BMC Microbiology

RESEARCH Open Access

Microbiome study of Murrah buffalo mastitis milk with emphasis on *Acinetobacter* species

Damini Sharma[1†], Hemlata Valmiki[1†], Pankaj Chayal[1], Sanjay Kumar[1] and Supriya Chhotaray[1*]

Abstract
Mastitis has been a major challenge in dairy industry incurring heavy loss to dairy farmers. Although targeted antibiotic regime is successful for culturable microbes, the non-culturable and novel microbes are often overlooked. Hence, the present study employed a whole-metagenome profiling to investigate and compare the microbial diversity among the milk samples of healthy and affected buffaloes. 16 Milk samples were collected from Murrah buffaloes and classified into three groups based on the somatic cell count and California Mastitis Test score, i.e., Clinical (3), Sub-clinical (6), and Healthy (7). DNA extraction from milk was followed by Whole Meta-Genome Shotgun (WGS) sequencing to study microbial diversity. The study revealed that *Proteobacteria* as the most abundant phylum in the clinical mastitis cases, which could be related to the severity of the disease, whereas *Firmicutes* were strongly associated with the healthy group of buffaloes. The genus *Acinetobacter* was most abundant in clinical cases (79%) and least in healthy animals (33%). The present study provides important insights into the microbial population and composition in the milk of mastitis-affected buffaloes. The findings will aid in investigating potential therapeutic ways for reducing antibiotic resistance and treatment costs.

Keywords *Acinetobacter*, Mastitis, Microbiome, Whole metagenome

Introduction

Livestock farming is continuously facing the severity of mastitis and its consequences are affecting the health of dairy herds and the economy worldwide. Mastitis can be categorized into Clinical mastitis (CM) and Sub-clinical mastitis (SCM). Visible signs and symptoms can be easily noticed in CM cases, for instance, swelling in the udder, watery and discoloured milk [2]. In contrast, it is not the same in the case of SCM as symptoms are not evident, and are often detected by Somatic cell counting (SCC), California mastitis test (CMT), and Infrared Thermography (IRT). Over time, incidences of mastitis have drastically increased. Globally, mastitis is associated with approximately $35 billions of annual economic losses [16]. In their 1962 study, Dandha and Sethi were the first to estimate the annual economic loss due to mastitis in India, which was 529 million INR. Over the decades, this amount increased significantly, with Bansal and Gupta 4 reporting 716.551 million INR in 2007.

Buffalo (*Bubalus bubalis*), or "Black Gold" is a vital livestock species contributing to milk, meat, draught power, manure, and fuel. According to Basic Animal Husbandry Statistics (BAHS) 2024, India's milk production rose by 3.78% in 2023–24 compared to 2022–23 estimates. The Food and Agriculture Organization (FAO) 2022 states that buffalo milk accounts for about 15% of global production. Buffaloes are less prone to mastitis than cattle.

†Damini Sharma and Hemlata Valmiki contributed equally to this paper and share 1st authorship.

*Correspondence:
Supriya Chhotaray
supriya.chhotaray@gmail.com
[1]Animal Genetics and Breeding Division, ICAR – Central Institute for Research on Buffaloes, Hisar, India

However, the mortality rate of young calves was observed to be high due to mother buffalo mastitis [1]. Bacterial infections are the primary cause of mastitis in buffaloes, and these infections can be exacerbated by poor management practices, environmental conditions, and nutritional deficiencies. Bardhan [5] estimated an average monthly loss of INR 1,934.78 per buffalo due to mastitis. Singh et al. [27] reported an annual economic loss of INR 24,175 per buffalo, with treatment costs being a major contributor. Das et al. [11] noted a 2.58 L daily milk yield reduction in subclinical mastitis cases, with INR 647.36 (50.87% of total loss) spent on treatment and INR 625 (49.12%) attributed to production losses.

Shotgun metagenomics (shotgun means untargeted and metagenome means microbial genome) [54] is a powerful approach to studying the whole genome, which includes the genome of all microbial communities ranging from archaea to viruses, which will further allow the identification of the strain level and detailed functional characterization. Whereas, 16 s rRNA gene sequencing provides details about the bacterial community based on the sequence of 16 s rRNA, a conserved region. Metagenome analysis can help diagnose and treat mastitis in buffalo. For instance, Jiang et al. [15] studied microbial diversity, which provides a way for treatment options, as they investigated the water extracts of Chinese herbal that can effectively inhibit the growth of *Escherichia coli, Staphylococcus aureus,* and *Klebsiella pneumoniae*. By analyzing the entire genetic material of a microbial community, metagenomics enables the identification of functional genes and metabolic pathways that are critical for microbial-host interactions.

This study aimed to analyze and compare the microbial communities in milk samples from Murrah buffaloes under three conditions: healthy, subclinical mastitis, and clinical mastitis (Fig. 1).

Materials and methods

Sample collection

A total of 16 milk samples were collected from the Murrah breed housed at the ICAR-Central Institute for Research on Buffaloes (CIRB), located in Hisar, Haryana, India (latitude 29.1492° N, longitude 75.7217° E). The samples were collected during the post-monsoon season. More description of milk samples collected from animals with clinical mastitis (MC), subclinical mastitis (MS), and healthy condition (HG), including animal information (Parity, Lactation Stage, and Yield), is provided in the supplementary data (Table S1). This seasonal period was characterized by moderate temperature and high humidity, creating favorable microbial growth and activity conditions. At ICAR-CIRB, Murrah buffaloes are fed with seasonal green and dry fodder along with oilseed cakes and salts for minerals that maintain the physical health of buffaloes and fulfil nutrient needs for physiological balance.

Before the sample collection, the udder of each animal was well-cleaned to remove dirt and fecal matter. After

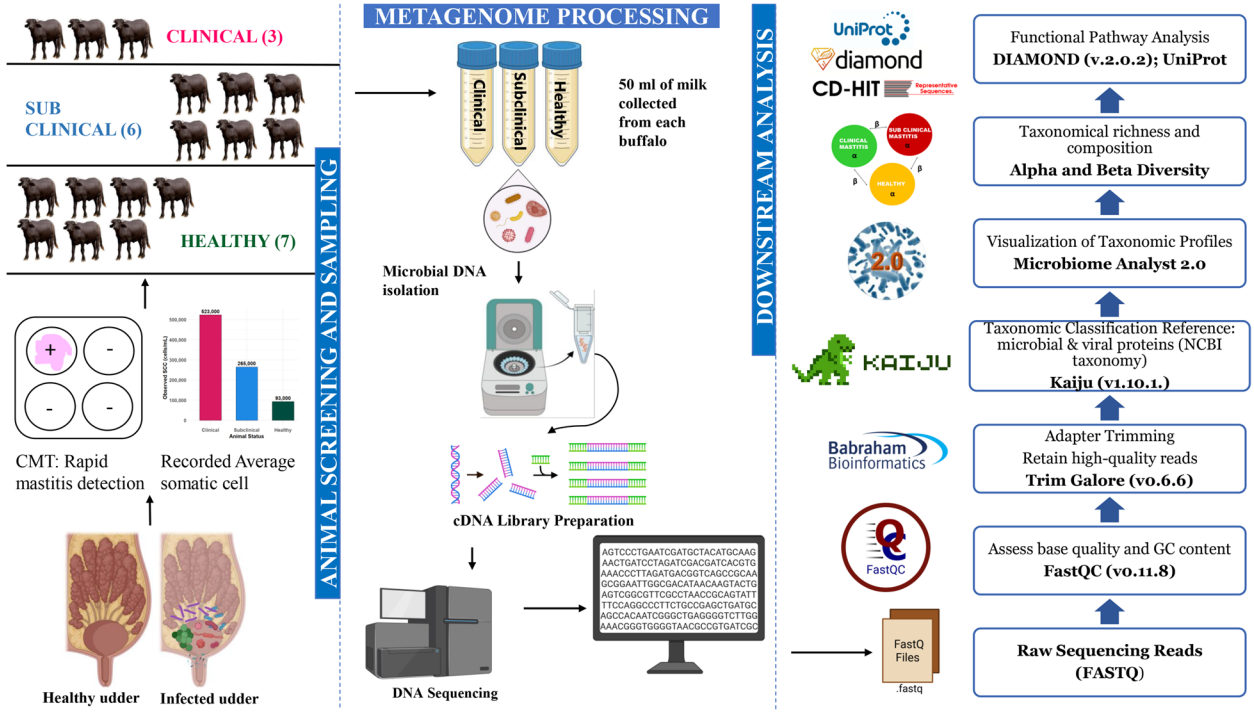

Fig. 1 Overall experimental workflow depicting animal screening and sampling, metagenome processing, and downstream analysis

cleaning up, the teats and udder were washed with 70% ethanol to disinfect. Milk samples were visually assessed to check the irregularities in color, texture, and consistency. 50 ml of each milk sample was collected aseptically in 50 ml sterile autoclaved Falcon tubes. After collection, the samples placed on ice, were brought into the laboratory.

Sample screening for mastitis

Buffaloes were grouped into three categories based on their mastitis status, and their health was evaluated through physical observation, the California Mastitis Test (CMT), and Somatic Cell Counts (SCC). This classification was based on the study by Dhakal [31], which found that normal Somatic Cell Counts (SCC) in Murrah buffaloes are typically ≤ 200,000 cells/ml, with more than 64% of healthy quarters having SCC ≤ 100,000/ml. Higher SCC values, particularly ≥ 200,000 cells/ml, indicate subclinical mastitis. Clinical cases were initially assessed based on observable symptoms. Udder abnormalities included swelling, redness, and a watery appearance of the milk with clots, which were later confirmed using the California Mastitis Test (CMT). The average somatic cell count (SCC) for cases of mastitis were observed to be more than 4,00,000. Table 1 provides the average SCC for the categorized animals. Using these criteria, the buffaloes in this study were divided into healthy, subclinical, and clinical groups based on their SCC, CMT, and clinical signs.

The CMT was performed by mixing approximately half a teaspoon of milk with an equal amount of CMT reagent and gently rotating the mixture to ensure thorough blending. The score was recorded within ten seconds while still rotating, as the reaction tends to break down after about 20 s.

Bacterial DNA isolation and assessment

The DNA was extracted by DNeasy® PowerFood® Microbial Kit (QIAGEN, Germany) following the manufacturer's instructions. The purity and concentration of DNA were checked by a micro-ultraviolet spectrophotometer (USA, ThermoFisher). The purity of DNA was determined based on the OD260/OD280 ratio (1.8–2.0) and was adjusted to a DNA concentration of 20–200 ng/µL. All extracted DNA samples were stored at – 20 °C for further analysis.

Table 1 Information on sampled buffaloes

Animal status	Threshold SCC (cells/mL) of the samples	Observed SCC (cells/mL) of the samples	Clinical symptoms	CMT
Healthy (7)	≤ 100,000	93,000	Normal milk	Negative
Subclinical (6)	200,000–400,000	2,65,000	No visible symptoms	Positive
Clinical (3)	≥ 400,000	5,23,000	Swelling & redness of udder and teats, with clots and abnormal milk	Positive

Sequence library preparation

The samples were outsourced for sequencing by NGB Diagnostics Pvt. Ltd, India. 1 µg DNA per sample was used for library preparation. Sequencing libraries were generated using NEBNext® Ultra™ DNA Library Prep Kit for Illumina (NEB, USA) following the manufacturer's recommendations. Index codes were added to attribute sequences to each sample. Briefly, DNA libraries of fragments (approximately 400 bp long) were prepared separately for each sample. The samples were sequenced on an Illumina HiSeq Xten instrument (Illumina, San Diego, California, USA). Paired-end reads (151 bp) were generated in both forward and reverse directions.

Metagenome data analysis

Read quality check

Raw sequencing reads were first subjected to quality control checks using FastQC (version 0.11.8) http://www.bioinformatics.babraham.ac.uk/projects/fastqc. This tool was used to evaluate the metrics. Important qualities to ensure the reliability of the data before downstream analysis, including base quality score distribution, sequence quality score distribution, average base content per read, and GC distribution in the reads were checked.

Adapter trimming

After quality assessment, raw reads were processed to remove adapter sequences using TrimGalore (version 0.6.6). Illumina's universal adapter was specifically targeted for trimming to ensure that contamination of the adapter is minimized. After pruning, only high-quality reads without adapter sequences were kept for further analysis.

Taxonomical classification

The trimmed reads were then mapped to the Kaiju database https://bioinformatics-centre.github.io/kaiju/, which is a fast and sensitive taxonomic classification of high-throughput sequencing reads from metagenomic whole genome sequencing and meta-transcriptomic experiments. Each sequencing read was assigned to a taxon in the NCBI taxonomy by comparing it to a reference protein database containing microbial and viral protein sequences. By using protein-level classification, Kaiju achieves a higher sensitivity compared with methods based on nucleotide comparison [29]. The Krona chart was then plotted from the kaiju mapping file [21].

Taxonomical profile plotting

After taxonomic classification, we performed in-depth statistical visualization of microbial data using MicrobiomeAnalyst (https://www.microbiomeanalyst.ca/), a web-based tool for comprehensive analysis of microbial datasets. The tool offers a variety of statistical techniques and visualization options for data interpretation [30]. Bar plots were created to visualize the relative abundance of taxa at different taxonomic levels (e.g., phylum, genus, and species). LefSE analysis (Linear discriminant analysis effect size) was used to identify classes with different abundances between samples. This method combines statistical significance testing and effect size estimation to help identify biomarkers associated with different conditions (e.g., healthy vs. mastitis). Heat maps were created to visualize the abundance of classes in a sample, which allowed us to identify clustering patterns among samples with similar taxonomic structures. This reveals possible relationships with specific experimental conditions. Hierarchical grouping of samples is done using dendrograms for a tree-like visualization of the relationships between samples based on their microbial profiles.

Bacterial taxonomic diversity

Alpha diversity

Microbial community diversity was estimated with both alpha and beta diversity metrics. Alpha diversity (diversity within samples) was estimated using the Shannon index and Simpson index for all three groups of samples [45]. Shannon's index H estimates taxa diversity by combining richness and evenness, measuring the uncertainty in predicting a species of an individual from a community. It ranges from 0 to a maximum based on the number of species and their relative abundances. A higher value indicates greater diversity, with the index emphasizing taxa richness. The value increases with more species and a more even distribution of individuals [46].

$$H\prime = -\sum_{i=1}^{s} p_i log\,(p_i)$$

where, s is the number of OTUs and p_i is the proportion of the community represented by OTU_i.

Simpson's index D measures taxa diversity, combining richness and evenness but emphasizes evenness. It indicates taxa dominance and the likelihood of two randomly chosen individuals belonging to the same taxa, ranging from 0 to 1. The index increases as diversity decreases. It reflects species dominance and the chance of two individuals being from the same species, with higher values indicating lower diversity [47].

$$D = \frac{1}{\sum_{i=1}^{s} P_i^2}$$

where, s is the total number of species in the community and pi is the proportion of community represented by OTU_i.

Beta diversity

Customized R scripts were employed to evaluate the disparities in microbial community composition between clinical mastitis and healthy samples at the genus level. Bray–Curtis dissimilarity matrices were computed based on normalized genus abundance data utilizing the *vegan* R package [48]. Principal Coordinate Analysis (PCoA) was conducted to illustrate beta diversity patterns among samples. The statistical significance of group differences was assessed using Permutational Multivariate Analysis of Variance (PERMANOVA), executed through the adonis2 function in the vegan package with 9,999 permutations.

Statistical analysis

Pairwise t-test was applied to determine statistically significant differences in microbial diversity between the clinical, subclinical, and healthy sample groups. To correct the multiple testing, the p-value was adjusted using the Benjamini–Hochberg method to control the False discovery rate (FDR). Adjusted p-values below 0.05 were considered statistically significant. All analyses were conducted using R version 4.4.2.

Functional analysis

Adapter-trimmed sequencing data were assembled into a de novo assembly using SPAdes (v.3.15) software for metagenomic assembly with default parameters. Gene prediction on these contigs was performed with MetaGeneMark, which employed an existing lineage-specific database to accurately identify potential coding regions. CD-hit (v.4.6) was then used to generate a non-redundant set of nucleotide genes based on sequence similarity [51]. This non-redundant gene set was aligned with the updated NR (Non-Redundant) database using DIAMOND (v.2.0.2), facilitating rapid and sensitive homology searches for functional annotation [49]. To gain further functional understandings, the predicted genes were annotated through the UniProt database, using RefSeq (NRID) identifiers to retrieve detailed annotations, including gene names, functions, and associated pathways [50].

Results

Sequence results

The raw FastQ sequences upon quality check revealed the dataset comprised 5.5 million to 11.6 million paired-end

Table 2 Read statistics in different groups

Group	No. of samples	Average reads per sample	Average GC content
Healthy	7	76,72,151	44.86%
Subclinical	6	97,62,219	45.17%
Clinical	3	82,71,659	43.33%

reads per sample with an average read length of 151 bp. In terms of total average reads, the subclinical group had the highest count (9.76 million), followed by the clinical group (8.27 million) and the healthy group (7.67 million). The average GC content for the groups was 45.17% for subclinical, 44.86% for healthy, and 43.33% for clinical (Table 2).

Dominant microbial community and taxonomy assessment

Phylum level distribution

Several microbial phyla showed varied abundance, some specific to a single group, while others were shared across multiple groups. The buffalo's milk comprises various microbes belonging to different phyla, including *Proteobacteria*, *Firmicutes*, *Actinobacteria*, *Deinococcus-Thermus*, *Bacteroidetes*, *Cyanobacteria*, *Candidatus sp.*, and others. The milk microbiota composition revealed significant differences in the relative abundance among clinical (Group I), sub-clinical (Group II), and healthy (Group III) buffalo groups. *Proteobacteria* were the most prevalent across all categories, peaking in clinical mastitis samples at 79.61%, 62.45% in sub-clinical, and 44.25% in healthy animals (Fig. 2a). Conversely, the abundance of *Firmicutes* and *Actinobacteria* was lower in clinical mastitis samples (18.39% and 1.52%, respectively) than in subclinical (33.68% and 2.84%) and healthy samples (47.02% and 2.89%), suggesting an inverse relationship with disease severity. The abundance patterns were also depicted by heat maps. The phylum-level heatmaps (Fig. 3a and S2a) utilized abundance data and statistical significance from three distinct groups, respectively. Group 2, which indicates cases of subclinical mastitis, demonstrated a higher relative abundance across several phyla, including *Chlorobi*, *Chlamydiae*, *Chloroflexi*, *Proteobacteria*, *Candidatus_Gracilibacteria*, *Actinobacteria*, *Acidobacteria*, *Cyanobacteria*, and *Bacteroidetes*, among others. A highly significant difference in *Deinococcus* abundance was observed in healthy cases (5.05%) compared to both subclinical abundance, 0.017% (FDR = 2.66×10^{-14}) and clinical groups, 0.008% (FDR = 7.27×10^{-10}), emphasizing its potential as a marker for healthy mammary microbiota. This pattern is also well explained in Table S2, which

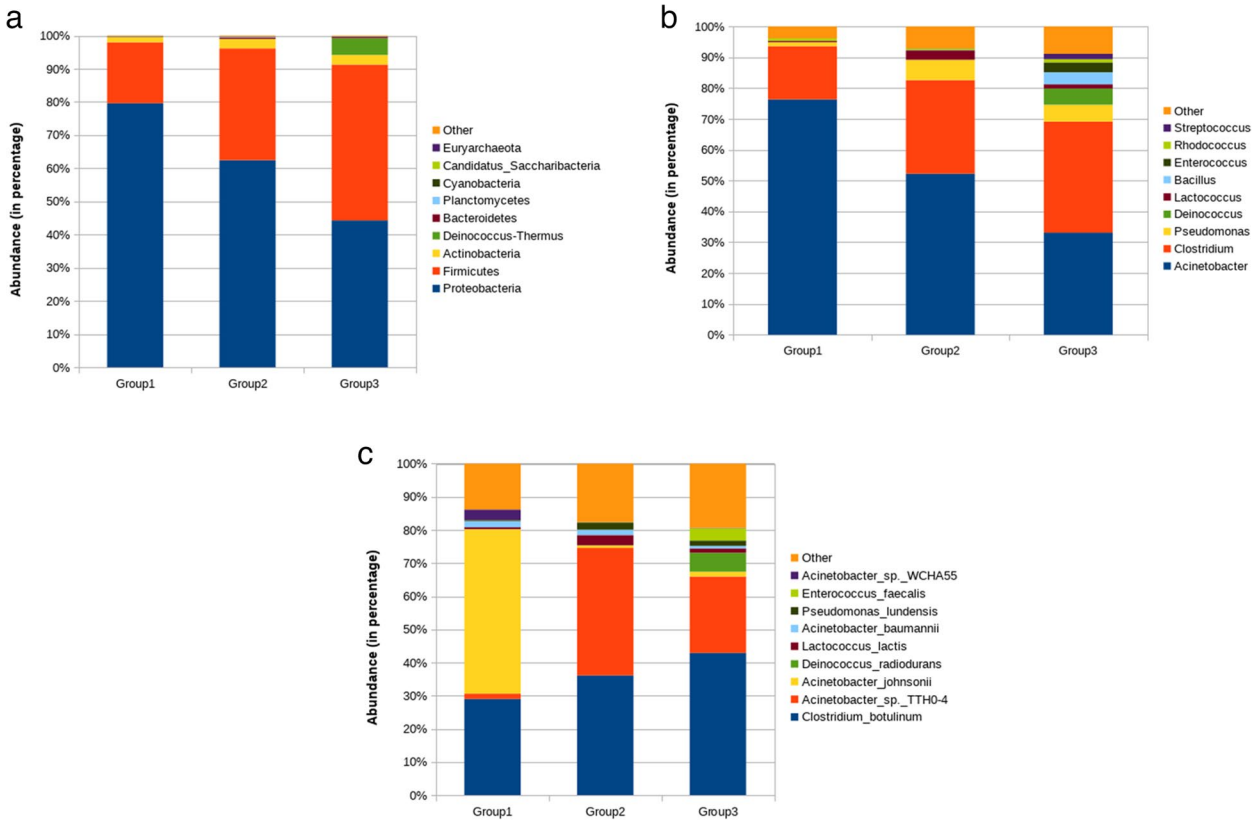

Fig. 2 Microbial abundance plot at the (**a**) phylum level, (**b**) genus level, (**c**) species level (where Group1: Clinical mastitis, Group2: Subclinical mastitis, and Group3: Healthy)

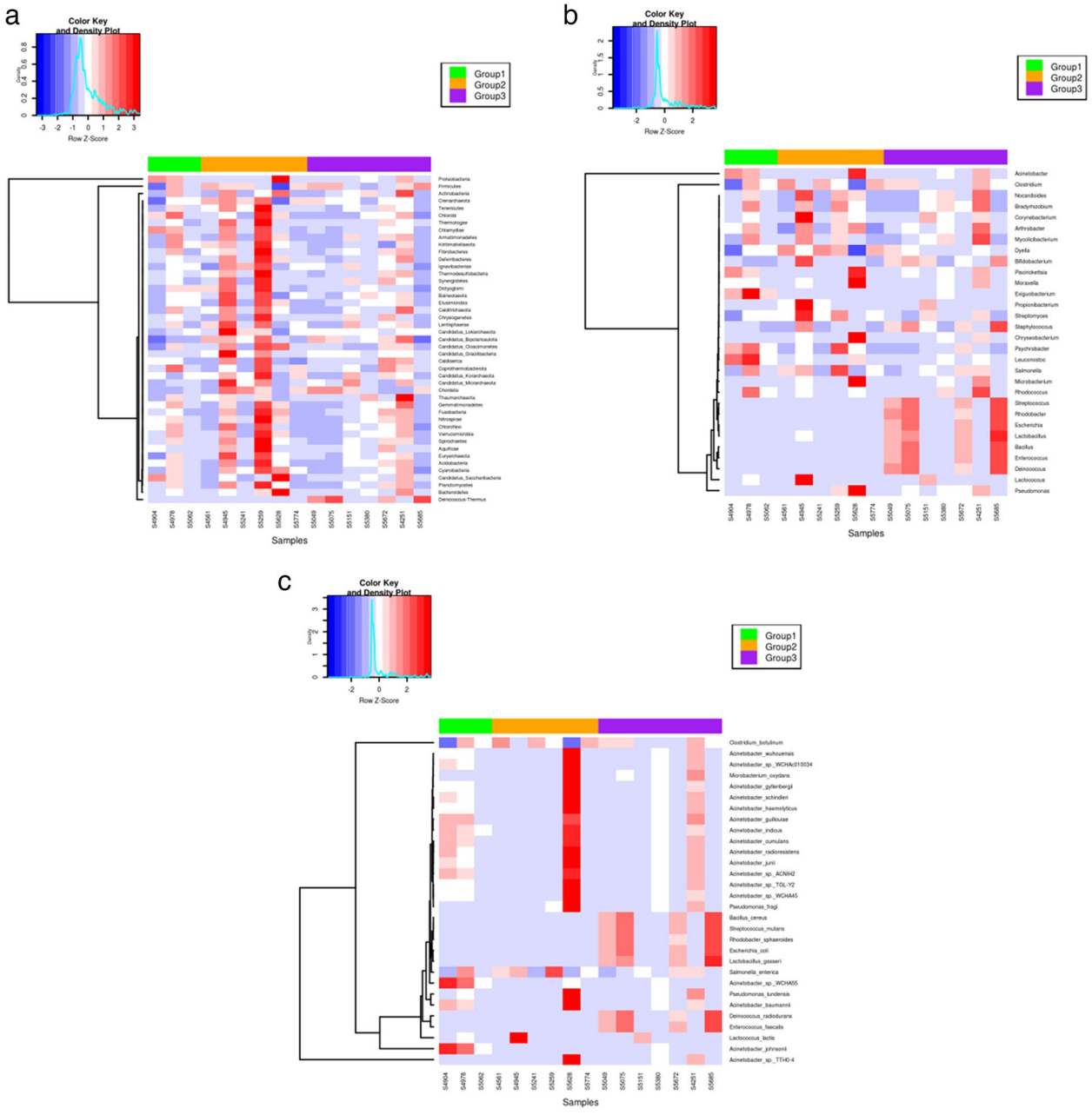

Fig. 3 Representation of taxonomical composition in individual samples through heat map (**a**) at the phylum level, (**b**) at the genus level, (**c**) at the species level (where Group1: Clinical mastitis, Group2: Subclinical mastitis, and Group3: Healthy)

gives a detailed account of the relative abundance of each phylum along with the results of statistical tests comparing different groups, provided in the supplementary material.

Genus level abundance

Considering the genus distribution, among the three samples, a range of microbes was present in the healthy group of buffaloes. *Clostridium* (36%) predominantly forms a large proportion of healthy microbiota, whereas *Acinetobacter* (33%) was less abundant as compared to the CM (76%) and SCM (52%) groups. *Acinetobacter* was found in the clinical mastitis cases. Some genera were specific to only healthy groups, for instance, *Deinococcus*, *Bacillus*, *Enterococcus*, *Streptococcus*, *Escherichia*, *Rhodobacter,* and others; however, they were less abundant (Fig. 2b). Abundance-based and significance heatmap at genus level (Fig. 3b and S2b) highlighted that *Exiguobacterium* exhibited greater abundance in the clinical group than in the subclinical and healthy groups. It was shown to be statistically significant when compared to the subclinical group (FDR = 7.19×10^{-9}) and the healthy group

(FDR = 2.5613×10^{-4}). For healthy animals, the bacterial genera that were relatively abundant and statistically significant include *Deinococcus, Streptococcus, Bacillus, Enterococcus, Rhodobacter,* and *Lactobacillus*, among others, when assessed individually against both clinical and subclinical cases.

Species level distribution

A comparative analysis of bacterial species revealed that *Clostridium botulinum* was the most abundant species across all groups, accounting for 42.9% in healthy animals and 36.1% in subclinical cases. It was also present prominently in clinical mastitis cases, where it constituted 29% of the microbial profile. However, in clinical mastitis samples, *Acinetobacter johnsonii* emerged as the dominant species, representing 49.6%, in stark contrast to its lower prevalence in healthy (1.4%) and subclinical animals (0.7%). Subclinical mastitis cases were primarily characterized by *Acinetobacter sp. TTH0-4*, which accounted for 38.5% of the microbial community, along with *C. botulinum*. There was also an increase in the representation of *Lactococcus lactis* (3.1%) and *Pseudomonas lundensis* (2.16%). In healthy animals, the microbial diversity was the highest, with no single species dominating. While *C. botulinum* and *Acinetobacter sp. TTH0-4* contributed notably; there were also unique taxa such as *Deinococcus radiodurans* (5.7%), *Enterococcus faecalis* (3.7%), *Rhodobacter sphaeroides* (1.5%), and *Streptococcus mutans* (1.3%), contributing to a more balanced microbial composition (Fig. 2c). In clinical mastitis cases, the microbial community was skewed towards dominant taxa, with *A. johnsonii* and *C. botulinum* exceeding 78% of the total community. Subclinical mastitis showed moderate diversity, primarily featuring two major species: *Acinetobacter. sp. TTH0-4* and *C. botulinum*. This trend was further validated by a sample-wise heatmap at the species level (Fig. 3c). These findings underscore a decline in microbial diversity from healthy animals to those with clinical mastitis, highlighting the increasing dominance of opportunistic or pathogenic taxa.

Acinetobacter as a dominant genus

Acinetobacter was the most prominent and diverse bacterial genus across clinical, subclinical, and healthy groups. In clinical mastitis, *A. johnsonii* was the dominant species, accounting for nearly half of the microbial community (49.6%). This marked difference in relative abundance was significantly associated with clinical mastitis compared to healthy and subclinical groups (FDR = 0.0389). A similar trend was observed in the case of *Acinetobacter. sp. WCHA55* with FDR value of 0.0449 (Table 3). Other species, including *A. baumannii* (1.9%), *Acinetobacter sp. TTH0-4* (1.6%), and *A. guillouiae* (1.0%), contributed at lower levels, indicating a broad representation of *Acinetobacter* species in diseased samples. In subclinical mastitis, *Acinetobacter sp. TTH0-4* (38.5%) replaced *A. johnsonii* as the major species, while *A. baumannii* (1.6%), *A. haemolyticus* (0.7%), *A. schindleri* (0.7%), and *A. indicus* (0.6%) were detected, suggesting greater species richness under subclinical conditions. In healthy animals, *Acinetobacter sp. TTH0-4* (23.1%) persisted at a relatively high level, alongside low proportions of *A. johnsonii* (1.4%) and other species (Figure S1). This distribution was also represented in the form of a chord diagram for all three categories (Figure S3). These findings suggest that Acinetobacter is a diverse and adaptable genus, with *Acinetobacter johnsonii* standing out as the most clinically significant species in mastitis cases.

Phylum- and genus-level biomarkers

Linear Discriminant Analysis (LDA) scores identified specific bacterial taxa that serve as biomarkers, distinguishing and associating bacterial communities with the three health status groups. At the phylum level, Proteobacteria exhibited the highest LDA score, indicating a strong association with clinical mastitis cases. Similarly,

Table 3 The abundance of different strains of acinetobacter species

Species	Relative abundance frequency			FDR		
	Group I: clinical	Group II: Sub-clinical	Group III: healthy	SC vs C	H vs C	H vs SC
Acinetobacter_johnsonii	49.62	0.73	1.45	*	*	
Acinetobacter_sp._WCHA55	3.09	0.08	0.11	*	*	
Acinetobacter_oleivorans	0.004	0.009	0.006			
Acinetobacter_bereziniae	0.14	0.18	0.12			
Acinetobacter_dispersus	0.027	0.072	0.048			
Acinetobacter_chinensis	0.039	0.058	0.038			
Acinetobacter_nosocomialis	0.135	0.168	0.11			
Acinetobacter_lactucae	0.011	0.017	0.010			
Acinetobacter_sp._185	0.098	0.120	0.079			
Acinetobacter_sp._TGL-Y2	0.142	0.356	0.223			

*indicates significance at FDR < 0.05

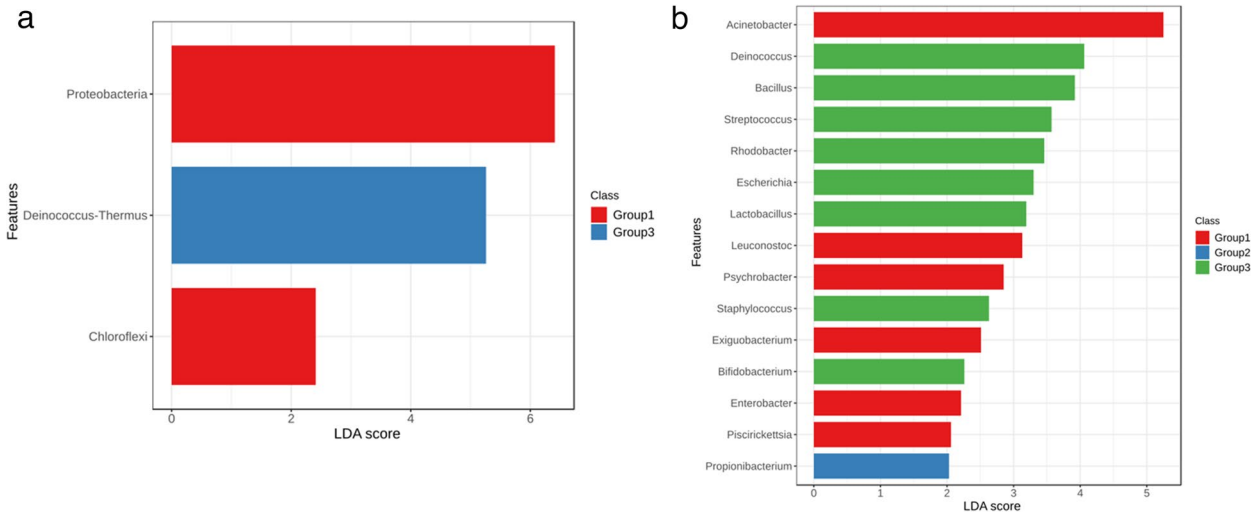

Fig. 4 Linear Discriminant Analysis (**a**) Phylum-level analysis (**b**) Genus-level analysis (where Group1: Clinical mastitis, Group2: Subclinical mastitis, and Group3: Healthy)

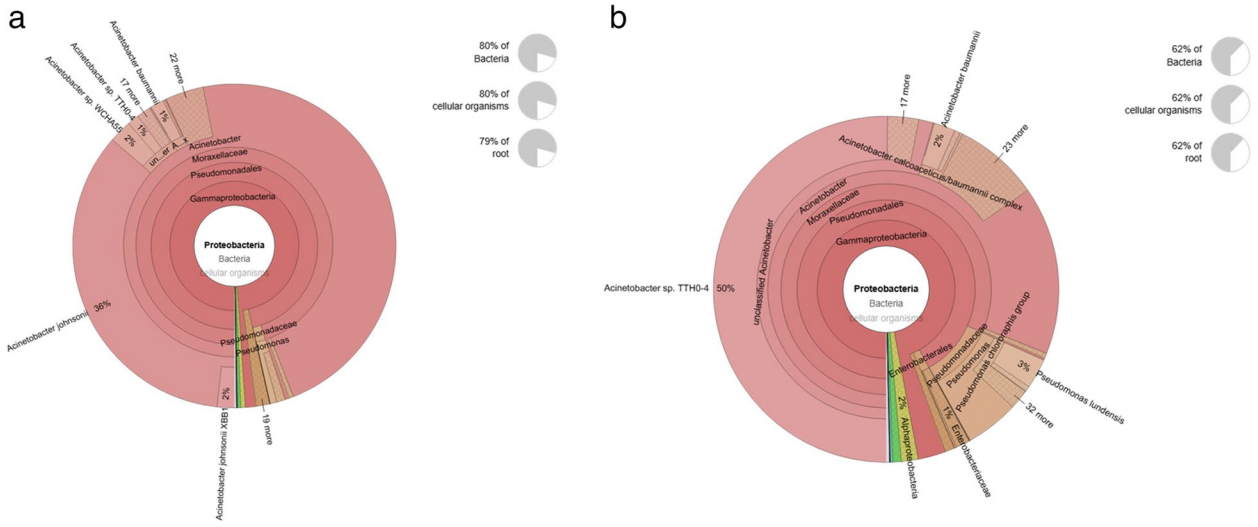

Fig. 5 Krona Chart illustrating taxonomic composition of *Proteobacteria* (**a**) clinical group and (**b**) in a subclinical mastitis group

the high LDA score of *Deinococcus-Thermus* underscores its strong association with healthy animals (Fig. 4a).

The relative abundance of the phylum *Deinococcus-Thermus* was significantly higher in healthy individuals compared to clinical mastitis cases (FDR = 7.27×10^{-10}), suggesting a commensal role of this taxon in maintaining udder health. Although the relative abundance of *Acinetobacter* was visibly higher in clinical mastitis cases compared to subclinical and healthy groups, the differences were not statistically significant (subclinical vs. clinical: $p > 0.05$; healthy vs. clinical: $p > 0.05$) (Fig. 4b). The relative abundance of *Propionibacterium* was higher in subclinical individuals compared to those with clinical and subclinical mastitis buffalo.(Supplementary Table S3).

The Krona plot provides a detailed hierarchical characterization of the microbial community composition, highlighting the taxonomic distribution and relative abundances of various taxa in clinical mastitis cases in Murrah buffalo. At the root level, the microbial community is dominated by the domain Bacteria, which is further divided into prominent phyla such as *Proteobacteria* and *Firmicutes*. Figure 5 illustrates the phylogenetic tree, which represents the taxonomic breakdown and proportion of the phylum *Proteobacteria* in both clinical and subclinical animal samples. The microbial profile of clinical mastitis samples was dominated by *Proteobacteria* (approximately 80%), with *Gammaproteobacteria* as the largest group. *Acinetobacter* comprised about 36% of the community, along with *Pseudomonas* and related genera (Fig. 5a). This prevalence of opportunistic pathogens

indicates a disrupted microbiome that likely contributes to inflammation and tissue damage in the udder. In contrast, subclinical mastitis samples showed a more diverse microbial composition, with *Proteobacteria* making up 62% of the community. *Acinetobacter* represented 50%, while *Pseudomonas* and *Moraxellaceae* were present in smaller amounts (~2% each) (Fig. 5b). This diversity suggests an earlier or less severe infection, where the host's immune defenses may still control microbial growth.

Group-wise comparison of specific taxa

The box and whisker plot (Fig. 6a) shows that the abundance of *Proteobacteria* is the highest in group 1, clinical mastitis cases, suggesting a link between *Proteobacteria* and more severe mastitis conditions. Whereas, log-transformed count explains the higher abundance of *Deinococcus-Thermus* in group 3, i.e., healthy, compared to CM and SCM (Fig. 6b). Groups 1 and 2 show near-zero counts, suggesting a depletion in mastitis cases (both clinical and subclinical). *Firmicutes* and *Chlorobi* revealed the opposite trend. *Chlorobi*'s highest abundance was observed in the CM group, as shown in the log-transformed counts (Figs. 6c and 5d), suggesting its possible correlation with the severity of clinical mastitis. The lower abundance of SCM and the markedly reduced levels in the healthy group linked its presence to disease progression. The highest abundance of *Firmicutes*

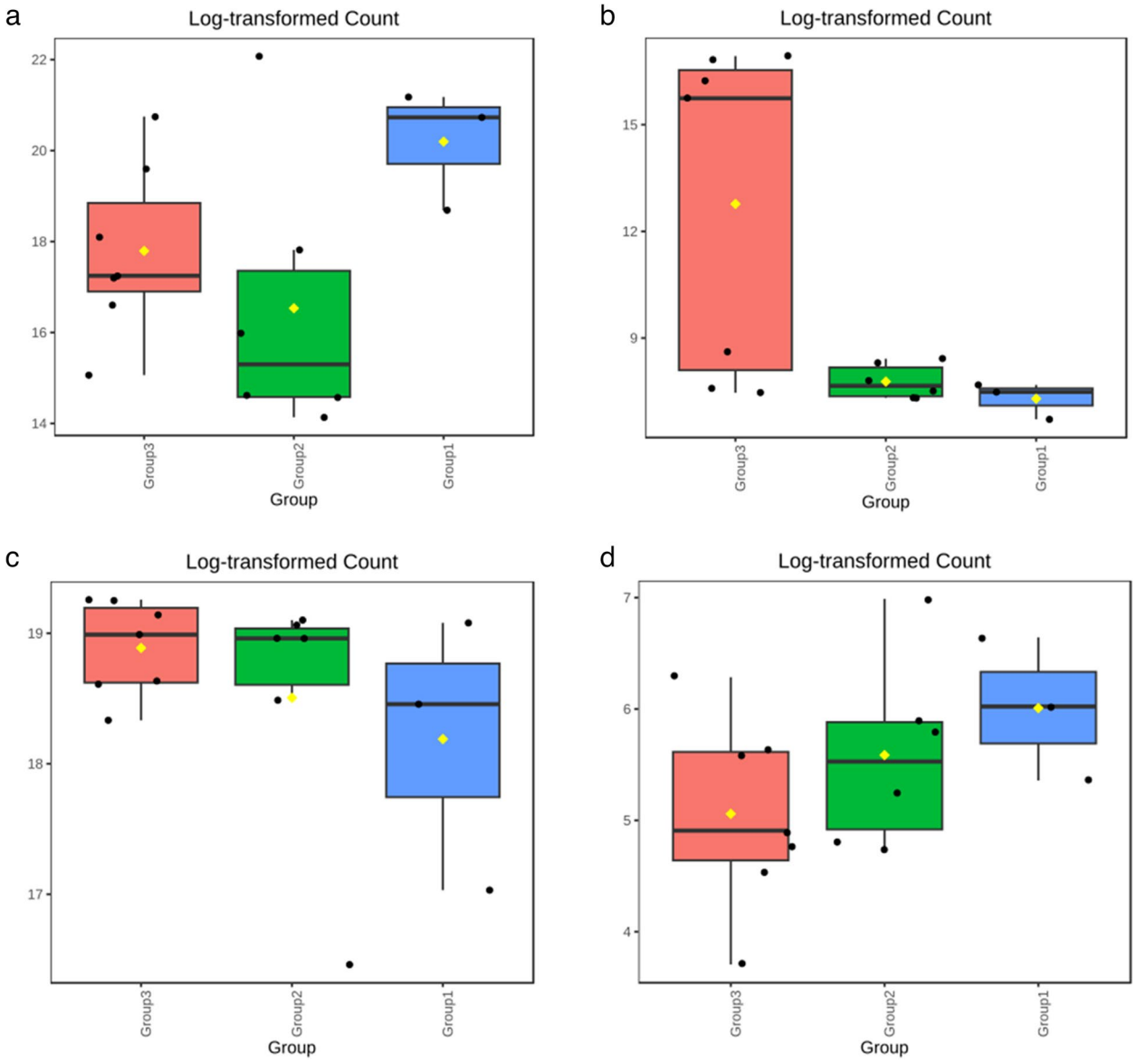

Fig. 6 Log-Transformed Abundance for different phyla: (**a**) *Proteobacteria*, (**b**) *Deinococcus-Thermus*, (**c**) *Firmicutes*, (**d**) *Chlorobi* (where Group1: Clinical mastitis, Group2: Subclinical mastitis, Group3: Healthy)

suggests a potential protective or stabilizing role in healthy animals.

Log-transformed counts of various genera showed a similar trend in the case of *Enterobacter*, *Exiguobacterium*, *Ferrimonas*, *Thalassotalea*, and *Virgibacillus*, as these genera were most abundant in Group 1, i.e., clinical mastitis. They represent microbiota that are directly associated with the severe mastitis condition and highlight their potential role as biomarkers for the disease. On the other hand, their abundance in subclinical cases varies. For instance, the levels of *Enterobacter* and *Ferrimonas* were lower in SCM than in healthy animals, whereas *Exiguobacterium* was nearly equal in both SCM and healthy animals. *Thalassotalea* and *Virgibacillus* were more prevalent in Group 2 (SCM) than in healthy ones (Fig. 7b, c, d, e and f). Whereas, a different pattern was observed for *Bifidobacterium*, where the healthy group (Group 3) showed higher log-transformed counts, pointing to its potential role in maintaining good health (Fig. 7a).

Shared microbial taxa among the groups

Among the total 45 detected phyla, 41 were shared by all three groups. Three phyla were shared by group 2 (Subclinical) and group 3 (Healthy), namely, *Candidatus_Gracilibacteria*, *Candidatus_Korarchaeota*, and *Candidatus_Micrarchaeota*. *Deinococcus-Thermus* was the unique phylum that was associated with the healthy group (Fig. 8a). This interpretation indicates its significant role in maintaining udder health and microbiome stability. Moreover, a large number of genera were reported in all three groups of animals. Of these, 1301 genera were shared by healthy, subclinical, and clinical groups. The clinical group consists of one distinctive genus, *Aliiarcobacter* (Fig. 8b).

Microbial diversity in mastitis and healthy samples

The dendrogram (Fig. 9) illustrates the hierarchical clustering of microbial profiles from animals categorized into three groups: clinical mastitis, subclinical mastitis, and healthy animals. The clustering at the genus level demonstrates distinct groupings based on microbial composition.

Clinical mastitis samples cluster tightly, merging at lower heights, which indicates high similarity within this group and suggests a shared microbial signature, representing pathogenic dominance and infection severity. Subclinical mastitis samples form separate clusters, with slightly greater variability, reflecting intermediate microbial states between health and disease. For instance, the distance matrix value between a subclinical case and a clinical case was relatively low at 0.241412, indicating

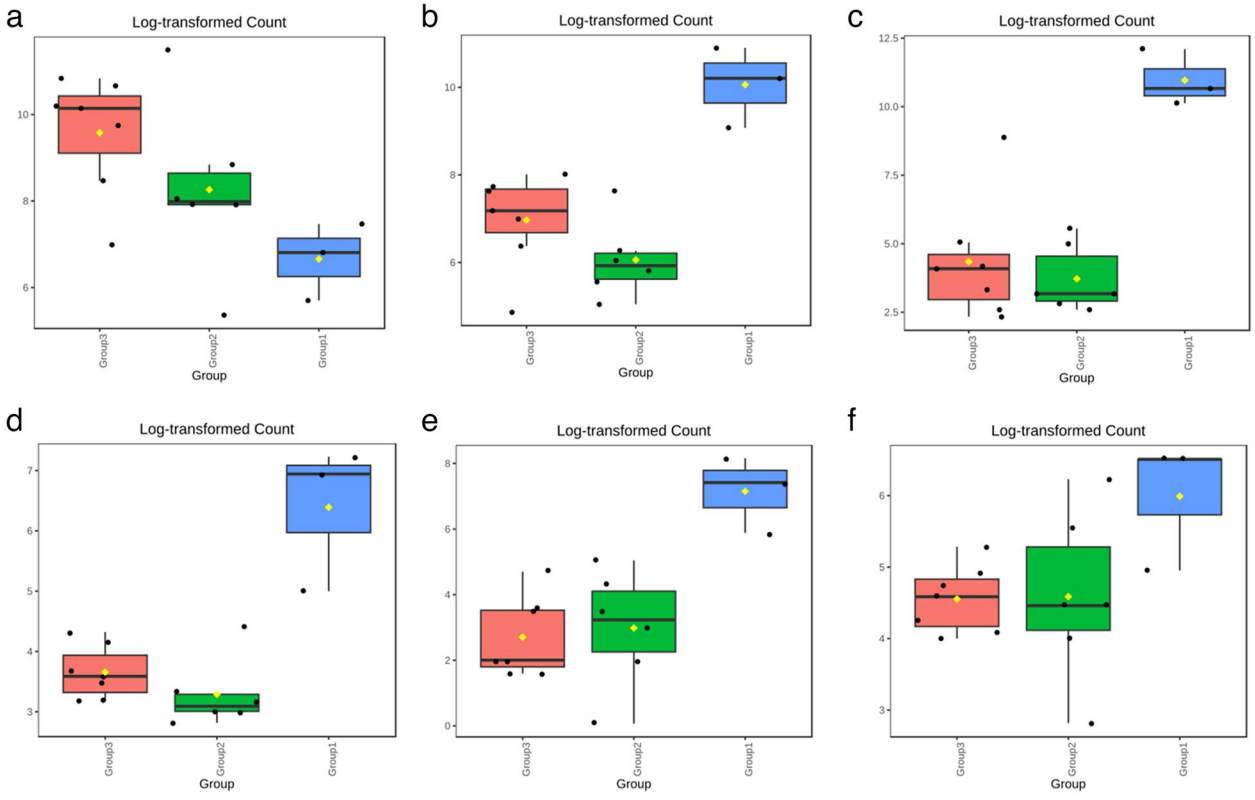

Fig. 7 Log-Transformed Abundance of different genera: (**a**) *Bifidobacterium*, (**b**) *Enterobacter*, (**c**) *Exiguobacterium*, (**d**) *Ferrimonas*, (**e**) *Thalassotalea*, (**f**) *Virgibacillus* (where Group1: Clinical mastitis, Group2: Subclinical mastitis, and Group3: Healthy)

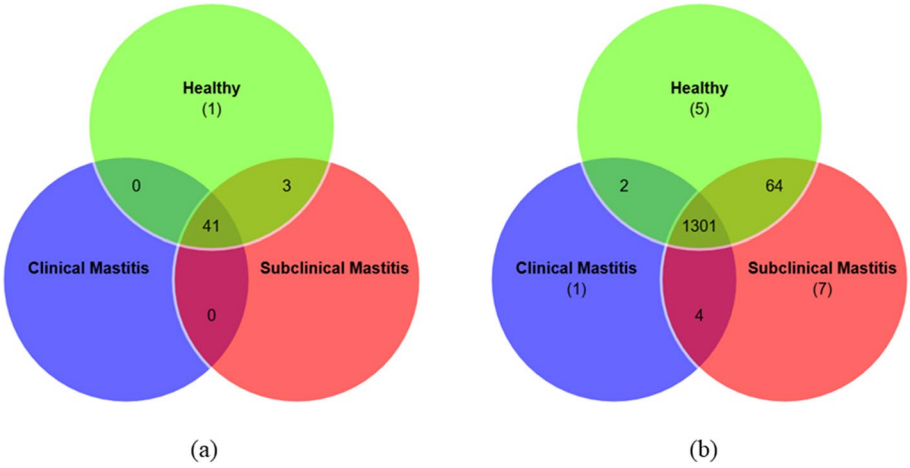

Fig. 8 Venn Diagram representing Shared and Unique microbial taxa at different levels (**a**) phylum, (**b**) genus, in the milk microbiome composition and udder status

a high level of microbial similarity. This similarity is reflected in their clustering together in the dendrogram, suggesting that the microbial communities of these two cases share certain similarities despite differing health statuses.

Healthy samples cluster separately and show greater diversity, reflecting a balanced microbial community. The clear separation between clinical, subclinical, and healthy groups highlights distinct microbial community structures associated with disease severity and health status.

OTU (Operational Taxonomic Unit) diversity

Alpha diversity metrics for healthy and diseased groups

Shannon and Simpson diversity indices were used to measure community diversity, which explains the richness (number) and evenness (distribution) of specific groups or taxa in a population. They provide complementary insights into microbial diversity.

Shannon diversity indices

At the phylum level, the healthy group exhibited the highest Shannon diversity, with a median index of approximately 1.0, indicating a stable and diverse microbial community. In contrast, clinical and subclinical mastitis groups showed reduced diversity, with a median index of 0.8 and 0.9, respectively, suggesting microbial dysbiosis in these conditions. A consistent trend was observed at the genus level, with the healthy group showing the highest diversity, clinical mastitis cases showing the lowest, and subclinical cases displaying intermediate diversity with notable variability.

Similarly, at the species level, the healthy group maintained the highest diversity (median index ~ 2.5), reflecting a balanced and resilient microbial ecosystem. The clinical group displayed a markedly lower diversity (median index ~ 2.0), likely dominated by a few pathogenic species. The subclinical group demonstrated high variability in species-level diversity, ranging from extremely low to relatively high values, indicating transitional states or heterogeneity in microbial composition (Fig. 10).

Simpson diversity indices

The Simpson diversity index plots at the genus, phylum, and species levels demonstrate the overall diversity trends across the three groups (Group 1, Group 2, and Group 3). At all taxonomic levels, Group 3 (healthy individuals) consistently exhibits the highest Simpson index values, indicating the greatest microbial evenness and diversity. Clinical mastitis cases show moderately high diversity with some variability, while subclinical mastitis cases consistently have the lowest diversity and most compact distributions across all taxonomic levels. These results highlight a clear decline in microbial diversity and evenness in the mastitis-affected groups, particularly in subclinical cases, suggesting that mastitis influences microbial composition and richness (Fig. 11).

Beta diversity metrics for healthy and diseased groups

Beta diversity examines variations in microbial composition across different samples or groups. Principal coordinate analysis (PCoA) utilizing the Bray–Curtis dissimilarity index assesses similarity among various populations through both quantitative and qualitative analysis of OTU occurrences. The two-dimensional PCoA plot portrays nearly 89.31% of the total variance at the genus level. The clinically affected samples were clustered separately with its centroid visibly distant from the centroids of the sub-clinical and healthy groups (Fig. 12). Statistical analysis using PERMANOVA (adonis2) supported these findings, indicating a significant effect of group on microbial community structure ($R^2 = 0.303$, $p = 0.049$),

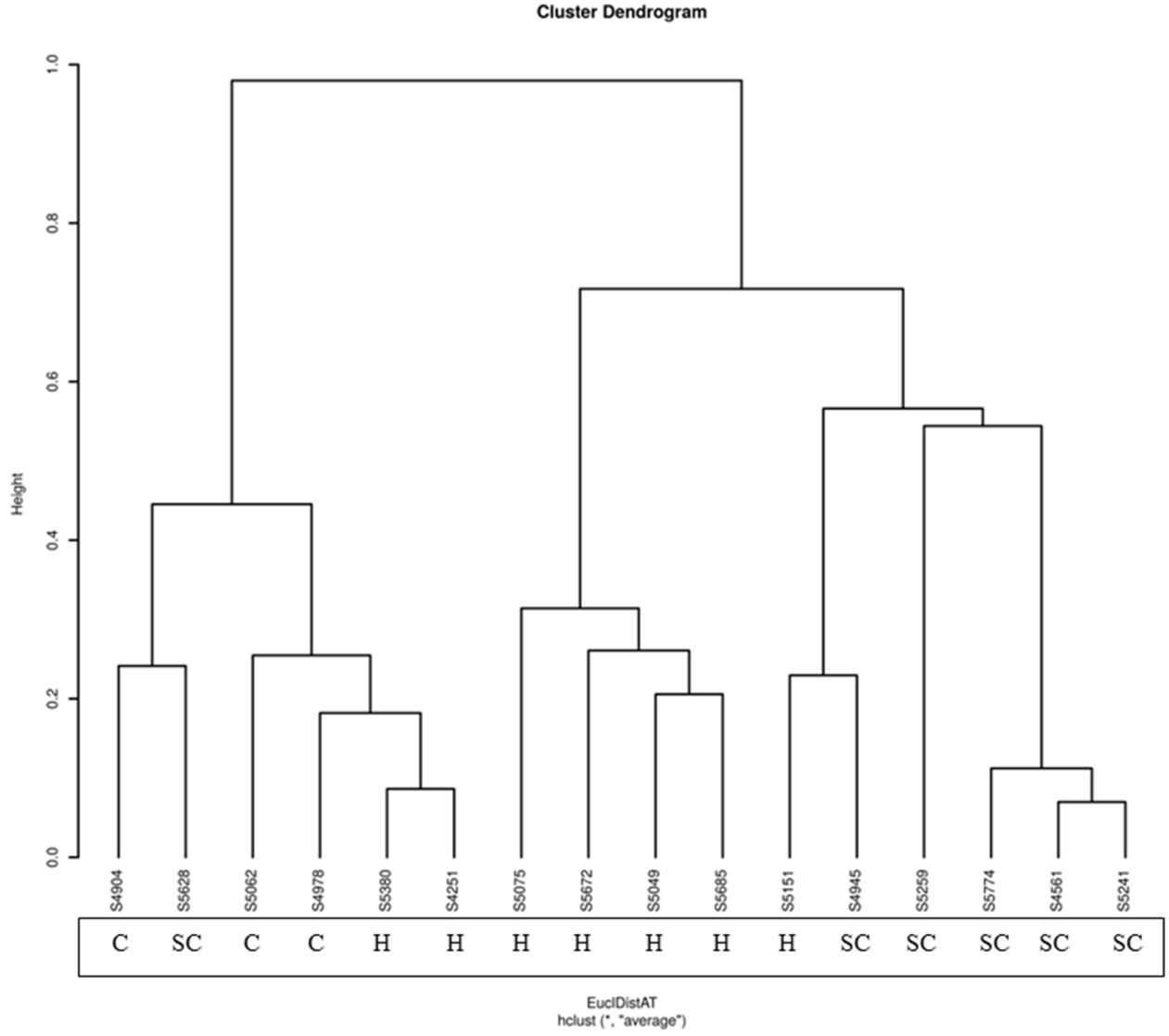

Fig. 9 Genus-level microbial clustering (where C represents Clinical mastitis samples, SC represents Subclinical mastitis samples, and H represents Healthy samples.)

accounting for about 30.3% of the overall beta diversity variation. The PCoA plot organizes samples based on total microbial abundance. At the genus level, analysis showed that the two healthy samples displayed high levels of *Acinetobacter* and *Clostridium*, which are often linked to clinical mastitis cases. Similarly, these healthy samples showed the significant abundance of *Acinetobacter* and *Clostridium*. The similarities in microbial profiles may explain the proximity of these two healthy samples to the clinical cluster in the PCoA, suggesting that these animals could be harboring subclinical infections or early-stage dysbiosis despite being classified as healthy.

Functional analysis

Among the total of 58,837 identified genes in the overall milk metagenome, there were 14,887 GO annotated genes that were assigned into the biological processes (8,751), cellular components (9,365), and molecular functions (12,523). Most prominent molecular functions were catalytic activity (15.2%), binding (12.1%), and transporter activity (2.9%). Metabolic (11.3%) and cellular processes (11.4%) were the prominent terms of the biological processes (Fig. 13a). Biosynthesis of amino acids (515 genes), carbohydrate metabolism (110 genes), lipid metabolism (158 genes), and energy-related processes such as the TCA cycle (76 genes), glycolysis (47 genes), and oxidative phosphorylation (16 genes) were evident for the pathway analysis. Notably, the cell wall biogenesis process associated with the peptidoglycan biosynthetic

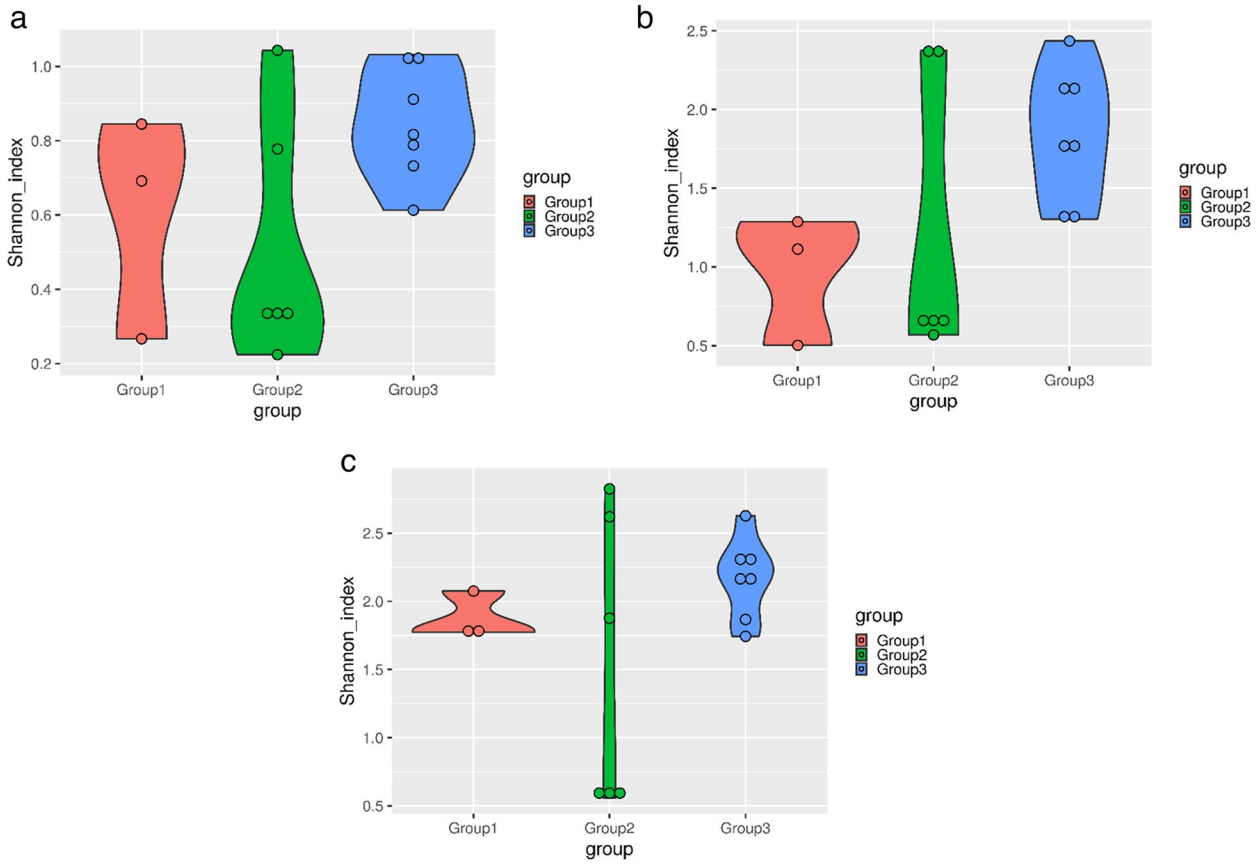

Fig. 10 Shannon plots representing alpha diversity indices at (**a**) phylum level, (**b**) genus level, (**c**) species level (where Group1: Clinical mastitis, Group2: Subclinical mastitis, and Group3: Healthy)

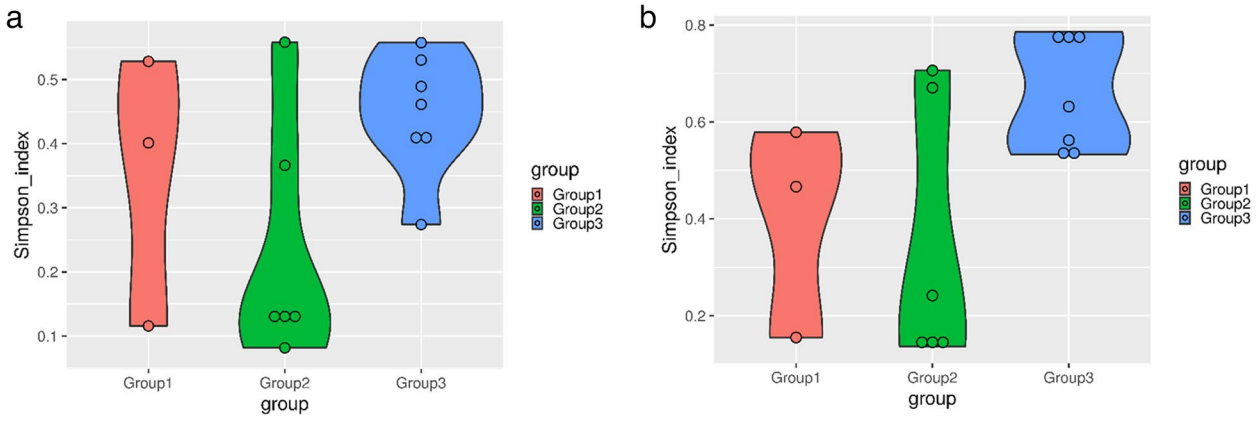

Fig. 11 Simpson plots representing alpha diversity indices at (**a**) phylum level, (**b**) genus level, (**c**) species level (where Group1: Clinical mastitis, Group2: Subclinical mastitis, and Group3: Healthy)

pathways (94 genes) and membrane lipid metabolism was prominent and indicative of their roles toward the structural integrity of the cell wall (Fig. 13b and c). In addition, cofactor biosynthetic, nucleotide metabolism, and carbohydrate degradation were prominent and indicative of an active microbial input toward the utilization of energy, cell restoration, and interaction with the host. Reads associated with *Deinococcus* demonstrated an enrichment of pathways related to the metabolism of galactose, xylulose, and glucose. Additionally, there was notable enrichment in the biosynthesis of coenzyme A and cofactors, as well as in sodium-proton and potassium-proton antiporters. Other enriched pathways included phosphoryl response regulators, DNA repair enzymes like helicases, and NADPH-dependent antioxidant systems (Figure S4).

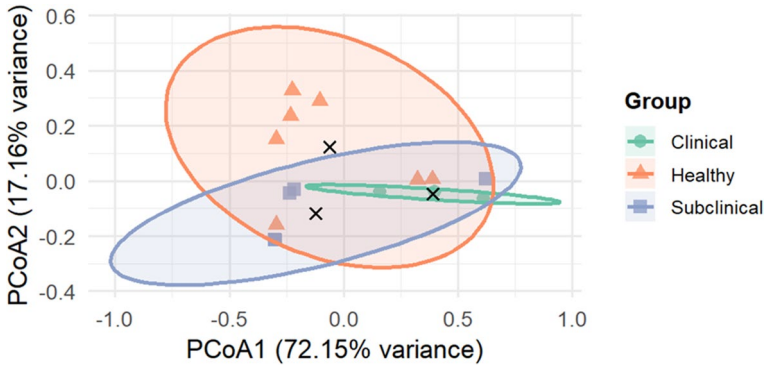

Fig. 12 PCoA plot using Bray–Curtis matrices representing beta diversity indices of microbial communities at the genus level

Discussion

The study of mastitis, a common inflammatory condition in dairy animals, is often hindered by the inability to culture certain bacterial pathogens using conventional methods. Only about 1% of bacteria are culturable, leaving the majority of microbial diversity unexplored [22]. This limitation leaves gaps in understanding the complexity of mastitis microbes. Shotgun metagenomic sequencing addresses this challenge by allowing the direct sequencing of DNA from milk samples, without the need for culturing. Several studies have utilized metagenomic approaches to investigate the microbiome of bovine milk and various buffalo breeds, linking it to udder health status. However, the association of microbial diversity and its practical application remain limited and underexplored in the Murrah buffalo breed.

The sequencing data obtained for the milk metagenome study were sufficient to capture the microbial profile and significant taxonomic and functional information associated with mastitis. In a study, Hillmann et al. [55] demonstrated that shallow sequencing provides nearly the same accuracy at both the species and functional levels as deep whole metagenome sequencing for known species and genus in five key aspects of microbiome analysis: beta and alpha diversity, species composition, functional composition, and clinical biomarker discovery. The same study also demonstrated that shallow sequencing cannot be used for strain-level characterization; however, in our research, several strains have been detected for numerous genera. For instance, approximately 20 strains of the key genus, *Acinetobacter*, have been spotted.

A small sample size poses challenges for fully validating disease-associated microbial changes and may pose a lower statistical power in many cases. The sampling strategy in the present study was completely dependent on the natural occurrence of clinical and sub-clinical mastitis cases in the buffaloes. As clinical mastitis has a lower rate of incidence in Buffaloes, only a few buffaloes exhibiting clinical symptoms of mastitis could only be included in the study. Stringent test statistics were applied to check for false positives and avoid the overestimation of abundances of some taxa due to low sample size. Furthermore, functional analysis was also conducted to support the findings, however, validation of the reported abundance taxa in relation to disease progression may be further explored in large samples to establish its true role in causing mastitis in Murrah buffaloes.

The microbiota present in milk is diverse, comprising both beneficial and harmful bacteria from the buffalo's udder environment. Some of these bacteria can lead to intramammary infections or mastitis in buffalo, making it crucial to identify the pathogens for effective treatment. In addition to traditional microbiological methods, metagenomics provides an in-silico approach to analyze the complete microbial community, covering even non-culturable species, which aids in pathogen detection and assessing udder health. Initial metagenomic research on mastitis used 16S rRNA gene pyrosequencing in cattle, exposing a complex variety of microbes linked with both subclinical and clinical mastitis [6, 32]. Jiménez et al. [33] used shotgun metagenomics in human mastitis cases to reveal a wider array of microbial and viral communities present in milk. One of the earliest uses of whole-genome sequencing (WGS) metagenomics in mastitis research was performed by Hoque et al. [14], who discovered significant changes in microbial communities and functional gene profiles in bovine milk.

The present study was mainly focused on the identification and comparison of microbiota present in clinical, subclinical, and healthy Murrah. The milk metagenome of Nilli Ravi has been explored using whole genome shotgun (WGS) sequencing [24], however, similar studies in Murrah buffaloes for microbial diversity have not been conducted regarding mastitis. Salman et al. [26] reported that *Proteobacteria* were the predominant phylum in the healthy group and *Firmicutes* in the subclinical and clinical groups. The same team also explored the microbiota of mastitic Sahiwal cows, and *Proteobacteria* were dominant in the healthy group and subclinical mastitis group (56.48% and 48.77%, respectively) as compared to the

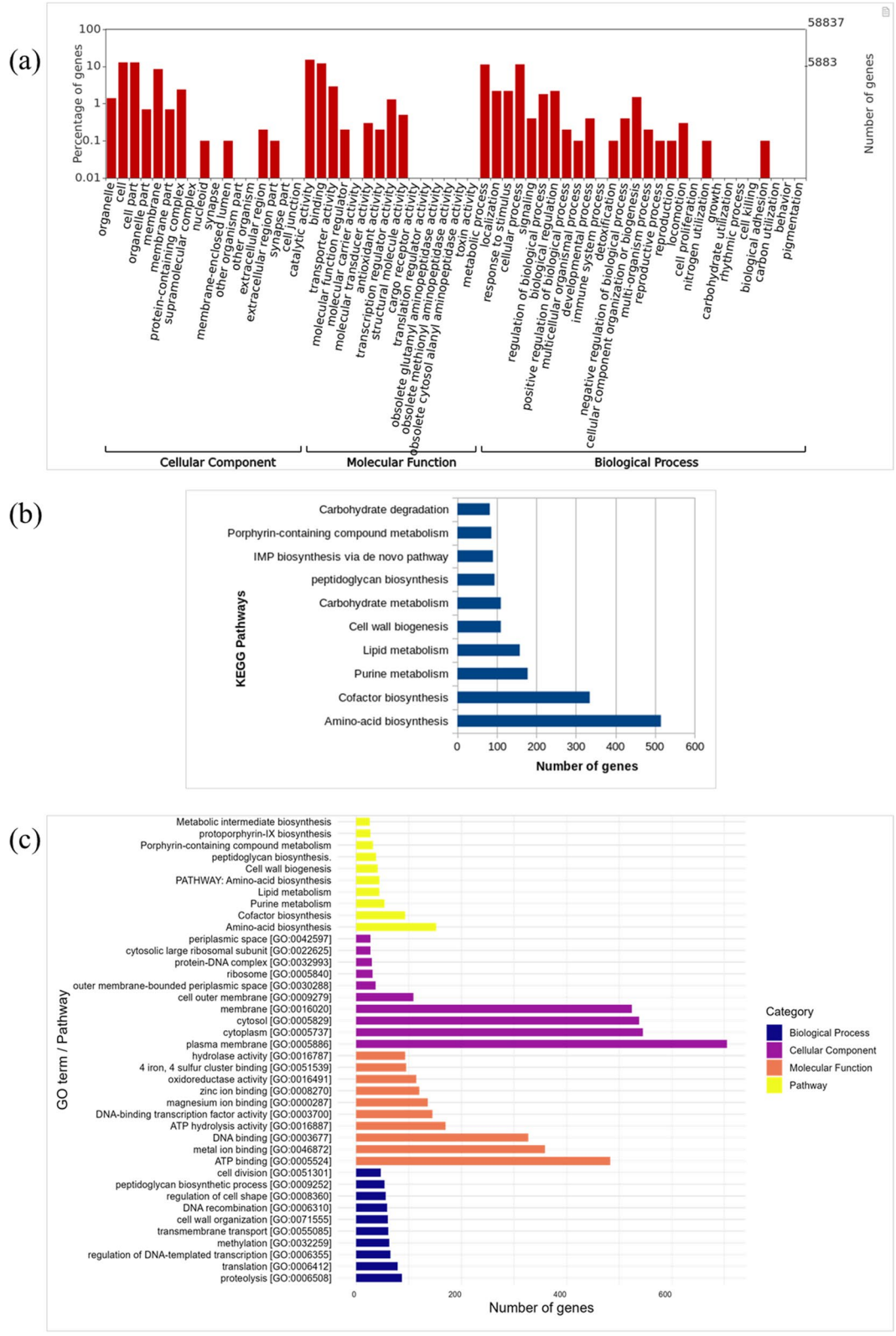

Fig. 13 Functional pathway analysis representing (**a**) Top gene ontology terms for biological process, molecular function, and cellular components enriched for the complete microbiome (**b**) KEGG pathways enriched for the complete microbiome (**c**) Top 10 GO terms for biological process, molecular function, cellular components, and pathways associated with *Acinetobacter spp*

clinical mastitis group (2.68%). In contrast, *Firmicutes* were abundant in the clinical mastitis group (64%) as compared to the healthy and subclinical mastitis groups (15.87% and 38.98%, respectively), however, our studies depicted that *Proteobacteria* were most abundant in clinical cases and *Firmicutes* in the healthy group. Patel et al. [23] conducted a comparative analysis of the metagenome utilizing 16S rRNA techniques in clinical and subclinical cases of Surti and Mehsana buffaloes. Their findings indicated that *Firmicutes* (57%) was the predominant phylum, followed by *Proteobacteria* (16%). At the family level, they observed lower abundances of *Moraxellaceae*, whereas our study demonstrated an opposite trend, with *Moraxellaceae* becoming the most prominent bacterial family in cases of clinical mastitis (76.7%). In their research, *Staphylococcaceae* (23%) and *Enterococcaceae* (27%) dominated the subclinical samples, while *Streptococcaceae* (31%) was more prevalent in clinical mastitis. Conversely, our subclinical mastitis cohort was predominantly characterized by *Moraxellaceae* (52.4%) and *Clostridiaceae* (30.2%), with *Streptococcaceae* (3.1%) and *Enterococcaceae* (0.11%) detected at significantly lower levels compared to their research. These differences suggest breed-specific and methodological factors influence microbial dominance, with *Moraxellaceae* key in Murrah mastitis and *Firmicutes* in Surti and Mehsana buffaloes. A similar study was done on cow breeds from different regions and revealed that in the taxon-phylum, bacteria are dominated by *Actinobacteria* (36.38%), and the dominant genus was *Corynebacterium_1* (20.53%) in mastitis cases [17]. According to metagenomic research on the microbiota of water buffalo milk during sub-clinical and clinical mastitis, Catozzi et al. [10] correlated two major genera with subclinical mastitis, i.e., *Acinetobacter* and *Pseudomonas,* whereas no genera were associated with the clinical cases. Braem et al. [56] used a culture-independent technique to analyze the teat apex microbiota of dairy cows using Denaturing Gradient Gel Electrophoresis (DGGE), targeting the V3 region of the 16S rRNA gene. They found that *Acinetobacter* was the most prevalent genus, colonizing 52% of all teat apices. This aligns with our findings, where *Acinetobacter* was also an abundant genus in both clinical and subclinical mastitis cases.

Our findings are consistent with prior research focused on the metagenomic analysis of bovine milk. Hoque et al. [14] examined the correlation between microbiome signatures and functional biases in bovine mastitis, identifying *Proteobacteria* as the most prevalent phylum in both clinical and subclinical cases, with *Acinetobacter* (60.14%) being the dominant genus. Similarly, our deep sequencing results revealed a comparable trend, showing the highest abundance of *Acinetobacter* in the clinical mastitis (CM) group (76%) and the subclinical mastitis (SCM) group (52%). Our metagenomic analysis of Murrah milk microbiota across healthy, subclinical mastitis (SCM), and clinical mastitis (CM) groups shows distinct microbial patterns that confirm and extend the findings from Jiang et al. [15] 16S rDNA study of the same breed's milk. Both studies indicate that microbial diversity is largely stable between healthy and mastitis-affected samples. At the phylum level, *Proteobacteria*, *Firmicutes*, and *Actinobacteria* are dominant in both studies. Clinical mastitis samples display a significant increase in *Proteobacteria* (~ 80%), while healthy milk shows a higher proportion of *Firmicutes* (47.02%) during our research. Moreover, *Acinetobacter* species play different roles in disease progression at the genus and species levels. Our results indicate that many *Acinetobacter* species are significantly more abundant in the subclinical mastitis group, suggesting their early involvement in the development of mastitis. Conversely, *Acinetobacter johnsonii* is mainly associated with the clinical mastitis group, accounting for nearly half (49.6%) of the microbial community in CM samples, but only small fractions in SCM (0.73%) and healthy groups (1.45%). This trend aligns with earlier research on bovines. For instance, Hoque et al. [34] identified *A. johnsonii XBB1* as the most common strain in clinical mastitis milk (59.92%), compared to a significantly lower presence in healthy milk (36.2%) and nearly absent in SCM and recurrent mastitis (< 0.5%). This indicates that *A. johnsonii* could be a key pathogen in transitioning to clinical mastitis. While they identified *Acinetobacter* in high-SCC bovine milk, they did not differentiate between species-level differences or explicitly distinguish subclinical from clinical cases, underscoring the superior resolution provided by metagenomics in our research. *Acinetobacter* was isolated with a frequency also considered high of 12.5%, in the study performed by Andrea Vásquez-García et al. [57], which mainly focused on the isolation and phenotypic characterization of the main microorganisms that cause subclinical mastitis in *Bubalus bubalis*. Comparable trends were identified in cows with clinical mastitis samples in Korea [19]. The association of this bacterium with subclinical mastitis in buffaloes presents a hurdle for the treatment and control of the disease.

The findings of this study underscored that the milk microbiota from healthy quarters exhibited notably greater OTU diversity compared to samples from clinical and subclinical mastitis cases. Shannon and Simpson's plots for alpha diversity revealed that the healthy group exhibited greater microbial diversity compared to clinical and subclinical cases. For example, the Shannon index value for one of the healthy samples was 1.03, and for the clinical sample was 0.26. These findings are consistent with Oikonomou et al. [20], who observed a reduced Shannon index in clinically affected quarters in cow's milk. Moreover, Falentin et al. [12] also reported that the

Shannon index of healthy (7.8) quarters was significantly higher compared to clinical mastitis (6.5) quarters during bovine teat microbiome analysis. Another study of milk microbiota in bovines by Hoque et al. [14] explained the same trend of the Shannon index.

In the course of our research, a beta diversity analysis was performed utilizing Bray–Curtis dissimilarity and was visualized through principal coordinate analysis (PCoA). It explained the clear progression from stability (healthy) to variability (subclinical) and then to pathogenic dominance (clinical). This analysis elucidated distinct clustering patterns, effectively distinguishing clinical mastitis (CM) samples from subclinical mastitis (SCM) and healthy samples. Such findings align with the study conducted by Urrutia-Angulo et al. [28], which indicated that the observed beta diversity differences between CM and other groups were predominantly influenced by the prominence of microbial dysbiosis in CM samples. Remarkably, our observation of overlapping clusters between SCM and healthy samples validates the results found by M.M. Salman et al. [25], suggesting that the microbiota of healthy and SCM animals might mirror transitional phases from infection to recovery, with shared bacterial taxa among the groups contributing to this overlap. Similarly, Salman et al. [26] recognized mixed diversity patterns between healthy and subclinical mastitis samples in Sahiwal cattle. In addition, Hoque et al. [14] demonstrated that the majority of CM and healthy samples formed distinct clusters with significant differences ($p = 0.001$), through metagenomic deep sequencing of bovine milk microbiota. Overall, these findings imply that while CM samples typically exhibit pronounced microbial shifts associated with dysbiosis, the microbial communities of healthy and SCM buffaloes may be influenced by factors such as subclinical infection, early-stage dysbiosis, or recovery dynamics, resulting in their closeness within the PCoA plot.

This report revealed that the species of *Proteobacteria*, *Firmicutes*, and *Actinobacteria* were shared among all three groups. Likewise, Salman et al. [26] also reported these phyla common to all groups. Genus *Acinetobacter* was detected in all three groups with the highest percentage abundance detected in the clinical mastitis group (76%) followed by the subclinical mastitis group (52%) and the least abundant in the healthy group (33%) in our study, however, the same metagenomic study by this team in Sahiwal cattle showed *Caviibacter* genus showed the significantly highest abundance in clinical mastitis group (31.80%) as compared to other and it was not detected in other groups.

Metagenomic analysis of mastitis can significantly enhance both diagnosis and treatment strategies. Our study compared the abundance of bacterial genera in three different groups: healthy, subclinical, and clinical mastitis cases in Murrah. It would help implement antimicrobial therapies for specific genera. Catozzi et al. (2019) examined the microbial diversity in the milk microbiota of water buffalo after the intramammary treatment with inactivated cultures of *Lactobacillus rhamnosus* of mammary gland quarters naturally affected by subclinical mastitis as compared to antibiotic therapy. After treatment with *Lactobacillus rhamnosus*, the milk microbiota showed a reduction in the dominance of *Firmicutes* (from 60.9% to 45.2%) and *Proteobacteria* (from 18.8% to 12.4%). The treatment promoted the growth of beneficial lactic acid bacteria (LAB), leading to a more balanced microbial composition in the udder.

The exploration of host immune responses and mastitis, metagenomic studies further reveal how commensal bacteria in the udder play a crucial role in reducing mastitis risk. Bacteria like non-Aureus *Staphylococci* (NAS) and *Corynebacterium* produce bacteriocins that inhibit harmful pathogens. Healthy udder microbiomes are diverse, which helps maintain balance and prevent infections. Dysbiosis is linked to higher mastitis cases, with clinical milk samples showing lower microbial diversity. Metagenomics helps identify key bacterial species that influence inflammation and may protect against mastitis, providing a way for treatment [8].

Metagenomics offers significant insights into microbial compositions and the specific functions of microbes under various conditions. In mastitis cases, metagenomic studies illustrate changes in diversity that can directly affect disease progression and interactions between hosts and pathogens. Jiang et al. [15] identified *Escherichia coli*, *Staphylococcus aureus*, and *Klebsiella pneumoniae* as the main pathogens in high SCC animals, which were found through both microbiological and metagenomic analyses. Catozzi et al. [10] recognized 15 dominant genera in buffalo milk, with several genera also noted in our study as prevalent in healthy samples, such as *Staphylococcus*, *Clostridium*, and *Aerococcus*. These genera were identified as part of the core microbiota in healthy water buffalo milk samples by Catozzi et al. [10]. They were consistently found in healthy animals, but their abundance decreased progressively in cases of subclinical and clinical mastitis. Our results support this trend, indicating that *Staphylococcus*, *Clostridium*, and *Aerococcus* may have similar function in preserving udder health, and that their disruption might contribute to the progression of mastitis. Additionally, metagenomic approaches enable the detection of bacteriophages linked to particular bacteria. While investigating the milk microbiome signatures in subclinical mastitis in cattle, Bhatt et al. [6] identified specific bacteriophages associated with *Escherichia coli*, *Staphylococcus aureus*, *Enterobacter*, and *Yersinia*, which contribute to the cattle's natural resistance. This research indicates that bacteriophages could

be utilized for targeted therapies against specific bacteria responsible for mastitis. One strategy implemented is the production of antimicrobial substances. Typically, *Staphylococcus aureus* is a primary bacterium causing mastitis,however, our findings revealed it was linked to healthier buffaloes. This unusual pattern may relate to the impact of certain antimicrobial compounds on the Staphylococcus population during mastitis. Machan et al. [18] reported that *Pseudomonas aeruginosa* synthesizes antimicrobial agents like pseudomonic acid that inhibit *Staphylococcus aureus* growth, explaining its prevalence over *Staphylococcus* in the milk of Kankrej and Gir cattle. Some previous insights also found *Pseudomonas aeruginosa* to be associated with subclinical cases, for instance, Banerjee et al. [3] screened 422 milk samples from bovines. They identified *P. aeruginosa* in 5.4% of subclinical mastitis cases. This aligns with our findings, where *Pseudomonas aeruginosa* was more abundant than *Staphylococcus aureus* in subclinical cases, suggesting that the antimicrobial properties of *Pseudomonas aeruginosa* could be responsible for the reduced prevalence of major mastitis-causing bacteria in buffaloes as well. Additionally, the ability of some non-aureus staphylococci (NAS) species, such as *S. chromogenes*, to produce a wide range of bacteriocins capable of inhibiting the growth of major mastitis pathogens. For instance, Braem et al. [7] and Carson et al. [9] exemplify mechanisms by which commensal microbiota may contribute to the modulation of mastitis susceptibility. This aligns with our finding of elevated *S. chromogenes* levels in subclinical cases, suggesting it may serve a protective role by limiting the proliferation of more virulent bacteria. The apparent contradiction of *S. aureus* being linked to healthier buffaloes in our study may thus reflect a homeostasis where bacteriocin-producing NAS species or antimicrobial-producing microbes like *Pseudomonas* limit pathogen growth and development. Gill et al. [13] previously discussed using Staphylococcus phages to treat subclinical mastitis dominated by *S. aureus*, reporting a bacteriological cure rate of 16.7%. These complex microbial processes highlight the need for further investigation into the causal and commensal interactions between mastitis-associated pathogens and commensals within the buffalo mammary system.

Our research reported the phylum *Deinococcus–Thermus* in buffalo milk, with significantly higher levels in healthy animals (~5%) compared to those with subclinical and clinical conditions. Prior findings have also reported the presence of this phylum in milk. For instance, Luo et al. [36] found *Thermus* (3.16%) among the core phyla in raw milk from various animal species, including cow, sheep, goat, donkey, horse, camel, and yak in Xinjiang. Moreover, Zhao et al. [35] identified *Deinococcus–Thermus* with a relative abundance of 14.25% in fresh camel milk from Inner Mongolia. *Deinococcus–Thermus* may appear as a common component of the raw milk microbiome across various host species although its role as an environmental contaminant can't be negated. Salter et al. [37] identified *Deinococcus* as part of the phylum *Deinococcus–Thermus*, and its presence in negative (blank) controls suggests it may be a potential contaminant from reagents. Furthermore, Dean et al. [38] highlighted that contaminants from DNA extraction kits, teat skin, and sampling methods accounted for up to 75% of the sequence reads in bovine milk samples. The non-random distribution of *Deinococcus* in the present study therefore may be considered as a differentially abundant taxa in healthy animals rather than a uniform contaminant across all samples. Few *Chordata* sequences were detected at a very low relative abundance of less than 0.02% across all groups, indicating the host DNA contamination, a common technical interference in metagenomics study, and do not represent the actual microbial community. Hoque et al. [14] investigated the metagenome in bovine mastitis and revealed functional pathways associated with bacterial colonization, chemotaxis, invasion, and immune modulation, emphasizing the role of microbial metabolism in disease progression. Such analysis generates differences between CM and H cases,genes related to specific pathways were upregulated in clinical cases as compared to those of healthy individuals. For instance, they found similar metabolic trends in the milk microbiome during bovine mastitis, with overexpression of *gltA* and *fumA* linked to increased energy requirement and altered host–pathogen interactions. The higher abundance of *motB*, involved in chemotaxis and membrane structure, indicates bacterial strategies for colonization and persistence. Our study's detection of *fumA*, *fumC*, *gltA*, and *motB* highlights the enrichment of metabolic pathways such as carbohydrate metabolism, the TCA cycle, and bacterial chemotaxis.

Functional analysis of *Acinetobacter*, a gram-negative bacterium, highlighted the peptidoglycan biosynthesis pathway as a key biological process. Peptidoglycan, a crucial part of the bacterial cell wall, provides structural strength, preserves cell integrity, and shields the cell from external stress. Pavelka et al. [40] offered detailed insights into the genes and enzymatic steps involved in bacterial cell wall formation, focusing especially on peptidoglycan synthesis. Our study identified the same genes, *mraY, ftsI, glmU, uppS, murA, murD, murF, murE, nagZ,* and *LysM*, which are essential to this pathway and emphasize their role in maintaining bacterial integrity under conditions like host immune response. Host transmembrane proteins sense bacterial components in the cytosol, with PRRs, TLRs, and NLRs serving as core sensors. NOD1 recognizes iE-DAP, a dipeptide found mainly in Gram-negative bacterial peptidoglycan, while NOD2 detects

MDP, present in all bacterial peptidoglycans [41, 42]. NOD1 activation triggers the NF-κB pathway, resulting in neutrophil recruitment, cytokine release, and bacterial clearance during intramammary infections in cows [44]. Wang et al. [43] corroborate this by showing that peptidoglycan fragments released during bacterial growth activate innate immune responses via TLR2 and NOD2 in mammary tissue, leading to increased pro-inflammatory cytokines such as TNF-α, IL-1β, and IL-6. This suggests that heightened peptidoglycan synthesis may significantly contribute to inflammation in mastitis.. On the other hand, some species of *Deinococcus*, a non-pathogenic bacterium, i.e., *Deinococcus radiodurans*, first described as a radiation-resistant environmental bacterium, is not known to be pathogenic to humans or animals [52, 53]. The higher abundance of *Deinococcus radiodurans* (nearly 6%) in healthy milk compared to mastitis samples suggests it is better adapted to the non-inflamed mammary environment and may act as a commensal organism rather than a pathogen. Our functional analysis shows enrichment in pathways related to carbohydrate metabolism, coenzyme and cofactor biosynthesis, membrane transporters, DNA repair enzymes, and antioxidant pathways, indicating resilience and metabolic independence instead of host invasion or toxin production. This reinforces the non-pathogenic role of *Deinococcus* in the mammary microbiome.

A comparative assessment of milk microbiome studies shows that environmental factors, feeding practices, and sampling methods significantly affect outcomes. Bhatt et al. [6] noted differences in microbiome composition between indigenous and crossbred cattle in Gujarat, highlighting the impact of breed and management. Hoque et al. [14] examined mastitis in various Bangladeshi districts, where breed and farm practices likely caused variability. Conversely, Urrutia-Angulo et al. [28] focused on a large Spanish Holstein–Friesian herd under strict management, differing from smaller Asian farms. Salman et al. [24] controlled management bias in Nili Ravi buffalo farms in Pakistan, although environmental heat stress was still a factor. In the current study at ICAR-CIRB, Hisar, India, post-monsoon conditions favorable for microbial activity were examined in Murrah buffaloes, following strict hygiene during sampling. This study used aseptically collected milk samples per animal and classified buffaloes into healthy, subclinical, and clinical mastitis groups. As noted by Oikonomou et al. [39], host factors, farm management, and sampling designs are crucial in determining microbiome profiles, explaining inconsistencies in milk microbiome research across studies. Our study provides important information on the microbial population and composition in the milk samples from the mastitis buffaloes. However, the actual clinical role of the differentially abundant taxa such as *Acinetobacter* may be further studied and validated with a larger sample size to implicate its causality for clinical mastitis in Murrah buffaloes. This information can be further used to devise farm and species-specific therapeutic management. Furthermore, the metagenomics study in future can be integrated with transcriptomics and proteomics, which would help study the interaction with the host.

Conclusion

The present study provides important insight into the milk microbiota of the Murrah buffaloes in our herd. The study revealed the differences and comparison among the microbial profiles of healthy, clinical mastitis, and subclinical mastitis cases by using a culture-independent metagenomics approach. The *Acinetobacter spp* that was significantly abundant in the clinically affected buffaloes may have a causal role in the disease progression in our herd which may be further validated with more clinical cases in the buffaloes reared in similar environmental and managemental conditions. The present study provides an preliminary account of microbial diversity in udder health, probable causative organisms for clinical and sub-clinical mastitis promoting the development of targeted therapies to restore balance and improve treatment outcomes.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12866-025-04458-7.

Supplementary Material 1.

Acknowledgements

We sincerely express our gratitude to the Director, ICAR-Central Institute for Research on Buffaloes, Hisar, and Network Project on Buffalo Improvement for the invaluable support and for providing the necessary infrastructure to facilitate this research.

Authors' contributions

S.C. and S.K. contributed to the conception and design of the study. D.S., P.C., and H.V. collected the samples and performed the experiments. D.S. and H.V. analyzed the data and prepared the draft of the manuscript. S.C. and S.K. revised the manuscript. All authors read and approved the final manuscript.

Funding

The authors received no additional funding to carry out the present work.

Data availability

The associated data is available at Sequence Read Archive (SRA) under BioProject titled: Milk microbiome study on clinical and sub clinical mastitis in Murrah buffaloes with accession no. PRJNA1227304.

Declarations

Ethics approval and consent to participate

The present work is a part of research activity approved by the institute ethical committee with approval no. IAEC/2025/12/05 dated 15.01.2025.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 17 June 2025 / Accepted: 6 October 2025
Published online: 31 October 2025

References

1. Akhtar S, Ali S. Monitoring of bovine health problems of small dairy herds in Islamabad Capital Territory: design, data, and disease frequencies. Trop Anim Health Prod. 1994;26(4):193–8. https://doi.org/10.1007/BF02240379.
2. Argaw, Amare. Review on epidemiology of clinical and subclinical mastitis on dairy cows. Food Science And Quality Management. 2016;52: ISSN 2224–6088 (Paper), ISSN 2225–0557. https://www.iiste.org/Journals/index.php/FSQM/article/view/31012.
3. Banerjee S, Batabyal K, Joardar SN, Isore DP, Dey S, Samanta I, et al. Detection and characterization of pathogenic *Pseudomonas aeruginosa* from bovine subclinical mastitis in West Bengal, India. Vet World. 2017;10(7):738–42. https://doi.org/10.14202/vetworld.2017.738-742.
4. Bansal BK, Gupta D. Economic analysis of bovine mastitis in India and Punjab: a review. Indian J Dairy Sci. 2009;62:337–45.
5. Bardhan D. Estimates of economic losses due to clinical mastitis in organized dairy farms. Indian J Dairy Sci. 2013;66(2):168–72.
6. Bhatt VD, Ahir VB, Koringa PG, Jakhesara SJ, Rank DN, Nauriyal DS, et al. Milk microbiome signatures of subclinical mastitis-affected cattle analysed by shotgun sequencing. J Appl Microbiol. 2012;112(4):639–50. https://doi.org/10.1111/j.1365-2672.2012.05244.x.
7. Braem G, Stijlemans B, Van Haken W, De Vliegher S, De Vuyst L, Leroy F. Antibacterial activities of coagulase-negative staphylococci from bovine teat apex skin and their inhibitory effect on mastitis-related pathogens. J Appl Microbiol. 2014;116(5):1084–93. https://doi.org/10.1111/jam.12447.
8. Bronzo V, et al. The role of innate immune response and microbiome in resilience of dairy cattle to disease: the mastitis model. Animals. 2020;10(8):1397. https://doi.org/10.3390/ani10081397.
9. Carson DA, Barkema HW, Naushad S, De Buck J. Bacteriocins of non-aureus staphylococci isolated from bovine milk. Appl Environ Microbiol. 2017;83(17):e01015-e1017. https://doi.org/10.1128/AEM.01015-17.
10. Catozzi C, et al. The microbiota of water buffalo milk during mastitis. PLoS One. 2017;12(9):e0184710. https://doi.org/10.1371/journal.pone.0184710.
11. Das D, Panda SK, Jena B, Sahoo AK. Economic impact of subclinical and clinical mastitis in Odisha, India. Int J Curr Microbiol Appl Sci. 2018;7(3):3651–4. https://doi.org/10.20546/ijcmas.2018.703.422.
12. Falentin H, et al. Bovine teat microbiome analysis revealed reduced alpha diversity and significant changes in taxonomic profiles in quarters with a history of mastitis. Front Microbiol. 2016;7:480. https://doi.org/10.3389/fmicb.2016.00480.
13. Gill JJ, Pacan JC, Carson ME, Leslie KE, Griffiths MW, Sabour PM. Efficacy and pharmacokinetics of bacteriophage therapy in treatment of subclinical *Staphylococcus aureus* mastitis in lactating dairy cattle. Antimicrob Agents Chemother. 2006;50(9):2912–8. https://doi.org/10.1128/AAC.01630-05.
14. Hoque MN, et al. Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis. Sci Rep. 2019;9(1):13536. https://doi.org/10.1038/s41598-019-49468-4.
15. Jiang H, et al. Revealing microbial diversity in Buffalo milk with high somatic cell counts: implications for mastitis diagnosis and treatment. Vet Res Commun. 2024;48(4):2537–53. https://doi.org/10.1007/s11259-024-10438-5.
16. Kebede N, Tilahun A. Review on dairy cow mastitis and its economic impact. Int J Adv Res Biol Sci. 2023;10(2). College of Veterinary Medicine and Animal Sciences, University of Gondar, Gondar, Ethiopia. https://doi.org/10.22192/ijarbs.2023.10.02.XXX.
17. Kusumawati A, Mustopa AZ, Wibawan IWT, Setiyono A, Sudarwanto MB. Metagenomic analysis of pathogen mastitis in cow's milk from Cicurug, Sukabumi, West Java, Indonesia. In IOP Conference Series: Earth and Environmental Science. 2021;762(1):012064. IOP Publishing.
18. Machan ZA, Pitt TL, White W, Watson D, Taylor GW, Cole PJ, et al. Interaction between *Pseudomonas aeruginosa* and *Staphylococcus aureus*: description of an anti-staphylococcal substance. J Med Microbiol. 1991;34(4):213–7. https://doi.org/10.1099/00222615-34-4-213.
19. Nam HM, et al. Prevalence and antimicrobial susceptibility of gram-negative bacteria isolated from bovine mastitis between 2003 and 2008 in Korea. J Dairy Sci. 2009;92(5):2020–6. https://doi.org/10.3168/jds.2008-1739.
20. Oikonomou G, et al. Microbiota of cow's milk; distinguishing healthy, sub-clinically and clinically diseased quarters. PloS One. 2014;9(1):e85904. https://doi.org/10.1371/journal.pone.0085904.
21. Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a Web browser. BMC Bioinformatics. 2011;12(1):385.
22. Orsvik V, Øvreås L. Microbial diversity and function in soil: from genes to ecosystems. Curr Opin Microbiol. 2002;5(3):240–5. https://doi.org/10.1016/s1369-5274(02)00324-7.
23. Patel R, et al. Metagenomic approach to study the bacterial community in clinical and subclinical mastitis in Buffalo. Meta Gene. 2017;12:4–12. https://doi.org/10.1016/j.mgene.2016.12.014.
24. Salman, MM et al. Exploring the milk microbiota of healthy and mastitic Nili Ravi Buffalo using 16S rRNA Gene Base Metagenomic Analysis. Animals : an open access journal from MDPI . 2023b;13(14):2298. https://doi.org/10.3390/ani13142298.
25. Salman MM, Nawaz M, Yaqub T, Mushtaq MH. Milk microbiota of Holstein Friesian cattle reared in Lahore: association with mastitis. Saudi J Biol Sci. 2024;31(6):103984. https://doi.org/10.1016/j.sjbs.2024.103984.
26. Salman MM, et al. Investigation of milk microbiota of healthy and mastitic sahiwal cattle. BMC Microbiol. 2023a:304. https://doi.org/10.1186/s12866-023-03051-0.
27. Singh D, et al. Economic losses due to important diseases of bovines in Central India. Vet World. 2014;7:579–85. https://doi.org/10.14202/vetworld.2014.579-585.
28. Urrutia-Angulo L, Ocejo M, Oporto B, Aduriz G, Lavín JL, Hurtado A. Unravelling the complexity of bovine milk microbiome: insights into mastitis through enterotyping using full-length 16S-metabarcoding. Anim Microbiome. 2024;6(1):58. https://doi.org/10.1186/s42523-024-00345-0.
29. 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.
30. 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. https://doi.org/10.1093/nar/gkx295.
31. Dhakal IP. Normal somatic cell count and subclinical mastitis in Murrah buffaloes. J Vet Med Ser B. 2006;53(2):81–6. https://doi.org/10.1111/j.1439-0450.2006.00918.x.
32. Oikonomou G, Machado VS, Santisteban C, Schukken YH, Bicalho RC. Microbial diversity of bovine mastitic milk as described by pyrosequencing of metagenomic 16S rDNA. PLoS One. 2012;7(10):e47671. https://doi.org/10.1371/journal.pone.0047671.
33. Jiménez E, de Andrés J, Manrique M, Pareja-Tobes P, Tobes R, Martínez-Blanch JF, et al. Metagenomic analysis of milk of healthy and mastitis-suffering women. J Hum Lact. 2015;31(3):406–15. https://doi.org/10.1177/0890334415585078.
34. Hoque MN, Istiaq A, Rahman MS, Islam MR, Anwar A, Siddiki AMAMZ, et al. Microbiome dynamics and genomic determinants of bovine mastitis. Genomics. 2020;112(6):5188–203. https://doi.org/10.1016/j.ygeno.2020.09.039.
35. Zhao J, Fan H, Kwok LY, Guo F, Ji R, Ya M, et al. Analyses of physicochemical properties, bacterial microbiota, and lactic acid bacteria of fresh camel milk collected in Inner Mongolia. J Dairy Sci. 2020;103(1):106–16. https://doi.org/10.3168/jds.2019-17023.
36. Luo B, Dong F, Liu Y, Du J, Sun H, Ni Y, et al. Insights into the microbiota of raw milk from seven breeds animals distributing in Xinjiang China. Front Microbiol. 2024;15:1382286. https://doi.org/10.3389/fmicb.2024.1382286.
37. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, 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.
38. Dean CJ, Deng Y, Wehri TC, Pena-Mosca F, Ray T, Crooker BA, et al. The impact of kit, environment, and sampling contamination on the observed microbiome of bovine milk. mSystems. 2024;9(6):e0115823. https://doi.org/10.1128/msystems.01158-23.
39. Oikonomou G, Addis MF, Chassard C, Nader-Macias MEF, Grant I, Delbès C, et al. Milk microbiota: what are we exactly talking about? Front Microbiol. 2020;11:60. https://doi.org/10.3389/fmicb.2020.00060.

40. Pavelka MS Jr, Mahapatra S, Crick DC. Genetics of peptidoglycan biosynthesis. Microbiol Spectr. 2014;2(4):2–2013. https://doi.org/10.1128/microbiolspec.MGM2-0034-2013.
41. Sirard JC, Vignal C, Dessein R, Chamaillard M. Nod-like receptors: cytosolic watchdogs for immunity against pathogens. PLoS Pathog. 2007;3(12):e152. https://doi.org/10.1371/journal.ppat.0030152.
42. Philpott DJ, Sorbara MT, Robertson SJ, Croitoru K, Girardin SE. NOD proteins: regulators of inflammation in health and disease. Nat Rev Immunol. 2014;14(1):9–23. https://doi.org/10.1038/nri3565.
43. Wang H, Yu G, Yu H, Gu M, Zhang J, Meng X, et al. Characterization of TLR2, NOD2, and related cytokines in mammary glands infected by *Staphylococcus aureus* in a rat model. Acta Vet Scand. 2015;57:25. https://doi.org/10.1186/s13028-015-0116-0.
44. Wei LJ, Tan X, Fan GJ, Jiang YN, Shah QA. Role of the NOD1/NF-κB pathway on bovine neutrophil responses to crude lipopolysaccharide. Vet J. 2016;214:24–31. https://doi.org/10.1016/j.tvjl.2016.02.006.
45. Kim BR, Shin J, Guevarra R, Lee JH, Kim DW, Seol KH, et al. Deciphering diversity indices for a better understanding of microbial communities. J Microbiol Biotechnol. 2017;27(12):2089–93. https://doi.org/10.4014/jmb.1709.09027.
46. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27(3):379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
47. Simpson E. Measurement of diversity. Nature. 1949;163:688. https://doi.org/10.1038/163688a0.
48. Bray JR, Curtis JT. An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr. 1957;27(4):325–49. https://doi.org/10.2307/1942268.
49. Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021;18:366–8. https://doi.org/10.1038/s41592-021-01101-x.
50. Ahmad S, da Costa Gonzales LJ, Bowler-Barnett EH, Rice DL, Kim M, Wijerathne S, Luciani A, Kandasaamy S, Luo J, Watkins X, Turner E, Martin MJ, UniProt Consortium The UniProt website API: facilitating programmatic access to protein knowledge Nucleic Acids Research. 2025:gkaf394
51. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics (Oxford, England). 2006;22(13):1658–9. https://doi.org/10.1093/bioinformatics/btl158.
52. Slade D, Radman M. Oxidative stress resistance in *Deinococcus radiodurans*. Microbiol Mol Biol Rev. 2011;75(1):133–91. https://doi.org/10.1128/MMBR.00015-10.
53. Cox MM, Battista JR. Deinococcus radiodurans - the consummate survivor. Nat Rev Microbiol. 2005;3(11):882–92. https://doi.org/10.1038/nrmicro1264.
54. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nature biotechnology. 2017;35(9):833–844. https://doi.org/10.1038/nbt.3935.
55. Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Gohl, DM, Beckman KB, Knight R, Knights D. Evaluating the Information Content of Shallow Shotgun Metagenomics. mSystems. 2018;3(6): e00069-18. https://doi.org/10.1128/mSystems.00069-18.
56. Braem G, De Vliegher S, Verbist B, Heyndrickx M, Leroy F, De Vuyst L. Culture-independent exploration of the teat apex microbiota of dairy cows reveals a wide bacterial species diversity. Veterinary microbiology. 2012;157(3-4):383–390. https://doi.org/10.1016/j.vetmic.2011.12.031.
57. Vásquez-García A, Silva TD S, Almeida-Queiroz S. RD, Godoy SH, Fernandes AM, Sousa RL, Franzolin R. Species identification and antimicrobial susceptibility profile of bacteria causing subclinical mastitis in buffalo. Pesquisa Veterinária Brasileira. 2017;37(5):447-452.

Publisher's Note