LAB diversity, as evident from higher Shannon, Simpson, and Chao 1 indices, indicating a relatively even and complex lactic acid bacterial community. In contrast, products such as *mesu* from Sikkim were characterized by strong dominance of LAB, particularly *Limosilactobacillus*, with comparatively lower richness but greater ecological selection towards specialized fermentative taxa. In addition, FBS products from Tripura, Arunachal Pradesh, Nagaland, and Manipur exhibited intermediate yet clearly product-specific LAB community structures. Differences in LAB diversity and dominance are likely to influence functional potential, metabolite profiles, and probiotic attributes, providing an ecological con-

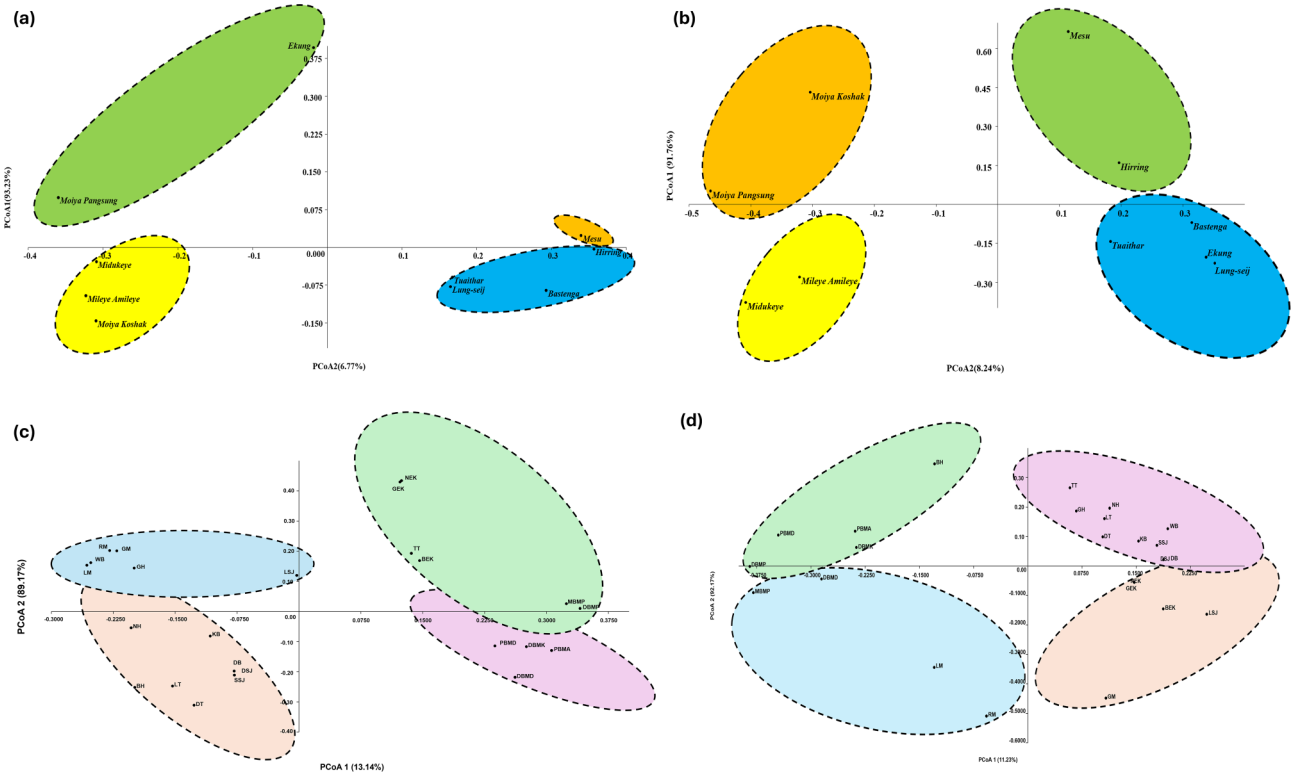

Figure 3 Beta diversity analysis of fermented bamboo shoot microbiomes based on genus-level taxonomic profiles. (a) Principal coordinates analysis (PCoA) based on Bray–Curtis dissimilarity showing differences in taxonomic composition among samples. (b) PCoA based on Jaccard distance showing presence/absence-based taxonomic differences. (c) Bray–Curtis clustering showing grouping patterns among samples based on taxonomic abundance. (d) Jaccard-based clustering illustrating qualitative differences in microbial community composition.

text for the functional patterns discussed in the subsequent sections.

Functional profiling

Mapping of the metagenomic ORFs to the KEGG database revealed diverse and significantly enriched functional pathways across the fermented bamboo shoot samples. Overall, the number of Prodigal-predicted ORFs varied widely among samples, ranging from 13 483 to 50 296 ORFs, reflecting substantial differences in functional gene content across fermentation types and geographic origins (Supplementary File 1: ST5). About 69% of the total ORFs were mapped to KEGG functional pathways. These pathways were further classified into three major functional categories based on KEGG hierarchy. In Level 1, metabolism (64.73%) was found to be most abundant followed by environment information processing (9.66%), genetic information processing (7.50%), human disease (6.73%), cellular process (6.71%), and lastly organismal systems (4.62%) (Fig. 4a). At Level 2, 54 super-pathways and Level 3, 219 sub-pathways from different categories were identified. From Super-pathways, carbohydrate metabolism was found to be the most abundant (15.49%) and in sub-pathways, starch and sucrose metabolism was found to be most abundant (5.49%). Human disease-associated KEGG pathways such as Staphylococcus aureus infection and Central carbon metabolism in cancer were retained only because multiple conserved orthologous genes representing substantial portions of these pathways were detected in the metagenomic dataset, rather than isolated single-gene hits. For example, the *Staphylococcus aureus* infection pathway contained several adhesion, capsules, and virulence-associated genes, including *clfA, clfB, fnbA, fnbB, spa, hla, hlb, capA-capC,* and *agrA/agrC*, indicating broad functional representation rather than sporadic annotation artifacts. Similarly, the "Central carbon metabolism in cancer" pathway included a nearly complete bacterial central carbon metabolism module comprising glycolytic and TCA cycle genes such as *pgi, pfkA, gapA, eno, pyk, aceE, aceF, gltA,* and *icd* (Supplementary File: ST6).

Correlation network analysis revealed selective and distinct associations between dominant probiotic bacteria and key KEGG functional pathways related to metabolism, stress response, and cellular processes. The refined network reduced edge density and enabled clearer identification of taxa-specific functional contributions. Core taxa, including *Lactiplantibacillus plantarum*, *Limosilactobacillus fermentum*, *Lactococcus lactis*, and *Levilactobacillus brevis*, exhibited predominantly positive correlations with specific pathways related to carbohydrate utilization, amino acid metabolism, oxidative phosphorylation, ABC transporters, quorum sensing, and two-component systems. Notably, these associations were not uniformly distributed across all taxa but were concentrated within key functional contributors, highlighting their distinct metabolic roles. This pattern is biologically consistent, as these LAB are metabolically specialized for efficient carbohydrate fermentation and nutrient uptake, which are essential processes in fermented food ecosystems. The enrichment of

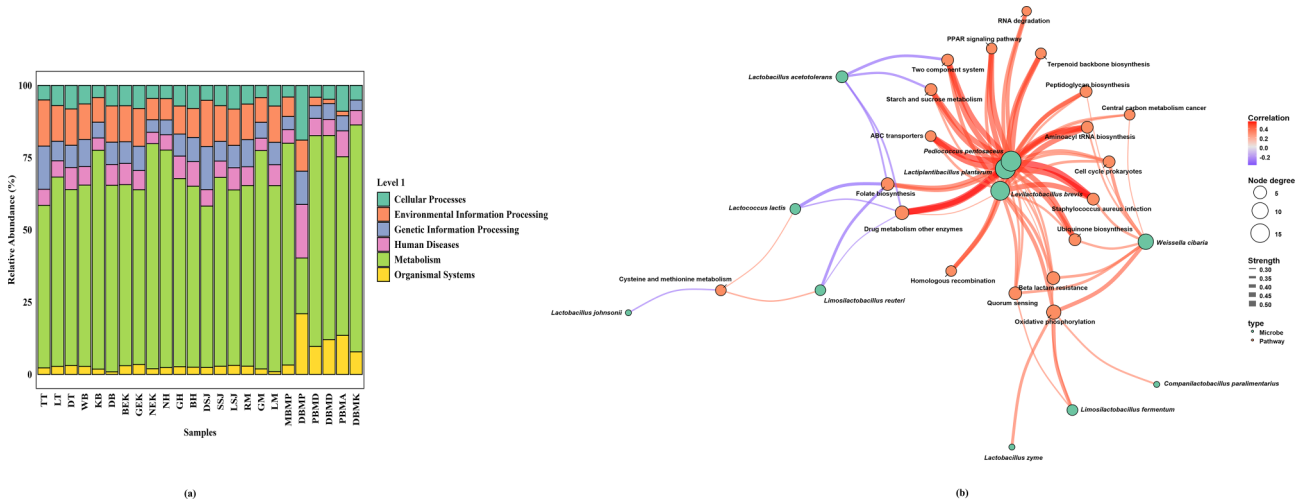

Figure 4 Predictive Functionality of the traditional fermented bamboo shoot (a) Relative abundance of predicted functional categories annotated against the KEGG database at Level 1, displaying the contribution of major categories such as Metabolism, Genetic Information Processing, Environmental Information Processing, Cellular Processes, Human Diseases, and Organismal Systems across samples. (b) Microbe–pathway association network illustrating correlations between dominant microbial taxa and predicted metabolic pathways. Green nodes represent microbial taxa, while orange nodes indicate metabolic pathways. Node size is proportional to node degree, reflecting the number of associations. Edge colour denotes correlation direction and strength, with red indicating positive correlations and blue/purple indicating negative correlations; colour intensity represents correlation magnitude. Edge thickness corresponds to interaction strength. Highly connected hub nodes indicate taxa or pathways with central roles in the network. The network highlights key functional relationships among LAB and pathways related to carbohydrate metabolism, transport systems, biosynthesis, resistance mechanisms, and cellular processes. .

signaling pathways such as quorum sensing and two-component systems further suggests their ability to adapt to dynamic environmental conditions and coordinate community behavior. Additionally, the positive association of major LAB species with folate biosynthesis suggests a potential role in cofactor production that supports microbial growth and metabolic activity . Correlations with peptidoglycan biosynthesis reflect active cell wall formation and turnover, which are essential for cell division and structural stability in rapidly growing microbial populations (Typas et al. 2010). Furthermore, the linkage with biofilm-associated pathways indicates the ability of these taxa to adhere to surfaces and form structured communities, providing protection against environmental stress and facilitating persistence (Shree et al. 2023). Together, these functional traits mechanistically explain how these microbes maintain growth, structural integrity, and long-term survival in fermented environments (Fig. 4b). Carbohydrate metabolism emerged as a dominant functional category across all FBS samples, consistent with previous observations from Tripura-based fermented products (Das et al. 2023). This pattern reflects the polysaccharide-rich composition of bamboo shoots, which provide substrates such as cellulose, hemicellulose, and soluble sugars that shape microbial functional potential (Li et al. 2022). Functional annotation further highlighted the concurrent predominance of amino acid metabolism and energy production pathways, indicating coordinated metabolic networks underlying fermentation (Hu et al. 2021). Rather than merely reflecting substrate availability, the predominance of carbohydrate-active pathways suggests selective adaptation of microbial communities toward efficient carbon utilization and energy generation. These processes drive the production of fermentation end-products, including organic acids and other metabolites, which contribute to environmental acidification, microbial stability, and characteristic flavour development (Malekijahan et al. 2025).

Probiotic gene annotation in FBS of NER

Distinct variation in the distribution of probiotic-associated genes was observed across samples. Core functional genes, including ATP synthase subunits (*atpA, atpB, atpC, atpD, atpE, atpF, atpG, atpH*) and NADH dehydrogenase (*ndh*), adhesion-related factors (*efaA*), and bile salt hydrolase (*bsh*), were consistently present across most samples, indicating conserved functional potential. In contrast, genes associated with stress response (*groEL, phoR, phoP, liaS, liaR, bceS, bceR, walK, walR*), acid tolerance (*gad*), and antimicrobial activity (*bacA*) exhibited sample-specific variability. Additionally, predominance of quorum-sensing genes (*luxS, agrA, agrC*), peptide transport systems (*oppA, oppB, oppC, oppD, oppF; dppA–dppF*), and multidrug/bile resistance modules (*bceA, bceB*) along with bacteriocin exporters (*lanT, nisT*) suggests enhanced adaptive, competitive, and antimicrobial capabilities. Functional genes involved in folate biosynthesis (*folA–folE*), sulfur amino acid metabolism (*cysE, cysK, metE, metH*), and carbohydrate utilization (*amyA, malE, malE, malG, malK, glgP*) further highlight the metabolic versatility of the microbial community (Fig. 5a). Such stress-response (*phoR/phoP, liaS/liaR*) and quorum-sensing (*luxS*) genes have been widely reported in LAB from fermented foods such as pearl millet slurry and plant-based fermentations, where they support environmental adaptation, biofilm formation, and microbial interactions (Turpin et al. 2011, Almeida et al. 2021). Similarly, transport and metabolic genes (*opp, bce, fol, cys, atp, ndh*) are commonly identified in LAB from fermented products like *kinema, dahi*, and *kefir*, reflecting their roles in nutrient utilization,

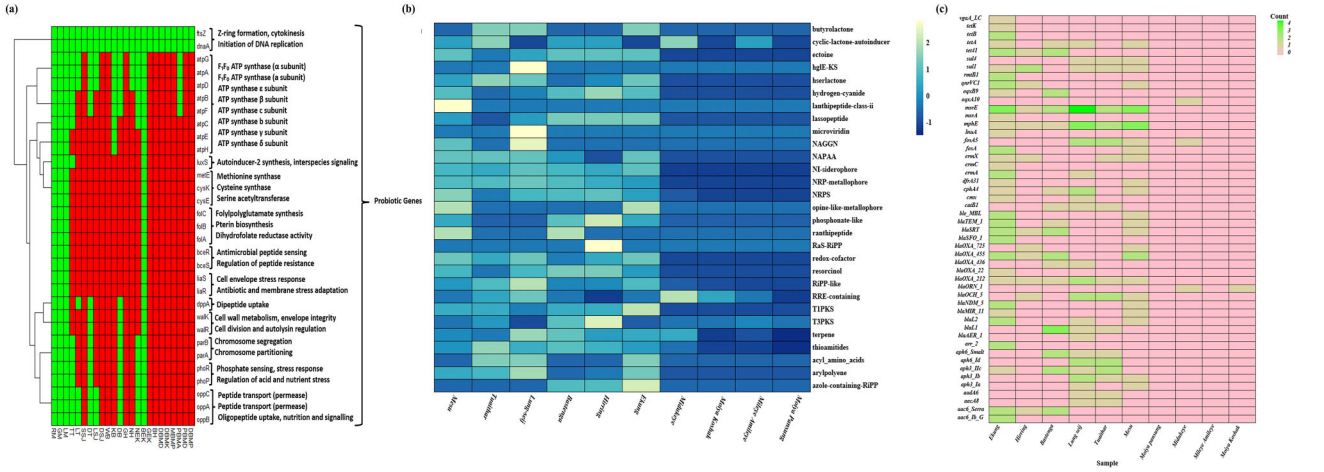

Figure 5 Probiotic traits, secondary metabolite potential, and antimicrobial resistance profiles of fermented bamboo shoot (a) Presence–absence heatmap of predicted probiotic and stress-resistance genes across all strains, showing clustering of genes related to Z-ring formation, ATP synthase, amino acid and folate biosynthesis, antimicrobial peptide sensing, cell envelope stress response, division and autolysin regulation, chromosome partitioning, phosphate and acid-stress response, and oligopeptide/peptide transport systems. (b) Heatmap of biosynthetic gene clusters predicted by antiSMASH, depicting distribution and relative abundance of secondary metabolite classes (e.g. butyrolactone, cyclic-lactone autoinducer, ectoine, NRPS, RiPP-like, terpenes, and thioamides) across representative genomes. (c) AMR gene heatmap showing the occurrence of individual ARGs (e.g. *van, tet, erm, bla, aac/aph* genes) in each sample, highlighting low but heterogeneous AMR gene loads in the probiotic candidates.

antimicrobial activity, and survival under acidic and gastrointestinal conditions (Clemens et al. 2018, Talib et al. 2019, Tamang et al. 2024).

Secondary metabolite and AMR gene analysis

Comparative analysis of secondary metabolite biosynthetic gene clusters (BGCs) revealed product-specific variation among the ten fermented bamboo shoot products. *Mesu, tuaithar, lung-seij, bastenga, hirring,* and *ekung* contained a broader diversity of BGC profiles, whereas *midukeye, moiya koshak, mileye amileye,* and *moiya pansung* showed limited representation. RiPP-related clusters were most abundant in *lung-seij* (29), followed by *tuaithar* and *ekung* (19 each), *mesu* (12), and *hirring* (10). Thioamitide clusters were observed in *tuaithar* (31), *bastenga* (25), *hirring* (22), *lung-seij* (21), and *ekung* (18), with lower counts in the remaining products (2–11). NRPS clusters were present in *mesu* (17), *hirring* (14), *bastenga* (13), and *lung-seij* (12), while *midukeye, moiya koshak, mileye amileye,* and *moiya pansung* showed minimal presence (0–2). Siderophore-associated clusters (NI-siderophores and NRP-metallophores) were detected in the first six products (NI-siderophores: 5–7; NRP-metallophores: 4–7) and were absent in the remaining four. Terpene clusters were distributed across products, with counts in *lung-seij* (11), *bastenga* (9), and *ekung* (7). Polyketide synthase clusters showed variable distribution, including T3PKS in *hirring* (12) and *bastenga* (7). Quorum-sensing and other specialized clusters were infrequent and product-specific (Fig. 5b and Supplementary File 1: ST7). Comparative analysis of secondary metabolite biosynthetic gene clusters (BGCs) revealed product-specific variation among fermented bamboo shoot samples. *Mesu, tuaithar, lung-seij, bastenga, hirring,* and *ekung* showed greater diversity, whereas *midukeye, moiya koshak, mileye amileye,* and *moiya pansung* exhibited limited representation. Similar trends have been reported in other fermented foods, where vegetable-based fermentations such as *kimchi, sauerkraut* and *gundruk* exhibit higher RiPP and NRPS diversity, while dairy and cereal fermentations show different BGC distributions (Jung et al. 2011, Gautam and Sharma 2015, Ismail et al. 2018, Veeranan Arun Giridhari et al. 2025).

Diverse AMR genes were detected across different FBS products of Northeast India. Genes conferring resistance to aminoglycosides, β-lactams, macrolides, tetracyclines, sulfonamides, and fluoroquinolones were identified, with marked variation among samples. *Lung-seij, tuaithar,* and *bastenga,* showed comparatively higher AMR gene abundance, whereas *moiya-pansung, midukeye, mileye-amileye*, and *moiya-koshak* exhibited minimal or no detectable AMR genes. β-lactam resistance genes, particularly *blaOXA* variants, *blaTEM-1, blaL1/L2,* and *blaNDM-5,* were predominantly observed in *ekung, bastenga, lung-seij,* and *mesu*. Aminoglycoside resistance genes such as *aac, aad*, and *aph* families were frequently detected, with *aph* (3′) and *aph* (6) variants showing higher occurrence in *lung-seij* and *tuaithar.* Macrolide resistance genes, including *mph*(E), *msr*(E), and *erm* genes, were widely distributed, with the highest counts observed in *Lung-seij*. Tetracycline resistance genes *tet*(41), *tet*(A), *tet*(B), and *tet*(K) were mainly confined to *ekung, bastenga,* and *hirring*. Sulfonamide resistance genes *sul1* and *sul4* were sporadically detected, while fluoroquinolone resistance gene *qnr* VC1 occurred at low frequency (Fig. 5c and Supplementary File 1: ST8). Comparative analysis of AMR gene profiles revealed product-specific variation among fermented bamboo shoot samples. *Lung-seij, tuaithar, bastenga, ekung,* and *mesu* exhibited relatively higher AMR gene diversity and abundance, whereas *midukeye, moiya-pansung, mileye-amileye,* and *moiya-koshak* showed limited or negligible representation. Similar patterns have been documented in other fermented foods, where dairy fermentations such as *cheese* and *yogurt* are frequently enriched in multidrug and β-lactam resistance genes, while brine-fermented vegetables including *kimchi*, pickled vegetables, and fermented bamboo shoots predominantly harbor tetracycline-associated resistance determinants; in contrast,

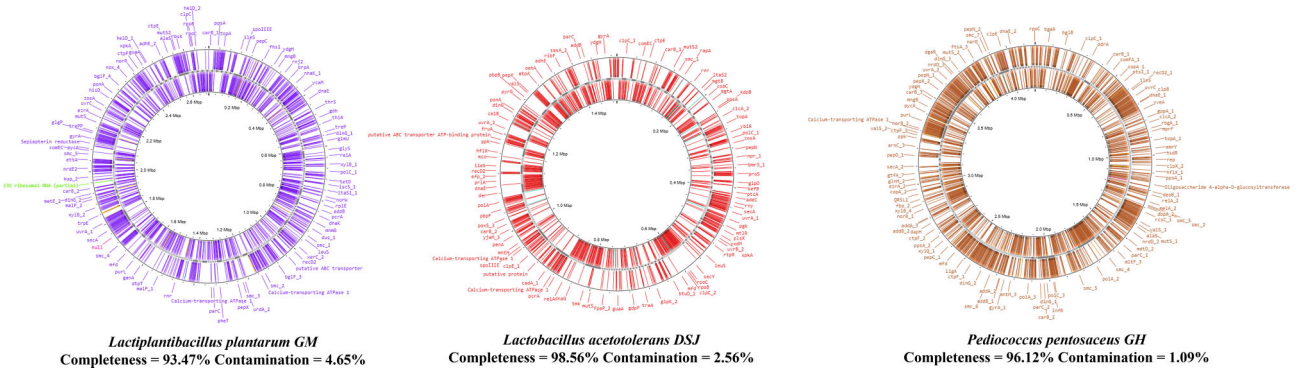

Figure 6 Genome architecture and functional potential of dominant probiotic MAGs from fermented bamboo shoots (a) Circular genome map of *Lactobacillus acetotolerans* DS3 showing GC content, GC skew, coding sequences on forward and reverse strands, and functional annotation blocks grouped into probiotic properties, short-chain fatty acid production, immunomodulation, anti-tumor traits, stress tolerance, essential amino acid and vitamin biosynthesis, and biosynthetic peptides. (b) Circular genome map of *Pediococcus pentosaceus* GH highlighting GC features, coding sequence distribution, and gene clusters associated with probiotic properties, short-chain fatty acids, immunomodulatory functions, anti-tumor potential, stress tolerance, essential amino acid and vitamin biosynthesis, and biosynthetic peptide production. (c) Circular genome map of *Lactiplantibacillus plantarum* GM depicting GC content, GC skew, coding sequences, and annotated loci linked to probiotic traits, short-chain fatty acid pathways, immunomodulation, anti-tumor mechanisms, stress tolerance, essential amino acid and vitamin biosynthesis, and biosynthetic peptides.

sugar-based fermentations such as *kombucha* and *water kefir* often display minimal or undetectable AMR gene prevalence (Leech et al. 2020). Overall, the AMR profiling indicates that while certain fermented bamboo shoot products, such as *Lung Sej* from Meghalaya, Tuaithar from Manipur, and Bastenga from Nagaland, harbour comparatively higher AMR gene abundance, products from Sikkim and Tripura exhibit the lowest AMR prevalence, suggesting relatively safer resistome profiles among these fermentations.

High-quality MAGs

As part of a comprehensive metagenomic study of FBS products from NER, MAG analysis was performed to obtain strain-level insights into the dominant LAB and to evaluate their potential suitability for probiotic applications through genome-based assessment of functional and safety-related traits. A total of 16 high quality MAGs were extracted from 6 different samples of fermented bamboo shoots. *Tuaithar*, *bastenga*, *ekhung*, *hirring*, *mesu*, and *lung-siej* were able to produce MAGs. Out of 16 MAGs extracted, 7 belonged to *Lactiplantibacillus plantarum,* 7 belonged from *Lactobacillus acetotolerans,* and 2 MAGs belonged to *Pediococcus pentosaceus,* their completeness and contamination are represented in Supplementary File 1: ST9. MAG of *Lactiplantibacillus plantarum* strain GM, *Lactobacillus acetotolerans* strain DSJ, and *Pediococcus pentosaceus* strain GH had the highest number of genes and are illustrated in Fig. 6 a–c. Other MAGs have been illustrated in Supplementary File 1: SF2-13. The functional potential of MAGs was annotated using KEGG, and gene counts were classified into three hierarchical levels (Level 1–3) based on pathway assignment. In level 1, metabolism was found to be most abundant followed by genetic information processing, environmental information processing, cellular process, organismal systems, and human disease. At Level 2, 37 super-pathways and at Level 3, 102 different sub-pathways from different category were identified (Supplementary File 1: ST10a, b and c). Similar LAB MAGs prevail in other fermented foods *L. plantarum* and *P. pentosaceus* in *jiang-sun* a fermented bamboo shoots from Taiwan (Chen et al. 2010), *nor mai som* a FBS from Thailand (Botthoulath et al. 2024), *sauerkraut, kimchi, gundruk, suancai,* and *hamei* (fermented fish) (Swain et al. 2014, Pasolli et al. 2020, Jain and Rajagopal 2024), often linked to acid tolerance, bacteriocin production, and fermentation stability.

Probiotic attributes of MAGs

Genome-resolved metagenomic analysis of MAGs revealed a functionally conserved yet quantitatively variable repertoire of probiotic-associated genes across samples and product types. Core probiotic traits including acid tolerance (*atp* operon, *clp* genes), bile resistance (*bsh*), adhesion (*eno, gtfA*), and stress adaptation (*dnaK, groEL, hrcA*) were consistently detected across all samples, indicating a stable functional backbone of LAB communities. However, relative gene abundance and pathway completeness varied across samples (e.g. GEK, NEK, BH), suggesting product-specific functional predominance. For instance, samples such as GEK and NEK exhibited higher representation of stress-response and adhesion-related genes, while BH showed comparatively enriched metabolic pathways linked to amino acid and vitamin biosynthesis. This indicates functional differentiation driven by fermentation conditions and substrate composition, as previously reported in traditional fermented foods (Zheng et al. 2020)

Members of *Lactiplantibacillus* plantarum and *Pediococcus pentosaceus* MAGs displayed greater predominance of adhesion (*eno, scpA/B*) and bile resistance (*bsh*) genes, whereas *Lactobacillus acetotolerans* showed relatively lower representation of these traits but retained genes linked to carbohydrate metabolism (*bgl* genes) and flavour formation (*mmuM*). Such functional partitioning within LAB taxa aligns with previous comparative genomic studies of LAB, where strain-level variation strongly influences probiotic potential (Tenea et al. 2025). Key metabolic functions, including short-chain fatty acid production (pta, ldh), exopolysaccharide biosynthesis (*epsH*), and antioxidant systems (*trxA, nrd* genes), were widely conserved, supporting their essential role in host interaction and gut persistence. Notably, complete biosynthetic pathways for essential amino acids (e.g. arginine, methionine, tryptophan) and vitamins (folate, riboflavin, thiamine) were

identified in several MAGs, with higher pathway completeness observed in dominant Lactobacillaceae members, suggesting their contribution to nutritional enhancement. In comparison with previously characterized LAB genomes, such as *Lactiplantibacillus plantarum* WCFS1 and *Lacticaseibacillus rhamnosus* GG, the detected gene repertoire shows high concordance in probiotic determinants, including *bsh*-mediated bile resistance, *eps*-driven immunomodulation, and stress-response systems (*clp, groEL*) (Önning et al. 2022). However, the present dataset demonstrates greater diversity in secondary metabolic and peptide-processing genes (*pep* family, *htpX*), likely reflecting adaptation to complex plant-based fermentation substrates. Among CAZymes, glycoside hydrolases (300) and glycosyltransferases (251) predominated across all MAGs, with additional auxiliary activity enzymes, carbohydrate esterases, and carbohydrate-binding modules also detected (Supplementary File 1: SF14).

In silico-Safety assessment of MAGs

In total, eight types of cas genes were identified, *Cas1-TypeI, cas2-TypeI, cas3-TypeI, cas5-TypeI, cas6-TypeI, cas7-TypeI, cas9-TypeII,* and *cas1-TypeII.* Genes for antibiotic resistance were also detected in 13 MAGs out of 16 belonging to the family of vancomycin resistance (*vanT, vanY,* and *vanH*). *Lactiplantibacillus plantarum* GM, *Lactobacillus acetotolerans* DSJ, and *Pediococcus pentosaceus* GH didn't consist of any antibiotic resistance gene. No mobile genetic elements, phage sequence and virulence genes were detected in any 16 MAGs. While *van* family genes (vancomycin resistance) were detected in 13 MAGs, this is likely an intrinsic, non-transmissible trait common in *Lactobacillaceae* due to D-Ala-D-Lac peptidoglycan precursors (Kruse et al. 2014). Crucially, the absence of mobile genetic elements (MGEs) and virulence factors, coupled with the identification of CRISPR-Cas systems (Types I and II), suggests high genomic stability (Pinilla-Redondo et al. 2022). The presence of diverse cas genes indicates an active adaptive immune system that protects the integrity of the bacterial genome against bacteriophage-mediated lysis, which is a frequent cause of fermentation failure. The total absence of AR genes in *L. plantarum* GM, *L. acetotolerans* DSJ, and *P. pentosaceus* GH marks these specific strains as superior candidates for development as safe starter cultures. Plasmid replicon analysis of the reconstructed MAGs revealed the presence of five plasmid-associated replicons belonging to the Rep3 and RepA_N families across lactic acid bacterial species. The Rep3 family was the most abundant, with four replicons detected. A rep31 replicon was identified in the MAG LPGEK (*Lactiplantibacillus plantarum*) showing 97.84% identity over 1155 bp, corresponding to plasmid LCKp400005 (pLCK4) located on contig-100_314. Another rep31 replicon was detected in PPGH (*Pediococcus pentosaceus*) with 98.01% identity across 1158 bp on NODE_82_length_75 956_cov_60.65. Similarly, *Lactobacillus acetotolerans* (LAKB) also harbored a rep31 replicon showing 96.98% identity over 1158 bp. In addition, a rep28 replicon belonging to the Rep3 family was identified in LADSJ (*Lactobacillus acetotolerans*) with 99.68% identity across 930 bp, corresponding to repA (pCIS4). Furthermore, a RepA_N family replicon (rep38) was detected in LAKB (*Lactobacillus acetotolerans*) with 98.46% identity over 1104 bp, associated with plasmid pLBUC03.

Plasmid-associated replicons were detected in 5 out of 14 reconstructed MAGs (∼35.7%), indicating a notable prevalence of putative plasmid elements among the lactic acid bacterial genomes analysed. In total, five replicons were identified, predominantly belonging to the Rep3 family ($n = 4$), with a single replicon assigned to the RepA_N family. These replicons were distributed across *Lactiplantibacillus plantarum*, *Pediococcus pentosaceus*, and *Lactobacillus acetotolerans*, suggesting a conserved presence of plasmid-associated elements within these taxa (Supplementary File 1: ST11).

Genome-resolved analysis revealed a diverse repertoire of antiviral defense systems across the reconstructed MAGs. Multiple defense modules, including restriction–modification (RM) systems, CRISPR—Cas, Gabija, CBASS, BREX, Thoeris, Abi systems, PD-T4-9, DdmDE, DS-6, ShosTA, and Gao—Her were detected, indicating that the microbial populations possess multiple layers of protection against bacteriophage infection. Among the analysed MAGs, PPGH encoded the highest diversity of antiviral defense systems, with several modules detected across multiple contigs. These included RM Type I and Type III systems, Gabija, AbiD, AbiAlpha, and the DdmDE defense system, suggesting the presence of a complex and redundant antiviral defense network. Such diversity of defense modules within a single genome is often associated with strong phage predation pressure, enabling microbes to deploy multiple strategies to inhibit viral replication. The coexistence of abortive infection systems (Abi) and restriction–modification systems indicates both intracellular viral detection and direct degradation of invading phage DNA (Antine et al. 2024). Similarly, the PPBH MAG harbored multiple defense mechanisms, including Thoeris type I, CRISPR—Cas Class I subtype I-E, and the Gao—Her defense system, representing both adaptive and innate immune strategies. The presence of CRISPR–Cas systems indicates that these microorganisms retain genetic memory of prior phage infections, allowing sequence-specific targeting of invading viral genomes. The Thoeris system, which is known to function through cyclic nucleotide signaling, further strengthens antiviral defense by triggering cellular responses upon phage detection (Ibrahim and Aranjani 2026).

In contrast, several MAGs including LPGM, LPGEK, LPDB, and LPTT were primarily characterized by restriction–modification systems, particularly RM Types I and IV. RM systems are among the most widespread bacterial defense mechanisms and function by recognizing and cleaving foreign DNA lacking host-specific methylation patterns (Birkholz et al. 2022). The LPGEK MAG additionally encoded a BREX type I system, which provides phage resistance through epigenetic modification of host DNA and inhibition of phage replication. Other innate defense systems were distributed across multiple genomes (Xu and Gu 2024). Gabija and PD-T4-9 systems were identified in MAGs such as LAKB, LANH, LADSJ, LALM, and PPGH, indicating their widespread occurrence within the microbial community. The CBASS type I system detected in the LAKB MAG represents a cyclic oligonucleotide-based signaling pathway that activates antiviral responses following phage infection. Furthermore, abortive infection systems such as AbiD and AbiAlpha were detected in several MAGs including LAWB, LAKB, LALM, LADSJ, and PPGH, suggesting that infected cells may sacrifice themselves to prevent phage propagation within the population (Egido et al. 2022). Adaptive immunity was represented by CRISPR–Cas systems in LANH, LPLSJ, and PPBH, including Class II subtype II-A and Class I subtype I-E systems. These systems provide sequence-specific immunity and are often associated

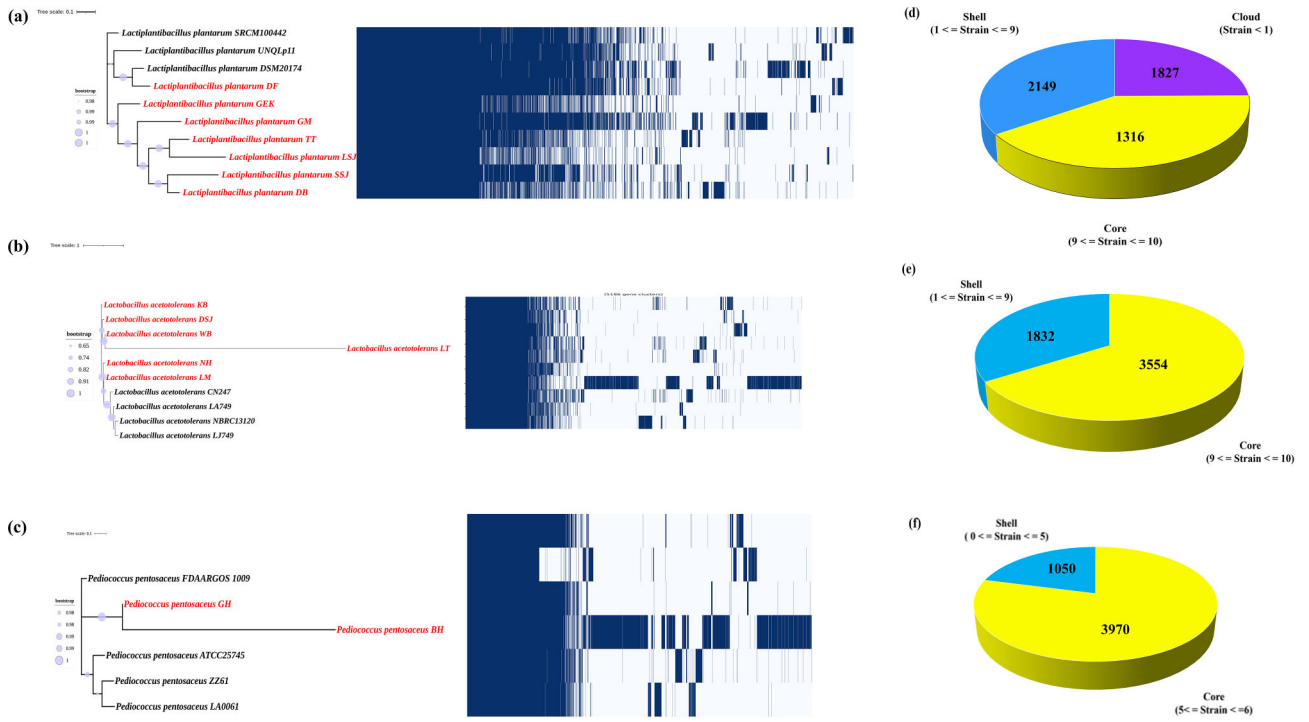

Figure 7 Pangenome analysis of selected LAB. (a–c) Phylogenetic trees and gene presence–absence heatmaps of *Lactiplantibacillus plantarum*, *Lactobacillus acetotolerans*, and *Pediococcus pentosaceus*. In the trees, strain names highlighted in red represent MAGs obtained in this study, while black labels indicate reference genomes. The heatmaps show gene distribution across strains. (d–f) Pangenome composition of the respective species, showing the distribution of core, shell, and cloud genes. Core genes are shared by most strains, shell genes are present in multiple but not all strains, and cloud genes represent strain-specific genes identified during pangenome analysis (i.e. genes present in only one genome). Cloud genes do not represent genes absent from all genomes; rather, they correspond to singleton gene clusters derived during pangenome construction.

with microbial populations experiencing frequent viral exposure (Supplementary File 1:ST11) (Joseph et al. 2025).

Pangenome analysis of high-quality MAGs

Pangenome analysis of the highlighted strains revealed distinct genomic features among the dominant LAB. The red-highlighted *Lactiplantibacillus plantarum* isolates strains GM TT LSJ, SSJ DB clustered closely with reference genomes, sharing a large, conserved core genome while displaying moderate accessory gene variability, indicating strain-level adaptation within the species. The blue-highlighted *Lactobacillus acetotolerans* strains DSJ, KB, WB, LM, LT, NH formed a tight phylogenetic cluster with minimal accessory genome divergence, reflecting high genomic conservation and specialization for acidic fermentation environments. In contrast, the pink-highlighted *Pediococcus pentosaceus* strains BH and GH exhibited greater variability in the accessory genome, with distinct gene blocks unique to individual strains, suggesting enhanced genomic flexibility and functional diversification. These patterns underscore species-specific strategies of genomic conservation and adaptation among the highlighted fermentation-associated bacteria (Fig. 7 a–c). The selection of reference genomes for phylogenetic reconstruction reflects the broad ecological versatility of LAB, spanning both host-associated and food-associated environments. For instance, *Lactiplantibacillus plantarum* DSM 20174, originally isolated from human saliva, represents the capacity of this species to colonize mucosal niches, whereas strains such as SRCM100442 and UNQLp11, derived from plant-based and dairy fermentations, respectively, highlight its well-established role in diverse food systems. This dual ecological distribution is consistent with previous studies demonstrating the remarkable genomic plasticity and niche adaptability of *L. plantarum*, enabling survival across gastrointestinal and fermentation environments (Kleerebezem et al. 2003, Zheng Jinshui et al. 2020). *Lactobacillus acetotolerans* strains (e.g. CN247, LA749, LJ749, NBRC 13120) are predominantly associated with acidic, carbohydrate-rich fermented foods, reflecting a more specialized ecological adaptation. Their frequent occurrence in plant- and cereal-based fermentations suggests a functional role in acid tolerance, carbohydrate metabolism, and fermentation stability, which has been widely reported in traditional fermented food microbiomes (Park and Mannaa 2025). *Pediococcus pentosaceus* reference strains (ATCC 25745, ZZ61, LA0061) are consistently linked to plant-derived fermentation systems, particularly vegetables and cereals. This ecological association supports their established contribution to lactic acid production, food preservation, and flavour development, as well as their probiotic potential in plant-based fermented foods (Qi et al. 2021). The pangenome of *Lactiplantibacillus plantarum* was composed of 1316 core genes shared across nearly all strains, along with a substantial shell genome (2149 genes) and a large cloud genome (1827 genes), indicating high genomic diversity and an open pangenome structure. In contrast, *Lactobacillus acetotolerans* exhibited a more conserved pangenome, with 3554 core genes and 1832 shell genes, reflecting strong genomic stability and limited strain-specific variation.

Pediococcus pentosaceus displayed a pangenome dominated by shell genes (3970 genes) and a smaller core genome (1050 genes), suggesting pronounced genomic plasticity and strain-level functional diversification (Fig. 7 d–f). Pangenome analysis revealed divergent evolutionary trajectories. The open pangenome of *L. plantarum* confirms its status as a genomic generalist, capable of niche-filling through the acquisition of accessory genes (Martino et al. 2016). In contrast, the highly conserved pangenome of *L. acetotolerans* suggests a specialist strategy, optimized specifically for the high-acid environment. We hypothesize that the genomic flexibility of *P. pentosaceus* (high shell gene count) allows it to act as a metabolic bridge, adapting its functional output to the specific phytochemical variations of different bamboo species. These species-specific strategies conservation in specialists versus diversification in generalists ensure the ecological stability of the fermented bamboo shoot ecosystem.

Conclusion

This study provides a comprehensive comparative metagenomic overview of traditionally FBS products consumed across Northeast India, elucidating their microbial diversity, functional potential, biosynthetic capacity, probiotic attributes, and AMR profiles. The analysis revealed distinct product- and region-specific microbiome signatures, while *Lactiplantibacillus*, *Levilactobacillus, Lactobacillus*, *Lactococcus*, and *Pediococcus* formed the core microbiome in these FBS products. Several products, notably *mesu* (Sikkim), *tuaithar* (Manipur), *lung-seij* (Meghalaya), and *bastenga* (Nagaland), exhibited enriched probiotic-associated genes, diverse secondary metabolite biosynthetic gene clusters, and higher functional versatility compared to other products. Functional annotations highlighted the central role of carbohydrate metabolism, stress-response systems, energy metabolism, and antimicrobial peptide biosynthesis in shaping fermentation dynamics, microbial resilience, and probiotic potential. The presence of diverse biosynthetic gene clusters, including bacteriocins, non-ribosomal peptides, terpenes, and siderophores, underscores the capacity of FBS microbiomes to contribute to food safety, microbial competitiveness, and potential health benefits. AMR profiling revealed a generally low resistome burden dominated by intrinsic resistance determinants, with limited evidence of mobile genetic elements, suggesting ecological adaptation rather than strong anthropogenic antibiotic pressure. Among all products, *mesu* from Sikkim demonstrated the most balanced functional, biosynthetic, and AMR-related profile. Nevertheless, the study is limited by its reliance on metagenomic predictions, which do not confirm gene expression or metabolite production, and by the absence of longitudinal sampling and culture-based validation. Future work integrating metatranscriptomics, metabolomics, targeted culturomics, and controlled fermentation studies will be crucial to link genetic potential with functional activity and to evaluate safety. Overall, this work establishes a foundational framework for understanding FBS microbiomes and supports their potential exploitation in functional food development and sustainable biotechnological applications.

Supplementary material

Supplementary material is available at *FEMSMC Journal* online.

Acknowledgements

The author (Rohit Das) is grateful to the Council of Scientific and Industrial Research for providing CSIR SRF-Direct (File No: 09/1114(18395)/2024-EMR-I) Fellowship during this research work.

Conflicts of 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.

Funding

This research did not receive intramural or extramural grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

The data of all the metagenomic shotgun sequences is publicly available in NCBI with BioProject bearing PRJNA1201339 for *tuaithar* (TT, DT, and LT), PRJNA1152108 for *lung-siej* (LSJ, DSJ, and SSJ), PRJNA1150940 for *mesu* (RM, LM, and GM), PRJNA1150661 for *bastenga* (WB, DB, and KB), and PRJNA1478174 for *hirring* (NH, GH, and BH), and *ekung* (BEK, NEK, and GEK). The MAGs extracted were also submitted under the BioProject bearing PRJNA1265206.

References

Alcock BP, Raphenya AR, Lau TTY *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2020;**48**:D517–25. https://doi.org/10.1093/nar/gkz935

Almeida OGGd, Vitulo N, Martinis ECPD *et al.* Pangenome analyses of LuxS-coding genes and enzymatic repertoires in cocoa-related lactic acid bacteria. *Genomics* 2021;**113**:1659–70. https://doi.org/10.1016/j.ygeno.2021.04.010

Anjana A, Adhikary R, Bhavana MV *et al.* Two Case Reports of Ochrobactrum anthropi Bacteremia: an Overlooked Pathogen. *J Lab Phys* 2023;**15**:166–8. https://doi.org/10.1055/s-0042-1757235

Antine SP, Johnson AG, Mooney SE *et al.* Structural basis of Gabija anti-phage defence and viral immune evasion. *Nature* 2024;**625**:360–5. https://doi.org/10.1038/s41586-023-06855-2

Badwaik LS, Borah PK, Borah K *et al.* Influence of fermentation on nutritional compositions, antioxidant activity, total phenolic and microbial load of bamboo shoot. *Food Sci Technol Res* 2014;**20**:255–62. https://doi.org/10.3136/fstr.20.255

Barge SR, Bhattacharya A, Kumar A *et al.* A traditional fermented bamboo shoot reduces intracellular fat accumulation and promotes fat browning in differentiated 3T3-L1 adipocyte cells through the activation of the AMPK signaling pathway. *Food Front* 2024;**5**:1642–57. https://doi.org/10.1002/fft2.405

Birkholz N, Jackson SA, Fagerlund RD *et al.* A mobile restriction-modification system provides phage defence and resolves an epigenetic conflict with an antagonistic endonuclease. *Nucleic*

Acids Res 2022;**50**:3348–61. https://doi.org/10.1093/nar/gkac147

Blin K, Shaw S, Augustijn HE *et al.* AntiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 2023;**51**:W46–50. https://doi.org/10.1093/nar/gkad344

Blin K, Shaw S, Vader L *et al.* AntiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. *Nucleic Acids Res* 2025;**53**:W32–8. https://doi.org/10.1093/nar/gkaf334

Botthoulath V, Dalmacio IF, Elegado FB. Physico-chemical and functional properties of the Lao fermented bamboo shoots (Nor Mai Som) inoculated with potential probiotic bacteria, Pediococcus pentosaceus BBS1 and Lactiplantibacillus plantarum BBS13. *Food Chem Adv* 2024;**5**:100803. https://doi.org/10.1016/j.focha.2024.100803

Brown CL, Mullet J, Hindi F *et al.* mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. *Appl Environ Microb* 2022;**88**:e00991–22. https://doi.org/10.1128/aem.00991-22

Cantalapiedra CP, Hernandez-Plaza A, Letunic I *et al.* eggNOG-mapper v2: functional Annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 2021;**38**:5825–9. https://doi.org/10.1093/molbev/msab293

Carattoli A, Zankari E, Garciá-Fernández A *et al.* In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 2014;**58**:3895–903. https://doi.org/10.1128/AAC.02412-14

Chalita M, Kim YO, Park S *et al.* EzBioCloud: a genome-driven database and platform for microbiome identification and discovery. *Int J Syst Evol Microbiol* 2024;**74**:006421. https://doi.org/10.1099/ijsem.0.006421

Chaumeil PA, Mussig AJ, Hugenholtz P *et al.* GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 2022;**38**:5315–6. https://doi.org/10.1093/bioinformatics/btac672

Chen C, Cheng G, Liu Y *et al.* Correlation between microorganisms and flavor of Chinese fermented sour bamboo shoot: roles of Lactococcus and Lactobacillus in flavor formation. *Food Biosci* 2022;**50**:101994. https://doi.org/10.1016/j.fbio.2022.101994

Chen YS, Wu HC, Liu CH *et al.* Isolation and characterization of lactic acid bacteria from Jiang-sun (fermented bamboo shoots), a traditional fermented food in Taiwan. *J Sci Food Agric* 2010;**90**:1977–82. https://doi.org/10.1002/jsfa.4034

Chi H, Lu W, Liu G *et al.* Physiochemical property changes and mineral element migration behavior of bamboo shoots during traditional fermentation process. *J Food Process Preserv* 2020;**44**:e14784. https://doi.org/10.1111/jfpp.14784

Clemens R, Zaschke-Kriesche J, Khosa S *et al.* Insight into two ABC transporter families involved in lantibiotic resistance. *Front Mol Biosci* 2018;**4**:91. https://doi.org/10.3389/fmolb.2017.00091

Couvin D, Bernheim A, Toffano-Nioche C *et al.* CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res* 2018;**46**:W246–51. https://doi.org/10.1093/nar/gky425

Das AJ, Deka SC. MiniReview Fermented foods and beverages of the North-East India. *Int Food Res J* 2012;**19**:377.

Das R, Pradhan D, Kishore A *et al.* Identification of potential histone deacetylase 3 inhibitors from fermented bamboo shoots using GC-MS metabolomics, molecular docking, pharmacokinetics and molecular dynamics simulations approaches. *Food Biosci* 2025;**69**:106849. https://doi.org/10.1016/j.fbio.2025.106849

Das R, Tamang B, Najar IN *et al.* First report on metagenomics and their predictive functional analysis of fermented bamboo shoot food of Tripura, North East India. *Front Microbiol* 2023;**14**:1158411. https://doi.org/10.3389/fmicb.2023.1158411

Das R, Tamang B, Najar IN *et al.* Probiotic yeast characterization and fungal amplicon metagenomics analysis of fermented bamboo shoot products from Arunachal Pradesh, northeast India. *Heliyon* 2024;**10**:e39500. https://doi.org/10.1016/j.heliyon.2024.e39500

Das S, Tamang JP. Changes in microbial communities and their predictive functionalities during fermentation of toddy, an alcoholic beverage of India. *Microbiol Res* 2021;**248**:126769. https://doi.org/10.1016/j.micres.2021.126769

Deka P, Mehetre GT, Lalnunmawii E *et al.* Metagenomic analysis of bacterial diversity in traditional fermented foods reveals food-specific dominance of specific bacterial taxa. *Fermentation* 2021;**7**:167. https://doi.org/10.3390/fermentation7030167

Denton M, Kerr KG. *Microbiological and Clinical Aspects of Infection Associated with Stenotrophomonas Maltophilia. Clinical microbiology reviews* 1998; **11** :57–80.

Dong X, Strous M. An integrated pipeline for annotation and visualization of metagenomic contigs. *Front Genet* 2019;**10**:999. https://doi.org/10.3389/fgene.2019.00999

Drula E, Garron ML, Dogan S *et al.* The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res* 2022;**50**:D571–7. https://doi.org/10.1093/nar/gkab1045

Egido JE, Costa AR, Aparicio-Maldonado C *et al.* Mechanisms and clinical importance of bacteriophage resistance. *FEMS Microbiol Rev* 2022;**46**:fuab048. https://doi.org/10.1093/femsre/fuab048

Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;**20**:238. https://doi.org/10.1186/s13059-019-1832-y

Gautam N, Sharma N. A study on characterization of new bacteriocin produced from a novel strain of Lactobacillus spicheri G2 isolated from Gundruk—a fermented vegetable product of North East India: a novel bacteriocin production from Lactobacillus spicheri G2. *J Food Sci Technol* 2015;**52**:5808–16. https://doi.org/10.1007/s13197-015-1710-x

Goyal AK, Usha T, Kalita M *et al.* Tradition meets technology: an overview of fermented bamboo shoots. *Adv Bamboo Sci* 2024;**7**:100078. https://doi.org/10.1016/j.bamboo.2024.100078

Grant JR, Enns E, Marinier E *et al.* Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res* 2023;**51**:W484–92. https://doi.org/10.1093/nar/gkad326

Guo X, Xu D, Li F *et al.* Current approaches on the roles of lactic acid bacteria in crop silage. *Microb Biotechnol* 2023;**16**:16.67–87. https://doi.org/10.1111/1751-7915.14184

Gurevich A, Saveliev V, Vyahhi N *et al.* QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;**29**:1072–5. https://doi.org/10.1093/bioinformatics/btt086

Hu Y, Chen X, Zhou J *et al.* Metagenomic analysis of suansun, a traditional chinese unsalted fermented food. *Processes* 2021;**9**:1669. https://doi.org/10.3390/pr9091669

Huerta-Cepas J, Szklarczyk D, Heller D *et al.* EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 2019;**47**:D309–14. https://doi.org/10.1093/nar/gky1085

Hyatt D, Chen GL, Locascio PF *et al.* Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification. *Bioinformatics* 2010;**11**:119. http://www.biomedcentral.com/1471-2105/11/119

Ibrahim R, Aranjani JM. Bacterial defense mechanisms against bacteriophages: an evolutionary arms race. *Arch Microbiol* 2026;**208**:229. https://doi.org/10.1007/s00203-026-04785-x

Ismail YS, Yulvizar C, Mazhitov B. Characterization of lactic acid bacteria from local cows milk kefir. *IOP Conf Ser: Earth Environ Sci* 2018;**130**:012019. https://doi.org/10.1088/1755-1315/130/1/012019

Jain C, Rodriguez-R LM, Phillippy AM *et al.* High throughput ANI analysis of 90 K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;**9**:5114. https://doi.org/10.1038/s41467-018-07641-9

Jain PM, Rajagopal K. Role of traditional fermented foods containing probiotics in combating COVID-19. *J Antibiot Antimicrob Agents* 2024;**1**:1–19.

Joseph CE, Jain A, Yaqub MO *et al.* CRISPR-Cas systems: bridging bacterial immunity and host interactions. *Appl Microbiol (Switz)* 2025;**5**:118. https://doi.org/10.3390/applmicrobiol5040118

Jung JY, Lee SH, Kim JM *et al.* Metagenomic analysis of kimchi, a Traditional Korean fermented food. *Appl Environ Microb* 2011;**77**:2264–74. https://doi.org/10.1128/AEM.02157-10

Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 2000;**28**:27–30. http://www.genome.ad.jp/kegg/

Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 2016;**428**:726–31. https://doi.org/10.1016/j.jmb.2015.11.006

Kikon D. Bamboo shoot in our blood fermenting flavors and identities in Northeast India. *Curr Anthropol* 2021;**62**:S376–87. https://doi.org/10.1086/715830

Kleerebezem M, Boekhorst J, Kranenburg Rv *et al.* Complete genome sequence of Lactobacillus plantarum WCFS1. *Proc Natl Acad Sci USA* 2003;**100**:1990–5. https://doi.org/10.1073/pnas.0337704100

Kruse T, Levisson M, Vos WMd *et al.* vanI: a novel d-Ala-d-Lac vancomycin resistance gene cluster found in Desulfitobacterium hafniense. *Microb Biotechnol* 2014;**7**:456–66. https://doi.org/10.1111/1751-7915.12139

Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 2004;**32**:11–6. https://doi.org/10.1093/nar/gkh152

Latif A, Shehzad A, Niazi S *et al.* Probiotics: mechanism of action, health benefits and their application in food industries. *Front Microbiol* 2023;**14**. https://doi.org/10.3389/fmicb.2023.1216674

Lee I, Chalita M, Ha SM *et al.* ContEst16S: an algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *Int J Syst Evol Microbiol* 2017;**67**:2053–7. https://doi.org/10.1099/ijsem.0.001872

Leech J, Cabrera-Rubio R, Walsh AM *et al.* Fermented-food metagenomics reveals substrate-associated differences in taxonomy and health-associated and antibiotic resistance determinants. *MSystems* 2020;**5**:10–1128. https://doi.org/10.1128/msystems.00522-20

Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;**49**:W293–6. https://doi.org/10.1093/nar/gkab301

Li D, Liu CM, Luo R *et al.* MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 2015;**31**:1674–6. https://doi.org/10.1093/bioinformatics/btv033

Li P, Lin Q, Song D *et al.* Microbial succession drives flavor development in low-salt fermented bamboo shoots: a correlation analysis of physicochemical dynamics and metabolic pathways. *Food Chem: X* 2025;**32**:103283. https://doi.org/10.1016/j.fochx.2025.103283

Li Z, Xu X, Yang K *et al.* Multifaceted analyses reveal carbohydrate metabolism mainly affecting the quality of postharvest bamboo shoots. *Front Plant Sci* 2022;**13**:1021161. https://doi.org/10.3389/fpls.2022.1021161

Liu Y, Han R, Zhou L *et al.* Comparative performance of the GenoLab M and NovaSeq 6000 sequencing platforms for transcriptome and LncRNA analysis. *Bmc Genom [Electronic Resource]* 2021;**22**:829. https://doi.org/10.1186/s12864-021-08150-8

Malekijahan F, Razavi SH, Shafiepour M *et al.* Advances in microbial metabolism for flavor development: exploring the roles of bacteria, yeasts, and molds in food applications. *Microbe (Netherlands)* 2025;**9**:100623. https://doi.org/10.1016/j.microb.2025.100623

Martino ME, Bayjanov JR, Caffrey BE *et al.* Nomadic lifestyle of Lactobacillus plantarum revealed by comparative genomics of 54 strains isolated from different habitats. *Environ Microbiol* 2016;**18**:4974–89. https://doi.org/10.1111/1462-2920.13455

Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 2016;**7**:11257. https://doi.org/10.1038/ncomms11257

Morais MMB, Santos RM, Morais CD *et al.* Academic Editors: Alcina Metabolic Variations in Bamboo Shoot Boiled Liquid During Pediococcus Pentosaceus B49 Fermentation. 2025. https://doi.org/10.3390/foods

Naik S, Shashank PR, Rajgopal NN *et al.* Linnean shortfall still a major concern in India: as evident by the micromoth family Tortricidae (Lepidoptera). *Curr Sci* 2021;**120**:712–5. https://doi.org/10.18520/cs/v120/i4/712-715

Nirmala C, Bisht MS, Laishram M. Bioactive compounds in bamboo shoots: health benefits and prospects for developing functional foods. *Int J Food Sci Technol* 2014;**49**:1425–31. https://doi.org/10.1111/ijfs.12470

Oh SC, Lee SJ, Ding K *et al.* Kimchi fermentation-driven detoxification of flaxseed: impact on physicochemical quality and antioxidant potential. *Foods* 2026;**15**:632. https://doi.org/10.3390/foods15040632

Önning G, Palm R, Linninge C *et al.* New Lactiplantibacillus plantarum and *Lacticaseibacillus rhamnosus* strains: well tolerated and improve infant microbiota. *Pediatr Res* 2022;**91**:1849–57. https://doi.org/10.1038/s41390-021-01678-1

Paganin P, Alisi C, Dore E *et al.* Microbial diversity of bacteria involved in biomineralization processes in mine-impacted freshwaters. *Front Microbiol* 2021;**12**:778199. https://doi.org/10.3389/fmicb.2021.778199

Page AJ, Cummins CA, Hunt M *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;**31**:3691–3. https://doi.org/10.1093/bioinformatics/btv421

Park I, Mannaa M. Fermented foods as functional systems: microbial communities and metabolites influencing gut health and systemic outcomes. *Foods* 2025;**14**:2292. https://doi.org/10.3390/foods14132292

Pasolli E, Filippis FD, Mauriello IE *et al.* Large-scale genome-wide analysis links lactic acid bacteria from food with the gut microbiome. *Nat Commun* 2020;**11**:2610. https://doi.org/10.1038/s41467-020-16438-8

Pinilla-Redondo R, Russel J, Mayo-Muñoz D *et al.* CRISPR-Cas systems are widespread accessory elements across bacterial and archaeal plasmids. *Nucleic Acids Res* 2022;**50**:4315–28. https://doi.org/10.1093/nar/gkab859

Pokrzywa R, Polanski A. BWtrs: a tool for searching for tandem repeats in DNA sequences based on the Burrows–Wheeler transform. *Genomics* 2010;**96**:316–21. https://doi.org/10.1016/j.ygeno.2010.08.001

Price MN, Dehal PS, Arkin AP. Fasttree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;**26**:1641–50. https://doi.org/10.1093/molbev/msp077

Qi Y, Huang L, Zeng Y *et al.* Pediococcus pentosaceus: screening and Application as Probiotics in Food Processing. *Front Microbiol* 2021;**12**:762467. https://doi.org/10.3389/fmicb.2021.762467

Romi W, Ahmed G, Jeyaram K. Three-phase succession of autochthonous lactic acid bacteria to reach a stable ecosystem within 7 days of natural bamboo shoot fermentation as revealed by different molecular approaches. *Mol Ecol* 2015;**24**:3372–89. https://doi.org/10.1111/mec.13237

Roux S, Enault F, Hurwitz BL *et al.* VirSorter: mining viral signal from microbial genomic data. *Peer J* 2015;**3**:e985. https://doi.org/10.7717/peerj.985

Salter SJ, Cox MJ, Turek EM *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;**12**:87. https://doi.org/10.1186/s12915-014-0087-z

Satya S, Bal LM, Singhal P *et al.* Bamboo shoot processing: food quality and safety aspect (a review). *Trends Food Sci Technol* 2010;**21**:181–9. https://doi.org/10.1016/j.tifs.2009.11.002

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;**30**:2068–9. https://doi.org/10.1093/bioinformatics/btu153

Seemann T. *ABRicate: Mass Screening of Contigs for Antiobiotic Resistance Genes*. 2016. https://github.com/tseemann/abricate (16 December 2025, date last accessed).

Shree P, Singh CK, Sodhi KK *et al.* Biofilms: understanding the structure and contribution towards bacterial resistance in antibiotics. *Med Microecol* 2023;**16**:100084. https://doi.org/10.1016/j.medmic.2023.100084

Singhal P, Satya S, Naik SN. Fermented bamboo shoots: a complete nutritional, anti-nutritional and antioxidant profile of the sustainable and functional food to food security. *Food Chem: Mol Sci* 2021;**3**:100041. https://doi.org/10.1016/j.fochms.2021.100041

Starikova EV, Tikhonova PO, Prianichnikov NA *et al.* Phigaro: high-throughput prophage sequence annotation. *Bioinformatics* 2020;**36**:3882–4. https://doi.org/10.1093/bioinformatics/btaa250

Sundriyal RC, Upreti TC, Varuni R. Bamboo and cane resource utilisation and conservation in the Apatani plateau, Arunachal Pradesh, India: implications for management. *J Bamboo Ratt* 2002;**1**:205–46. https://doi.org/10.1163/156915902760184277

Swain MR, Anandharaj M, Ray RC *et al.* Fermented fruits and vegetables of asia: a potential source of probiotics. *Biotechnol Res Int* 2014;**2014**:1–19. https://doi.org/10.1155/2014/250424

Talib N, Mohamad NE, Yeap SK *et al.* Isolation and characterization of Lactobacillus spp. From kefir samples in Malaysia. *Molecules* 2019;**24**:2606. https://doi.org/10.3390/molecules24142606

Tamang B, Tamang J, Tamang B *et al.* Lactic acid bacteria isolated from indigenous fermented bamboo products of arunachal pradesh in india and their functionality lactic acid bacteria in ethnic fermented bamboo products. *Food Biotechnol* 2009;**23**:133–47. https://doi.org/10.1080/08905430902875945

Tamang B, Tamang JP, Schillinger U *et al.* Phenotypic and genotypic identification of lactic acid bacteria isolated from ethnic fermented bamboo tender shoots of North East India. *Int J Food Microbiol* 2008;**121**:35–40. https://doi.org/10.1016/j.ijfoodmicro.2007.10.009

Tamang JP, Kharnaior P, Halami PM. Lactic acid bacteria in some Indian fermented foods and their predictive functional profiles. *Braz J Microbiol* 2024;**55**:1745–51. https://doi.org/10.1007/s42770-024-01251-y

Tenea GN, Hidalgo J, Pircalabioru GG *et al.* Genomic and functional insights into Lactiplantibacillus plantarum UTNGt3 from Ecuadorian Amazon Chrysophyllum oliviforme: a safe and promising probiotic. *Front Microbiol* 2025;**16**:1634475. https://doi.org/10.3389/fmicb.2025.1634475

Tesson F, Planel R, Egorov AA *et al.* A Comprehensive resource for exploring antiphage defense: defensefinder webservice, wiki and databases. bioRxiv, 26 Jan 2024. Preprint. https://doi.org/10.1101/2024.01.25.577194

Tiwari S, Nanda M. Bacteremia caused by comamonas testosteroni an unusual pathogen. *J Lab Phys* 2019;**11**:087–90. https://doi.org/10.4103/jlp.jlp_116_18

Turpin W, Humblot C, Guyot JP. Genetic screening of functional properties of lactic acid bacteria in a fermented pearl millet slurry and in the metagenome of fermented starchy foods. *Appl Environ Microb* 2011;**77**:8722–34. https://doi.org/10.1128/AEM.05988-11

Typas A, Banzhaf M, Berg Van Saparoea BVD *et al.* Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell* 2010;**143**:1097–109. https://doi.org/10.1016/j.cell.2010.11.038

Veeranan Arun Giridhari V, Uma Maheswari T, Vanniarajan C *et al.* Probing the metagenome and nutritional composition of idli batter fortified with electrolytic iron for addressing anaemia. *J Food Sci Technol* 2025;**62**:1481–90. https://doi.org/10.1007/s13197-024-06119-5

Vernikos GS, Parkhill J. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the Salmonella pathogenicity islands. *Bioinformatics*

2006;**22**:2196–203. https://doi.org/10.1093/bioinformatics/btl369

Wang X, Yu P, Wang T *et al.* Novel insights into p-cresol production in fermented bamboo shoots: core microbiota and metabolic signatures. *LWT* 2025;**239**:118940. https://doi.org/10.1016/j.lwt.2025.118940

Wishart DS, Han S, Saha S *et al.* PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res* 2023;**51**:W443–50. https://doi.org/10.1093/nar/gkad382

Wu JW, Qin R, Lu YY *et al.* Dynamic changes in microbiota during fermentation of bamboo shoots under varying conditions in Guangxi, China. *Front Microbiol* 2025;**16**. https://doi.org/10.3389/fmicb.2025.1529935

Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 2016;**32**:605–7. https://doi.org/10.1093/bioinformatics/btv638

Xu X, Gu P. Overview of phage defense systems in bacteria and their applications. *Int J Mol Sci* 2024;**25**:13316. https://doi.org/10.3390/ijms252413316

Yang J, Sung J, Kim Y *et al.* Inhibitory effects of butein on adipogenesis through upregulation of the Nrf2/HO-1 Pathway in 3T3-L1 Adipocytes. *Prev Nutr Food Sci* 2017;**22**:306–11. https://doi.org/10.3746/pnf.2017.22.4.306

Yoon SH, Ha SM, Kwon S *et al.* Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;**67**:1613–7. https://doi.org/10.1099/ijsem.0.001755

Zhang H, Yohe T, Huang L *et al.* DbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2018;**46**:W95–101. https://doi.org/10.1093/nar/gky418

Zheng J, Ge Q, Yan Y *et al.* DbCAN3: automated carbohydrate-Active enzyme and substrate annotation. *Nucleic Acids Res* 2023;**51**:W115–21. https://doi.org/10.1093/nar/gkad328

Zheng J, Wittouck S, Salvetti E *et al.* A taxonomic note on the genus Lactobacillus: description of 23 novel genera, emended description of the genus Lactobacillus beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *Int J Syst Evol Microbiol* 2020;**70**:2782–858. https://doi.org/10.1099/ijsem.0.004107

Received: 8 January 2026. **Revised:** 25 May 2026. **Accepted:** 25 May 2026