J Appl Biotechnol Rep. 2025 December;12(4):1840-1851

Journal of
Applied Biotechnology Reports

doi 10.30491/jabr.2025.494841.1821

Original Article

Probiotic Niche Specificity and Microbiome Analysis of Different Non-Dairy based Traditional Food Samples Using Metagenomic Approach

Bijayanta Sircar [1], **Manisha Mandal** [2], **Shyamapada Mandal** [1*]

[1] Laboratory of Microbiology and Experimental Medicine, Department of Zoology, University of Gour Banga, Malda-732103, West Bengal, India
[2] Department of Physiology, MGM Medical College and LSK Hospital, Kishanganj, Bihar-855107, West Bengal, India

Corresponding Author: Shyamapada Mandal, PhD, Professor, Laboratory of Microbiology and Experimental Medicine, Department of Zoology, University of Gour Banga, Malda-732103, West Bengal, India. Tel: +9831279239, E-mail: sam.micro11@ugb.ac.in

Received December 19, 2024; **Accepted** May 19, 2025; **Online Published** December 30, 2025

Abstract

Introduction: People consume different types of traditional dairy and non-dairy-based fermented foods worldwide. The literature review suggests that there are scientifically unexplored diverse non-dairy-based fermented foods. The present study was conducted to determine the taxonomic and metabolic status of beneficial bacteria in various non-dairy-based traditional food samples using a metagenomics approach.

Materials and Methods: Metagenomic analysis of food samples may reveal the presence of many beneficial culturable or non-culturable bacteria through Next Generation Sequencing techniques. In this study, we used six non-dairy-based food products, including two types of idli batter, vegetable fermented mixture, two types of vegetable pickles, and the Indian traditional food Shinni. Phylogenetic relationships, taxonomic plot analysis, presence of enriched genera, core microbiome comparative analysis, species richness in terms of α and β diversity, and functional profiling of the microbiome were also determined in this research work.

Results: All the samples showed a good abundance of probiotic bacterial presence. *Specifically*, *Pediococcus*, *Weissella*, *Lactococcus*, *Acetobacter*, and *Lactobacillus* were statistically more viable genera than others. Functional profiling of the six metagenome samples displayed the top nine KEGG Orthology metabolism and the top fourteen clusters of orthologous genes metabolism.

Conclusions: Variation in the richness and diversity of bacteria in all six samples, along with variations in metabolic functions, indicates the need for further investigation to explore the nutritional value and bioactivity of beneficial bacteria in our region.

Keywords: Non-Dairy Fermented Foods, Next Generation Sequencing, Metagenomic Analysis, Core Microbiome, Bacterial Metabolic Pathway, PROBIOTICS

Citation: Sircar B, Mandal M, Mandal S. Probiotic Niche Specificity and Microbiome Analysis of Different Non-Dairy based Traditional Food Samples Using Metagenomic Approach. J Appl Biotechnol Rep. 2025;12(4):1840-1851. doi:10.30491/jabr.2025.494841.1821

Introduction

The relationship between the human host and an enormous number of microbial symbionts is clearly understandable due to the establishment of the Human Microbiome Project (HMP), which served as a bridge between the real physiological setup and the Human Genome Project (HGP).[1] The microbial ecological parameters of the human gastrointestinal (GI) tract in terms of microbial niche specificity have significant functional significance and have been associated with various diseases such as digestive ailments, metabolic syndrome, diabetes, obesity, psoriatic arthritis, celiac malabsorption, psychiatric illnesses, and more.[2] The survival ability and retention capacity of microbes, specifically commensals or probiotics, in the GI tract in relation to virulence and colonization factors ultimately depend on niche factors.[3] Naturally, foods may contain microbes, whether purposefully added or prepared exclusively as a dietary addition. The definition of probiotics according to the International Scientific Association for Probiotics and Prebiotics (ISAPP) includes products made up of an appropriate viable count of distinct strains of live microorganisms that provide benefits for the well-being of consumers. Nowadays, some well-known probiotic strains include *Lactococcus*, *Lactobacillus*, *Pediococcus*, *Bifidobacterium*, and *Streptococcus* species, either as a combination or individual isolates.[4,5] Researchers are using culture-independent methods, such as next-generation sequencing (NGS), to define the contents of various types of probiotic foods and determine the types and relative amounts of beneficial bacteria present in them.[6,7]

Currently, researchers are interested in non-dairy products, including vegetable pickles and traditional fermented foods (such as Kimchi: Korean traditional fermented vegetables, Sha'a: a maize-based fermented beverage, Ganjang: traditional Korean fermented soy sauce), to explore the diversity of microbes

and associated health benefits.[7-10] Isolation, identification, and characterization of probiotic lactic acid bacteria from cereal and legume-based foods like idli batter and tempeh are also a principal interest of investigators worldwide.[11,12]

The 16S gene, responsible for the small ribosomal subunit (16S rRNA), is the most widely used phylogenetic marker for bacterial sequences of various existing species. Through NGS sequencing targeting the V3-V4 region, researchers can analyze microbiomes up to the genus level.[13] However, in this region, there is no available data on microbiome studies using NGS metagenomic approaches in non-dairy-based probiotics. In the current study, we used a culture-independent technique involving NGS to determine the contents of various types of traditional and homemade probiotic foods, including two types of fermented idli batter, vegetable pickle (prepared with oil and without oil), fermented vegetable mixture, and shinni, followed by unique traditional preparations. All the selected samples are Indian traditional foods with unique compositions, and no such explorations have been made previously. This approach will provide significant outcomes to determine the identity and relative presence of common beneficial bacteria, as well as the occurrence of bacteria not reported earlier.

Materials and Methods

Raw Sample Preparation

Six samples were prepared, including two types of fermented idli batter, two types of vegetable pickle, a vegetable fermented mixture, and shinni, using unique traditional methods while maintaining laboratory conditions as shown in Figure 1A. Each sample had different ingredients to create specific bacterial microbiome ecosystems for niche-specific analysis. The ingredients and amounts of each sample used in this study are listed in Table 1.

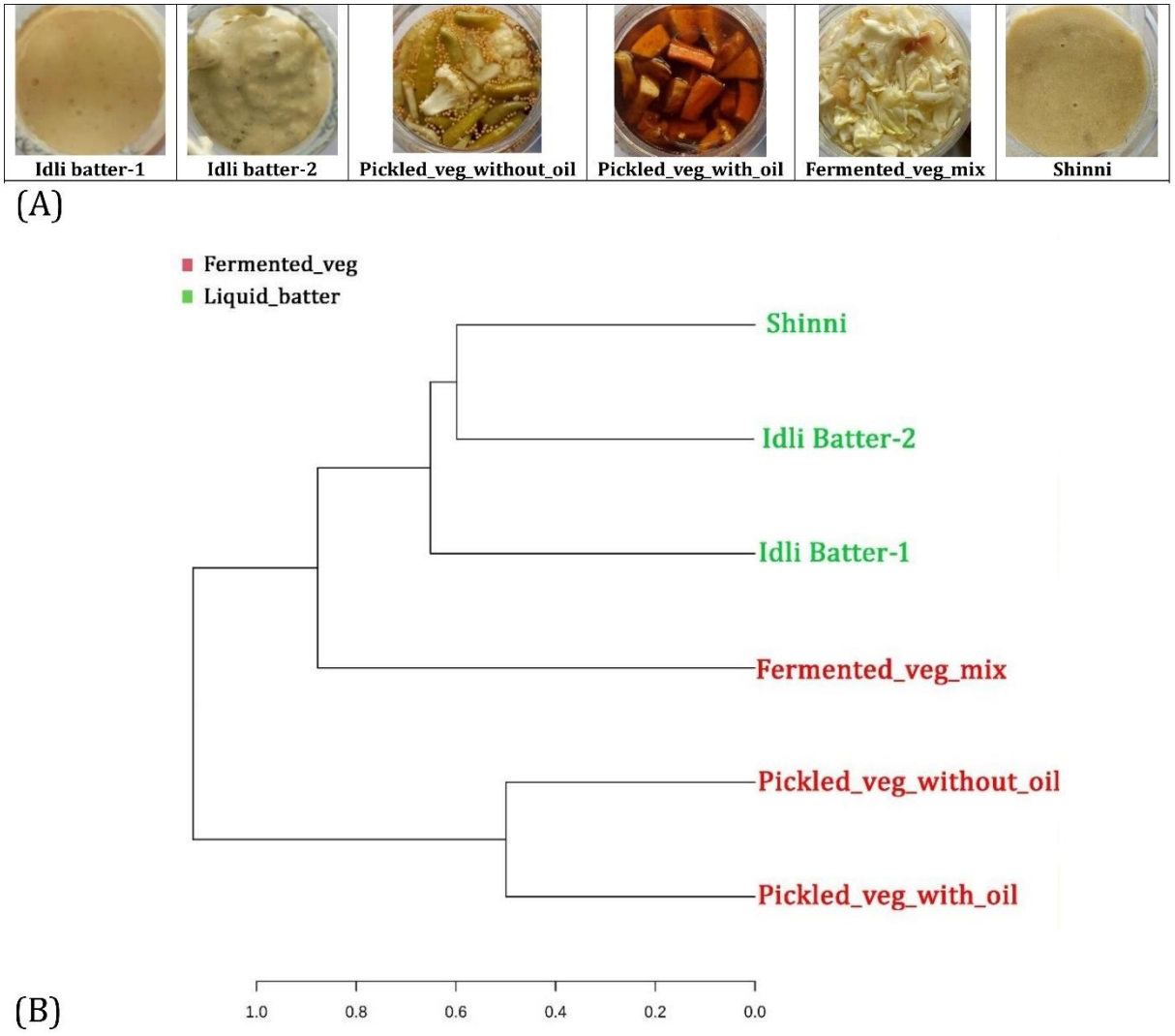

Figure 1. A) Prepared Six Non-Dairy based Traditional Fermented Food Samples; B) Phylogenetic Relationship among Selected Microbiomes.

Table 1. All the Samples Utilized in this Study Enlisted with Types, Ingredients and Amounts

Sample Type	Pickled / Non-Pickled	Sample Name	Ingredients	Scientific/Extended Name	Amount
Liquid _batter	Non-Pickled	Idli_batter-1	Rice	*Oryza sativa*	300 gm
			Mung dal	*Vigna radiata* L.	100 gm
			Methi dana	Trigonella foenum-graecum	5 gm
			Distilled water	-	100 ml
		Idli_batter-2	Rice	*Oryza sativa*	300 gm
			Urad dal	Vigna mungo L.	100 gm
			Meethi dana	*Trigonella foenum-graecum*	5 gm
			Distilled water	-	100 ml
		Shinni	Wheat flour	*Triticum aestivum*	100 gm
			Banana	Musa acuminata	2 nos.
			Dusted Jaggery	*Saccharum officinarum* L. product	5%
			Distilled water	-	250 ml
			Sodium chloride	NaCl	6.50%
Fermented_veg		Fermented_veg_mix	Cabbage	*Brassica oleracea*, var. capitata	200 gm
			Radish	*Raphanus sativus* var. Longipinnatus	200 gm
			Ginger	*Zingiber officinale*	25 gm
			Lemon	*Citrus limon* L.	2 nos.
			Distilled water	-	650 ml
			Sodium chloride	NaCl	7.50%
	Pickled	Pickled_veg_without_oil	Cauliflower	*Brassica oleracea*	250 gm
			Carrot roundels	*Daucus carota* L.	250 gm
			Cucumber diced	Cucumis sativus L.	250 gm
			Cloves	*Syzygium aromaticum*	4 nos.
			Garlic peeled	*Allium sativum*	6 nos.
			Ginger sliced	*Zingiber officinale*	6 tbsps.
			Fresh green Chilies	*Capsicum frutescens*	6 nos.
			Mustard seeds	*Brassica nigra*	3 tsp.
			Indian Bay leaf	*Cinnamomum Tamala*	1 no.
			Distilled water	-	100 ml
			Sodium chloride	NaCl	3tsp.
			Vinegar	Acetic Acid	100 ml
			Sugar	Soluble carbohydrates	25 gm
		Pickled_veg_with_oil	Cauliflower	*Brassica oleracea*, var. botrytis	250 gm
			Carrots	Daucus carota L.	250 gm
			Turnips/shalgam	*Brassica rapa* subsp. rapa	250 gm
			Cumin seeds	Cuminum cyminum	10 gm
			Black cardamoms	*Amomum subulatom* Roxburgh	4 nos.
			One inch Cinnamon	Cinnamomum verum	2 nos.
			Cloves	*Syzygium aromaticum*	8 nos.
			Mustard seed oil	*Brassica nigra*	350 ml
			Coarsely grated Ginger	*Zingiber officinale*	6 tbsps.
			Coarsely grated Garlic	*Allium sativum*	4 tbsps.
			Red chilli powder	*Capsicum annuum*	1.5 tbsps.
			Mustard seed powder	*Brassica nigra*	1.5 tbsps.
			Sodium chloride	NaCl	1.5 tbsps.
			Vinegar	Acetic Acid	10 ml
			Grated Jaggery	*Saccharum officinarum* L. products	150 gm
			Ginger-garlic paste	*Zingiber officinale* and *Allium sativum* paste	1 tbsp.

tbsps.: Tablespoons; tps.: Teaspoons.

DNA Extraction and DNA Quality Control for NGS

DNA extraction was performed using a suitable method for all the samples with commercially available DNA extraction kits (QIAGEN) according to the manufacturer's recommendations. The extraction was outsourced to Biokart India Pvt. Ltd., Bangalore, India. Extracted DNA from the samples was assessed using NanoDrop and gel electrophoresis before proceeding to PCR amplification. The NanoDrop readings of the optical density ratio (at 260 and 280 nm wavelengths) in the range of approximately 1.8 to 2 were used to evaluate the DNA quality.

PCR Amplification of V3-V4 Region of 16s Gene

The composition of the TAQ Master MIX includes High-Fidelity DNA Polymerase, 0.5 mM dNTPs, 3.2 mM $MgCl_2$, PCR enzyme buffer with forward (V13F) and reverse primer (V13R) (sequences 5'-AGAGTTTGATGMTGGCTCAG-3' and 5'-TTACCGCGGCMGCSGGCAC-3', respectively). For each sample, 40 ng of extracted DNA was used for amplification along with 10 pM of each primer for up to 25 cycles, maintaining conditions including a denaturation temperature of 95 °C for 15 sec, an annealing temperature of 60 °C for 15 sec, an elongation temperature of 72 °C for 2mins, a final extension at 72 °C for 10mins, and holding at 4 °C.

Library Preparation and Sequencing Data Generation

The amplicons from each sample were decontaminated with

Ampure beads to remove unexploited primers, and an additional 8 cycles of PCR were performed using Illumina barcoded adapters to prepare the sequencing libraries. Library purification was done using Ampure beads and quantitated through the Qubit dsDNA high sensitivity assay kit. Sequencing was executed using Illumina MiSeq administered with a 2x300PE v3 sequencing kit.

All the generated sequencing data (fastq files) submitted to NCBI-SRA (National Centre for Biotechnology Information - Sequence Read Archive as BioProject (https://www.ncbi.nlm.nih.gov/sra/PRJNA1236850) with specific accession numbers: Idli_batter-1 (https://www.ncbi.nlm.nih.gov/biosample/47406561), Idli_batter-2 (https://www.ncbi.nlm.nih.gov/biosample/47406562), Shinni (https://www.ncbi.nlm.nih.gov/biosample/47406563), Fermented_veg_mix (https://www.ncbi.nlm.nih.gov/biosample/47406564), Pickled_veg_without_oil (https://www.ncbi.nlm.nih.gov/biosample/47406565) and Pickled_veg_with_oil (https://www.ncbi.nlm.nih.gov/biosample/47406566).

Data Filtering, Normalization, Bioinformatic and Statistical Analysis

To improve downstream statistical analysis, data filtering has been performed to remove low-quality or uninformative features.[14] The metagenomic sequences underwent data filtering to eliminate features with a prevalence below 20% and a threshold count of 4, as well as the removal of features with low variance (10% or n = 3) based on standard deviation. After filtering, a total of 136 features were considered for further statistical and bioinformatic analysis. Normalization was also conducted to address inconsistencies in sampling depth and data sparsity, enabling more biologically meaningful comparisons.[14] The total sequence read counts for six samples were 114,336, with an average of 19,056 and a range of 2,294 to 30,150. The data were rarefied to a minimum library size of 2,294 and scaled using Total Sum Scaling (TSS) to standardize all samples for comparison.

For bioinformatic analysis, a well-constructed pipeline based on the non-commercial package QIIME 2 (Quantitative Insights into Microbial Ecology 2, https://qiime2.org) in the Miniconda3 environment was followed.[15] The overall steps included trimming and excluding low-quality reads, assembling short reads into overlapping contigs (consensus sequences), aligning sample sequences, identifying phylogenetic relationships and/or functions, visualizing data, and conducting statistical analysis.[16] The samples were clustered using the Ward clustering algorithm based on the Euclidean distance measure.[17]

The MicrobiomeAnalyst (https://www.microbiomeanalyst.ca) was used for statistical investigation and visualization of data.[18] Differential abundance analysis with analysis of covariance (ANCOVA) was determined using different statistical parameters such as log2-fold changes (log2FC), false discovery rates (FDRs), standard error and p-value where statistical significance was estimated at $p < 0.05$.[19-22] To analyze α and β-diversity in terms of community structure, the Chao1 index at the OTU level and analysis of group similarities (ANOSIM) were used in this study. The analysis of core microbiome, clustering analysis with heatmap (using the Euclidean distance method) for taxon level abundance comparison, dendrogram (following the Bray-Curtis distance method), and random forest model (as a structural and functional classifier) were applied following methodologies stated elsewhere.[15] The functional diversity profiling using metagenomics data of all the test samples /microbiome was predicted based on the KEGG (using categories of pathways and modules) or COG (Cluster of Orthologous Groups) annotation systems.[18] The whole analysis was performed on Ubuntu 20.04.2 LTS 64-bit OS, 3.36.8 Gnome version, NVIDIA® GeForce RTX™ 4060, 16 GB memory, Intel® Core™ i7 processor.

Results

Taxonomic Profiling and the Bacterial Prevalence

The dendrogram plot analysis revealed the phylogenetic differences/similarities between the microbiomes of the test samples, as represented in Figure 1B. Comparatively, the Liquid_batter food type exhibited more species richness than Fermented_veg (Figure 2A). Shinni had the maximum species-rich niche, whereas Fermented_veg_mix and Pickled_veg_with_oil occupied approximately the same position on the graph. The taxonomic plot analysis clearly shows the occurrence and bacterial prevalence from phylum to genus level. Phylum-level taxonomic plot analysis (Figure 2B) revealed that most of the bacteria in the test samples belong to the phyla Proteobacteria and Firmicutes, although some specific bacterial abundance of the phylum Bacteroidetes was found in Fermented_veg_mix and in pickled_veg (without_oil and with_oil). Bacteria belonging to the class Bacilli are abundantly present in both fermented idli batter and in shinni. Bacteria belonging to Gammaproteobacteria are dominantly present in Fermented_veg_mix, as well as in Idli_batter-2, and pickled_veg (without_oil and with_oil) showed class Alphaproteobacteria dominance (Figure 2C).

According to the order-level taxonomic plot analysis (Figure 2D), bacteria belonging to the order Lactobacillales showed a higher prevalence in both fermented idli batter samples and shinni. On the other hand, bacteria belonging to the order *Enterobacterales* were dominantly present in Fermented_veg_mix and also showed a mixed occurrence of different orders of bacteria, including *Lactobacillales*, *Pseudomonadales*, *Xanthomonadales*, and Caulobacterales. Both pickled_veg (without_oil and with_oil) microbiomes exhibited a dominant presence of bacteria of the order *Lactobacillales* as well as *Rhodospirillales*. All test samples of microbiome analysis showed different culturable and non-

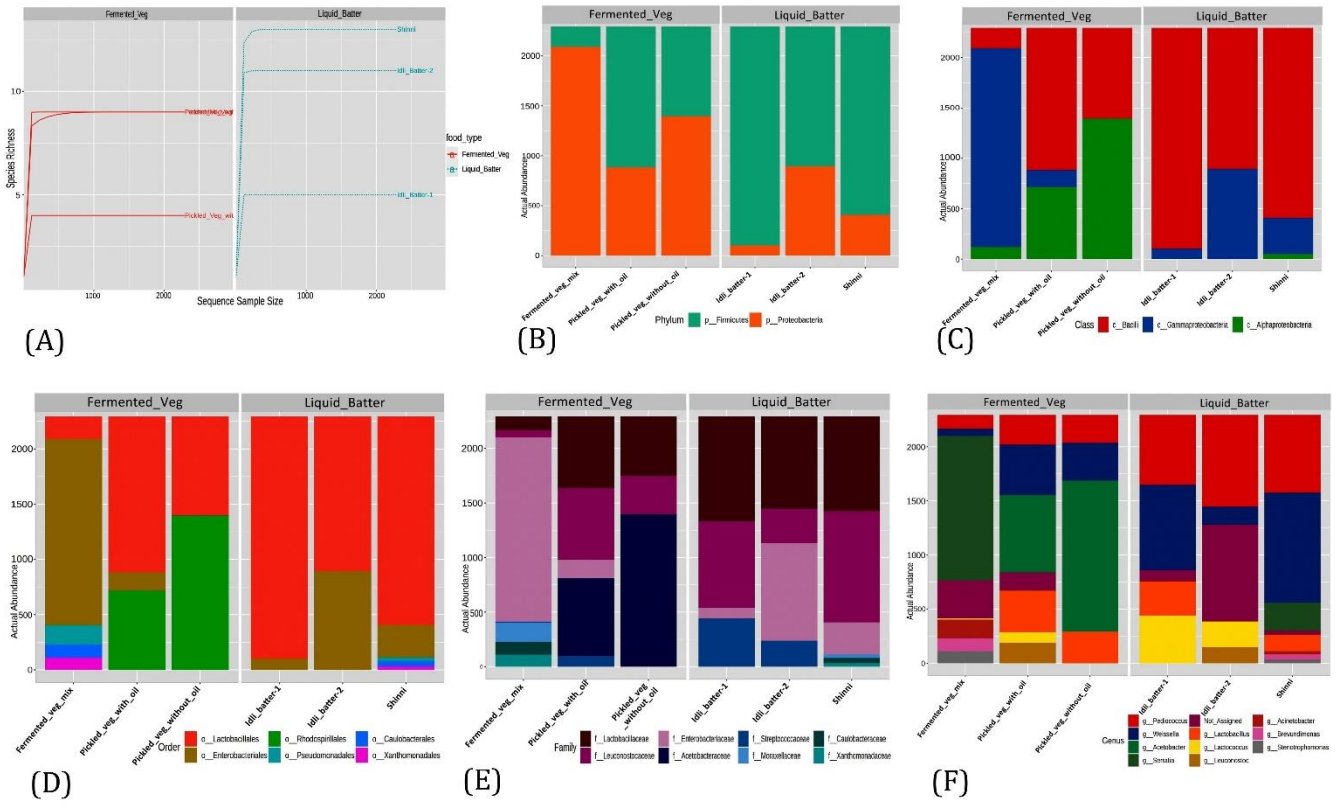

Figure 2. A) Species richness between Fermented_veg and Liquid_batter food types. Taxonomic plot analysis B) Phylum level; C) Class level; D) Order level; E) Family level; F) Genus level.

culturable bacterial presence belonging to several families like *Acetobacteraceae*, *Actinomycetaceae*, *Lactobacillaceae*, *Leuconostocaceae*, and *Streptococcaceae* (Figure 2E).

The genus-level taxonomic analysis (Figure 2F) displayed that bacterial genera *Lactobacillus*, *Lactococcus*, *Weissella*, and *Pediococcus* were highly dominant in both fermented idli batter. *Acetobacter*, *Lactobacillus*, *Weissella*, and *Pediococcus* were dominant in both Pickled_veg (without_oil and with_oil), while Serratia, *Pediococcus*, and *Acetobacter* were dominant in Fermented_veg_mix. The genera *Weissella*, *Lactobacillus*, *Pediococcus*, and *Serratia* were abundant in Shinni. Among the top enriched genera, two lactic acid bacterial genera, *Pediococcus* and *Weissella*, were found in all test microbiome structures. Some well-known probiotic genera like Lactobacillus, *Lactococcus*, and Enterococcus were also found among the top enriched genera for all the test samples except Fermented_veg_mix. The species-level taxonomic analysis revealed the presence of bacteria such as *Serratia marcescens*, *Acinetobacter rhizosphaerae*, *Brevundimonas diminuta*, and *Stenotrophomonas geniculata* in the Fermented_veg_mix as well as Shinni.

Core Microbiome Comparative Analysis

The α-diversity measured by Chao1 at the OTU (Operational Taxonomic Unit) level showed a diverse difference in species richness between Fermented_veg and Liquid_batter (food types), with the sample Shinni showing the maximum level of α-diversity among all test samples (Figure 3A). The β-diversity of the microbiome profile at the genus level depicts dissimilarities between microbial communities in pickled and non-pickled sample clusters (Figure 3B). The Venn diagram in Figure 3 represents how similar or dissimilar components and taxa are present in the studied microbiome (C and D, respectively). A total of 3, 6, 9, 5, and 3 unique taxa were found in Idli_batter-1, Idli_batter-2, Pickled_veg_with_oil, Fermented_veg_mix, and Shinni, respectively.

The core microbiome refers to the set of taxa that are detected in a high fraction of the population above a given abundance threshold. In order to perform core microbiome analysis, the count data has been transformed to compositional (relative) abundance. Core microbiome analysis displayed that approximately ten genera were shared in all the test samples at the lowest detection threshold of 0.01% (Figure 4A). The shared core microbiome (i.e., detected throughout all of the test food samples) included *Pediococcus*, *Weissella*, *Lactobacillus*, *Lactococcus*, *Acetobacter*, *Serratia*, *Leuconostoc*, *Acinetobacter*, *Brevundimonas*, and *Stenotrophomonas* bacterial genera at the detection threshold from 0.01 to 0.1%. Statistical analysis of covariance reflects significant positive covariance evidence (Figure 4B), where the *p*-value ranged between

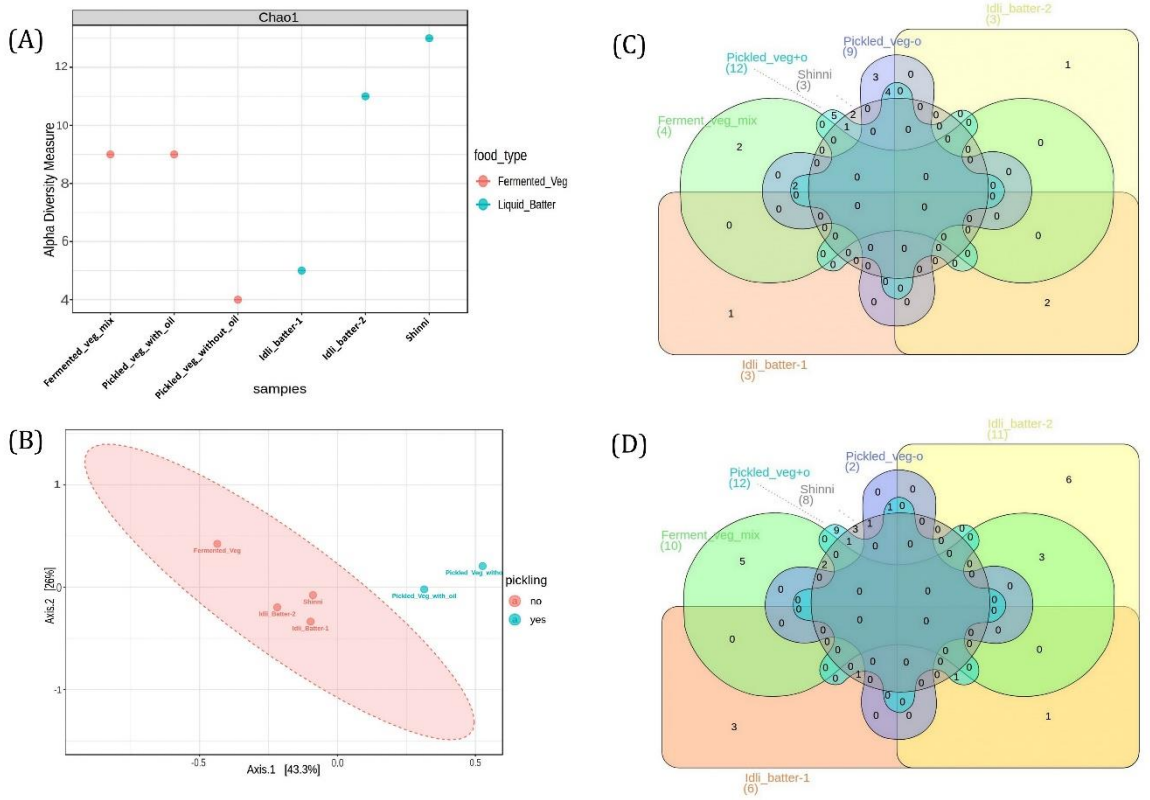

Figure 3. A) α-diversity and B) β-diversity based on core microbiome comparative analysis of six samples; (C-D) Venn diagram showing core microbiome similarities/dissimilarities among test samples; C) Composition; D) Taxa.

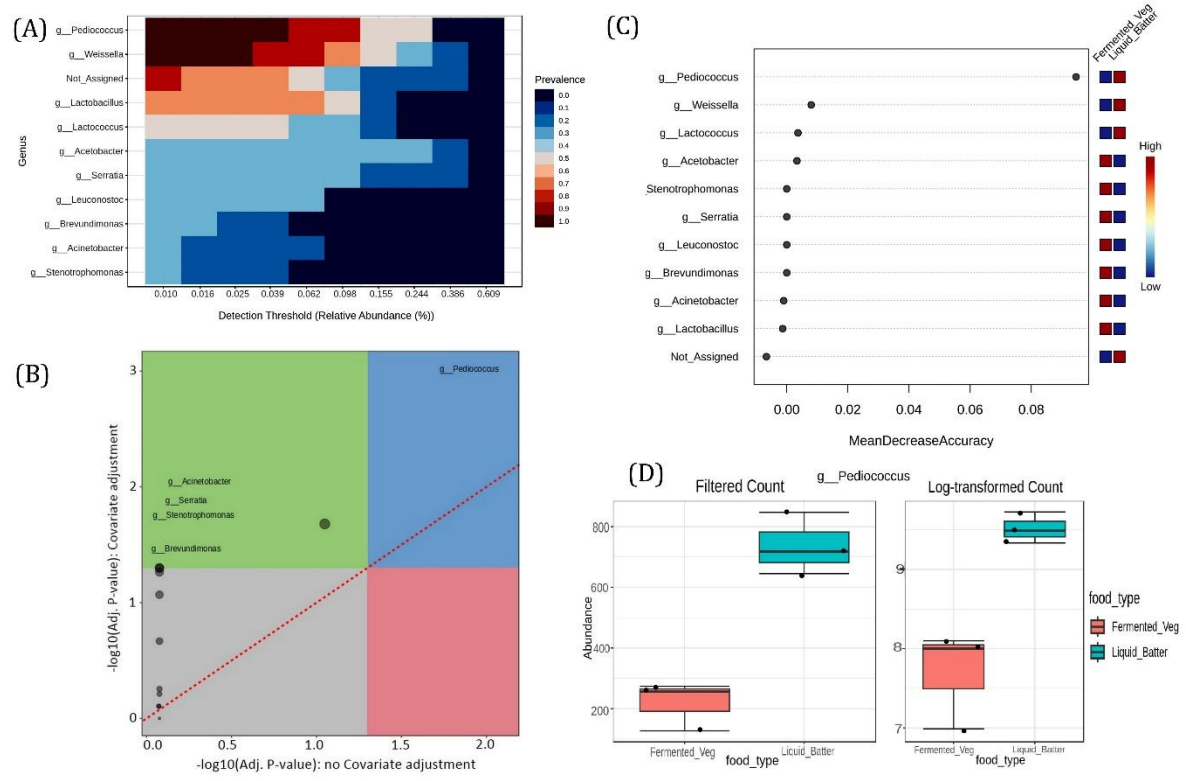

Figure 4. A) Core Microbiome comparative analysis using relative abundance (%) correlated with the prevalence pattern of six test samples; B) Statistical significance as per *p*-value analysis; C) Mean Decrease accuracy pattern; D) Genus Pediococcus abundance among Fermented_veg and Liquid_batter food types (represented in filtered and log transformed count).

0.000952 and 0.0388 for genera *Pediococcus*, *Acinetobacter*, *Stenotrophomonas*, *Brevundimonas* and *Serratia*. The mean decrease accuracy plot (Figure 4C) was created to analyze how much accuracy the model loses by the removal of each component. The *Pediococcus* genus showed the highest value as mean decrease accuracy, hence it has more importance in determining variables for the structural arrangement. The single-factor statistical comparison between test samples has been studied for the genus *Pediococcus*, where the Liquid_batter food type showed more diversified bacterial abundance than Fermented_veg, as represented through the box plot in Figure 4D.

Hierarchical clustering of taxa with a heat tree (from phylum to genus level) and a heatmap (only at the genus level) was constructed based on the abundance pattern of taxonomic levels with the detailed view mode of <1500 features (Figure 5A and C, respectively). In Figure 5B, the phylogenetic tree demonstrates the evolutionary relationships among bacterial phyla, highlighting various specific subclades, including *Pediococcus*, *Weissella*, *Lactobacillus*, *Lactococcus*, *Acetobacter*, and *Serratia* with high abundance.

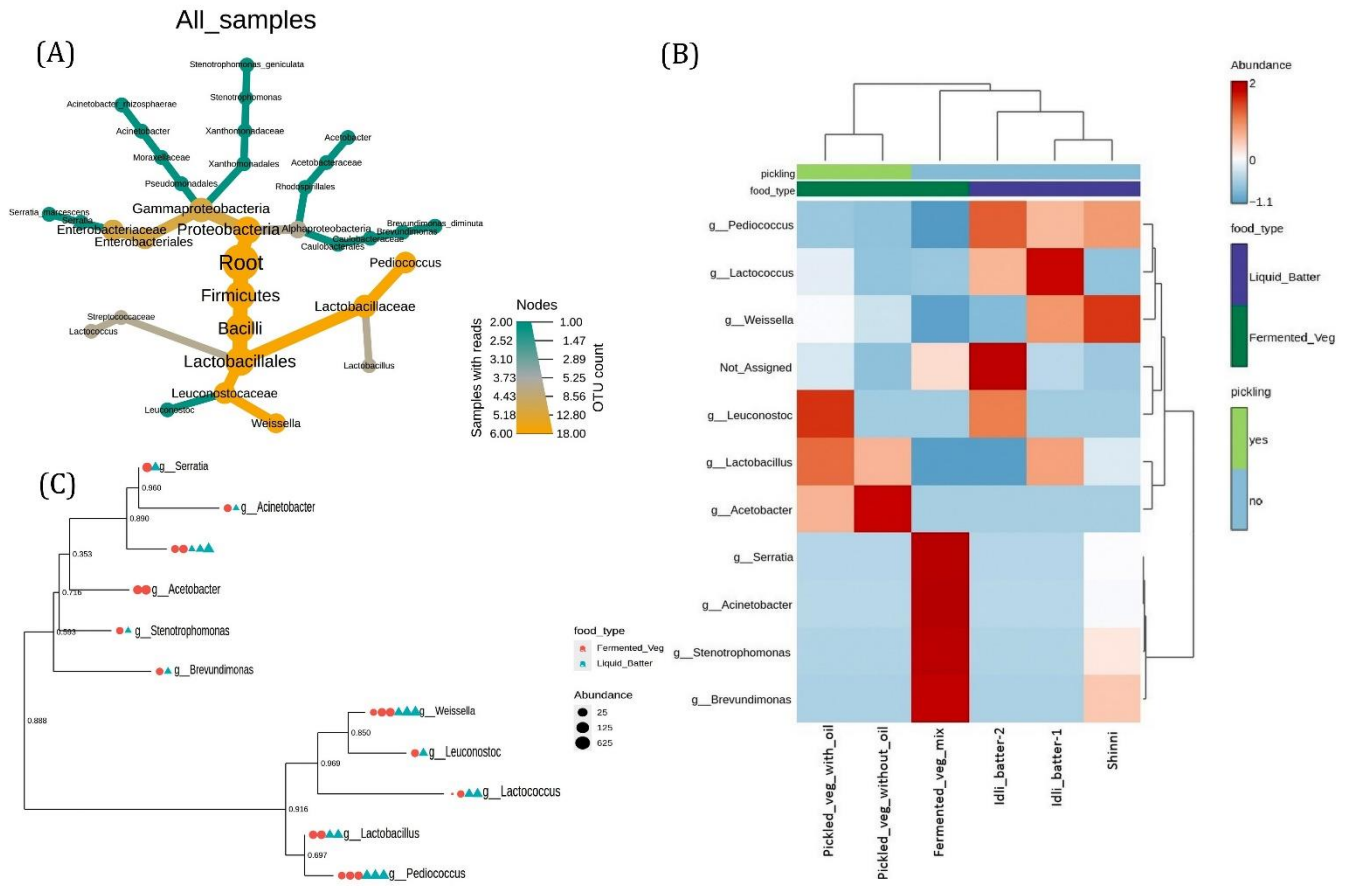

Figure 5. A) Heat tree showing the abundance of organisms in all the microbiome communities where the abundance pattern correlated with the size and color of nodes and edges based on OTU and read counts; B) Phylogenetic tree representing the evolutionary relationships of bacterial phyla, showing different distinct subclades and their close relatedness based on recent common ancestral sharing. The colored points and shapes specify the food types, and the size of black circles represents the abundance intensity of respective phyla; C) Heat map showing a comparative analysis of genus-level taxonomic units.

Functional Profile Analysis of Microbial Community

Different food microbiomes have their own unique metabolic activities due to specific microenvironments/ niches that allow specific types of microbial genera to be the principal components. Results of functional profiling through KEGG Orthology (KO) gene-mediated metagenomic analysis revealed various metabolic activities such as metabolism of cofactors and vitamins, amino acid metabolism, nucleotide, carbohydrate, nitrogen, lipid, terpenoid, and energy metabolism, xenobiotics degradation and metabolism, and production of antibiotic-like compounds (secondary metabolites) for all the test samples, which were expressed in relative abundance (Figure 6A). Clusters of orthologous genes (COG) functional potential (Figure 6B) included amino acid transport and metabolism, carbohydrate transport and metabolism, cell cycle control, cell division and chromosome partitioning, cell wall membrane envelope biogenesis, coenzyme transport and metabolism, energy production and

conversion, inorganic ion, lipid, nucleotide transport and metabolism, post-translational modification of protein turnover, chaperones, replication, recombination and repair, transcription, and secondary metabolites biosynthesis transport and catabolism. There were 19 KEGG modules (Figure 6C) including *De novo* purine biosynthesis (M00048), Fatty acid biosynthesis (M00083), bacterial C10-C20 isoprenoid biosynthesis (M00364), Leucine and Isoleucine biosynthesis (M00432 and M00535), and Galactose degradation (M00632).

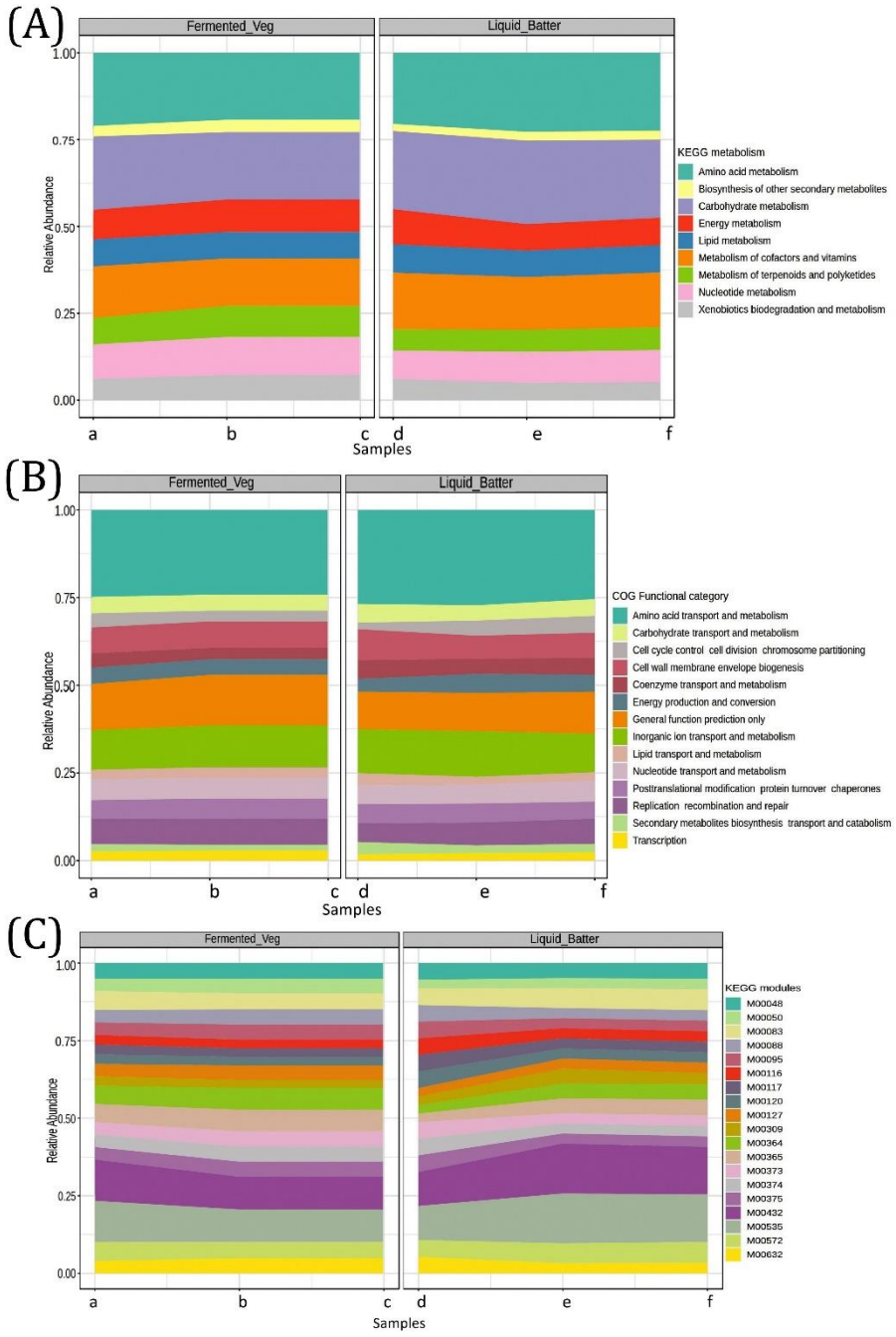

Figure 6. Functional Profiling of the Microbiome among Six Metagenome Samples, with Two Different Main Subtypes (Fermented_veg and Liquid_batter) in Terms of Relative Abundance Stated as a Value between Zero (lowest abundance) and One (highest abundance) predicted at: A) KEGG level metabolism showing the distribution of the top nine metabolic functions; B) COG level showing the distribution of the top fourteen clusters of orthologous genes; C) KEGG modules showing the distribution of 19 significant common modules between the studied samples. M00048: *De novo* purine biosynthesis, M00050: Guanine ribonucleotide biosynthesis, M00083: Fatty acid biosynthesis, M00088: Ketone body biosynthesis, M00095: C5 isoprenoid biosynthesis, mevalonate pathway, M00116: Menaquinone biosynthesis, M00117: Ubiquinone biosynthesis, M00120: Coenzyme A biosynthesis, M00127: Thiamine biosynthesis, M00309: Non-phosphorylative Entner-Doudoroff pathway, M00364: C10-C20 isoprenoid biosynthesis in bacteria, M00365: C10-C20 isoprenoid biosynthesis in archaea, M00373: Ethylmalonyl pathway, M00374: Dicarboxylate-hydroxybutyrate cycle, M00375: Hydroxypropionate-hydroxybutylate cycle, M00432: Leucine biosynthesis, M00535: Isoleucine biosynthesis, M00572: Pimeloyl-ACP biosynthesis, M00632: Galactose degradation (Leloir) pathway. a: Fermented_veg_mix. b: Pickled_veg_with_ oil. c: Pickled_veg_without_oil. d: Idli_batter-1. e: Idli_batter-2. f: Shinni.

Discussion

According to Zhang et al.[23] study, a diversity of non-culturable bacteria may exist in different foods, including fruits, vegetables, dairy products, meat products, rice and flour-related products. With the help of Next-generation sequencing (NGS) technique, we can identify and characterize the diverse microorganisms from different sample types with more precision and accuracy compared to traditional micro biological techniques. NGS can also be used to identify closely related bacterial species like the *Lactobacillus casei* group, which are not easy to discriminate based on traditional methods.[24]

The range of GC content may correspond with the ecological niche, genomic size, and structure of bacteria. Occasionally, it may vary even for strains.[24,25] According to our study, the GC contents (V3-V4 amplicon region) were in the range between 51% to 54.5%, with approximately 53% as the mean value. Previous research supports our results that Lactobacillus have GC contents ranging from 32% to 52%.[26] Bacteria belonging to different phyla, including *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, and *Acidobacteriota*, showed a varied range of GC contents, approximately 30% to 70%, 55% to 75%, 30% to 75%, 30% to 55%, and 52% to 65%, respectively.[27] Earlier studies on different isolates of probiotic *Bifidobacterium* showed varied GC contents of the genome ranging between 59.27% to 62.77%.[25]

In this study, we have found a diverse range of potential probiotic bacteria present in the test samples. Approximately 99%, 97%, 70%, 90%, 30%, and 75% of potential probiotic bacteria were found among the top enriched genera of Idli_batter-1, Idli_batter-2, Pickled_veg_without_oil, Pickled _veg_with_oil, Fermented_veg_mix, and Shinni, respectively. Commonly known pathogenic bacteria such as Pseudomonas and Staphylococcus were identified in 1-8% of Idli_batter-1, Pickled_veg_with_oil, Fermented_veg_mix, and Shinni, while approximately 30% presence was confirmed in Pickled_veg_without_oil among the top enriched genera. Recent studies from various regions around the globe have revealed that certain species within these genera possess good probiotic potential.

According to the study by Hu et al.,[28] plant growth capacity and nutrient assimilation can be enhanced by using probiotic *Pseudomonas communities*, which contain beneficial microbes. Another study highlighted the probiotic nature of *Staphylococcus gallinarum* FCW1 MCC4687, isolated from naturally fermented coconut water. Only Pickled_veg_with_ oil showed a high percentage abundance of two genera: *Bacteroides* and *Prevotella* (of the Phylum Bacteroidetes). Various factors, including the nutritional constituents present in food, may impact the conformation of niche-specific microbiota. Based on the published literature, bacteria belonging to the genus Bacteroides are mostly clinical pathogens that cause anaerobic infections. However, they have a positive association with the host,[29] and microbiomes constructed by *Prevotella* are capable of fermenting carbohydrates more quickly than Bacteroides-mediated niches.[30] All the samples showed the presence of a significant level of *Proteobacteria*, which is believed to prepare the gut for colonization with a large number of anaerobes essential for good gut function by lowering the redox potential and utilizing oxygen in the gut micro environment.[31] The Phylum Firmicutes was also abundant in all test food samples except Fermented_ veg_mix. Most potential probiotic bacteria belong to the phylum Firmicutes, which may enhance the composition of the microbiota by fermenting dietary fiber and could also interact with the gut mucosa to maintain the host's healthy gut.[32]

The comparative report on taxonomy revealed the presence of a single common bacterial genus, *Lactococcus*, in fermented Idli_batter-1, Idli_batter-2, and Pickled_veg_ with_oil whereas the *Acetobacter* genus is common between Pickled_veg_without_oil and Pickled_veg_with_oil. Although, bacteria such as *Cystobacter sp., Gemmiger formicilis, Megamonas sp., Prevotella copri, Lactobacillus zeae, Roseburia sp., Bacteroides sp.,* and *Acinetobacter johnsonii* were exclusively found in Pickled_veg_with_oil, possibly due to the presence of five distinct food compositions (*Brassica rapa*, *Cuminum cyminum*, *Amomum subulatum*, *Cinnamomum verum*, *Capsicum annuum*). On the other hand, Shinni had two dissimilar food compositions (*Triticum aestivum* and *Musa acuminata*) which exclusively constituted *Alcaligenes faecalis, Stenotrophomonas geniculata,* and *Weissella paramesenteroides* bacteria (Figure 3C and D). Two completely different food microbiomes of Idli_batter-2, Fermented_veg_mix shared three bacteria *Sphingomonas* sp., *Stenotrophomonas* sp., and *Acinetobacter rhizosphaerae*. Comparison between Idli_batter-1 and Shinni showed only one common genus *Lactobacillus*. Different types of traditional pickles are well-known foods that can be prepared using popular and ancient techniques, which offer tremendous health benefits. According to Behera et al.[33], different species of *Lactobacillus*, *Leuconostoc* and *Pediococcus* were identified from vegetable-based pickles (Nozawana-Zuke pickle or Tursu), radish pickle, cucumber pickle, Chinese pickle (Paocai), cabbage and carrot pickle, and Korean traditional fermented vegetables (Kimchi). One similar taxon and thirteen different taxa are present between two fermented idli batter samples (Idli_batter1 and 2), though Idli_batter-1 showed a higher beneficial bacterial load (including *Lactobacillus* sp., *Lactococcus* sp., *Pediococcus* sp.), possibly due to the change of a single food constituent (*Vigna radiata* in place of *Vigna mungo*) only (Figure 3C and D). According to previous studies, fermented idli batter can be a good source of lactic acid bacteria (LAB) and all the isolated LAB displayed a diverse range of

antibacterial activity against known pathogenic bacteria.[34-36]

Microbiome instabilities can be reduced through prebiotic - supplemented probiotic treatment. Prebiotics are characteristically oligosaccharides or more complex saccharides that are resistant to digestion, absorption, and adsorption by the host before fermentation, and can be utilized by commensal bacteria, specifically beneficial for the human host.[38] Some types of naturally occurring prebiotics include galactooligo saccharides (GOS), fructooligosaccharides (FOS), and inulin. GOS can be found in cow's milk, yogurt, and colostrum, whereas FOS and inulin are mainly found in more than 3,600 fruits and vegetables, such as asparagus, beans, cereals, and mostly belonging to the Cichorium family, such as chicory, bananas, large onions, and garlic.[39] In the current study, we have used similar types of vegetables and ingredients to prepare traditional fermented foods to ensure prebiotic properties, which can stimulate the growth of good bacteria and preferably contribute to creating a beneficial microbiome.

The KEGG metabolism experiment based on relative abundance showed that amino acid metabolism, carbohydrate metabolism, and metabolism of cofactors and vitamins were higher in Liquid_batter food type whereas Fermented_veg food type displayed high bacterial relative abundance related to the metabolism of polyketides, terpenoids, nucleotides, energy, xenobiotics, and xenobiotics biodegradation. Hormones, short-chain fatty acids (SCFAs), neurotransmitters and anti-inflammatory cytokines producing psychobiotic bacteria,[40] including *Bifidobacteria*, *Lactobacilli*, *Enterococci*, *Streptococci,* and *Escherichia* were most frequently abundant in our samples. Literature review suggests that SCFAs producing bacteria, including *Prevotella*, *Bacteroides*, *Ruminococcaceae*, and *Lachnospiraceae* are good for a healthy gut environment,[41] which could help humans by maintaining a balance through reducing the Firmicutes to Bacteroidetes ratio as well as decreasing the presence of lipopolysaccharide (LPS) producing bacteria, and facultative anaerobic bacteria.[42]

Explicitly, amino acid metabolism was higher in Fermented_veg_mix, Idli_batter-2 and shinni, whereas both types of idli batter and shinni displayed carbohydrates, co-factors and vitamin metabolism. *Prevotella* found in Pickled _veg_with_oil is normally found in specified food enriched with monounsaturated fatty acids (such as the “Mediterranean diet”) that correlates with the reduction of traditional risk factors defining metabolic syndrome (MetS) and cardio vascular disease (CVD).[43,44] The detailed analysis of KEGG modules revealed different pathways such as leucine, isoleucine, fatty acid, C5 isoprenoid, pimeloyl-ACP biosynthesis and mevalonate, non-phosphorylative Entner-Doudoroff pathway found more in fermented idli batter (both 1 and 2) as well as shinni. This advocates that these food microbiomes are connected with the production of different essential amino acids and important vitamins like biotin (vitamin B7).[45] Fermented_veg_mix and Pickled_veg (with_oil and without _oil) showed guanine ribonucleotide biosynthesis, which is a method related to purine nucleotide biosynthesis and can eventually synthesize GMP to GTP, which helps to form RNA.[46] The C10-C20 isoprenoid biosynthesis was also found to be enhanced in samples of vegetable ingredients, reflecting that these microbiomes are capable of producing isoprenoid bioactive compounds like different terpenes.[47] Comparable to our study, food-derived LAB also displayed C10-C20 isoprenoid biosynthesis and the pentose phosphate pathway.[48]

Finally**,** although the study showed that NGS-based metagenomic approach is an acceptable technique to identify different types of bacterial presence in food microbiomes, it has some limitations. The study revealed the presence of several known beneficial as well as pathogenic bacterial genera in the test microbiomes. However, it is not possible to determine the specific characteristics of beneficial or pathogenic bacteria based on the detection method used. Further in vitro isolation might be a future scope to address this issue. Another limitation of the study was the species-level taxonomical analysis. Only a few species-level identifications were confirmed, as mentioned in the results section.

Conclusion

It is a novel NGS-based culture free research work to explore diversified bacterial microbiomes including beneficial probiotics from non-dairy based traditional food products such as fermented idli batter, vegetable pickles, fermented vegetables and shinni. Thus, metagenomic analysis revealed probiotic bacterial prevalence in each test sample, which confirms all tested locally available non-dairy based foods can be used as good sources of probiotics. Metabolic pathway and functional analysis displayed all the tested microbiomes can play diverse metabolic activities and are crucial for human metabolism. Based on the maximum presence of probiotic beneficial bacteria, all the samples can be represented in increasing order:

Fermented_veg_mix<Pickled_veg_without_oil< Idli_batter-2 < Shinni < Pickled_veg_with_oil < Idli_batter-1.

Authors’ Contributions

BS: Performed experimental works and wrote the paper; MM: Executed software work on metagenomics and statistical analysis; SM: Designed the study, analysed, discussed and approved the paper.

Conflict of Interest Disclosures

The authors declare that they have no conflicts of interest.

References

1. The Human Microbiome Project Consortium. Structure,

function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207-14. doi:10.1038/nature11234
2. Dei-Cas I, Giliberto F, Luce L, Dopazo H, Penas-Steinhardt A. Metagenomic analysis of gut microbiota in non-treated plaque psoriasis patients stratified by disease severity: development of a new Psoriasis-Microbiome Index. Sci Rep. 2020;10(1):12754. doi:10.1038/s41598-020-69537-3
3. Bienenstock J, Gibson G, Klaenhammer TR, Walker WA, Neish AS. New insights into probiotic mechanisms. Gut Microbes. 2013;4(2):94-100. doi:10.4161/gmic.23283
4. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, et al. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol. 2014;11(8):506-14. doi:10.1038/nrgastro.2014.66
5. Patro JN, Ramachandran P, Barnaba T, Mammel MK, Lewis JL, Elkins CA. Culture-Independent Metagenomic Surveillance of Commercially Available Probiotics with High-Throughput Next-Generation Sequencing. mSphere. 2016;1(2). doi:10.1128/mSphere.00057-16
6. Lugli GA, Mangifesta M, Mancabelli L, Milani C, Turroni F, Viappiani A, et al. Compositional assessment of bacterial communities in probiotic supplements by means of metagenomic techniques. Int J Food Microbiol. 2019;294:1-9. doi:10.1016/j.ijfoodmicro.2019.01.011
7. Kaktcham PM. Antimicrobial and Safety Properties of Lactobacilli Isolated from Two Cameroonian Traditional Fermented Foods. Sci Pharm. 2012;80(1):189-203. doi:10.3797/scipharm.1107-12
8. Hwang CE, Haque MdA, Hong SY, Kim SC, Cho KM. Origin of lactic acid bacteria in *mulkimchi* fermentation. J Appl Biol Chem. 2019;62(4):441-6. doi:10.3839/jabc.2019.060
9. Jeong DM, Yoo SJ, Jeon MS, Chun BH, Han DM, Jeon CO, et al. Genomic features, aroma profiles, and probiotic potential of the *Debaryomyces hansenii* species complex strains isolated from Korean soybean fermented food. Food Microbiol. 2022:104011. doi:10.1016/j.fm.2022.104011
10. Sawada K, Koyano H, Yamamoto N, Yamada T. The effects of vegetable pickling conditions on the dynamics of microbiota and metabolites. PeerJ. 2021;9:e11123. doi:10.7717/peerj.11123
11. Mandhania MH, Paul D, Suryavanshi MV., Sharma L, Chowdhury S, Diwanay SS, et al. Diversity and Succession of Microbiota during Fermentation of the Traditional Indian Food Idli. Appl Environ Microbiol. 2019;85(13). doi:10.1128/AEM.00368-19
12. Ganguly S. Cereal-based Fermented Foods for Enhanced Nutritional Attributes and Better Gut Health. Int J Fermented Foods. 2021;10(1). doi:10.30954/2321-712X.01.2021.1
13. Hiergeist A, Ruelle J, Emler S, Gessner A. Reliability of species detection in 16S microbiome analysis: Comparison of five widely used pipelines and recommendations for a more standardized approach. PLoS One. 2023;18(2):e0280870. doi:10.1371/journal.pone.0280870
14. Paulson JN, Chen CY, Lopes-Ramos CM, Kuijjer ML, Platig J, Sonawane AR, et al. Tissue-aware RNA-Seq processing and normalization for heterogeneous and sparse data. BMC Bioinformatics. 2017;18(1):437. doi:10.1186/s12859-017-1847-x
15. Mandal, Mandal S. Cross-biome metagenomic analyses of the impact of pollutants on taxonomic and functional diversity of bacterial communities from different geographical regions. Gene Rep. 2022;29:101690. doi:10.1016/j.genrep.2022.101690
16. Godlewska U, Brzoza P, Kwiecień K, Kwitniewski M, Cichy J. Metagenomic Studies in Inflammatory Skin Diseases. Curr Microbiol. 2020;77(11):3201-12. doi:10.1007/s00284-020-02163-4
17. Jain AK, Dubes RC. Algorithms for Clustering Data. USA: Prentice-Hall, Inc.; 1988.
18. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res. 2017;45(W1):W180-8. doi:10.1093/nar/gkx295
19. Mishra P, Singh U, Pandey C, Mishra P, Pandey G. Application of student's t-test, analysis of variance, and covariance. Ann Card Anaesth. 2019;22(4):407. doi:10.4103/aca.ACA_94_19
20. McCarthy DJ, Smyth GK. Testing significance relative to a fold-change threshold is a TREAT. Bioinformatics. 2009;25(6):765-71. doi:10.1093/bioinformatics/btp053
21. Stephens M. False discovery rates: a new deal. Biostatistics. 2016;kxw041. doi:10.1093/biostatistics/kxw041
22. McHugh M. Standard error: meaning and interpretation. Biochem Med (Zagreb). 2008;7-13. doi:10.11613/BM.2008.002
23. Zhang J, Yang H, Li J, Hu J, Lin G, Tan BK, et al. Current Perspectives on Viable but Non-Culturable Foodborne Pathogenic Bacteria: A Review. Foods. 2023;12(6):1179. doi:10.3390/foods12061179
24. Ullah M, Rizwan M, Raza A, Zhao X, Sun Y, Gul S, et al. Comparative Genomic and Functional Characterization of *Lactobacillus casei* Group (LCG) Probiotic Strains Isolated from Traditional Yogurts by Next-Generation Sequencing. Pak J Zool. 2023;55(4). doi:10.17582/journal.pjz/20210711190719
25. Korzhenkov AA, Tepliuk A V, Sidoruk K V, Voyushin KE, Patrushev M V, Kublanov I V, et al. A dataset of four probiotic Bifidobacterium strains genome assemblies. Data Brief. 2021;34:106710. doi:10.1016/j.dib.2020.106710
26. Mendes-Soares H, Suzuki H, Hickey RJ, Forney LJ. Comparative Functional Genomics of Lactobacillus spp. Reveals Possible Mechanisms for Specialization of Vaginal Lactobacilli to Their Environment. J Bacteriol. 2014;196(7):1458–70. doi:10.1128/JB.01439-13
27. Teng W, Liao B, Chen M, Shu W. Genomic Legacies of Ancient Adaptation Illuminate GC-Content Evolution in Bacteria. Microbiol Spectr. 2023;11(1). doi:10.1128/spectrum.02145-22
28. Dhanya Raj CT, Kandaswamy S, Suryavanshi MV, Ramasamy KP, Rajasabapathy R, Arthur James R. Genomic and metabolic properties of *Staphylococcus gallinarum* FCW1 MCC4687 isolated from naturally fermented coconut water towards GRAS assessment. Gene. 2023;867:147356. doi:10.1016/j.gene.2023.147356
29. Cho S, Kim D, Lee Y, Kil EJ, Cho MJ, Byun SJ, et al. Probiotic Lactobacillus Paracasei Expressing a Nucleic Acid-Hydrolyzing Minibody (3D8 Scfv) Enhances Probiotic Activities in Mice Intestine as Revealed by Metagenomic Analyses. Genes (Basel). 2018;9(6):276. doi:10.3390/genes9060276
30. Flint HJ, Duncan SH, Scott KP, Louis P. Links between diet, gut microbiota composition and gut metabolism. Proc Nutr Soc. 2015;74(1):13-22. doi:10.1017/S0029665114001463
31. Shin NR, Whon TW, Bae JW. Proteobacteria: microbial signature of dysbiosis in gut microbiota. Trends

Biotechnol. 2015;33(9):496-503. doi:10.1016/j.tibtech.2015.06.011
32. Sun Y, Zhang S, Nie Q, He H, Tan H, Geng F, et al. Gut firmicutes: Relationship with dietary fiber and role in host homeostasis. Crit Rev Food Sci Nutr. 2022;1-16. doi:10.1080/10408398.2022.2098249
33. Behera SS, El Sheikha AF, Hammami R, Kumar A. Traditionally fermented pickles: How the microbial diversity associated with their nutritional and health benefits? J Funct Foods. 2020;70:103971. doi:10.1016/j.jff.2020.103971
34. Iyer BK, Singhal RS, Ananthanarayan L. Characterization and *in vitro* probiotic evaluation of lactic acid bacteria isolated from idli batter. J Food Sci Technol. 2013;50(6):1114-21. doi:10.1007/s13197-011-0445-6
35. Katepogu H, Wee YJ, Almaary KS, Elbadawi YB, Gobinath R, Chinni SV, et al. Isolation and Characterization of *Pediococcus* sp. HLV1 from Fermented Idly Batter. Fermentation. 2022;8(2):61. doi:10.3390/fermentation8020061
36. Sircar B, Mandal S. Exploring the probiotic potentiality and antibacterial activity of idli batter isolates of lactic acid bacteria from West Bengal, India. Futur J Pharm Sci. 2023;9(1):54. doi:10.1186/s43094-023-00506-z
37. Ma C, Wasti S, Huang S, Zhang Z, Mishra R, Jiang S, et al. The gut microbiome stability is altered by probiotic ingestion and improved by the continuous supplementation of galactooligosaccharide. Gut Microbes. 2020;12(1):1785252. doi:10.1080/19490976.2020.1785252
38. Preidis GA, Versalovic J. Targeting the Human Microbiome with Antibiotics, Probiotics, and Prebiotics: Gastroenterology Enters the Metagenomics Era. Gastroenterology. 2009;136(6):2015-31. doi:10.1053/j.gastro.2009.01.072
39. Gupta D, Saini S, Saxena S, Saxena S, Varshney T. Current Research and Future Trends in Prebiotics. nt J Innov Res Sci Eng Technol. 2023;10(4):21-8.
40. Choudhary S, Shanu K, Devi S. Psychobiotics as an Emerging Category of Probiotic Products. In: Probiotics, Prebiotics, Synbiotics, and Postbiotics. Singapore: Springer Nature Singapore; 2023. pp. 361-91. doi:10.1007/978-981-99-1463-0_19
41. Sakamoto M, Takagaki A, Matsumoto K, Kato Y, Goto K, Benno Y. *Butyricimonas synergistica* gen. nov., sp. nov. and *Butyricimonas virosa* sp. nov., butyric acid-producing bacteria in the family "Porphyromonadaceae" isolated from rat faeces. Int J Syst Evol Microbiol. 2009;59(7):1748-53. doi:10.1099/ijs.0.007674-0
42. Chen J, Xiao Y, Li D, Zhang S, Wu Y, Zhang Q, et al. New insights into the mechanisms of high-fat diet mediated gut microbiota in chronic diseases. iMeta. 2023;2(1). doi:10.1002/imt2.69
43. Gillingham LG, Harris-Janz S, Jones PJH. Dietary Monounsaturated Fatty Acids Are Protective Against Metabolic Syndrome and Cardiovascular Disease Risk Factors. Lipids. 2011;46(3):209-28. doi:10.1007/s11745-010-3524-y
44. Piccioni A, Covino M, Candelli M, Ojetti V, Capacci A, Gasbarrini A, et al. How Do Diet Patterns, Single Foods, Prebiotics and Probiotics Impact Gut Microbiota? Microbiol Res (Pavia). 2023;14(1):390-408. doi:10.3390/microbiolres14010030
45. Hu Y, Cronan JE. α-proteobacteria synthesize biotin precursor pimeloyl-ACP using BioZ 3-ketoacyl-ACP synthase and lysine catabolism. Nat Commun. 2020;11(1):5598. doi:10.1038/s41467-020-19251-5
46. Sekiguchi T, Ito R, Hayakawa H, Sekiguchi M. Elimination and Utilization of Oxidized Guanine Nucleotides in the Synthesis of RNA and Its Precursors. J Biol Chem. 2013;288(12):8128-35. doi:10.1074/jbc.M112.418723
47. Pazouki L, Niinemets ь. Multi-Substrate Terpene Synthases: Their Occurrence and Physiological Significance. Front Plant Sci. 2016;7. doi:10.3389/fpls.2016.01019
48. Jin H, Quan K, You L, Kwok LY, Ma T, Li Y, et al. A genomic compendium of cultivated food-derived lactic acid bacteria unveils their contributions to human health. Sci Bull (Beijing). 2024:S2095-9273(24)00885-5. doi:10.1016/j.scib.2024.12.002