Longitudinal study of the udder microbiome of Norwegian red dairy cows using metataxonomic and shotgun metagenomic approaches: Insights into pathogen-driven microbial adaptation and succession

Vinicius da Silva Duarte, Fiona Valerie Franklin, Alicja Monika Krysmann & Davide Porcellato*.

Faculty of Chemistry. Biotechnology and Food Science. Norwegian University of Life Sciences. NMBU. P.O. Box 5003. N-1432 Ås. Norway

Abstract

Bovine mastitis remains the most significant disease affecting dairy herds globally, driven by its multi-etiological nature and the complex dynamics of udder immunity and infection. While research addressing the microbial and immunological aspects of the bovine udder is limited, optimizing the udder microbiome has emerged as a promising strategy for preventing mastitis. This longitudinal study aimed to investigate the udder microbiome throughout lactation using both metataxonomic and shotgun metagenomic approaches, including analysis at the metagenome-assembled genome (MAG) level. The use of such methodologies provide a deeper understanding of the microbial composition and dynamics within the udder, bridging critical gaps in knowledge and revealing potential interactions between the microbiota and host. Milk samples were collected from 342 individual quarters of 24 Norwegian Red dairy cows. Significant variations in somatic cell count and microbiota composition were observed across lactation stages. Quarters with low somatic cell count (< 100,000 cells/mL) were notably higher during early lactation (80%) and mid-lactation (78.9%) compared to dry-off (53.1%) and late lactation (53%), with high somatic cell count (> 100,000 cells/mL) observed in 20–47% of samples. Diversity indices based on Shannon and Simpson metrics indicated significant effects of lactation stage, somatic cell count, and individual animal variability on microbial diversity. PERMANOVA analyses confirmed that individual animal variability (15.73%) and lactation period (5.52%) were the strongest factors influencing microbiota structure, with dysbiosis linked to mastitis-causing pathogens contributing 7.17% of the variance. Key pathogens identified included *E. faecalis*, *S. aureus*, *S. uberis*, and *S. chromogenes*, with persistent infections observed for *S. aureus* and *S. chromogenes*, but not *S. uberis*. Samples with low somatic cell count were enriched in beneficial genera such as *Corynebacterium*, *Bradyrhizobium*, and *Lactococcus*, while *Staphylococcus predominated* in milk samples with high somatic cell count. Dimensionality reduction with t-SNE and integration with culturomics enhanced milk microbiota classification, providing novel insights into udder microbiota dynamics and their role in bovine mastitis. Metagenomic shotgun sequencing revealed pathogen-specific metabolic signatures in the bovine mammary gland, identifying 289 MetaCyc pathways. Genome-centric analysis reconstructed 142 metagenome-assembled genomes, including 26 from co-assembly and 116 from individual assembly. Multi-locus sequence typing, virulence factors, and antimicrobial resistance gene profiling

provided insights into pathogen adaptation and persistence in the bovine mammary gland. Notably, 27 bacteriocin gene clusters and 322 biosynthetic gene clusters were predicted using genome mining tools. Our findings contribute to the establishment of pathogen-based therapies and enhance our understanding of the milk microbiome, which remains far from fully characterized. Such insights are vital for developing effective strategies to combat mastitis and improve dairy cattle health.

Keywords: Bovine mastitis, metataxonomic analysis, shotgun sequencing, udder health, milk microbiome

1. Introduction

Bovine mastitis remains the leading disease affecting dairy herds worldwide [1]. Due to its multi-etiological nature, it is a challenging condition to eradicate and is primarily managed through husbandry practices and the use of antibiotics to treat bacterial infections [2], which are the primary cause of intramammary infections (IMIs).

Mastitis can be classified into clinical and subclinical forms based on clinical features. Clinical mastitis (CM) is marked by distinct symptoms, such as the presence of flakes, clots, or watery secretions in milk. Affected quarters typically show swelling, warmth, and pain. In acute cases, systemic signs like hyperthermia, anorexia, and depression may also occur. The consequences of CM can be severe, often leading to cow mortality, agalactia, or premature culling [3]. Subclinical mastitis (SCM), on the other hand, is primarily identified by an increased somatic cell count (SCC) and reduced milk production, with SCC being considered the diagnostic gold standard for detecting SCM [4].

Microbiological culture of milk remains the gold standard for pathogen identification, despite some limitations [5]. Few studies have investigated the microbial composition of milk samples classified as "mixed growth," which are often considered contaminated during sampling [6], culture-negative samples from clinical mastitis [7], and samples from cows or quarters with high SCC (subclinically affected quarters) [8]. Overall, both mixed growth and culture-negative samples constitute a substantial proportion (30% or more) of the milk samples submitted for mastitis diagnostics [9].

Recent advancements in high-throughput next-generation sequencing (NGS) technology and bioinformatics tools over the past decade have facilitated a shift from traditional clinical microbiology to ecological and genomic characterization of the microbiome associated with infections in different research fields [10], including udder health [11]. This transition does not aim to replace traditional clinical microbiology, but rather emerges as a complementary diagnostic tool and a means to study the interactions between the host and its microbiome [12,13]. For example, studies emloying metataxonomic analysis have shown that culture-based diagnoses generally align with the most prevalent organisms identified through metagenomic sequencing, although additional potential pathogens have been detected in culture-negative samples [14].

The dynamics of udder immunity and infection are complex, influenced by factors such as the cow's health status, environmental conditions, and the resident microbiota [15]. Research has identified distinct changes in microbiome composition

between healthy and mastitic milk, with key phyla such as *Proteobacteria*, *Bacteroidetes*, *Firmicutes*, and *Actinobacteria* playing significant roles [16]. Metagenomic analysis has further revealed previously unreported opportunistic strains in clinical mastitis samples, along with functional pathways related to bacterial colonization and antibiotic resistance [17]. However, the etiology of udder dysbiosis remains unclear, particularly whether it serves as a cause or consequence of disturbances in host-specific factors, health status, environmental conditions, xenobiotic exposure, or inadequate hygiene. While comprehensive studies directly addressing this question in bovine udders are limited, research on lactating mothers can provide valuable insights [18]. Additionally, as reported by Urrutia-Angulo et al. (2024)[19] , the clinical definition of udder health status does not consistently align with the microbial profile.

Evidence suggests that SCM may have a distinct pathological origin compared to CM, with unique microbial signatures and sensory protein profiles [20]. These findings highlight the significant potential of metagenomic approaches to improve mastitis diagnosis and advance our understanding of its pathophysiology. Metagenomic studies provide valuable insights into the temporal dynamics of microbial communities and [21], when paired with metagenome reconstruction, reveal possible interactions within the microbiota (ecological studies) and between the microbiota and the host [22,23]. Furthermore, when integrated with community proteogenomics, these approaches become instrumental in identifying potential biomarkers for the early diagnosis of subclinical mastitis and in elucidating the pathophysiology of both subclinical and clinical mastitis, which remains incompletely understood [24,25].

Optimizing the udder microbiome has been proposed as a promising strategy for preventing mastitis in dairy cattle, as a healthy microbiome contributes to protection against pathogen colonization and the overgrowth of opportunistic pathobionts [26,27]. Therefore, this longitudinal study aimed to investigate the udder microbiome throughout lactation using metataxonomic and shotgun sequencing approaches. Studies of this nature contribute to establishing pathogen-based therapies and provide insights into the composition of the milk microbiome, which remains far from fully characterized. This gap in knowledge motivated the use of shotgun metagenomics and analysis at the metagenome-assembled genome (MAG) level in this study since such comprehensive investigations remain relatively scarce.

2. Material and methods

2.1. Experimental design and sample collection

Twenty-four Norwegian Red cows were chosen from the “Centre for Livestock Production” at the Norwegian University of Life Sciences. The farm adheres to the regulations set by the Norwegian Food Safety Authority for food production and animal welfare. Sample collection and the use of related information were approved by the farm owners, and no invasive procedures were involved in the study. The dairy cows were kept in freestalls equipped with rubber mats and raw wood chips for bedding. Their diet included silage available at all times and pelleted feed, adjusted according to each cow's milk production.

The experimental period took place between the autumns of 2022 and 2023. Milk samples were collected at four different time points, covering two lactation cycles: i) before dry-off in 2022 (DO-2022), ii) early lactation (EL-2023), iii) mid-lactation (ML-2023), and iv) late lactation (LL-2023). On each sampling day, hindmilk was aseptically collected from each quarter at the end of the regular milking routine, as described by Porcellato et al. (2020) [28]. In short, after the milking apparatus was removed, the teats were cleaned with iodine and then 70% ethanol, and approximately 350 mL of milk was manually collected. The procedure followed the guidelines of the National Mastitis Council (NMC; www.nmconline.org). Once collected, the milk samples were transported on ice to the laboratory, where they were divided into polystyrene tubes and immediately frozen at -20°C.

2.2. BacSomatic. culturing. and identification of isolates

Milk samples from all quarters were sent to the TINE Mastitis Laboratory in Molde, Norway, for microbiological analysis and species identification (gold standard diagnostic). Bacteriological culturing followed standard protocols 2.3. In summary, 10 µL of milk was plated on cattle blood agar with esculin and incubated at 37°C. The plates were examined after 24 and 48 hours. Species identification was then carried out using MALDI-Tof MS (Microflex LT system. Bruker Daltonics).

Concomitantly, the same samples were also plated in four different media at SciFood (KBM. NMBU) and incubated at 37°C for 24, 48 and 72 hours as follows: Tryptic Soy Agar (TSA, Sigma-Aldrich) plus 0.01% Tween 20, Blood agar (Thermo Scientific), brilliance CRE agar (detection of *Enterobacteriaceae* Producing KPC- and Metallo-Carbapenemase, ThermoFisher Scientific). and brilliance ESBL agar

157 (detection of extended-spectrum beta-lactamase-producing *Enterobacteriaceae*,
158 Thermo Scientific). Lastly, the samples were analyzed for somatic cell count (SCC)
159 and individual bacterial count (IBC) using the BacSomatic instrument (Foss Electric,
160 Hillerød, Denmark). In this study, samples were classified as either negative or positive
161 for mastitis-causing pathogens based on the official mastitis report.
162

163 **2.3. Metagenomic DNA extraction and 16S rRNA gene amplicon sequencing**

164 **2.3.1. Bacterial pellet preparation and metagenomic DNA extraction with**
165 **DNeasy PowerFood Microbial Kit**

166 For metagenomic DNA (mgDNA) extraction, a bacterial pellet was isolated from
167 40 mL of milk as outlined by Winther et al. (2022) [21]. A total of 345 milk samples were
168 processed. Initially, 40 mL of milk was thawed overnight at 4°C and centrifuged at
169 8,000 × g for 10 minutes at 4°C. The whey fraction was discarded, and the bacterial
170 pellet along with the fat layer was washed on an orbital shaker at 250 rpm for 15
171 minutes using a solution of 2% citrate water and 0.1% Tween 20. The mixture was
172 then centrifuged again at 8,000 × g for 10 minutes at 4°C, after which the supernatant
173 was removed, and the fat layer was gently collected using a sterile cotton swab soaked
174 in 70% ethanol. The bacterial pellet was then harvested and washed with 2% citrate
175 water (16,200 × g for 3 minutes).

176 The metagenomic DNA (mgDNA) was extracted using the DNeasy PowerFood
177 Microbial Kit (Qiagen. Düsseldorf. Germany). The bacterial pellet was transferred
178 directly into a PowerBead tube, placed in a FastPrep-24 5G Instrument (MP
179 Biomedicals European HQ) for three rounds of bead-beating at (*Lactococcus* program,
180 30 seconds with a 5-minute cooling period between each round). The samples were
181 then processed following the manufacturer's instructions. Finally, the mgDNA was
182 eluted in 50 µL of elution buffer and stored at −20°C until further processed.
183

184 **2.3.2. Library preparation and 16S rRNA gene amplicon sequencing**

185 Library preparation for amplicon sequencing followed the method previously
186 described. Specifically, the V3 and V4 regions of the 16S rRNA genes were amplified
187 using the primers Uni340F (CCTACGGGRBGCASCAG) and Bac806R
188 (GGACTACYVGGGTATCTAAT). The PCR reagents and amplification conditions
189 matched those outlined by Porcellato et al. (2020)[28]. Negative controls were included
190 to check for contamination during DNA extraction and library preparation. Sub-libraries

191 were cleaned and normalized using the SequalPrep Normalization Plate (96) Kit
192 (Thermo Fisher Scientific. USA) and then pooled. The final library concentration was
193 measured with Qubit 2 using the dsDNA HS kit (Thermo Fisher Scientific. USA), and
194 sequencing was performed on an Illumina NovaSeq 6000 platform (Illumina) using the
195 2 × 250 bp V3 kit (Illumina) at Novogene (Cambridge. UK). A total of four sequencing
196 batches were completed.
197

198 **2.4. Metagenomic DNA extraction and shotgun metagenomic sequencing**

199 **2.4.1. Metagenomic DNA extraction with MolYsis complete5 kit**

200 For metagenomic shotgun sequencing, 75 samples were selected based on
201 their SCC levels and their representation of all the major mastitis-causing pathogens
202 identified among the samples included in the study. The bacterial pellet was obtained
203 as outlined in section 2.3.1 and processed using the manufacturer's protocol (Molzym
204 GmBH & Co. KG. Bremen. Germany). In brief, cells were resuspended in 1 mL of
205 buffer SU, and host cells were lysed by adding 250 µL of chaotropic buffer (buffer CM).
206 The released nucleic acids were degraded using an enzyme (MolDNase B). Microbial
207 cells were then pelleted and lysed with specific reagents and proteinase K. The
208 mgDNA was isolated using spin columns, and 50 µL of DNA was eluted and stored at
209 −20°C for further processing.
210

211 **2.4.2. Multiple displacement amplification by Phi29 DNA polymerase**

212 Following the manufacturer's instructions, whole metagenome amplification
213 was conducted on 75 samples using multiple displacement amplification (MDA) with
214 the REPLI-g Single Cell kit (Qiagen: 150345). Briefly, 5 µL of metagenomic DNA
215 (mgDNA) was added to a microcentrifuge tube along with 5 µL of buffer D1 and mixed
216 by vortexing. After a 3-minute incubation at room temperature (15–25°C), 10 µL of
217 buffer N1 (stop solution) was added. To the 20 µL of denatured mgDNA, 30 µL of
218 master mix (comprising 29 µL REPLI-g mini reaction buffer and 1 µL REPLI-g mini
219 DNA polymerase) was added. The final mixture (50 µL) was gently mixed, incubated
220 at 30°C for 16 hours, and then heat-inactivated at 65°C for 3 minutes using the
221 SimpliAmp (Applied Biosystems). The mgDNA concentration was measured using the
222 Qubit HS dsDNA assay Fluorometer (Thermo Fisher Scientific. USA) and stored at
223 −20°C for further analysis.
224

225 **2.4.3. Library preparation and sequencing**

226 For Illumina sequencing library preparation, the mgDNA concentration of the
227 MDA-treated samples was measured using the Qubit HS dsDNA kit. If the extracted
228 DNA concentration was below 0.5 ng/µL, 5 µL of the undiluted sample was used.
229 Otherwise, the DNA was diluted to 0.2 ng/µL. DNA preparation for sequencing followed
230 the guidelines of the Illumina Nextera XT Library Preparation Kit. The DNA
231 concentration was measured using the Qubit HS dsDNA assay, and the concentration
232 was calculated before being diluted and pooled at equimolar ratios. The DNA library
233 was sequenced on the Illumina NovaSeq 6000 system (2 × 150 bp read length) at
234 Novogene (Cambridge. UK).

235

236 **2.5. Bioinformatics processing data**

237 **2.5.1. 16S rRNA gene amplicon sequencing**

238 The demultiplexed raw paired-end reads obtained from sequencing were
239 uploaded and processed in QIIME2 (version 2021.4) using the Casava 1.8 paired-end
240 pipeline [31]. DADA2 was selected for its ability to enhance taxonomic resolution through
241 precise identification and error correction of sample sequences that differ by as little
242 as a single nucleotide. This process involved several sequential steps, including
243 filtering, trimming, denoising, dereplicating, merging paired reads, and removing
244 chimeric sequences [32]. Subsequently, amplicon sequence variants were used to
245 generate a phylogenetic tree through the align-to-tree-mafft-fasttree pipeline in the q2-
246 phylogeny plugin [33]. Taxonomy assignment for the 16S data was performed using a
247 Naïve Bayes pre-trained Silva-138–99-nb-classifier [34].

248 For downstream metataxonomic analysis, QIIME artifacts were imported into R
249 (R Core Team 3.6.2. 2019) using the qiime2R package v.099.20
250 (https://github.com/jbisanz/qiime2R; accessed on 5 November 2024). *In silico*
251 contaminant identification and removal of amplicon sequence variants (ASVs) were
252 conducted for each sequencing batch using the R package Decontam (version 1.12)
253 [35], employing the frequency method with a user-defined threshold of 0.5. After
254 decontaminating the samples with the help of negative controls and before calculating
255 any diversity indices, ASVs assigned to *g-Mitochondria*. *o-Chloroplast*, and *d-Archaea*,
256 anssigned as Unknown at the phylum level, were excluded from the dataset. To
257 improve the taxonomical assignment and make inferences at the species level when

258 appropriate, ASVs were manually verified using the MegaBLAST search against the
259 16S ribosomal RNA sequences (Bacteria and Archaea) database.
260

261 **2.5.2. Shotgun sequencing**

262 Short reads from the metagenomic datasets were processed using the metaWRAP
263 v1.3[36] pipeline, following the bioinformatics workflow for genome-centric
264 metagenomics of the bovine hindmilk microbiome as previously outlined by Duarte &
265 Porcellato (2024)[23]. In short, the quality of the raw sequence data was verified using
266 the FastQC tool (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The
267 raw reads were trimmed with Trim-galore v0.4.3 [37], and then, the bovine-derived reads
268 were removed with bmtagger v3.101 (default settings). The high-quality reads were
269 assembled using both the individual or co-assembly methods with MegaHit v1.0[38].
270 Subsequently, the assemblies were binned with metaBAT v2.12.1 (-m 1500 and --
271 unbinned parameters) [39], Maxbin v2.2.4 (-markerset 40 option) [40] and CONCOCT
272 v0.4.0 (default settings) [41] using the metaWRAP-Binning module. The resulting three
273 bin sets were consolidated with metaWRAP-Bin-refinement, and the completion and
274 contamination of the resulting bins were evaluated with CheckM v1.0.7 (default
275 settings) [42]. Taxonomic and functional profiles were obtained using short reads with
276 MetaPhlAn 3.0 [43] and HUMAnN 3.0 [43], respectively.

277 All bioinformatics analyses were performed with the Orion Cluster at NMBU. The
278 raw reads were deposited into the Sequence Read Archive database
279 (http://www.ncbi.nlm.nih.gov/sra) under the BioProject PRJNA950968.
280

281 **2.5.3. MAG data analysis**

282 Metagenome assembled genomes (MAGs) were classified into high, medium
283 and low quality according to the minimum information about metagenome-assembled
284 genome guidelines (MIMAG) [44]. GTDB-Tk (v2.2.5, reference data version r207_v2) [45]
285 was used for taxonomic classification. Identifiers were assigned to the MAGs based
286 on taxonomic level. Dereplicated high-quality MAGs were used to generate a
287 phylogenetic tree using FastTree (v2.1.11) [46] and their relationships drawn with iTOL
288 [47]. CoverM (v0.6.1) [48] was used to retrieve MAGs's relative abundance and read
289 counts.

290 MAGs were annotated using the NCBI Prokaryotic Genome Annotation Pipeline
291 (PGAP)[49] and protein sequences forwarded for functional annotation with EggnNOG-

292 mapper [50,51]. MAGs were also processed through TORMES [52] version 1.3.0 (options -
293 -gene_min_id 30 --gene_min_cov 30) to detect antimicrobial resistance (AMR) and
294 virulence genes. For AMR prediction, three databases were adopted by TORMES
295 (ResFinder [53], CARD [54] and ARG-ANNOT [55], whereas virulence genes were detected
296 through the VFDB [56]. Multi-locus sequence typing (MLST) was performed with the mlst
297 software (Seemann T, mlst Github https://github.com/tseemann/mls) and the
298 PubMLST database [57]. To investigate molecules potentially associated with microbial
299 and host interactions in the mammary gland, BAGEL4 [58] was employed to identify
300 genes predicted to encode bacteriocins in the MAGs, while antiSMASH [59], with the
301 strictness parameter set to "relaxed," was used to predict bioactive compounds as part
302 of their secondary or specialised metabolism.
303

304 **2.6. Statistical analysis**

305 **2.6.1. Generalized Additive Model (GAM)**

306 Two GA models were developed to investigate the influence of lactation period
307 (Period, DO-2022, EL-2023, ML-2023, LL-2023), somatic cell count classification
308 (SCC, “High” and “Low”), and individual animal variability (animal) on microbial
309 diversity metrics, specifically Shannon and Simpson indices. Both models were fitted,
310 treating each categorical variable as a random effect by including smooth terms (bs =
311 "re"). The REML method was selected for estimation to optimize the accuracy of
312 parameter estimates and random effect variances. Data handling, model fitting, and
313 visualization were conducted using R (version 4.4.1) with the R packages phyloseq [60],
314 mgcv [61], and dplyr [62].
315

316 **2.6.2. Metataxonomic analysis**

317 Statistical analysis and graph construction were performed with RStudio
318 (v.1.2.5033) using one or the combination of the following R packages: MicrobiomeR
319 (https://github.com/microbiome/microbiome), dplyr [62], ggplot2
320 (https://ggplot2.tidyverse.org), phyloseq [60], tidyr (https://github.com/tidyverse/tidyr),
321 vegan (https://github.com/vegandevs/vegan), and pairwise Adonis
322 (https://github.com/pmartinezarbizu/pairwiseAdonis).
323 Groups with parametric distribution were analyzed by one-way ANOVA followed
324 by Tukey’s *post hoc*. The Wilcoxon test was used to compare two unpaired groups.
325 Differences were considered significant at $P < 0.05$. Significant differences in alpha-

326 diversity between the groups with high and low SCC, as well as between the different
327 lactation periods, were determined using the alpha function in microbiome R package
328 v.2.1.24, adopting Wilcoxon's test. For beta diversity, weighted and unweighted
329 UniFrac distances were subjected to permutational multivariate analysis of variance
330 (ANOVA) to assess significant differences (pseudo-F test) in bacterial community
331 composition and structure among the groups with a permutation number of 999. Non-
332 metric multidimensional scaling (NMDS) plots were constructed with microViz [63]
333 To investigate the association between microbial taxa and key variables, we
334 employed the MaAsLin 2 [64] package with a negative binomial regression model with
335 Benjamini-Hochberg correction (max_significance = 0.25), incorporating both fixed
336 and random effects. The input data matrix included the absolute abundances of
337 microbial taxa at the phylum and genus level, while the metadata comprised variables
338 of interest for the model, including "SCC", "Period", and "sampling day", with "animal"
339 as a random effect.
340 The resilience of resident microbiota following dysbiosis was evaluated based
341 on a microbial dysbiosis index (MDI), calculated as follows: (MDI_{sample}=log(Relative
342 Abundance of non-pathobionts/Relative Abundance of pathobionts). The genera
343 *Staphylococcus*, *Streptococcus*, *Enterococcus*, *Trueperella*, *Klebsiella*, *Serratia*,
344 *Peptostreptococcus*, *Corynebacterium*, and *Escherichia-Shigella* were defined as
345 pathobionts. A quarter was defined as dysbiotic if it exhibited a Shannon diversity index
346 lower than the mean of quarters with somatic cell counts (SCC) below 100,000
347 cells/mL and an MDI greater than 0, indicating a predominance of dysbiotic
348 microorganisms within the specific sample.
349 To reduce the dimensionality of the metataxonomic data, t-distributed
350 stochastic neighbor embedding (t-SNE) was run using the R function Rtsne from the
351 Rtsne package [65] and selecting the three-dimensionality output option. A hierarchical
352 clustering method using the "hclust" function in R was then used on the output to
353 identify groups of samples with the same microbiota profile.
354
355 **2.6.3. Microbial differential abundance analyses based on microbial functional**
356 **profile**
357 To explore associations between microbial features and metadata variables,
358 microbial differential abundance analyses were conducted using the Maaslin2
359 package [64] with negative binomial regression models, incorporating fixed effects (e.g.,

360 “Pathogen”, “SCC”, and “Lactation period”), random effects (e.g., “animal”), and
361 Benjamini-Hochberg correction (max_significance = 0.25), without normalization or
362 transformation.
363
364 **3. RESULTS**
365 **3.1. Changes in SCC and pathogen distribution across lactation periods**
366 **reveal significant variability**
367 In this study, we collected 342 individual quarter-level milk samples from 24
368 Norwegian Red dairy cows (8 primiparous and 16 multiparous). These samples were
369 collected during four lactation periods and encompassed two lactation cycles. The pre-
370 drying period in 2022 (DO-2022) was used as the baseline. In 2023, milk samples
371 were collected at three distinct stages of the lactation cycle based on the days in milk
372 (DIM) of each animal. Samples were obtained during early lactation (EL-2023, median
373 DIM: 22 days), mid-lactation (ML-2023, median DIM: 126 days), and late lactation (LL-
374 2023, median DIM: 244 days), corresponding to the respective stages of the lactation
375 cycle.
376 Based on SCC, a total of 233 samples were classified as low SCC (< 100.000
377 cells/mL), while 109 samples were categorized as high SCC (> 100.000 cells/mL). The
378 percentage of quarters with low SCC was notably higher during EL-2023 (80%) and
379 ML-2023 (78.9%) compared to DO-2022 (53.1%) and LL-2023 (53%), with high SCC
380 percentages of 20%, 21.1%, 46.9%, and 47%, respectively. As depicted in **Figure 1A**,
381 there was a statistical difference between the different lactation periods, except
382 between DO-2022/LL-2023 and EL-2023/ML-2023. The distribution of individual
383 bacterial counts (IBC) exhibited a comparable pattern, revealing a significant
384 difference between samples classified as low SCC and those categorized as high
385 SCC. Specifically, low SCC samples had a mean IBC of 3.31 ± 0.475, while high SCC
386 samples showed a mean IBC of 3.74 ± 0.334.
387 The percentage of quarters testing positive or negative for mastitis pathogens,
388 identified via MALDI-Tof, also reflected the trends observed for SCC and IBC **(Figures**
389 **1B and 1C)**. The lowest detection of mastitis-causing pathogens was recorded during
390 EL-2023, with only 10.5% of samples testing positive. Among the mastitis-causing
391 pathogens identified in this study (from 60 samples), *E. faecalis* emerged as the
392 predominant etiological agent, accounting for 30% of the cases. *S. aureus*, *S. uberis*,
393 and *S. chromogenes* were detected in 25%, 20%, and 10% of the samples,

394 respectively. The bacterial species *S. epidermidis* (7%), *C. bovis* (3%), *S. dysgalactiae*
395 (3%), and *S. haemolyticus* (2%) were identified to a lesser extent.
396 The relationship between SCC and IBC ratio concerning specific mastitis-causing
397 pathogens was also investigated. Groups with fewer than three samples, specifically
398 *C. bovis*, *S. dysgalactiae*, and *S. haemolyticus*. were excluded from this analysis. Our
399 findings indicate that the level of somatic cells was significantly higher in samples
400 associated with *S. epidermidis* compared to those classified as negative, with an
401 adjusted p-value of 0.014. This suggests that a small number of *S. epidermidis* cells
402 may be sufficient to trigger a strong pro-inflammatory response in the mammary gland.
403 When we examined these same samples in relation to bacterial counts on blood
404 agar plates (after 48 hours of incubation at 37°C), alongside individual bacterial count
405 (IBC), somatic cell count (SCC), and the identified pathogen **(Figure 2A)**, we observed
406 that many samples classified as negative (i.e.. absence of mastitis-causing pathogens
407 according to the gold standard) still showed high plate counts. Specifically, 129
408 samples had colony counts ranging between 30 and 300 CFUs. From these samples,
409 57 showed SCC higher than 100,000 (median: 2.9 x 10^5 cells/mL) and a median IBC
410 of 7.6 x 10^3 cells/mL. Our findings indicate that these samples harbor microorganisms
411 distinct from the major mastitis-causing pathogens, which, although not routinely
412 identified, may still contribute to increased somatic cell counts in the mammary gland.
413

414 **3.2. Metataxonomic analysis**

415 **3.2.1. Alpha and beta diversity analyses show significant effects of lactation**
416 **period on diversity indices**

417 In this longitudinal study, we investigated the microbial composition of 306 milk
418 samples encompassing a whole lactation cycle. By conducting amplicon sequencing
419 of the hypervariable region V3-V4 of the 16S rRNA gene, we obtained a total of 1.3 ×
420 10^7 high-quality filtered sequences after the removal of chimeric, contaminants, and
421 spurious sequences (e.g.. mitochondrial and chloroplast sequences). The median
422 number of sequences per sample in the low SCC group was 37,589, while the high
423 SCC group had a median of 47,282 sequences. Across all groups, the overall median
424 number of sequences per sample was 42,970.
425 We first investigated alpha diversity analysis considering the variables SCC
426 (low and high) and lactation period (DO-2022, EL-2023, ML-2023, and LL-2023).
427 When we analyzed the variable SCC, statistical significance was observed for the

428 indices Shannon, Simpson, and inverse Simpson **(Figure 2B)**. Overall, samples with
429 low SCC exhibited greater microbial diversity and lower dominance of individual
430 species. When the different periods were considered, five out of seven indices showed
431 higher diversity measurements for the periods DO-2022 when compared to LL-2023.
432 All seven indices were statistically significant for the pairwise comparison between EL-
433 2023 and LL-2023 **(Figure 2C)**. As depicted in the volatility plots **(Figures 2D and 2E)**
434 for the Shannon and Simpson indices, there is a clear trend indicating a reduction in
435 both indices at the end of lactation in samples containing more than 100,000 cells/mL.
436 We also employed the generalized additive model (GAM) to analyze how the
437 Shannon and Simpson diversity indices were influenced by different predictor
438 variables. The GAM analysis revealed significant effects of the lactation period,
439 somatic cell count, and individual animal variability on both the Shannon and Simpson
440 indices, while bacterial count did not show a significant impact. The smooth term for
441 the lactation period indicated a modest but statistically significant effect on microbial
442 diversity (Shannon: edf = 2.10, F = 2.72, p = 0.039; Simpson: edf = 2.11, F = 2.66, p
443 = 0.047), suggesting variation in diversity across lactation stages. Somatic cell count
444 had a highly significant effect (Shannon: edf = 0.92, F = 28.95, p < 0.001; Simpson:
445 edf = 0.94, F = 38.44, p < 0.001). indicating that the degree of mammary gland
446 inflammation plays a critical role in shaping the microbiota. Additionally, the random
447 effect of individual animals was significant (Shannon: edf = 16.37, F = 2.25, p < 0.001;
448 Simpson: edf = 14.7, F = 1.72, p = 0.002), highlighting considerable inter-animal
449 variability in microbial diversity. However, the bacterial count was not a significant
450 predictor of diversity (Shannon: Estimate = 1.49e-08, p = 0.681; Simpson: Estimate =
451 4.353e-09, p = 0.387). The model explained 24.9% and 23.4% of the deviance, with
452 adjusted R-squared values of 19.5% and 18.4% for the Shannon and Simpson indices,
453 respectively, underscoring the relevance of the factors considered in the model.
454 The PERMANOVA analyses using unweighted and weighted UniFrac distances
455 revealed significant effects of various factors on the microbial community structure
456 **(Figure 3)**. For the unweighted UniFrac analysis, the lactation period explained 3.01%
457 of the variance in microbiota composition (F = 3.20, p = 0.01), indicating shifts in
458 microbial diversity across different stages of lactation. Individual animal variability also
459 contributed 9.42% to the observed variance (F = 1.25, p = 0.01), highlighting notable
460 inter-animal differences in microbial composition. However, somatic cell count and the
461 history of mastitis-causing pathogens identified in milk samples were not significant

predictors, accounting for only 0.30% and 2.85% of the variance, respectively. The majority of the variation, 83.97%, remained unexplained by the measured factors, suggesting other influences on the microbial community.

In contrast, the weighted UniFrac analysis indicated more pronounced effects of the measured variables. Somatic cell count had a small but significant effect, explaining 0.78% of the variance ($F = 3.22$, $p = 0.01$). The lactation period played a more substantial role, explaining 5.52% of the variance ($F = 7.57$, $p = 0.01$), reinforcing the influence of lactation stages on microbial composition. Individual animal variability was the largest contributor to variance, explaining 15.73% ($F = 2.69$, $p = 0.01$), while the identification of major mastitis-causing pathogens in milk samples contributed 7.17% ($F = 3.28$, $p = 0.01$), highlighting the impact of dysbiosis on microbial communities in the udder. The remaining 65.18% of the variance was attributed to residual (i.e., unmeasured factors), reflecting biological complexity and other unknown influences shaping the udder microbiota.

3.2.2. Differential abundance analysis and multivariable associations reveal shifts in *Actinobacteriota* across lactation stages

After clustering at a 99% similarity threshold and preprocessing steps (removal of *g-Mitochondria*, *o-Chloroplast*, and *d-Archaea*), we identified a total of 48,622 ASVs. Four phyla – *Firmicutes* (median: 63%; mean: 60.1%), *Actinobacteria* (median: 20%; mean = 25.2%), *Proteobacteria* (median: 7%; mean: 11.4%), and *Bacteroidota* (1%; mean: 2.3%) – comprised approximately 99% of the milk bacterial population. At the genus level, ten taxa stand out and comprise approximately 65% of the microorganisms at this taxonomic level: *Corynebacterium* (median: 12.9%; mean: 19.4%), *Staphylococcus* (median: 4.0%; mean: 17.0%), *Bradyrhizobium* (median: 1.8%; mean: 5.4%), *Streptococcus* (median: 0.2%; mean: 4.9%), *Aerococcus* (median: 2.6%. mean: 4.8%), *Romboutsia* (median: 3.4%; mean: 4.5%), *Enterococcus* (median: 0.2%; mean: 3.1%), *Kocuria* (median: 0.9%; mean: 2.0%), *Lactococcus* (median: 0.2%; mean: 1.7%), and *Weissella* (median: 0.6%; mean: 1.6%).

We performed a differential abundance analysis using MaAsLin2 to determine multivariable associations across samples grouped by somatic cell count (SCC), lactation period, sampling day, and microbial composition at both the phylum and genus levels. We found significant results for 15 phyla **(Supplementary Table 1)**. Samples with low SCC showed a significantly greater abundance of eight phyla, with

496 *Proteobacteria*, *Actinobacteriota*, and *Bacteroidota* particularly prominent based on
497 their prevalence in our dataset (belonging to the top four phyla). Additionally, shifts in
498 bacterial phyla were observed across different lactation periods. While the abundance
499 of the *Firmicutes* phylum remained relatively stable throughout the lactation periods
500 (q-value > 0.25), we observed an enrichment of the *Proteobacteria* phylum in the early
501 lactation period (EL-2023), which gradually decreased in the subsequent periods. In
502 contrast, members of the *Actinobacteriota* phylum were enriched in the mid- (ML-
503 2023) and late lactation (LL-2023) periods. When considering the continuous variable
504 "day", we observed a distinct fluctuation in *Actinobacteriota* abundance across the
505 studied periods. Specifically, *Actinobacteriota* was highly abundant during the baseline
506 period DO-2023, but this abundance decreased in the early lactation period (EL-2023),
507 followed by two subsequent increases during the mid- (ML-2023) and late lactation
508 (LL-2023) periods.

509 Our genus-level analysis identified 66 differentially abundant genera
510 **(Supplementary Table 2)**. Samples with low SCC showed enrichment in several
511 genera, with *Corynebacterium*, *Bradyrhizobium*, *Romboutsia*, and *Lactococcus*
512 standing out among the top 10 genera. Conversely, *Staphylococcus* was significantly
513 enriched in samples with high SCC. Examining differentially abundant genera by
514 lactation period, *Bradyrhizobium*, Unidentified *Lactobacillaceae* family. *Pediococcus*,
515 Unidentified *Aerococcaceae* Family, and *Weissella* were enriched in samples collected
516 during EL-2023 compared to DO-2022. Additionally, two genera from the families
517 *Aerococcaceae* and *Lactobacillaceae* were enriched in ML-2023. When analyzed by
518 sampling day, *Bifidobacterium* and *Bacillus* displayed positive coefficients, indicating
519 a progressive increase in abundance over time, whereas the abundance of
520 Unidentified *Lactobacillaceae* family decreased over time.

521

522 **3.3. The integration between t-SNE and culturomics improves milk microbiota**
523 **classification**

524 The microbial composition of the 295 samples was subjected to dimensionality
525 reduction t-SNE analysis. An optimal number of k-mean clusters was identified by
526 using the “Elbow method,” and the microbiota was divided into eighteen clusters
527 **(Supplementary Figure 1A-C)**. The eighteen clusters were then used to investigate
528 the microbial composition, which significantly differed in main ASVs and genera. Each
529 cluster was then used as a factor to map Chao1 and Shannon diversity indices and

530 other variables such as SCC and IBC. As depicted in **Figure 4**, cluster analysis
531 revealed that IBC levels varied significantly from the global mean in clusters 9, 10, 13,
532 14, 16, and 17. For SCC, statistically significant differences were observed in clusters
533 9, 10, 11, 13, 14, 15, and 17. The diversity indices Chao1 and Shannon also exhibited
534 substantial variation across clusters. For the Chao1 index, clusters 6, 7, 10, 13, 14,
535 15, 16, and 18 did not differ significantly from the global mean. In contrast, for the
536 Shannon index, all clusters except 2, 7, 9, 10, 15, 16, and 18 showed significant
537 differences from the global mean.

538 We conducted an analysis to examine the distribution of the predominant
539 bacterial genera across different clusters **(Figure 5)**. The genera *Corynebacterium*,
540 *Staphylococcus*, and *Aerococcus* show high prevalence and widespread distribution
541 across clusters, indicating that these genera are commonly found across diverse milk
542 samples used in this study and are not restricted to any specific group. In contrast,
543 *Enterococcus* and *Streptococcus* were majority limited to a single cluster each,
544 respectively, clusters 13 and 14.

545 Considering a minimum mean relative abundance of 10% per cluster
546 **(Supplementary Figure 2)**, we observed a notably high relative abundance of
547 *Staphylococcus* in five clusters (cluster 4: 80.50%, cluster 6: 77.88%, cluster 7:
548 58.85%, cluster 10: 56.96%, cluster 15: 52.69%). Combined with the culturomics'
549 analysis, cluster 4 was enriched for *S. chromogenes* (4 positive samples out of 12),
550 whereas cluster 10 was enriched for *S. aureus* (7 positive samples out of 9). Only one
551 species of *S. haemolyticus* was identified by MALDI-Tof (cluster 6).

552 We also conducted a MegaBLAST analysis on the top 20 ASVs to investigate
553 species-level identities within clusters with high prevalence of *Staphylococcus* and
554 *Corynebacterium* and *Streptococcus* **(Figure 6)**. For cluster 4, samples showed a
555 homogenous distribution of the four most abundant ASVs, here annotated as *S.*
556 *chromogenes* (identity scores ranging from 99.53 to 100%). Cluster 6 showed a high
557 abundance of an ASV that had as top hits (identity of 100%) with *Staphylococcus*
558 *borealis*, *Staphylococcus pragensis*, *Staphylococcus petrasii,* and *Staphylococcus*
559 *croceilyticus*.

560 Cluster 10 was primarily dominated by a single ASV with 100% identity,
561 annotated as *Staphylococcus aureus*, *Staphylococcus schweitzeri*, *Staphylococcus*
562 *argenteus*, *Staphylococcus roterodami*, and *Staphylococcus simiae*. Notably, sample
563 H193 displayed a microbial profile distinct from the other samples in this cluster,

diverging from the typical dominance of *S. aureus*. Cluster 7 was dominated by an ASV with 100% identity matching *Staphylococcus epidermidis*, *Staphylococcus caprae*, and *Staphylococcus capitis*. Additionally, four samples (H298, H300, H305, and H307) showed a high abundance of an ASV annotated with 100% identity as *Staphylococcus borealis*, *Staphylococcus pragensis*, *Staphylococcus petrasii*, and *Staphylococcus croceilyticus*. In the final *Staphylococcus* cluster analyzed (cluster 15), ASVs were annotated as *S. epidermidis, S. caprae, S. capitis,* and *Staphylococcus hominis*.

Regarding *Corynebacterium*, eleven out of eighteen clusters showed a relative abundance greater than 10%. Clusters 2 (37.68%), 8 (64.92%), and 18 (44.56%) exhibited particularly high abundances. It is important to note that from culturomics' data, the detection of *Corynebacterium* species was limited to only two samples (H087 and H109), both positive for *Corynebacterium bovis*. The ASV-level investigation provided insights into the composition of these clusters. Our analysis revealed that several ASVs had the best hits with uncultured bacteria. Cluster 2 stood out for its heterogeneity of *Corynebacterium* species, a pattern not observed in clusters 8 and 18, and overall enrichment for *C. bovis*. Cluster 8 showed a more homogeneous distribution of ASVs, and specifically two sequence variants annotated as uncultured bacterium/*C. bovis* (identities of 99.76% and 100%).

Lastly, clusters 3 and 14 showed a high abundance of members of the genus *Streptococcus*. Cluster 3 was enriched for three ASVs (100% identity) annotated as *Streptococcus uberis* /*Streptococcus porcinus*, *Streptococcus alactolyticus*/*Streptococcus macedonicus*/*Streptococcus pasteurianus*, whereas cluster 14 was dominated by a single ASV identified as *Streptococcus uberis* (identity: 100%).

3.3.1. Udder microbiota dynamics reveal persistent infection by *S. aureus* and *S. chromogenes*, but not *S. uberis*

Following t-SNE analysis to enhance sample classification and elucidate underlying patterns in the dataset, we conducted a quarter-level analysis to track temporal fluctuations in genus-level taxa before and after detecting the pathogen identified by the gold standard method. We focused on samples testing positive for Staphylococci, Enterococci, and Streptococci. Notably, these milk samples predominantly grouped into clusters 4, 10, 13, and 14. For robustness, only quarters

598 with a minimum of three sampling points were included in the analysis **(Figure 7)**. To
599 estimate the resilience of resident microbiota following a dysbiosis event, we
600 established a microbial dysbiosis index combined with the Shannon diversity index. A
601 quarter was classified as dysbiotic if it exhibited a Shannon index below 3.5 (the mean
602 value for quarters with somatic cell counts below 100,000 cells/mL) and an MDI
603 greater than 0, indicating a predominance of dysbiotic microorganisms within that
604 sample.

605 For cluster 10, which was identified as enriched with *S. aureus*, 5 quarters were
606 analyzed and the longitudinal analysis revealed a persistent state of infection by *S.*
607 *aureus*. Once an infection was established, *S. aureus* remained consistently present
608 in samples regardless of the lactation stage. Although Intramammary bacterial counts
609 varied over time, somatic cell counts consistently remained elevated, typically
610 exceeding, or in a few cases, hovering around 100,000 cells/mL. A marked depletion
611 of specific taxa, such as *Corynebacterium*, *Aerococcus*, and *Romboutsia*, was
612 observed, along with reductions in other low-abundance taxa grouped here as
613 "Others." Additionally, a co-infection with species of the genus *Streptococcus* was
614 detected at two specific time points (samples H195 and H349). When the MDI index
615 was calculated for samples in this cluster, it was evident that samples following an
616 episode of high *S. aureus* abundance did not return to a state of eubiosis.

617 Cluster 4 showed enrichment for *S. chromogenes*, where persistent infection
618 behavior was also observed. However, it's noteworthy that the infection-associated
619 somatic cell counts and intramammary bacterial counts in this cluster are significantly
620 lower than those found in cluster 10. In the case of quarter 7032_VF, both *S. aureus*
621 and *S. chromogenes* were isolated and identified. Interestingly, once this quarter was
622 infected by both species, the microbiota failed to revert to its original state. This
623 phenomenon is also clearly demonstrated in quarters 6691_HB and 6691_VB, which
624 suggests that *S. aureus* and *S. chromogenes* exhibit specific competitive mechanisms
625 that are reflected in a pattern of microbiota depletion. This pattern impedes the
626 recolonization of the native microbiota, leading to a more prolonged dysbiosis status,
627 as evidenced by the dysbiosis indices calculated for those samples.

628 The analysis of cluster 13, enriched for *E. faecalis*, revealed a comparatively
629 less severe dysbiosis than that observed in quarters infected with *S. aureus* and *S.*
630 *chromogenes*. In this cluster, a reduction in the relative abundance of taxa such as
631 *Corynebacterium*, *Aerococcus*, and *Romboutsia* was noted. However, the native

632 microbiota appears capable of re-establishing itself, a resilience not observed in
633 clusters 4 and 10. This phenomenon could be observed for the quarters 6415_VB
634 (sample H077: Shannon = 1.73, MDI = 1.03; sample H165: Shannon = 4.77, MDI = -
635 2.35), 7003_HB (sample H255: Shannon = 2.53, MDI = 1.23; sample H355: Shannon
636 = 5.17, MDI = -0.72), and 7003_HF (sample H156: Shannon = 2.04, MDI = 1.05;
637 sample H256: Shannon = 4.43, MDI = -0.49).

638 For cluster 14 (characterized by enrichment in *Streptococcus uberis*), we
639 observed that the reduction in *Corynebacterium* was not as pronounced as in clusters
640 dominated by *Staphylococcus* and *Enterococcus*. Although *S. uberis* induced a
641 dysbiotic state in the mammary gland, we identified five instances in which a dysbiotic
642 quarter returned to a eubiotic state, and one which we classified as in the borderline
643 of a eubiotic state (samples H084 and H164), including quarters such as 6611_HF
644 (sample H082: Shannon = 1.65, MDI = 2.33; sample H162: Shannon = 3.92, MDI = -
645 0.65), 6415_HB (sample H079: Shannon = 1.60, MDI = 2.04; H167 = Shannon = 4.58,
646 MDI = -1.94), 6611_HB (sample H083: Shannon, 2.44, MDI: 1.98), 6611_HF (sample
647 H084: Shannon = 2.36, MDI = 1.55; sample H164: Shannon = 2.96, MDI = -0.21), and
648 6941_VB (sample H141: Shannon = 3.38, MDI = 0.34; sample H241: Shannon = 4.10,
649 MDI = -0.65).

650

651 **3.4. Metagenomic shotgun sequencing uncovers pathogen-specific**
652 **metabolic signatures in the bovine mammary gland**

653 After trimming, quality filtering, and in silico separation of bacterial reads from
654 bovine sequences, we obtained approximately 269.5 GB of high-quality PE reads
655 across 73 samples, which were then processed for functional profiling using
656 HUMAnN3. This profiling identified 289 MetaCyc metabolic pathways, which were
657 then assessed for differential abundance with MaAsLin2 considering “SCC” and
658 “pathogen” as fixed effects. At first glance, the heatmap **(Figure 8A)** constructed with
659 the 289 predicted MetaCyc pathways revealed that the samples did not form well-
660 defined clusters when grouped by SCC or lactation period, highlighting a significant
661 degree of heterogeneity. However, four distinct clusters of pathogen-negative samples
662 are evident in the heatmap, characterized by a notable lack of metabolic pathway
663 enrichment. Additionally, clusters of samples identified as positive for *S. aureus* and
664 *E. faecalis* were also discernible.

665 Focusing on pathways that were differentially expressed (q-value < 0.25) based
666 on somatic cell content, we observed an enrichment of 134 microbial MetaCyc
667 pathways for samples with high SCC and four MetaCyc pathways for samples with
668 low SCC **(Figure 8BC)**. When pathways were stratified by bacterial species within the
669 high SCC group, many pathways were common across several bacterial species,
670 including Staphylococci and Streptococci, and were largely associated with central
671 metabolism. Notably, the pathway PWY-6507 (PWY-6507: 4-deoxy-L-threo-hex-4-
672 enopyranuronate degradation) was significantly enriched in *E. faecalis*. In contrast,
673 the four pathways enriched in the low SCC group (PWY-7185, PWY-7184, PWY-6545,
674 and PWY-7210) were primarily related to pyrimidine biosynthesis.
675 We also assessed the enrichment of metabolic pathways by bacterial species
676 identified via MALDI-ToF, compared to samples classified as negative. This analysis
677 aimed to identify potential metabolic distinctions associated with pathogen presence
678 and uncover infection-specific metabolic signatures unique to the mammary gland. In
679 total, five pathogens were analyzed: *Staphylococcus aureus*, *Staphylococcus*
680 *chromogenes*, *Enterococcus faecalis*, *Staphylococcus epidermidis*, and
681 *Streptococcus uberis* **(Figure 9)**.
682 For *S. aureus*, 37 pathways were significantly identified, with 36 pathways
683 enriched in this group and one pathway (ASPASN-PWY) more abundant in the
684 reference group. For *S. chromogenes*, 33 pathways were detected, including 32
685 enriched in this group and one (POLYISOPRENSYN.PWY) enriched in the reference
686 group. In *E. faecalis*, 38 pathways were identified, of which 29 were enriched in this
687 group, while 9 were highly represented in the negative group: PWY-1861, PWY-1042,
688 CITRULBIO-PWY, PWY-7200, SULFATE.CYS.PWY, HISDEG.PWY,
689 SO4ASSIM.PWY, PWY-6703, and PWY1ZNC.1. For *S. epidermidis*, only pathways
690 enriched in the negative group were found, including PWY66-429,
691 PEPTIDOGLYCANSYN.PWY, PWY-6124, PWY-7234, PENTOSE.P.PWY, PWY-3841,
692 PWY-6123, PWY-6122, PWY-6277, and PWY-6121. Lastly, for *S. uberis*, all seven
693 identified pathways were enriched in this group. Overall, these findings highlight
694 distinct metabolic profiles associated with specific pathogens, underscoring the utility
695 of pathway-based analyses to better understand pathogen-host interactions and their
696 metabolic underpinnings in mastitis.
697 Based on the previously identified pathways with a q-value < 0.25, a Venn
698 diagram was constructed to identify unique and shared MetaCyc pathways **(Figure**

699 **10)**. For *S. aureus*, 14 unique MetaCyc pathways were identified (PWY-5484, PWY-
700 5005, ARGSYNBSUB.PWY, GLYCOLYSIS, PWY-5840, PWY-7111, PWY-7953, PWY-
701 3001, PWY-7977, FERMENTATION.PWY, PWY-5189, P125-PWY, PWY-5918, and
702 RUMP.PWY). The enrichment of these pathways demonstrated a versatile metabolic
703 profile, encompassing central carbon metabolism, amino acid and nucleotide
704 biosynthesis, stress tolerance, and virulence factors, highlighting the adaptability of *S.*
705 *aureus* and the potential for pathogenicity in various host environments. *S.*
706 *chromogenes* exhibited unique metabolic pathways associated with amino acid
707 biosynthesis (PWY-7383, glutamate and glutamine biosynthesis), lipid metabolism
708 (PWY-6936, fatty acid biosynthesis), and energy production (PWY-6163,
709 fermentation), reflecting its ability to adapt to resource-limited environments and its
710 potential role in infection dynamics. The *S. uberis* group showed enrichment in
711 pathways related to carbohydrate metabolism, including glycogen biosynthesis
712 (GLYCOGENSYNTH.PWY), lactose and galactose utilization (P461-PWY), and
713 nucleotide sugar metabolism (PWY-5973), as well as processes involved in cell wall
714 biosynthesis (PWY-621), indicating its capacity to thrive in nutrient-rich environments,
715 such as the mammary gland during infection. For *S. epidermidis,* eight unique
716 metabolic pathways were annotated for this species and involved, for instance, in
717 energy production (glycolysis – PWY66-429), cell-wall structural maintenance
718 (peptidoglycan biosynthesis – PEPTIDOGLYCANSYN-PWY), stress tolerance
719 (polyamine biosynthesis – PWY-6122), and nutrient acquisition (siderophore
720 biosynthesis – PWY-7234**;** and arginine pathways – PWY-6123). These features
721 contribute to its resilience and success as a commensal and opportunistic pathogen.
722 Finally, *E. faecalis* displayed a metabolic profile enriched in pathways related to amino
723 acid biosynthesis (PWY-5971, PWY.7357), nucleotide metabolism (PWY.6317,
724 P164.PWY), and lipid-related processes (PWY0.1297, PWY0.301), as well as
725 pathways supporting secondary metabolite production (PWY0.1477) and cofactor
726 biosynthesis (PWY.7242, PWY.7791), underscoring its metabolic flexibility and
727 potential resilience in diverse environmental conditions.
728

729 **3.5. Genome-resolved metagenomic data analysis**

730 **3.5.1. MAG quality assessment**

731 The genome-centric approach applied to shotgun reads from milk microbiomes
732 facilitated the reconstruction of 142 metagenome-assembled genomes (MAGs), of

733 which 26 were obtained using a co-assembly method and 116 using individual
734 assembly method. Among these, 38 genomes were classified as high-quality
735 (completion > 90%, contamination < 5%), 50 as medium-quality (completion ≥ 50%,
736 contamination < 10%), and 39 as low-quality (completion < 50%, contamination >
737 10%) **(Supplementary Table 3)**. MAGs with contamination higher than 10% were
738 excluded from subsequent analyses.

739 Taxonomic assignments were consolidated at the species level when always
740 possible. For high-quality MAGS, species were annotated as *Corynebacterium bovis*
741 (n = 1), *Corynebacterium kroppenstedtii* (n= 6), *Enterococcus faecalis* (n =9),
742 *Staphylococcus aureus* (n = 5), *Staphylococcus chromogenes* (n = 4), *Staphylococcus*
743 *epidermidis* (n = 2), *Staphylococcus haemolyticus* (n = 2), *Streptococcus equinus* (n =
744 2), and *Streptococcus uberis* (n = 7). For medium-quality MAGs, *Corynebacterium*
745 *bovis* (n = 5), *Corynebacterium kroppenstedtii* (n = 12), *Corynebacterium urogenitale*
746 (n = 1), *Dermatophilus* spp. (n = 2), *Enterococcus faecalis* (n = 10), *Leuconostoc*
747 *mesenteroides* (n = 1), *Staphylococcus aureus* (n = 2), *Staphylococcus borealis* (n =
748 2), *Staphylococcus chromogenes* (n = 3), *Staphylococcus epidermidis* (n = 3),
749 *Staphylococcus hominis* (n = 2), *Streptococcus bovis* (n = 1), *Streptococcus*
750 *dysgalactiae* (n = 2), *Streptococcus equinus* (n = 1), *Streptococcus gallolyticus* (n =
751 1), and *Streptococcus uberis* (n = 2).

752 The highest diversity of MAGs were found among the low-quality ones, such as
753 *Aerococcaceae* (n = 2), *Cellulomonas shaoxiangyii* (n = 1), *Corynebacterium bovis* (n
754 = 2), *Corynebacterium glutamicum* (n = 2), *Corynebacterium kroppenstedtii* (n = 4),
755 *Corynebacterium rouxii* (n = 1), *Corynebacterium urogenitale* (n = 1), *Corynebacterium*
756 *vitaeruminis* (n = 1), *Enterococcus faecalis* (n = 2), *Enterococcus faecium* (n = 1),
757 *Facklamia miroungae* (n = 1), *Kocuria atrinae* (n = 2), *Kocuria carniphila* (n = 1),
758 *Lysobacter* (n = 1), *Micrococcaceae* bacterium (n = 1), *Peptoniphilus* (n = 1),
759 *Peptoniphilus nemausensis* (n = 1), *Ruminococcus flavefaciens* (n = 1), *Ruoffia*
760 *tabacinasalis* (n = 1), *Staphylococcus aureus* (n = 1), *Staphylococcus chromogenes*
761 (n = 4), *Staphylococcus haemolyticus* (n = 1), *Staphylococcus hominis* (n = 1),
762 *Streptococcaceae* (n = 1), *Streptococcus equinus* (n = 1), *Streptococcus uberis* (n =
763 1), and *Trichococcus flocculiformis* (n = 1).

764 After genome dereplication and the exclusion of MAGs with contamination
765 levels exceeding 10% (four MAGs assigned to the family *Aerococcaceae* and one to
766 the family *Micrococcaceae*), 26 MAGs were selected as representative of the

767 metagenomes. These MAGs were subsequently used to map the reads and estimate
768 their relative abundances across the samples **(Figure 11)**. The MAG assigned as
769 *Cutibacterium acnes* was identified as a putative contaminant in our dataset and
770 removed from downstream analysis.

771 To explore the functional potential of the microbiome in greater detail, we
772 predicted microbial genes from the MAGs of interest using the NCBI Prokaryotic
773 Genome Annotation Pipeline, enabling robust downstream comparisons. The resulting
774 protein sequences were annotated using EggNOG, and the COG categories were
775 normalized to the total number of proteins per MAG. Our analysis focused on the
776 bacterial families *Streptococcaceae*, *Aerococcaceae*, *Staphylococcaceae*, and
777 *Corynebacteriaceae* due to their significant representation within the dataset.

778 As shown in **Figure 12**, the top 20 COG categories primarily comprised proteins
779 associated with metabolism (E: Amino acid transport and metabolism, P: Inorganic ion
780 transport and metabolism, G: Carbohydrate transport and metabolism, C: Energy
781 production and conversion, F: Nucleotide transport and metabolism, H: Coenzyme
782 transport and metabolism, I: Lipid transport and metabolism, and Q: Secondary
783 metabolites biosynthesis, transport, and catabolism), cellular processes and signaling
784 (M: Cell wall/membrane/envelope biogenesis, V: Defense mechanisms, O: Post-
785 translational modification, protein turnover, and chaperones, D: Cell cycle control, cell
786 division, chromosome partitioning, T: Signal transduction mechanisms, and U:
787 Intracellular trafficking, secretion, and vesicular transport), and information storage
788 and processing (L: Replication, recombination, and repair, J: Translation, ribosomal
789 structure, and biogenesis, and K: Transcription). Notably, a substantial proportion of
790 proteins across all bacterial families fell into the “function unknown” group (COG
791 category S), highlighting the large fraction of poorly characterized proteins within these
792 genomes.

793 We aimed to determine which functional categories displayed statistically
794 significant differences in the abundance of annotated proteins across the bacterial
795 families. In total, four COG categories exhibited significant variation. For COG
796 category G (Carbohydrate transport and metabolism), the families *Aerococcaceae* and
797 *Streptococcaceae* demonstrated a higher abundance of orthologs compared to
798 *Corynebacteriaceae* and *Staphylococcaceae*. In COG category Q (Secondary
799 metabolites biosynthesis, transport, and catabolism), *Corynebacteriaceae* displayed
800 the highest abundance of orthologs among the four families. For COG category C

801 (Energy production and conversion), the family *Staphylococcaceae* was enriched
802 compared to *Corynebacteriaceae* and *Streptococcaceae*, although no significant
803 difference was observed when compared to *Aerococcaceae*. Finally, for COG category
804 U (Intracellular trafficking, secretion, and vesicular transport), the family
805 *Streptococcaceae* exhibited a greater abundance of orthologs compared to
806 *Staphylococcaceae*. Overall, the classification of COGs into functional categories
807 provided valuable insights into the metabolic preferences of these bacterial families,
808 highlighting differences in their functional potential and adaptation to specific
809 ecological niches.
810

811 **3.5.2. Multi-Locus Sequence Typing (MLST), Virulence Factors (VFs) and**
812 **Antimicrobial Resistance Genes (AMRGs)**

813 All metagenomes, along with high- and medium-quality MAGs, were subjected
814 to MLST analysis. Results revealed that four bacterial species were associated with a
815 single sequence type (*S. aureus*: ST133; *E. faecalis*: ST40; *S. dysgalactiae*: ST302;
816 *S. uberis*: ST1409). In contrast, two sequence types were identified for *S. epidermidis*
817 (ST100 and ST575), while *S. chromogenes* displayed greater diversity with three
818 sequence types (ST1, ST59, and ST104).

819 The distribution of virulence factors (VFs) across multiple bacterial taxa found
820 in milk samples was analyzed to assess their biosynthetic and pathogenic potential
821 **(Figure 13)**. With regards to *Corynebacterium* spp., *C. bovis* exhibited a predominant
822 focus on regulatory factors, representing 51.85% of its VF genes, followed by stress
823 survival and other factors (18.52% each), and nutritional/metabolic factors at 11.11%.
824 Similarly, *C. kroppenstedtii* showed a strong emphasis on regulatory factors (73.33%),
825 with other VF categories accounting for the remaining 26.67%. In *C. urogenitale*,
826 regulation (66.67%) and other factors (33.33%) were the only VF categories identified.
827 Bacterial species annotated as *E. faecalis* showed a broad distribution of VF
828 categories, with adherence genes comprising the largest fraction (43.90%).
829 Exoenzymes (15.70%), immune modulation (12.50%), and stress survival (12.50%)
830 were also well-represented. Other categories included biofilm formation (11.34%),
831 nutritional/metabolic factors (3.78%), and effector delivery systems (0.29%).

832 For *Staphylococcus* spp., *S. aureus* showed dominant VFs categories
833 associated to immune modulation (24.53%) and exoenzymes (14.34%), followed by
834 nutritional/metabolic factors (13.58%), exotoxins (12.08%), adherence (11.70%), and

835 effector delivery systems (11.70%). Other categories included biofilm formation
836 (6.79%) and stress survival (5.28%). *S. chromogenes* was predominantly associated
837 with immune modulation (51.02%), followed by stress survival (25.51%) and
838 exoenzymes (13.27%). Adherence (7.14%) and biofilm formation (3.06%) were less
839 frequent. In *S. epidermidis*, exoenzymes were most abundant (36.08%), followed by
840 adherence (22.68%), stress survival (19.59%), and immune modulation (18.56%).
841 Exotoxins accounted for a minor proportion (3.09%). *S. hominis* was heavily skewed
842 towards immune modulation (77.27%), with smaller contributions from stress survival,
843 adherence, and exoenzymes (13.64%, 4.55%, and 4.55%, respectively).

844 In species of *Streptococcus*, adherence, immune modulation, and stress
845 survival were equally represented (33.33% each) for *S. bovis*. *S. dysgalactiae*
846 demonstrated a focus on adherence (50%), followed by exoenzymes and stress
847 survival (14.29% each), with other categories such as immune modulation, invasion,
848 and nutritional/metabolic factors at 7.14% each. For *S. equinus*, exotoxins dominated
849 (71.11%), with smaller contributions from immune modulation (13.33%), adherence
850 (8.89%), and stress survival (6.67%). *S. gallolyticus* exhibited a strong preference for
851 immune modulation (63.64%), with stress survival and adherence making up 27.27%
852 and 9.09%, respectively. Finally, *S. uberis* showed a varied distribution, with immune
853 modulation (46.67%) and adherence (28%) being the most common categories,
854 followed by stress survival (13.33%), nutritional/metabolic factors (6.67%),
855 exoenzymes (4%), and exotoxins (1.33%).

856 Since the VFDB database lacks information on members of the *Aerococcaceae*
857 family (as it only includes experimentally verified virulence factors), the prediction of
858 virulence factors for microorganisms associated with this taxon was performed using
859 category V (Defense mechanisms) from the COG database as a reference. This
860 analysis identified six protein-coding genes that may enable bacteria of this taxon to
861 colonize the bovine mammary gland and enhance their pathogenic potential: hrtA
862 (putative hemin import ATP-binding protein), hsdM (site-specific DNA-
863 methyltransferase, adenine-specific), ylbB (efflux ABC transporter, permease protein),
864 yheH (ABC transporter transmembrane region), ywjA (ABC transporter
865 transmembrane region), and XK27_06785 (putative ABC transport system ATP-
866 binding protein, ABC.CD.A).

867 In the analysis of antimicrobial resistance (AMR) genes, a total of 144 genes
868 were annotated using three databases: ResFinder, CARD, and ARG-ANNOT **(Figure**

869 **14)**. The top five antibiotic classes represented in the dataset were fluoroquinolones
870 (15%), macrolides (10.2%), tetracyclines (9.6%), aminoglycosides (7.6%), and
871 rifamycins (5.8%).

872 When stratified by bacterial species, distinct patterns of AMR gene distribution
873 were observed. *Corynebacterium* spp. exhibited a high abundance of genes
874 associated with rifamycins (20.4%), macrolides (14.6%), penams (14.0%),
875 aminoglycosides (12.7%), and aminocoumarins (12.1%). For *Staphylococcus* spp.,
876 the most abundant AMR genes were associated with fluoroquinolones (18.2%),
877 acridine dyes (11.2%), tetracyclines (10.8%), aminoglycosides (8.3%), and phenicols
878 (7.9%). Similarly, *Streptococcus* spp. displayed a predominance of genes linked to
879 fluoroquinolones (24.2%), macrolides (14.2%), beta-lactams (11.1%), rifamycins
880 (10.4%), and tetracyclines (6.6%). For *Enterococcus faecalis*, AMR genes were most
881 frequently associated with macrolides (13.6%), fluoroquinolones (11.3%),
882 lincosamides (8.1%), tetracyclines (8.1%), and aminoglycosides (8.0%).
883 *Dermatophilus* spp. demonstrated a high prevalence of genes for aminocoumarins
884 (22.9%), fluoroquinolones (20%), macrolides (14.3%), ciprofloxacin (8.6%), and
885 oleandomycin (8.6%). In *Kocuria atrinae*, AMR genes were predominantly annotated
886 to rifamycins (35.7%) and fluoroquinolones (28.6%), with additional annotations to
887 aminocoumarins, macrolides, oleandomycin, oxazolidinones, and peptides (each at
888 7.1%). Lastly, *Leuconostoc mesenteroides* had only one annotated AMR gene, which
889 was associated with rifamycins. These findings underscore the diversity of AMR gene
890 profiles across MAGs reconstructed from metagenomes obtained from milk samples,
891 highlighting species-specific adaptations to antimicrobial pressures.
892

893 **3.5.3. Prediction of bacteriocins and bioactive compounds**

894 We used BAGEL4 to identify gene clusters in the MAGs DNA involved in the
895 biosynthesis of Ribosomally synthesized and Post translationally modified Peptides
896 (RiPPs) and (unmodified) bacteriocins. In total, 27 gene clusters were predicted
897 **(Figure 15A)**. Bacterial species annotated as *S. aureus* the highest number of different
898 classes such as 159.2; Pediocin (only predicted in the MAG H249_bin.1, 100% identity
899 with bacteriocin immunity protein from *Pediococcus pentosaceus*), 185.1;BacCH91,
900 putative_bacteriocin (Lactococcin_972 – PF09683), 314.1; Auto_Inducing_Peptide_I,
901 491.1; Lantibiotic, 96.1; putative_lantibiotic (BsaA2), and Sactipeptides. Four areas of
902 interest were identified in MAGs assigned to *E. faecalis* (62.3;enterolysin_A, 63.3;

903 Enterolysin_A, Lasso_peptide, and Sactipeptides), *S. chromogenes* (145.1;
904 Subtilosin_A , 314.1;Auto_Inducing_Peptide_I, 315.1;Auto_Inducing_Peptide_II, and
905 Sactipeptides), *S. uberis* (157.1; uberolysin, 163.2; Penocin_A,
906 42.2;Bovicin_255_peptide, and 61.3;Dysgalacticin), and *C. kroppenstedtii*
907 (189.2;protease-activatedantimicrobialprotein(PAMP), 42.2;Bovicin_255_peptide,
908 84.1;Nukacin_A_(NukacinISK-1), and 97.2;Enterocin_X_chain_beta). Three gene
909 clusters were identified in *S. epidermidis* (Sactipeptides,
910 314.1;Auto_Inducing_Peptide_I, and 64.2;Delta-lysinI), *Dermatophilus* spp.
911 (Sactipeptides, 64.1;Microbisporicin_(NAI-107), and 93.3;Zoocin_A), *S. hominis*
912 (Sactipeptides, 314.1;Auto_Inducing_Peptide_I, and
913 315.1;Auto_Inducing_Peptide_II), and *S. dysgalactiae* (13.1;Butyrivibriocin,
914 219.1;Streptolysin, and 58.1;Macedocin). Two gene clusters were identified in *S.*
915 *equinus* (42.2;Bovicin_255_peptide and 84.1;Nukacin_A_(NukacinISK-1)) and *S.*
916 *haemolyticus* (315.1;Auto_Inducing_Peptide_II and Sactipeptides). Only one gene
917 cluster was predicted for *S. gallolyticus* (LAPs), *S. bovis* (159.2;Pediocin), *S. borealis*
918 (Sactipeptides), and *L. mesenteroides* (97.2;Enterocin_X_chain_beta).
919 Genome mining using AntiSMASH identified 322 biosynthetic gene clusters
920 (BGCs) across all analyzed MAGs. These BGCs were classified into 26 distinct types
921 of biosynthetic pathways associated with the production of secondary/specialized
922 metabolites. The ten most abundant pathways were non-ribosomal peptide synthetase
923 (NRPS) (12.7%), terpene (11.1%), non-ribosomal peptide siderophore (NI-
924 siderophore) (10.8%), type III polyketide synthase (T3PKS) (10.2%), ribosomally
925 synthesized and post-translationally modified peptides (RiPP-like) (9.6%), cyclic
926 lactone autoinducer (9.3%), opine-like metallophore (5.4%), NRPS-like (4.8%),
927 aminopolycarboxylic acid (3.9%), and type I polyketide synthase (T1PKS) (3.0%).
928 When the analysis was stratified by bacterial species and limited to hits with
929 similarity associated with known clusters, several species-specific putative
930 metabolites were identified **(Figure 15B)**. In *C. bovis*, the annotated metabolites
931 included Œµ-Poly-L-lysine, 5-isoprenylindole-3-carboxylate β-D-glycosyl ester (here
932 abbreviated as 5-IPI-β-D-Gly), SF2575, and rhizomide A/B/C. For *C. kroppenstedtii*,
933 the identified metabolites were phenazine SA/SB/SC, [S,S]-ethylenediamine-
934 disuccinate ([S,S]-EDDS), trichrysobactin/cyclic
935 trichrysobactin/chrysobactin/dichrysobactin (here abbreviated as
936 TCB/cTCB/CB/DCB), griseobactin, and ε-Poly-L-lysine. *C. urogenitale* showed the

937 presence of SF2575 and carotenoid. Metabolites detected in *Dermatophilus* spp.
938 included carotenoid, hedamycin, microtermolide A, rakicidin A/B, and ε-Poly-L-lysine.
939 The compound ε-Poly-L-lysine was the sole metabolite identified in *K. atrinae*. For *S.*
940 *bovis*, [S,S]-EDDS was annotated, while macedocin was specific to *S. dysgalactiae*.
941 *S. equinus* displayed streptococcin A-FF22, whereas mutanocyclin/leuvalin/tyrvalin
942 and streptide were annotated for *S. gallolyticus*. In *S. uberis*, the detected putative
943 metabolites included uberolysin, ubericin K, and iturin. *S. aureus* was associated with
944 aureusimine A/B/C, hyicin 3682, staphyloferrin A, staphyloferrin B, staphylopine, and
945 kijanimicin. For *S. borealis*, staphyloferrin A was identified, while *S. chromogenes*
946 yielded staphyloferrin A, staphylopine, and kijanimicin. Staphyloferrin A and
947 staphylopine were predicted in *S. epidermidis*, and *S. haemolyticus* showed
948 staphyloferrin A. Lastly, in *S. hominis*, kijanimicin and staphyloferrin A were annotated.
949 The AntiSMASH analysis revealed a group of bioactive compounds shared among
950 several Staphylococci species, including staphyloferrin A, staphylopine, and
951 kijanimicin. In contrast, unique metabolic signatures were observed in specific
952 species, such as *S. aureus*, *C. kroppenstedtii*, *S. uberis*, and *C. bovis*. These findings
953 suggest a potential common production of certain bioactive compounds, here
954 predominantly identifed among Staphylococci, while other microorganisms, such as
955 *C. kroppenstedtii*, appear to exhibit a more species-specific metabolite spectrum.
956

957 **4. Discussion**

958 Bovine mastitis is widely recognized as the most economically significant disease
959 affecting dairy cattle [66]. Its multifactorial nature poses substantial challenges to
960 effective management through zootechnical and veterinary interventions [67]. In Norway,
961 where the disease is relatively well-controlled and posses one of the lowest
962 antimicrobial use in livestock in the world, studies on microbial ecology offer valuable
963 insights into potential taxa associated with udder health, contributing to advancements
964 in prevention and management strategies [68].

965 Somatic cell count (SCC) is globally used as a proxy for diagnosing intramammary
966 infections and monitoring mastitis control throughout lactation . Elevated SCC values
967 are typically associated with a recent or ongoing infection, often confirmed by a
968 positive bacterial culture or, more recently, by PCR targeting specific taxa [69]. The
969 beginning and end of lactation are generally considered critical periods for udder
970 health due to increased susceptibility to infections. In this study, the lactation period

971 significantly influenced SCC levels, with lower counts observed during early and mid-
972 lactation, while higher counts were recorded near the dry-off period in 2022 and during
973 late lactation in 2023, which is in agreement with the scientific literature [70].

974 Mastitis-causing pathogens are a primary driver of elevated SCC in the udder, with
975 various bacterial species implicated. These pathogens are conventionally classified
976 into two main groups: major and minor mastitis-causing pathogens [3]. In our study,
977 major pathogens such as *S. aureus*, *Enterococcus* spp., and environmental
978 streptococci were prominent, whereas minor pathogens like *S. chromogenes*, *S.*
979 *haemolyticus*, *S. epidermidis*, and *C. bovis* were less frequently reported. Apart from
980 the elevated culturing and detection of *E. faecalis* in milk samples from the mid-
981 lactation period of 2023, our findings align with a recent prevalence study of udder
982 pathogens in Norwegian dairy cows between 2019 and 2020 [71]. The high occurrence
983 of *E. faecalis* coincides with the period when animals transition from freestalls to
984 pasture. As reported elsewhere, *Enterococcus* is considered both a commensal
985 organism and an opportunistic pathogen, possessing a remarkable ability to survive
986 under adverse conditions. It is commonly found as part of the normal gut microbiota
987 in animals, which may facilitate its dissemination under specific environmental or
988 management practices [72].

989 Bacteriological culturing of milk samples aids in guiding mastitis treatment
990 decisions by identifying the causative bacteria of intramammary infections. However,
991 as not all mastitis-causing bacteria require treatment, this approach, from a microbial
992 ecology perspective, overlooks the presence of a substantial fraction of
993 microorganisms within a given sample due to the fastidious nature of some microbial
994 taxa [73]. In our study, two groups of samples classified as negative (i.e., absence of
995 mastitis-causing pathogens according to the gold standard) were identified. The first
996 group (n = 72) exhibited high colony counts (30–300 CFUs) but had somatic cell
997 counts (SCC) below 100,000 cells/mL. The second group (n = 57) also showed high
998 colony counts but had a median SCC of 2.9×10^5 cells/mL. Although these
999 microorganisms are not routinely identified, they may still contribute to elevated SCC
1000 levels. Notably, the genus *Lactococcus* has emerged as a potential group of emerging
1001 mastitis pathogens, yet it is rarely documented in conventional culturing reports [74,75].

1002 In an effort to overcome the limitations of culture-dependent methods and
1003 complement them, numerous studies have utilized metataxonomic analysis to explore
1004 the microbial composition of bovine milk in both health and disease [21,76–78]. In this

longitudinal study, we conducted a comprehensive investigation of the microbial composition at the quarter level in over 300 milk samples, aiming to uncover microbial succession patterns associated with bovine mastitis. The lower diversity indices observed in samples with high somatic cell counts are a common finding in ecological studies of the bovine mammary gland. This reduction in diversity reflects the dominance of specific microbial taxa, which disrupt the homeostasis of the milk microbial community, resulting in a dysbiotic state within the affected quarter [19,23,76].

In this study, we observed that both the composition and relative abundances of bacterial taxa at the phylum and genus levels varied across lactation periods. Specifically, an enrichment of *Proteobacteria* during the early lactation period (EL-2023) was associated with a lower frequency of samples exhibiting high somatic cell counts. This phylum is linked to cases of clinical mastitis, such as those caused by *Escherichia coli*. However, *Proteobacteria* are also frequently observed at high relative abundances in healthy milk samples. In contrast, cases of subclinical mastitis are more commonly associated with the phylum *Firmicutes*, including bacterial species such as *Staphylococcus* spp. [17]. The increase in the phylum *Actinobacteriota* in the milk microbiota during the mid- and late-lactation periods may be attributed to dietary changes and seasonal factors, as it aligns with rising temperatures and the transition to a pasture-based system. Recent studies suggest that both the ruminal microbiota and climatic conditions can significantly impact udder health [79,80].

The microbial diversity observed at the genus level in this study aligns with findings from previous research conducted on the same herd and globally, with taxa such as *Staphylococcus*, *Corynebacterium*, *Streptococcus*, *Enterococcus*, *Aerococcus*, *Kocuria* and *Romboutsia* among the most prevalent ones [28,81,82]. To further investigate the structure and dynamics of the milk microbiome community, we applied a dimensionality reduction method commonly used in human microbiome studies to explore associations with parameters of interest [83]. This approach provided valuable insights, particularly regarding sample classification, bridging the gap between culturomics and culture-independent methods. By combining these approaches, we increased the detection of specific pathogens in samples initially classified as negative, highlighting the potential of this methodology to enhance microbial detection and classification. Conventional bacteriological examination remains the most widely used method for diagnosing bovine intramammary infections, however, it is estimated that over 30% of milk samples yield negative culture results or fail to identify the

1039 causative agent [84,85]. According to Smistad et al. (2023) [71], the proportion of samples
1040 classified as culture-negative in subclinical mastitis cases can exceed 20% in Norway,
1041 while culture-negative samples from clinical mastitis cases account for approximately
1042 10%. Overall, culture-negative mastitis is often associated with bacterial species
1043 present below the detection threshold of culture methods or with fastidious
1044 microorganisms [11,86].

1045 The application of dimensionality reduction enabled the evaluation of the relative
1046 abundance of key bacterial genera within each cluster, offering valuable insights into
1047 their prevalence across various samples. *Staphylococcus*, *Corynebacterium*, and
1048 *Aerococcus* were consistently distributed among the clusters, bringing out their
1049 widespread occurrence and variability in the analyzed milk samples. This observation
1050 confirms the significant prevalence of these taxa in the milk microbiome, irrespective
1051 of whether the sample originates from a healthy udder or a mastitis case.
1052 *Staphylococcus* species stand out for their high contagious potential within herds and
1053 can be present in both healthy and mastitic quarters, with their prevalence being highly
1054 strain-dependent. Although the specific role of *Corynebacterium* spp. in udder health
1055 remains unclear [87], it has been hypothesized that natural infections may confer
1056 protection against major mastitis pathogens, such as *S. aureus*. This is consistent with
1057 our findings and may be explained by the production of antimicrobial compounds or
1058 modulation of the udder immune system, as reported in previous studies [88–90].
1059 Meanwhile, *Aerococcus* has been primarily associated with healthy milk samples and
1060 its occurrence with non-aureus staphylococci have been negatively correlated with *E.*
1061 *coli* in clinical mastitis [26,91].

1062 Within the bacterial genera evaluated in this study, the remarkable diversity of
1063 *Corynebacterium* and *Staphylococcus* species observed through taxonomic analysis
1064 stands out, aligning with findings from the scientific literature. According to Seshadri
1065 et al. (2022) [90], only 30%–50% of the projected actinobacterial phylogenetic diversity
1066 has genomic representation via isolates or metagenome-assembled genomes. This
1067 limited representation may account for the high number of amplicon sequence variants
1068 annotated as uncultured bacteria within the *Corynebacterium* genus in our study,
1069 highlighting the need for further metagenomic research utilizing shotgun sequencing
1070 to improve taxonomic resolution. In a study designed to survey and compare the
1071 microbial diversity of external and teat canal epithelial microbiomes using different
1072 sequencing approaches, Dean et al. (2021)[92] reported that the most diverse

1073 operational taxonomic units across all sampling locations were assigned to the
1074 *Corynebacterium* genus. Among these, *C. efficiens*, *C. marinum*, *C. maris*, and *C.*
1075 *xerosis* constituted the majority of species-level identifications (mean relative
1076 abundance exceeding 1%). Except for *C. bovis*, many *Corynebacterium* spp. present
1077 in milk samples can also be isolated from the housing environment and milking-related
1078 niches [93], reflecting the ubiquity of these microorganisms in the dairy environment.

1079 The majority of the ASVs annotated within the *Staphylococcus* genus belong to the
1080 Non-aureus Staphylococci and Mammaliicocci (NASM) group, one of the most
1081 frequently isolated bacterial groups from milk samples [94]. In this study, NASM species
1082 such as *S. borealis*, *S. chromogenes*, *S. hominis*, *S. epidermidis*, *S. haemolyticus*, *S.*
1083 *xylosus*, *S. saprophyticus*, and *S. simulans* were identified across the milk samples.
1084 Compared to coagulase-positive staphylococci, these species generally exhibit lower
1085 virulence but can occasionally cause mastitis [95,96]. Moreover, they are often
1086 overlooked in routine laboratory diagnostics, as recently reported for *S. borealis* and
1087 *S. rostri* [97]. Nonetheless, NASM has garnered increasing attention due to its genetic
1088 diversity and its ability to antagonize colonization by *S. aureus* [98,99]. This antagonistic
1089 effect is attributed to the production of various antibacterial molecules, such as
1090 bacteriocins, which may play a critical role in shaping microbial dynamics and udder
1091 health [100].

1092 We observed that the composition of milk microbiota varied both at the cow level
1093 and the quarter level. When an intramammary infection was established by
1094 *Staphylococcus*, a persistent dysbiosis was detected in subsequent samples, a
1095 phenomenon not observed for *Streptococcus*. While a considerable number of cross-
1096 sectional studies have examined the bovine mammary gland [68,77,101], there is limited
1097 knowledge regarding the dynamics of milk microbiota at the quarter level in
1098 longitudinal studies over a complete lactation cycle. Our findings align with the
1099 scientific literature, which indicates that once *Staphylococcus* establishes dominance
1100 within a quarter, it becomes challenging for the cow to eliminate it [21]. This aligns with
1101 the wide variety of mechanisms deployed by *S. aureus*, such as biofilm formation, the
1102 emergence of small colony variants, and its capability to invade professional and
1103 nonprofessional cells. These mechanisms protect *S. aureus* from the cow's innate and
1104 adaptive immune responses as well as from antimicrobial treatments [102]. In contrast,
1105 our findings regarding *Streptococcus* are consistent with a recent study by Urrutia-
1106 Angulo et a. (2024) [19] which reported a microbiome profile suggesting full recovery

1107 from dysbiosis caused by *Streptococcus* one and three months after the diagnosis of
1108 clinical mastitis. Based on the ASVs identified for the genus *Streptococcus* in this
1109 study, there is a high prevalence of sequences attributed to *S. uberis*, a primarily
1110 environmental pathogen [103]. This bacterium can cause either persistent subclinical
1111 inflammation or severe inflammation associated with transient mastitis, depending on
1112 the strain involved [104]. These results highlight distinct microbiota dynamics depending
1113 on the causative pathogen and underscore that *S. aureus* remains a major concern.

1114 Through shotgun sequencing and genome-centric metagenomics, we aimed to
1115 gain deeper insights into the functional potential of the microbial community and to
1116 reconstruct genomes to better understand microbial interactions within the udder.
1117 Despite progress in this field, a significant gap remains in the scientific literature, which
1118 can be addressed through more studies incorporating shotgun sequencing to
1119 complement metataxonomic analyses at the quarter level. One major challenge is the
1120 contamination with eukaryotic DNA from mammalian cells during metagenomic DNA
1121 extraction. Combined with the low microbial biomass typically present in milk samples,
1122 this reduces the proportion of microbial reads, resulting in poor resolution of the
1123 microbial community. Our research group recently developed a method to address this
1124 issue [23], although further optimization is required for milk samples with low somatic
1125 cell counts.

1126 The microbial diversity observed from the reconstructed microbial genomes in the
1127 metagenomic sequencing data included the most prevalent bacterial genera identified
1128 through amplicon sequencing (*Staphylococcus*, *Corynebacterium*, *Streptococcus*,
1129 *Enterococcus*, *Aerococcus*, *Kocuria*, and *Romboutsia*). Notably, a diverse array of
1130 draft genomes was detected, particularly within the genera *Staphylococcus* and
1131 *Corynebacterium*. Of particular interest is the high abundance of the MAG *C.*
1132 *kroppenstedtii* HC-IA_H069.1. This species belongs to the *C. kroppenstedtii*-like
1133 group, which includes *C. parakroppenstedtii* and *C. pseudokroppenstedtii* [105]. Species
1134 of this group have gained recognition for their involvement in breast abscesses and
1135 granulomatous mastitis in women [106]. However, the role of *C. kroppenstedtii*-like
1136 strains in bovine udder health and disease remains unclear and warrants further
1137 investigation. Species of the genus *Corynebacterium* are well-known for their ability to
1138 produce secondary metabolites and are routinely used as cell factories, exemplified
1139 by *C. glutamicum* [107]. A recent study conducted by Liu et at. (2024) [108] reports the
1140 discovery of a novel glycolipid involved in the the development of granulomatous

1141 lobular mastitis. This pathogenic factor called corynekropbactin is secreted by *C.*
1142 *parakroppenstedtii* and may chelate iron, cause the death of mammary cells and other
1143 mammary -gland-colonizing bacteria, and increase the levels of inflammatory
1144 cytokines. Studies exploring the adaptation of this species to the bovine mammary
1145 gland, as well as comparative genomics analyses with human isolates, are needed to
1146 better understand its potential implications in bovine mammary gland.

1147 The MAGs *Facklamia* bacterium HC-CA013, *Kocuria atrinae* HC-CA018, *S.*
1148 *equinus* HC-IA_356.3, *Staphylococcus* bacterium HC-IA_319.2, *S. chromogenes* HC-
1149 IA_H042.1, *Peptoniphilus* bacterium HC-CA024, *Corynebacterium* bacterium HC-
1150 CA026, and *Corynebacterium* bacterium HC-IA_212.3 were found at higher coverage
1151 in milk samples with low somatic cell counts. The genera *Facklamia* and *Kocuria* have
1152 been reported as part of the milk microbiota of healthy bovine quarters and human
1153 precolostrum [109,110]. *Facklamia* belongs to the family *Aerococcaceae*, which comprises
1154 nine genera; recently, some species of *Facklamia* were reclassified, such as *F.*
1155 *tabacinasalis*, now *Ruoffia tabacinasalis* [111]. Meanwhile, *Kocuria* is a member of the
1156 family *Micrococcaceae*, and the genus was established by subdividing the genus
1157 Micrococcus [112]. Metabolically, Aerococcaceae are classified as lactic acid bacteria
1158 and can produce lactic acid as the primary product of carbohydrate fermentation [113], a
1159 pathway highlighted in this study by the enrichment of COG category G (Carbohydrate
1160 transport and metabolism). Their metabolism reflects adaptation to low-oxygen,
1161 nutrient-rich environments, such as milk. Although these microorganisms are well-
1162 suited to milk, studies exploring the crosstalk between the identified MAGs and their
1163 roles in bovine udder health remain largely unexplored.

1164 A comprehensive understanding of the genetic variation among mastitis-causing
1165 pathogens in bovine populations, including temporal trends, is essential for the
1166 development of effective mastitis control strategies. This is particularly important
1167 because the severity of the disease and its impact on milk production are, at least in
1168 part, influenced by the genetic and phenotypic characteristics of specific pathogenic
1169 microorganisms. In this study, we conducted an *in silico* multilocus sequence typing
1170 (MLST) analysis on both metagenomic samples and the metagenome-assembled
1171 genomes (MAGs) derived from these datasets. Among the lineages identified, all *S.*
1172 *aureus* sequences were associated with the ST133, which has been recognized as a
1173 globally distributed sequence type predominantly associated with livestock such as
1174 large and small ruminants [114,115]. Regarding *E. faecalis*, all sequences analyzed in this

study were assigned to clonal complex 40 (ST40), the most prevalent clonal type of this species. ST40 is globally distributed, with isolates originating from various sources, including livestock, humans, and environmental reservoirs [116]. This lineage is of particular significance due to its widespread presence in clinical and non-clinical settings. It is frequently isolated from both commensal and pathogenic contexts, demonstrating its adaptability across different environments. *Streptococcus dysgalactiae* ST302 has been previously identified in cows, with evidence suggesting it is well-adapted to this host [117]. In contrast, limited information is available in the scientific literature regarding *Streptococcus uberis* ST1409. However, it appears closely related to sequence type 1405 and has been reported to be sensitive to penicillin, tetracycline, and erythromycin [118]. The ST100 identified in *S. epidermidis* has been described as having an animal origin and is reported to cause mastitis in both ewes and bovines [119]. In contrast, limited information is available regarding ST575. However, one study identified an antimicrobial compound known as epifadin, produced by nasal *S. epidermidis* strain IVK83 belonging to the ST575, which has demonstrated the ability to eliminate *S. aureus* [120]. For *S. chromogenes*, ST1 has been reported as one of the most predominant sequence types in a study investigating differences in the occurrence of intramammary infections in freshly calved heifers in Sweden [121]. In contrast, for ST59 and ST104, further research is needed to elucidate their clinical significance, antimicrobial resistance profiles, and associated virulence factors.

To successfully colonize a host, bacteria possess an array of proteins and other components collectively known as virulence factors [122], although this terminology for commensal bacteria is not appropriate and should be repleced by a more appropriate term such as “niche factors” [123]. Interestingly, many of these factors are not exclusively specialized for host-pathogen interactions but play broader roles in adapting to diverse ecological niches and for cooperation in complex biological ecosystems [124]. The analysis of virulence factors across bacterial taxa in milk samples revealed distinct patterns. According to Tauch & Burkovski (2015) [125], many members within the genus *Corynebacterium* developed into rare pathogens “by chance” and most of the species share a more general adhesion mechanism. In our study, regulatory factors such as iron-dependent repressor (IdeR) and PhoP seem to play an important role in niche adaptation. IdeR in *Mycobacterium tuberculosis* and its homolog, the diphtheria toxin repressor (DtxR) in *Corynebacterium diphtheriae*, are crucial iron-dependent

transcriptional regulators [126,127]. IdeR controls iron acquisition, storage, and virulence-related genes in *M. tuberculosis* [128], whereas PhoP gene is the response regulator of the PhoPR two-component systems, which plays a crucial role in the virulence of C. pseudotuberculosis [129]. In *E. faecalis*, adherence genes constituted the largest fraction of virulence factors in this study. Key colonization factors included the collagen adhesin Ace, the endocarditis antigen EfaA, and the endocarditis- and biofilm-associated pili (Ebp), all of which are commonly present in *E. faecalis* isolates [130,131]. The Esp surface protein is more frequently found in isolates from urinary tract infections and bacteremia [131]. Ebp pili play a major role in adherence to fibrinogen and a minor role in collagen binding, whereas Ace serves as the primary collagen adhesin [132]. Species of *Staphylococcus* employ diverse immune evasion strategies to persist within host organisms. In this study, immune modulation was identified as the predominant category of virulence factors among *Staphylococcus* species, particularly genes involved in capsule formation (*cap8*), *adsA*, and *sbi*. The *cap8* gene cluster, consisting of 16 open reading frames (*cap8A-P*), is responsible for the synthesis of type 8 capsular polysaccharides, with 11 genes directly contributing to its production [133]. Capsular polysaccharides 8, which is predominant among *S. aureus* isolates from humans, can elicit a robust inflammatory response. However, these capsular types are rapidly eliminated compared to capsule-negative *S. aureus*, which are more efficiently internalized by mammary gland epithelial cells and capable of establishing chronic infections in the host [134]. Other virulence factors, such as staphylococcal protein A (SpA), staphylococcal binder of immunoglobulin (Sbi), and adenosine synthase A (AdsA), also play crucial roles in immune evasion [135,136]. Sbi, which is both secreted and cell-envelope associated, enhances bacterial survival in human blood and prevents neutrophil-mediated opsonophagocytosisand modulate B-cell activation [137,138]. Collectively, these mechanisms enable *S. aureus* to evade host immune defenses effectively, facilitating persistent and recurrent infections. *Streptococcus* species possess various virulence factors that contribute to their pathogenicity. Biofilm production is a prevalent trait, observed in over 70% of *Streptococcus* isolates [139]. *S. agalactiae* strains frequently carry genes like *cyl* genes, *cfa*/*cfb*, and *hylB* [140]. The lactoferrin-binding protein is a crucial virulence factor found in *S. dysgalactiae*, *S. uberis* and *S. agalactiae*, with differing effects on cell invasion [141–143]. Notably, *S. agalactiae* strains carrying the *lmb* gene showed enhanced invasion ability and were associated with reduced antibiotic treatment efficacy [140].

Many of the identified virulence factors are directly linked to the high recalcitrance associated with antibiotic use in dairy farming. More recently, shotgun metagenomic sequencing has shown to be a valuable tool to quantify the dynamics of AMR transmission at the human–livestock interface [144], as well to implications for antimicrobial stewardship in livestock [145]. Norway stands out for its notably low rates of antibiotic use and bovine mastitis treatment globally, with benzylpenicillin procaine being the first-line therapy for mastitis [71]. Currently, antimicrobial treatments are primarily reserved for moderate to severe cases of mastitis during lactation and for addressing subclinical mastitis caused by *S. aureus*, *S. dysgalactiae*, *S. uberis*, or *S. agalactiae* at dry-off [71]. *Corynebacterium* species exhibit widespread antibiotic resistance, particularly to macrolides, lincosamides, and streptogramins [146,147]. Additionaly, resistance to β-lactams, clindamycin, erythromycin, azithromycin, ciprofloxacin, and gentamicin is also common among various *Corynebacterium* species. Most strains remain susceptible to vancomycin, minocycline, and linezolid [148]. Given the high prevalence and diversity of *Corynebacterium* species in milk samples, this may be linked to their common resistance to beta-lactams [149,150], potentially contributing to the persistence of these microorganisms within the herd enrolled in this study. Of food safety concern, *S. aureus* has developed resistance to various antibiotics through mechanisms including enzymatic inactivation, target alteration, antibiotic trapping, and efflux pumps [151]. Exposure to biocides and dyes can lead to overexpression of efflux pump genes like *mepA*, *mdeA*, *norA*, and *norC*, contributing to multidrug resistance [152]. Staphylococci of animal origin possess numerous resistance genes conferring resistance to multiple antibiotic classes, including fluoroquinolones, tetracyclines, aminoglycosides, and phenicols. These genes are often carried on mobile genetic elements, facilitating their spread among bacterial populations [153]. Species of *Streptococcus* have also developed resistance to multiple classes of antibiotics in both humans and various animal species, presenting significant challenges for effective treatment [154,155]. Key resistance mechanisms include genetic modifications, alterations in target sites, and the activity of efflux pumps. [156]. The dissemination of resistance genes among streptococci is primarily mediated by transposons and integrative conjugative elements [155]. While streptococci are typically regarded as sensitive to benzylpenicillin and are not routinely tested in Norway, there have been reports of structural modifications in penicillin-binding proteins caused by point mutations in *S. pneumoniae*, *S. agalactiae*, and *S.*

dysgalactiae subsp. equisimilis. These changes have been linked to reduced susceptibility to β-lactam antibiotics [157–159]. Macrolide resistance in *Streptococcus* arises through three primary mechanisms: ribosomal modifications (both post- and pre-transcriptional), active antibiotic expulsion via efflux pumps, and target protection [160], and studies have frequently linked this resistance to the presence of the *ermB* gene in *S. agalactiae*, *S. dysgalactiae*, and *S. uberis* [161,162]. Although fluoroquinolone resistance remains low, it is increasing due to genomic mutations and plasmid-encoded genes, while tetracycline resistance is mediated by ribosomal protection genes, with resistance determinants for macrolides and tetracyclines often co-occurring on mobile genetic elements [163,164]. *E. faecalis* has been identified as a significant pathogen in bovine mastitis and isolates can exhibit high resistance to tetracycline and erythromycin, with resistance rates of 78.6-87.7% and 28.7-79.0%, respectively [165,166]. The most prevalent resistance genes include *tetK*, *tetL*, and *tetM* for tetracycline, as well as *ermB* and *ermC* for erythromycin, although the latter were not identified in our study [165,166].

Bacteriocins produced by LAB play a significant role in shaping microbial communities and maintaining microbiome homeostasis [167]. In the context of bovine mastitis, LAB isolated from the mammary gland of cows have demonstrated potential as probiotic strains for prevention [168]. However, it remains debatable whether the use of LAB aligns with the biology of the mammary gland, as both the bacterial strain and the site of application are critical factors for ensuring effective action without triggering a strong pro-inflammatory response [169]. Given that coagulase-negative staphylococci are more naturally adapted to the mammary gland, they could represent a viable alternative for the prevention or treatment of mastitis. In this study, several gene clusters were identified across different bacterial species, reflecting the diverse range of molecules utilized by microorganisms in the mammary gland for both competition and pathogenesis. *S. aureus* can express lantibiotics like BsaA2, which provides a competitive advantage in community-acquired strains [170]. Nukacins are lantibiotics produced by *Staphylococci* and shows promising potential for controlling bovine mastitis [171,172]. *Staphylococcus* species also utilize peptide-based quorum sensing systems, such as the accessory gene regulator (agr) system, which employs auto-inducing peptides (AIPs) to regulate various cellular processes [173]. Enterolysin A, a bacteriocin produced by *E. faecalis*, exhibits antimicrobial activity against various Gram-positive bacteria, such as *S. aureus* and *Listeria monocytogenes* [174,175]. Another

toxin, the *E. faecalis* cytolysin, is a two-peptide lytic system active against both prokaryotic and eukaryotic, belongs to the lantibiotic family, and requires two post-translationally modified peptides for its activity cells [176,177]. The expression of the cytolysin is carefully controlled by a quorum-sensing mechanism, potentially balancing its defensive function against predators with maintaining stealth [177,178]. *S. agalactiae* strains isolated from bovine mastitis cases in Brazil demonstrated antimicrobial activity against other *Streptococcus* species, with the bacteriocin zoocin A identified in their genomes [179]. Similarly, bacteriocin-producing strains of *Lactobacillus fermentum* and *S. bovis* isolated from raw milk in Thailand exhibited inhibitory effects against *S. dysgalactiae* [180].

Staphylococcus species produce various secondary metabolites that play crucial roles in microbiome interactions and nutrient acquisition. Staphyloferrin A, a highly hydrophilic siderophore, was isolated from *Staphylococcus hyicus* DSM 20459 and found to be active in 37 other staphylococci [181]. This compound, along with staphyloferrin B, helps scavenge scarce iron in the environment. Another metallophore, staphylopine, is also produced by staphylococci to acquire transition metal ions [182]. Siderophores play a pivotal role in iron acquisition for bacterial pathogens associated with bovine mastitis, enabling their survival in iron-restricted environments. For example, *S. aureus* can utilize a variety of exogenous iron sources, including siderophores such as ferrichrome and desferrioxamine, as well as heme-containing proteins [183]. The ability to acquire iron through these mechanisms significantly enhances the competitive fitness of *Staphylococcus* species across diverse ecological niches, contributing to their adaptability and pathogenic potential. A genome-wide CRISPR interference sequencing (CRISPRi-seq) screen carried out by Mårli et al. (2024) [184] identified key fitness determinants essential for *S. aureus* growth and survival in milk. The study revealed several genes with differential fitness, highlighting specific adaptations for thriving in this environment, with metal acquisition emerging as particularly critical. In this study, ε-Poly-l-lysine, a bioactive secondary metabolite with potent antimicrobial properties, was identified across various *Corynebacterium* species. This compound is widely utilized in several Asian countries as a food preservative and is also implicated in shaping the microbial dynamics of cheese and skin ecosystems [185]. The indole-type compound 5-isoprenylindole-3-carboxylate β-D-glycosyl ester has its bioactivity exploited in developing new drugs and it is widely found in actinobacteria [186]. The compound [S,S]-ethylenediamine-

1345 disuccinate ([S,S]-EDDS) was predicted in *C. kroppenstedtii*. According to Costa et a.
1346 2008[187] , metal complexes of the biodegradable ligand EDDS exhibit antimicrobial
1347 activity against fungi and bacteria. Another two noteworthy compound predicted for *C.*
1348 *kroppenstedtii* using AntiSMASH analysis is a triscatecholamide siderophore and a
1349 catechol-peptide structure named griseobactin. These compounds may play a crucial
1350 role in iron acquisition under low-iron conditions in the bovine mammary gland, where
1351 iron availability varies depending on udder health, lactation stage, and daily milk
1352 production [188,189]. Phenazines, versatile secondary metabolites predicted for *C.*
1353 *kroppenstedtii*, are commonly produced by various bacteria, including *Pseudomonas*,
1354 *Burkholderia*, and *Streptomyces* species [190,191]. These compounds exhibit broad-
1355 spectrum antibiotic activity and play a significant role in enhancing the ecological
1356 fitness and pathogenic potential of the producing strains [192,193]. Overall, understanding
1357 the potential production of bacteriocins and bioactive compounds may lead to novel
1358 anti-infective strategies and microbiome modulation approaches to tackle bovine
1359 mastitis.
1360

1361 **5. Conclusion**

1362 Bovine mastitis continues to present a significant challenge to global dairy
1363 production, highlighting the need for innovative strategies to manage this complex,
1364 multi-etiological disease. This longitudinal study offers valuable insights into the udder
1365 microbiome dynamics across lactation stages using both metataxonomic and shotgun
1366 metagenomic approaches. Our results reveal substantial variations in somatic cell
1367 count and microbiota composition, with dominant species identified in high SCC
1368 samples. Through dimensionality reduction, we were able to identify health-promoting
1369 bacteria and examine microbial dynamics at the quarter level, both before and after
1370 pathogen establishment, demonstrating that *Staphylococcus* remains a primary
1371 concern in mastitis pathogenesis. Metagenomic analyses provided a deeper
1372 understanding by uncovering pathogen-specific metabolic pathways and
1373 reconstructing 142 metagenome-assembled genomes. These findings offer crucial
1374 insights into pathogen adaptation, virulence, and resistance mechanisms. The
1375 protective roles of *Aerococcus*, *Kocuria*, and *coryneform* bacteria warrant further
1376 investigation as health-promoting bacteria. Additionally, the identification of bacteriocin
1377 and biosynthetic gene clusters opens potential pathways for the development of
1378 targeted antimicrobial therapies. In conclusion, this study enhances our understanding

of the bovine udder microbiome and its role in mastitis pathophysiology. These insights provide a foundation for microbiome-informed strategies that can improve the health and productivity of dairy herds.

Acknowledgments

This work received economic support from the Norwegian Research Council (grant number 314733), the Faculty of Chemistry. Biotechnology. and Food Science at the Norwegian University of Life Sciences.

Declarations of competing interest

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

Declaration of generative AI in scientific writing

No generative AI has been used in this article.

References

1. Morales-Ubaldo, A. L. *et al.* Bovine mastitis, a worldwide impact disease: Prevalence, antimicrobial resistance, and viable alternative approaches. *Veterinary and Animal Science* vol. 21 Preprint at https://doi.org/10.1016/j.vas.2023.100306 (2023).
2. Diseases of the Mammary Gland. *Vet Med* 1904–2001 (2017) doi:10.1016/B978-0-7020-5246-0.00020-6.
3. Cobirka, M., Tancin, V. & Slama, P. Epidemiology and Classification of Mastitis. *Animals (Basel)* **10**, 2212 (2020).
4. Milk — Enumeration of somatic cells — Part 1: Microscopic method (Reference method). Preprint at https://www.iso.org/standard/40446.html (2008).
5. Adkins, P. R. F. & Middleton, J. R. Methods for Diagnosing Mastitis. *Veterinary Clinics of North America: Food Animal Practice* **34**, 479–491 (2018).
6. Astrup, L. B., Pedersen, K. & Farre, M. Microbiological Diagnoses on Clinical Mastitis—Comparison between Diagnoses Made in Veterinary Clinics versus in Laboratory Applying MALDI-TOF MS. *Antibiotics 2022, Vol. 11, Page 271* **11**, 271 (2022).
7. Taponen, S., Salmikivi, L., Simojoki, H., Koskinen, M. T. & Pyörälä, S. Real-time polymerase chain reaction-based identification of bacteria in milk samples from bovine clinical mastitis with no growth in conventional culturing. *J Dairy Sci* **92**, 2610–2617 (2009).
8. Bradley, A. J., Leach, K. A., Breen, J. E., Green, L. E. & Green, M. J. Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales. *Veterinary Record* **160**, 253–258 (2007).
9. Langhorne, C. *et al.* Bacterial culture and susceptibility test results for clinical mastitis samples from Australia's subtropical dairy region. *J Dairy Sci* **107**, 1151–1163 (2024).

10. Yadav, D. *et al.* Next-Generation sequencing transforming clinical practice and precision medicine. *Clinica Chimica Acta* **551**, 117568 (2023).
11. Oikonomou, G. *et al.* Microbiota of Cow's Milk; Distinguishing Healthy, Sub-Clinically and Clinically Diseased Quarters. *PLoS One* **9**, e85904 (2014).
12. Sakwinska, O. & Bosco, N. Host microbe interactions in the lactating mammary gland. *Front Microbiol* **10**, 448318 (2019).
13. Cheema, S. K., Li, R. & Cameron, S. J. S. Culturomics as a tool to better understand the human milk microbiota and host–microbiota interactions. *Microbiota and Host* **1**, (2023).
14. Oikonomou, G., Machado, V. S., Santisteban, C., Schukken, Y. H. & Bicalho, R. C. Microbial Diversity of Bovine Mastitic Milk as Described by Pyrosequencing of Metagenomic 16s rDNA. *PLoS One* **7**, e47671 (2012).
15. Srithanasuwan, A., Pangprasit, N., Mektrirat, R., Suriyasathaporn, W. & Chuammitri, P. Divergent Immune Responses to Minor Bovine Mastitis-Causing Pathogens. *Vet Sci* **11**, 262 (2024).
16. Hoque, M. N. *et al.* Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis. *Sci Rep* **9**, 13536 (2019).
17. Hoque, M. N. *et al.* Microbiome dynamics and genomic determinants of bovine mastitis. *Genomics* **112**, 5188–5203 (2020).
18. Oikonomou, G. *et al.* Milk Microbiota: What Are We Exactly Talking About? *Frontiers in Microbiology* vol. 11 Preprint at https://doi.org/10.3389/fmicb.2020.00060 (2020).
19. Urrutia-Angulo, L. *et al.* Unravelling the complexity of bovine milk microbiome: insights into mastitis through enterotyping using full-length 16S-metabarcoding. *Anim Microbiome* **6**, 1–14 (2024).
20. Bhar, S. & Bose, T. Evidence from metagenomic study indicate that subclinical mastitis may have a different pathological origin than clinical mastitis. *bioRxiv* 2024.05.24.595548 (2024) doi:10.1101/2024.05.24.595548.
21. Winther, A. R. *et al.* Longitudinal dynamics of the bovine udder microbiota. *Anim Microbiome* **4**, 26 (2022).
22. Winther, A. R., da Silva Duarte, V. & Porcellato, D. Metataxonomic analysis and host proteome response in dairy cows with high and low somatic cell count: a quarter level investigation. *Vet Res* **54**, (2023).
23. Duarte, V. da S. & Porcellato, D. Host DNA depletion methods and genome-centric metagenomics of bovine hindmilk microbiome. *mSphere* **9**, (2024).
24. Farkaš, V. *et al.* Biomarkers for subclinical bovine mastitis: a high throughput TMT-based proteomic investigation. *Vet Res Commun* **48**, 2069–2082 (2024).
25. Satheesan, L. *et al.* Comparative Profiling of Milk Somatic Cells Proteomes Revealed Key Players in Mammary Immune Mechanisms During Mastitis in Tropical Sahiwal (Bos indicus) Cows. *Proteomics Clin Appl* **18**, e202400054 (2024).
26. Jung, D., Park, S., Kurban, D., Dufour, S. & Ronholm, J. The occurrence of Aerococcus urinaeequi and non-aureus staphylococci in raw milk negatively correlates with Escherichia coli clinical mastitis . *mSystems* (2024) doi:10.1128/MSYSTEMS.00362-24.
27. Caballero-Flores, G., Pickard, J. M. & Núñez, G. Microbiota-mediated colonization resistance: mechanisms and regulation. *Nature Reviews Microbiology 2022 21:6* **21**, 347–360 (2022).

28. Porcellato, D., Meisal, R., Bombelli, A. & Narvhus, J. A. A core microbiota dominates a rich microbial diversity in the bovine udder and may indicate presence of dysbiosis. *Sci Rep* **10**, (2020).
29. Hope, A. Laboratory Handbook on Bovine Mastitis. *Aust Vet J* **78**, (2000).
30. Smistad, M., Sølverød, L., Inglingstad, R. A. & Østerås, O. Distribution of somatic cell count and udder pathogens in Norwegian dairy goats. *J Dairy Sci* **104**, (2021).
31. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* Preprint at https://doi.org/10.1038/s41587-019-0209-9 (2019).
32. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* (2016) doi:10.1038/nmeth.3869.
33. Katoh, K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* (2002) doi:10.1093/nar/gkf436.
34. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* **41**, (2013).
35. Davis, N. M., Proctor, Di. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 1–14 (2018).
36. Uritskiy, G. V., DiRuggiero, J. & Taylor, J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, (2018).
37. Krueger, F. Trim Galore!: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. *Babraham Institute* (2015).
38. Li, D. *et al.* MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**, 3–11 (2016).
39. Kang, D. D. *et al.* MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **2019**, (2019).
40. Wu, Y. W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, (2016).
41. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat Methods* **11**, (2014).
42. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**, (2015).
43. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3. *Elife* **10**, (2021).
44. Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology* vol. 35 Preprint at https://doi.org/10.1038/nbt.3893 (2017).
45. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, (2022).
46. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* (2010) doi:10.1371/journal.pone.0009490.
47. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res* **52**, W78–W82 (2024).
48. Aroney, S. T. N. *et al.* CoverM: Read coverage calculator for metagenomics. *Zenodo* (2024) doi:10.5281/ZENODO.10531254.

49. Tatusova, T. *et al.* NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* **44**, 6614–6624 (2016).
50. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**, D309–D314 (2019).
51. Cantalapiedra, C. P., Herṇandez-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol Biol Evol* **38**, 5825–5829 (2021).
52. Quijada, N. M., Rodríguez-Lázaro, D., Eiros, J. M. & Hernández, M. TORMES: an automated pipeline for whole bacterial genome analysis. *Bioinformatics* **35**, 4207–4212 (2019).
53. Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D. & Aarestrup, F. M. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microb Genom* **8**, 000748 (2022).
54. Alcock, B. P. *et al.* CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res* **51**, D690–D699 (2023).
55. Gupta, S. K. *et al.* ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* **58**, 212–220 (2014).
56. Liu, B., Zheng, D., Zhou, S., Chen, L. & Yang, J. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res* **50**, D912–D917 (2022).
57. Jolley, K. A. & Maiden, M. C. J. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11**, 1–11 (2010).
58. Van Heel, A. J. *et al.* BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res* **46**, W278 (2018).
59. Blin, K. *et al.* antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* **51**, W46–W50 (2023).
60. McMurdie, P. J. & Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* (2013) doi:10.1371/journal.pone.0061217.
61. Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc Series B Stat Methodol* **73**, (2011).
62. Wickham, H., François, R., Henry, L. & Müller, K. dplyr: A Grammar of Data Manipulation. R package version. *Media* Preprint at (2019).
63. Barnett, D., Arts, I. & Penders, J. microViz: an R package for microbiome data visualization and statistics. *J Open Source Softw* **6**, (2021).
64. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol* **17**, (2021).
65. Van Der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* **9**, 2579–2605 (2008).
66. Krishnamoorthy, P., Goudar, A. L., Suresh, K. P. & Roy, P. Global and countrywide prevalence of subclinical and clinical mastitis in dairy cattle and buffaloes by systematic review and meta-analysis. *Res Vet Sci* **136**, 561–586 (2021).

67. Ruegg, P. L. A 100-Year Review: Mastitis detection, management, and prevention. *J Dairy Sci* **100**, 10381–10397 (2017).
68. Derakhshani, H., Plaizier, J. C., De Buck, J., Barkema, H. W. & Khafipour, E. Composition and co-occurrence patterns of the microbiota of different niches of the bovine mammary gland: potential associations with mastitis susceptibility, udder inflammation, and teat-end hyperkeratosis. *Anim Microbiome* **2**, 11 (2020).
69. Ramuada, M., Tyasi, T. L., Gumede, L. & Chitura, T. A practical guide to diagnosing bovine mastitis: a review. *Frontiers in Animal Science* **5**, 1504873 (2024).
70. Hagnestam-Nielsen, C., Emanuelson, U., Berglund, B. & Strandberg, E. Relationship between somatic cell count and milk yield in different stages of lactation. *J Dairy Sci* **92**, 3124–3133 (2009).
71. Smistad, M., Bakka, H. C., Sølverød, L., Jørgensen, H. J. & Wolff, C. Prevalence of udder pathogens in milk samples from Norwegian dairy cows recorded in a national database in 2019 and 2020. *Acta Vet Scand* **65**, (2023).
72. Paschoalini, B. R. *et al.* The Emergence of Antimicrobial Resistance and Virulence Characteristics in Enterococcus Species Isolated from Bovine Milk. *Antibiotics* **12**, (2023).
73. Angelopoulou, A. *et al.* Bovine mastitis is a polymicrobial disease requiring a polydiagnostic approach. *Int Dairy J* **99**, 104539 (2019).
74. Rodrigues, M. X., Lima, S. F., Higgins, C. H., Canniatti-Brazaca, S. G. & Bicalho, R. C. The Lactococcus genus as a potential emerging mastitis pathogen group: A report on an outbreak investigation. *J Dairy Sci* **99**, 9864–9874 (2016).
75. Lin, Y. *et al.* Comparative Genomic Analyses of Lactococcus garvieae Isolated from Bovine Mastitis in China. *Microbiol Spectr* **11**, (2023).
76. Alessandri, G. *et al.* Metataxonomic analysis of milk microbiota in the bovine subclinical mastitis. *FEMS Microbiol Ecol* **99**, 1–10 (2023).
77. da Silva Duarte, V. *et al.* Milk microbial composition of Brazilian dairy cows entering the dry period and genomic comparison between Staphylococcus aureus strains susceptible to the bacteriophage vB_SauM-UFV_DC4. *Sci Rep* **10**, 5520 (2020).
78. Oultram, J. W. H., Ganda, E. K., Boulding, S. C., Bicalho, R. C. & Oikonomou, G. A metataxonomic approach could be considered for cattle clinical mastitis diagnostics. *Front Vet Sci* **4**, (2017).
79. Hu, X. *et al.* The Rumen Microbiota Contributes to the Development of Mastitis in Dairy Cows. *Microbiol Spectr* **10**, (2022).
80. Nguyen, T. T., Wu, H. & Nishino, N. An investigation of seasonal variations in the microbiota of milk, feces, bedding, and airborne dust. *Asian-Australas J Anim Sci* **33**, 1858 (2019).
81. Tong, J., Zhang, H., Zhang, Y., Xiong, B. & Jiang, L. Microbiome and Metabolome Analyses of Milk From Dairy Cows With Subclinical Streptococcus agalactiae Mastitis—Potential Biomarkers. *Front Microbiol* **10**, 474819 (2019).
82. Derakhshani, H. *et al.* Invited review: Microbiota of the bovine udder: Contributing factors and potential implications for udder health and mastitis susceptibility. *J Dairy Sci* **101**, 10605–10625 (2018).
83. Xu, X., Xie, Z., Yang, Z., Li, D. & Xu, X. A t-SNE Based Classification Approach to Compositional Microbiome Data. *Front Genet* **11**, 620143 (2020).

84. Schwaiger, K. *et al.* Hot topic: Bovine milk samples yielding negative or nonspecific results in bacterial culturing—The possible role of PCR-single strand conformation polymorphism in mastitis diagnosis. *J Dairy Sci* **95**, 98–101 (2012).
85. Taponen, S., Salmikivi, L., Simojoki, H., Koskinen, M. T. & Pyörälä, S. Real-time polymerase chain reaction-based identification of bacteria in milk samples from bovine clinical mastitis with no growth in conventional culturing. *J Dairy Sci* **92**, 2610–2617 (2009).
86. Kuehn, J. S. *et al.* Bacterial Community Profiling of Milk Samples as a Means to Understand Culture-Negative Bovine Clinical Mastitis. *PLoS One* **8**, e61959 (2013).
87. Lücken, A., Woudstra, S., Wente, N., Zhang, Y. & Krömker, V. Intramammary infections with Corynebacterium spp. in bovine lactating udder quarters. *PLoS One* **17**, e0270867 (2022).
88. Silva, V. M. *et al.* Milk lymphocyte profile and macrophage functions: new insights into the immunity of the mammary gland in quarters infected with Corynebacterium bovis. *BMC Vet Res* **17**, 1–8 (2021).
89. Lam, T. J. G. M. *et al.* Effect of natural infection with minor pathogens on susceptibility to natural infection with major pathogens in the bovine mammary gland. *Am J Vet Res* **58**, (1997).
90. Seshadri, R. *et al.* Expanding the genomic encyclopedia of Actinobacteria with 824 isolate reference genomes. *Cell Genomics* **2**, 100213 (2022).
91. Park, S. *et al.* A longitudinal census of the bacterial community in raw milk correlated with Staphylococcus aureus clinical mastitis infections in dairy cattle. *Anim Microbiome* **4**, 1–13 (2022).
92. Dean, C. J. *et al.* Investigating the cow skin and teat canal microbiomes of the bovine udder using different sampling and sequencing approaches. *J Dairy Sci* **104**, 644–661 (2021).
93. Woudstra, S. *et al.* Reservoirs of Corynebacterium spp. in the Environment of Dairy Cows. *Pathogens* **12**, 139 (2023).
94. Ruiz-Romero, R. A. & Vargas-Bello-Pérez, E. Non-aureus staphylococci and mammaliicocci as a cause of mastitis in domestic ruminants: current knowledge, advances, biomedical applications, and future perspectives – a systematic review. *Veterinary Research Communications* vol. 47 Preprint at https://doi.org/10.1007/s11259-023-10090-5 (2023).
95. Król, J. *et al.* Staphylococcus borealis – A newly identified pathogen of bovine mammary glands. *Vet Microbiol* **286**, 109876 (2023).
96. De Buck, J. *et al.* Non-aureus Staphylococci and Bovine Udder Health: Current Understanding and Knowledge Gaps. *Front Vet Sci* **8**, 658031 (2021).
97. Penati, M. *et al.* An update on non-aureus staphylococci and mammaliicocci in cow milk: unveiling the presence of Staphylococcus borealis and Staphylococcus rostri by MALDI-TOF MS. *Vet Res Commun* **48**, 2555–2561 (2024).
98. Toledo-Silva, B., Oliveira, A. C. D., Souza, F. N., Haesebrouck, F. & De Vliegher, S. Metabolites of non-aureus staphylococci affect the ability of Staphylococcus aureus to adhere to and internalize into bovine mammary epithelial cells. *Vet Res* **54**, 100 (2023).
99. Mahmmod, Y. S., Klaas, I. C., Svennesen, L., Pedersen, K. & Ingmer, H. Communications of Staphylococcus aureus and non-aureus Staphylococcus species from bovine

intramammary infections and teat apex colonization. *J Dairy Sci* **101**, 7322–7333 (2018).

100. Carson, D. A., Barkema, H. W., Naushad, S. & De Buck, J. Bacteriocins of non-aureus staphylococci isolated from bovine milk. *Appl Environ Microbiol* **83**, (2017).
101. Metzger, S. A. *et al.* A cohort study of the milk microbiota of healthy and inflamed bovine mammary glands from dryoff through 150 days in milk. *Front Vet Sci* **5**, 412587 (2018).
102. Zaatout, N., Ayachi, A. & Kecha, M. Staphylococcus aureus persistence properties associated with bovine mastitis and alternative therapeutic modalities. *Journal of Applied Microbiology* vol. 129 Preprint at https://doi.org/10.1111/jam.14706 (2020).
103. Kabelitz, T., Aubry, E., van Vorst, K., Amon, T. & Fulde, M. The Role of Streptococcus spp. in Bovine Mastitis. *Microorganisms* **9**, 1497 (2021).
104. Srithanasuwan, A., Schukken, Y. H., Pangprasit, N., Chuammitri, P. & Suriyasathaporn, W. Different cellular and molecular responses of Bovine milk phagocytes to persistent and transient strains of Streptococcus uberis causing mastitis. *PLoS One* **19**, e0295547 (2024).
105. Luo, Q. *et al.* Classification of 27 Corynebacterium kroppenstedtii -Like Isolates Associated with Mastitis in China and Descriptions of C. parakroppenstedtii sp. nov. and C. pseudokroppenstedtii sp. nov . *Microbiol Spectr* **10**, (2022).
106. Stevenson, D. R. *et al.* Corynebacterium kroppenstedtii breast abscesses in context, a retrospective cohort study. *J Med Microbiol* **71**, 001616 (2022).
107. Milke, L., Kallscheuer, N., Kappelmann, J. & Marienhagen, J. Tailoring Corynebacterium glutamicum towards increased malonyl-CoA availability for efficient synthesis of the plant pentaketide noreugenin. *Microb Cell Fact* **18**, 1–12 (2019).
108. Liu, R. *et al.* Corynebacterium parakroppenstedtii secretes a novel glycolipid to promote the development of granulomatous lobular mastitis. *Signal Transduction and Targeted Therapy 2024 9:1* **9**, 1–17 (2024).
109. Shinozuka, Y. *et al.* Examination of the microbiota of normal cow milk using MinION™ nanopore sequencing. *J Vet Med Sci* **83**, 1620 (2021).
110. Ruiz, L. *et al.* Microbiota of human precolostrum and its potential role as a source of bacteria to the infant mouth. *Scientific Reports 2019 9:1* **9**, 1–13 (2019).
111. Fotedar, R. *et al.* Reclassification of Facklamia ignava, Facklamia sourekii and Facklamia tabacinasalis as Falseniella ignava gen. nov., comb. nov., Hutsoniella sourekii gen. nov., comb. nov., and Ruoffia tabacinasalis gen. nov., comb. nov., and description of Ruoffia halotolerans sp. nov., isolated from hypersaline Inland Sea of Qatar. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology* **114**, 1181–1193 (2021).
112. Stackebrandt, E., Koch, C., Gvozdiak, O. & Schumann, P. Taxonomic dissection of the genus Micrococcus: Kocuria gen. nov., Nesterenkonia gen. nov., Kytococcus gen. nov., Dermacoccus gen. nov., and Micrococcus cohn 1872 gen. emend. *Int J Syst Bacteriol* **45**, 682–692 (1995).
113. Stackebrandt, E. The family aerococcaceae. *The Prokaryotes: Firmicutes and Tenericutes* **9783642301209**, 3–6 (2014).
114. Naushad, S., Nobrega, D. B., Naqvi, S. A., Barkema, H. W. & De Buck, J. Genomic Analysis of Bovine Staphylococcus aureus Isolates from Milk To Elucidate Diversity and Determine the Distributions of Antimicrobial and Virulence Genes and Their Association with Mastitis. *mSystems* **5**, (2020).

115. Merz, A., Stephan, R. & Johler, S. Staphylococcus aureus isolates from goat and sheep milk seem to be closely related and differ from isolates detected from bovine milk. *Front Microbiol* **7**, 184387 (2016).
116. Zischka, M. *et al.* Comprehensive molecular, genomic and phenotypic analysis of a major clone of Enterococcus faecalis MLST ST40. *BMC Genomics* **16**, 175 (2015).
117. Porcellato, D. *et al.* Whole genome sequencing reveals possible host species adaptation of Streptococcus dysgalactiae. *Scientific Reports 2021 11:1* **11**, 1–13 (2021).
118. Myrenås, M., Fasth, C., Waller, K. P. & Pedersen, K. Genomic analyses of Streptococcus uberis reveal high diversity but few antibiotic resistance genes. *Vet Microbiol* **300**, 110319 (2025).
119. Vasileiou, N. G. C. *et al.* Antimicrobial Agent Susceptibility and Typing of Staphylococcal Isolates from Subclinical Mastitis in Ewes. *https://home.liebertpub.com/mdr* **25**, 1099–1110 (2019).
120. Torres Salazar, B. O. *et al.* Commensal production of a broad-spectrum and short-lived antimicrobial peptide polyene eliminates nasal Staphylococcus aureus. *Nature Microbiology 2023 9:1* **9**, 200–213 (2023).
121. Persson Waller, K., Lundberg, Börjesson, S. & Nyman, A. K. Intramammary infections and risk factors in freshly calved heifers in Swedish dairy herds. *J Dairy Sci* **106**, 7893–7907 (2023).
122. Sharma, A. K. *et al.* Bacterial Virulence Factors: Secreted for Survival. *Indian J Microbiol* **57**, 1–10 (2017).
123. Hill, C. Virulence or Niche Factors: What's in a Name? *J Bacteriol* **194**, 5725 (2012).
124. Herzberg, C., van Meegen, E. N. & van Hasselt, J. G. C. Interplay of virulence factors shapes ecology and treatment outcomes in polymicrobial infections. *Math Biosci* **377**, 109293 (2024).
125. Tauch, A. & Burkovski, A. Molecular armory or niche factors: virulence determinants of Corynebacterium species. *FEMS Microbiol Lett* **362**, 185 (2015).
126. Schmitt, M. P., Predich, M., Doukhan, L., Smith, I. & Holmes, R. K. Characterization of an iron-dependent regulatory protein (IdeR) of Mycobacterium tuberculosis as a functional homolog of the diphtheria toxin repressor (DtxR) from Corynebacterium diphtheriae. *Infect Immun* **63**, (1995).
127. Marcos-Torres, F. J., Juniar, L. & Griese, J. J. The molecular mechanisms of the bacterial iron sensor IdeR. *Biochemical Society Transactions* vol. 51 Preprint at https://doi.org/10.1042/BST20221539 (2023).
128. Chou, C. J. *et al.* Functional studies of the Mycobacterium tuberculosis iron-dependent regulator. *Journal of Biological Chemistry* **279**, (2004).
129. Tiwari, S. *et al.* C. pseudotuberculosis Phop confers virulence and may be targeted by natural compounds. *Integrative Biology* **6**, 1088–1099 (2014).
130. Kafil, H. S., Mobarez, A. M. & Moghadam, M. F. Adhesion and virulence factor properties of Enterococci isolated from clinical samples in Iran. *Indian J Pathol Microbiol* **56**, (2013).
131. Creti, R. *et al.* Survey for virulence determinants among Enterococcus faecalis isolated from different sources. *Journal of Medical Microbiology* vol. 53 Preprint at https://doi.org/10.1099/jmm.0.05353-0 (2004).
132. Nallapareddy, S. R., Singh, K. V., Sillanpää, J., Zhao, M. & Murray, B. E. Relative contributions of Ebp Pili and the collagen adhesin ace to host extracellular matrix

protein adherence and experimental urinary tract infection by Enterococcus faecalis OG1RF. *Infect Immun* **79**, (2011).

133. Sau, S. *et al.* The Staphylococcus aureus allelic genetic loci for serotype 5 and 8 capsule expression contain the type-specific genes flanked by common genes. *Microbiology (N Y)* **143**, (1997).

134. Tuchscherr, L. P. N. *et al.* Capsule-negative Staphylococcus aureus induces chronic experimental mastitis in mice. *Infect Immun* **73**, 7932–7937 (2005).

135. Kim, H. K., Thammavongsa, V., Schneewind, O. & Missiakas, D. Recurrent infections and immune evasion strategies of Staphylococcus aureus. *Current Opinion in Microbiology* vol. 15 Preprint at https://doi.org/10.1016/j.mib.2011.10.012 (2012).

136. Darisipudi, M. N. & Bröker, B. M. How S. aureus blinds the inflammasome to escape immune control. *EBioMedicine* **71**, 103549 (2021).

137. Smith, E. J., Visai, L., Kerrigan, S. W., Speziale, P. & Foster, T. J. The Sbi protein is a multifunctional immune evasion factor of Staphylococcus aureus. *Infect Immun* **79**, (2011).

138. Dunphy, R. W. *et al.* Staphylococcal Complement Evasion Protein Sbi Stabilises C3d Dimers by Inducing an N-Terminal Helix Swap. *Front Immunol* **13**, 892234 (2022).

139. Kaczorek, E., Małaczewska, J., Wójcik, R. & Siwicki, A. K. Biofilm production and other virulence factors in Streptococcus spp. isolated from clinical cases of bovine mastitis in Poland. *BMC Vet Res* **13**, (2017).

140. Torres, G., Macias, D., Reyes-Velez, J., Rios-Agudelo, P. & Caraballo-Guzmán, A. Streptococcus agalactiae virulence factors isolated from bovine mastitis and antibiotic treatment response. *J Appl Microbiol* **134**, (2023).

141. O'Halloran, F., Beecher, C., Chaurin, V., Sweeney, T. & Giblin, L. Lactoferrin affects the adherence and invasion of Streptococcus dysgalactiae ssp. dysgalactiae in mammary epithelial cells. *J Dairy Sci* **99**, 4619–4628 (2016).

142. Moshynskyy, I. *et al.* Characterization of a bovine lactoferrin binding protein of Streptococcus uberis. *Microb Pathog* **35**, 203–215 (2003).

143. Rainard, P. Binding of bovine lactoferrin to Streptococcus agalactiae. *FEMS Microbiol Lett* **98**, 235–239 (1992).

144. Wee, B. A., Muloi, D. M. & van Bunnik, B. A. D. Quantifying the transmission of antimicrobial resistance at the human and livestock interface with genomics. *Clinical Microbiology and Infection* **26**, 1612 (2020).

145. Rubiola, S. *et al.* Shotgun metagenomic sequencing of bulk tank milk filters reveals the role of Moraxellaceae and Enterobacteriaceae as carriers of antimicrobial resistance genes. *Food Research International* **158**, 111579 (2022).

146. Ortiz-Pérez, A. *et al.* High frequency of macrolide resistance mechanisms in clinical isolates of Corynebacterium species. *Microbial Drug Resistance* **16**, (2010).

147. Olender, A. Antibiotic resistance and detection of the most common mechanism of resistance (MLSB) of opportunistic Corynebacterium. *Chemotherapy* **59**, (2014).

148. Barberis, C. M. *et al.* Comparison between disk diffusion and agar dilution methods to determine in vitro susceptibility of Corynebacterium spp. clinical isolates and update of their susceptibility. *J Glob Antimicrob Resist* **14**, (2018).

149. Soriano, F., Zapardiel, J. & Nieto, E. Antimicrobial susceptibilities of Corynebacterium species and other non-spore-forming gram-positive bacilli to 18 antimicrobial agents. *Antimicrob Agents Chemother* **39**, 208–214 (1995).

150. Feßler, A. T. & Schwarz, S. Antimicrobial Resistance in Corynebacterium spp., Arcanobacterium spp., and Trueperella pyogenes. *Microbiol Spectr* **5**, (2017).
151. Brdová, D., Ruml, T. & Viktorová, J. Mechanism of staphylococcal resistance to clinically relevant antibiotics. *Drug Resistance Updates* **77**, 101147 (2024).
152. Huet, A. A., Raygada, J. L., Mendiratta, K., Seo, S. M. & Kaatz, G. W. Multidrug efflux pump overexpression in Staphylococcus aureus after single and multiple in vitro exposures to biocides and dyes. *Microbiology (N Y)* **154**, (2008).
153. Schwarz, S. *et al.* Antimicrobial Resistance among Staphylococci of Animal Origin. *Microbiol Spectr* **6**, (2018).
154. Samanta, I. & Bandyopadhyay, S. Streptococcus. *Antimicrobial Resistance in Agriculture* 217–232 (2020) doi:10.1016/B978-0-12-815770-1.00017-1.
155. Haenni, M., Lupo, A. & Madec, J.-Y. Antimicrobial Resistance in Streptococcus spp . *Microbiol Spectr* **6**, (2018).
156. Alves-Barroco, C., Rivas-García, L., Fernandes, A. R. & Baptista, P. V. Tackling Multidrug Resistance in Streptococci – From Novel Biotherapeutic Strategies to Nanomedicines. *Front Microbiol* **11**, 579916 (2020).
157. Yu, D., Guo, D., Zheng, Y. & Yang, Y. A review of penicillin binding protein and group A Streptococcus with reduced-β-lactam susceptibility. *Front Cell Infect Microbiol* **13**, 1117160 (2023).
158. Fuursted, K. *et al.* Description and characterization of a penicillin-resistant Streptococcus dysgalactiae subsp. equisimilis clone isolated from blood in three epidemiologically linked patients. *Journal of Antimicrobial Chemotherapy* **71**, (2016).
159. Metcalf, B. J. *et al.* Short-read whole genome sequencing for determination of antimicrobial resistance mechanisms and capsular serotypes of current invasive Streptococcus agalactiae recovered in the USA. *Clinical Microbiology and Infection* **23**, (2017).
160. Berbel, D., González-Díaz, A., López de Egea, G., Càmara, J. & Ardanuy, C. An Overview of Macrolide Resistance in Streptococci: Prevalence, Mobile Elements and Dynamics. *Microorganisms* **10**, 2316 (2022).
161. Kamińska, D. *et al.* Macrolide and lincosamide resistance of Streptococcus agalactiae in pregnant women in Poland. *Scientific Reports 2024 14:1* **14**, 1–8 (2024).
162. Loch, I. M., Glenn, K. & Zadoks, R. N. Macrolide and lincosamide resistance genes of environmental streptococci from bovine milk. *Vet Microbiol* **111**, 133–138 (2005).
163. Malhotra-Kumar, S., Lammens, C., Piessens, J. & Goossens, H. Multiplex PCR for simultaneous detection of macrolide and tetracycline resistance determinants in streptococci. *Antimicrob Agents Chemother* **49**, (2005).
164. Cherazard, R. *et al.* Antimicrobial Resistant Streptococcus pneumoniae: Prevalence, Mechanisms, and Clinical Implications. *Am J Ther* **24**, (2017).
165. Yang, F. *et al.* Short communication: Antimicrobial resistance and virulence genes of Enterococcus faecalis isolated from subclinical bovine mastitis cases in China. *J Dairy Sci* **102**, (2019).
166. Goksel, E. *et al.* Distribution of antibiotic resistance genes in enterococcus spp. isolated from mastitis bovine milk. *Acta Vet Brno* **66**, (2016).
167. Zhang, D. *et al.* A systematically biosynthetic investigation of lactic acid bacteria reveals diverse antagonistic bacteriocins that potentially shape the human microbiome. *Microbiome* **11**, 1–20 (2023).

168. Espeche, M. C., Otero, M. C., Sesma, F. & Nader-Macias, M. E. F. Screening of surface properties and antagonistic substances production by lactic acid bacteria isolated from the mammary gland of healthy and mastitic cows. *Vet Microbiol* **135**, 346–357 (2009).
169. Rainard, P. & Foucras, G. A Critical Appraisal of Probiotics for Mastitis Control. *Front Vet Sci* **5**, (2018).
170. Daly, K. M. *et al.* Production of the Bsa lantibiotic by community-acquired Staphylococcus aureus strains. *J Bacteriol* **192**, (2010).
171. Sadaoka, N. *et al.* Opposing genetic polymorphisms of two ABC transporters contribute to the variation of nukacin resistance in Streptococcus mutans. *Appl Environ Microbiol* **90**, (2024).
172. Ceotto, H. *et al.* Nukacin 3299, a lantibiotic produced by Staphylococcus simulans 3299 identical to nukacin ISK-1. *Vet Microbiol* **146**, 124–131 (2010).
173. Thoendel, M., Kavanaugh, J. S., Flack, C. E. & Horswill, A. R. Peptide signaling in the Staphylococci. *Chem Rev* **111**, (2011).
174. Nilsen, T., Nes, I. F. & Holo, H. Enterolysin A, a cell wall-degrading bacteriocin from Enterococcus faecalis LMG 2333. *Appl Environ Microbiol* **69**, (2003).
175. Zhang, T. *et al.* Molecular cloning and antimicrobial activity of enterolysin A and helveticin J of bacteriolysins from metagenome of Chinese traditional fermented foods. *Food Control* **31**, 499–507 (2013).
176. Coburn, P. S. & Gilmore, M. S. The Enterococcus faecalis cytolysin: A novel toxin active against eukaryotic and prokaryotic cells. *Cellular Microbiology* vol. 5 Preprint at https://doi.org/10.1046/j.1462-5822.2003.00310.x (2003).
177. Cox, C., Coburn, P. & Gilmore, M. Enterococcal Cytolysin: A Novel Two Component Peptide System that Serves as a Bacterial Defense Against Eukaryotic and Prokaryotic Cells. *Curr Protein Pept Sci* **6**, (2005).
178. Van Tyne, D., Martin, M. J. & Gilmore, M. S. Structure, function, and biology of the Enterococcus faecalis cytolysin. *Toxins* vol. 5 Preprint at https://doi.org/10.3390/toxins5050895 (2013).
179. Amaral, J. R. V. *et al.* Bacteriocin Producing Streptococcus agalactiae Strains Isolated from Bovine Mastitis in Brazil. *Microorganisms* **10**, (2022).
180. Chaimanee, V., Sakulsingharoj, C., Deejing, S., Seetakoses, P. & Niamsup, P. Screening and characterisation of bacteriocin-producing bacteria capable of inhibiting the growth of bovine mastitis. *Maejo International Journal of Science and Technology* **3**, (2009).
181. Meiwes, J. *et al.* Isolation and characterization of staphyloferrin A, a compound with siderophore activity from Staphylococcus hyicus DSM 20459. *FEMS Microbiol Lett* **67**, (1990).
182. Torres Salazar, B. O., Heilbronner, S., Peschel, A. & Krismer, B. Secondary Metabolites Governing Microbiome Interaction of Staphylococcal Pathogens and Commensals. *Microbial Physiology* vol. 31 Preprint at https://doi.org/10.1159/000517082 (2021).
183. Diarra, M. S., Petitclerc, D. & Lacasse, P. Response of Staphylococcus aureus isolates from bovine mastitis to exogenous iron sources. *J Dairy Sci* **85**, (2002).
184. Mårli, M. T. *et al.* Genome-wide analysis of fitness determinants of Staphylococcus aureus during growth in milk. *bioRxiv* 2024.10.31.621336 (2024) doi:10.1101/2024.10.31.621336.

185. Jiang, X. *et al.* Distribution of ε-Poly-L-Lysine Synthetases in Coryneform Bacteria Isolated from Cheese and Human Skin. *Appl Environ Microbiol* **87**, 1–8 (2021).
186. P, V. *et al.* Streptomyces sp. VITGV100: An endophyte from Lycopersicon esculentum as new source of indole type compounds. *Biochem Syst Ecol* **105**, (2022).
187. Costa, N. J. S. *et al.* Antimicrobial Activity of Ethylenediaminedisuccinate Metal Complexes. Short Communication. *Chem Biodivers* **5**, 2156–2159 (2008).
188. Hyvönen, P. *et al.* Concentrations of bovine lactoferrin and citrate in milk during experimental endotoxin mastitis in early-versus late-lactating dairy cows. *Journal of Dairy Research* **77**, (2010).
189. Patzer, S. I. & Braun, V. Gene cluster involved in the biosynthesis of griseobactin, a catechol-peptide siderophore of Streptomyces sp. ATCC 700974. *J Bacteriol* **192**, (2010).
190. Mavrodi, D. V. *et al.* Diversity and evolution of the Phenazine Biosynthesis Pathways. *Appl Environ Microbiol* **76**, (2010).
191. Hendry, S. *et al.* Functional Analysis of Phenazine Biosynthesis Genes in Burkholderia spp. *Appl Environ Microbiol* **87**, (2021).
192. Guttenberger, N., Blankenfeldt, W. & Breinbauer, R. Recent developments in the isolation, biological function, biosynthesis, and synthesis of phenazine natural products. *Bioorganic and Medicinal Chemistry* vol. 25 Preprint at https://doi.org/10.1016/j.bmc.2017.01.002 (2017).
193. Fitzpatrick, D. A. Lines of evidence for horizontal gene transfer of a phenazine producing operon into multiple bacterial species. *J Mol Evol* **68**, (2009).

Captions

Figure 1 – Raincloud plot **(A)** illustrating somatic cell count (log10(SCC)) across different lactation periods (DO-2022, EL-2023, ML-2023, and LL-2023). Individual samples are represented by dots, color-coded as high SCC (red) or low SCC (light blue). Side histograms depict data distribution, while central bars indicate the median and interquartile range. Statistical comparisons between lactation periods are denoted by significance levels: ***$p < 0.001$, **$p < 0.01$, and NS (not significant). Density plot **(B)** showing the relationship between SCC and bacterial count (log10(IBC)) across lactation periods, stratified by bacterial species identified using MALDI-ToF. Density curves represent distinct pathogens (*E. faecalis*, *S. aureus*, *S. chromogenes*, *S. epidermidis*, and *S. uberis*) and samples classified as "negative", highlighting variations in bacterial prevalence and load among lactation periods.

Figure 2 –Bubble chart **(A)** displaying the relationship between somatic cell count (SCC) and individual bacterial count (IBC) across samples, categorized by bacterial species. Bubble sizes represent CFU/mL on blood agar, while colors indicate bacterial species (*E. faecalis*, *S. aureus*, *S. chromogenes*, *S. epidermidis*, *S. uberis*), and milk samples classified as "negative" for specific mastitis-causing pathogens. Box plots comparing Shannon **(B)** and Simpson **(C)** diversity indices between high and low SCC groups across different time points and lactation periods. Statistical significance is denoted as ***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$. Volatility plot of Shannon **(D)** and Simpson **(E)** diversity indices over time, showing trends across lactation periods (DO-2023, EL-2023, ML-2023, and LL-2023). Solid and dashed lines represent high and low SCC groups, respectively, with shaded areas indicating 95% confidence intervals with group-wise comparisons highlighted.

Figure 3 – Non-metric multidimensional scaling (NMDS) plot based on unweighted **(A)** and weighted **(B)** UniFrac distances, illustrating microbial community composition across different lactation periods (DO-2022, EL-2023, ML-2023, LL-2023) and SCC levels (high and low). Points represent individual samples, with colors indicating SCC levels and shapes representing lactation periods. Marginal density plots along the axes depict the distribution of samples by MDS dimensions.

Figure 4 – Box plots **(A–F)** showing the distribution of the relative abundance of the six dominant genera (*Corynebacterium*, *Staphylococcus*, *Enterococcus*, *Streptococcus*, *Lactococcus*, and *Aerococcus*) across the 18 microbiota clusters. Each box represents the interquartile range (IQR), with the median indicated by a horizontal line and whiskers extending to 1.5 times the IQR. Outliers are shown as individual points. Statistical significance of differences in relative abundance between clusters and the global mean is denoted above each box plot ($p < 0.05$; $p < 0.01$; *$p < 0.001$; **$p < 0.0001$; ns = not significant).

Figure 5 – Distribution of bacterial count **(A)**, somatic cell count **(B)**. and alpha diversity indices (**C** and **D**) across the 18 microbiota clusters. Each box represents the interquartile range (IQR), with the median indicated by a horizontal line and whiskers extending to 1.5 times the IQR. Outliers are shown as individual points. Statistical significance of differences in relative abundance between clusters and the global mean is denoted above each box plot ($p < 0.05$; $p < 0.01$; *$p < 0.001$; **$p < 0.0001$; ns = not significant).

Figure 6 – Stacked bar plots illustrating the relative abundance of the most abundant amplicon sequence variants (ASVs) across microbiota clusters. Each bar represents the proportional composition of ASVs within a specific cluster, with colors corresponding to distinct ASVs as shown in the legend.

Figure 7 – Udder microbiota dynamics. Stacked bar plots depicting the relative abundance of the 20 most abundant genera across different sampling quarters. Each bar represents the proportional composition of genera, with colors corresponding to specific genera as indicated in the legend. Genera with lower abundances are grouped under "Others." Bacterial and somatic cell counts are overlaid: triangles connected by dashed lines represent IBC levels, while dots connected by solid lines indicate SCC levels, illustrating changes in microbiota composition in relation to these metrics over sampling time. The analyzed animal and sampling quarter, along with microbiota clusters, are labeled above each panel, with sample identification numbers displayed below each plot. Samples highlighted in red indicate the detection of a mastitis-causing pathogen identified via MALDI-ToF analysis.

Figure 8 – Heatmap **(A)** showing the most abundant MetaCyc pathways detected across 73 metagenomic samples analyzed using HUMAnN3. Hierarchical clustering was applied to both samples and pathways, with samples annotated by lactation period, somatic cell count, and pathogen (identified by MALDI-Tof). Pathway abundance is represented by a red-to-blue color gradient, where darker red indicates higher abundance and darker blue indicates lower abundance. Bar plots **(B)** display the significant MetaCyc pathways (q-value < 0.25) associated with high and low SCC groups, with coefficients categorized as positive (blue) or negative (red) to illustrate their direction of association. Boxplots **(C)** depict the relative abundance of significant MetaCyc pathways stratified by SCC levels, highlighting pathway abundance differences between the two groups.

Figure 9 – Abundance distribution of significant MetaCyc pathways stratified by bacterial species: (A) *S. aureus*, (B) *S. chromogenes*, (C) *S. epidermidis*, (D) *E. faecalis*, and (E) *S. uberis*. Pathway abundances are shown as boxplots, grouped by species, with pathways color-coded according to their coefficients. Positive and negative coefficients are represented by gradients of blue and red, respectively, highlighting pathways significantly associated with each bacterial species. An empty white dot indicates the mean abundance for each pathway.

Figure 10 – Shared and unique MetaCyc pathways across microorganisms. The Venn diagram shows the number and percentage of pathways that are unique or shared among different species. The shading gradient reflects pathway abundance, with darker shades indicating higher values.

Figure 11 – Phylogenetic tree and relative abundance of the species. From left to right tree shows pie charts with species's abundance in samples with Low (blue) and High (red) somatic cell count, phylogenetic relationships colored at the family level, barplots reporting genome completeness/contamination and heatmap representing relative abundance of species per sample.

Figure 12 – Comparative analysis of clusters of orthologous groups categories among different bacterial families. Bar plots represent the top 20 COG categories across all samples for the bacterial families: *Aeroococcaceae* **(A)**, *Corynebacteriaceae* **(B)**,

Staphylococcaceae **(C)**, and *Streptococcaceae* **(D)**. Each color indicates a specific metagenome-assembled genome as shown in the legend. COG categories - (S: Function unknown, L: Replication, E: Amino acid transport and metabolism, J: Translation, ribosomal structure and biogenesis, K: Transcription, P: Inorganic ion transport and metabolism, G: Carbohydrate transport and metabolism, M: Cell wall/membrane/envelope biogenesis, C: Energy production and conversion, F: Nucleotide transport and metabolism, H: Coenzyme transport and metabolism, V: Defense mechanisms, O: Posttranslational modification, protein turnover, chaperones, I: Lipid transport and metabolism, D: Cell cycle control, cell division, chromosome partitioning, T: Signal transduction mechanisms, U: Intracellular trafficking, secretion, and vesicular transport, Q: Secondary metabolites biosynthesis, transport and catabolism – are represented on the x-axis, and their normalized proportion on the y-axis. Boxplots **(E–H)** represent COG categories that differed significantly among bacterial families based on the Kruskal-Wallis test. Significant pairwise differences are indicated with * ($p<0.05$) and ns (not significant).

Figure 13 – Sankey plot illustrating the relationship between bacterial taxa, virulence factor categories, and specific virulence factors. The thickness of the connecting lines represents the relationships between the taxa, virulence factor categories, and specific virulence factors. Color coding differentiates bacterial taxa and highlights the distribution of virulence traits across the species.

Figure 14 – Dot plot showing the distribution of antibiotic resistance genes across metagenome-assembled genomes. Dot sizes are proportional to the frequency of each gene as annotated by three databases: ResFinder, CARD, and ARG-ANNOT. Each dot is color-coded by MAG, highlighting the presence and prevalence of specific ARGs across different MAGs. Larger dots indicate higher consensus across the databases regarding the detection of that gene.

Figure 15 – Chord diagram illustrating the relationships between bacteriocin **(A)** gene clusters predicted using BAGEL4 across the different MAGs. Chord diagram depicting the diversity and distribution of secondary metabolite gene clusters identified with antiSMASH **(B)** across the same MAGs. Connections highlight shared features or

2040 pathways among the identified clusters, providing insights into the biosynthetic
2041 capabilities within each taxon
2042
2043

Figures

Figure 1

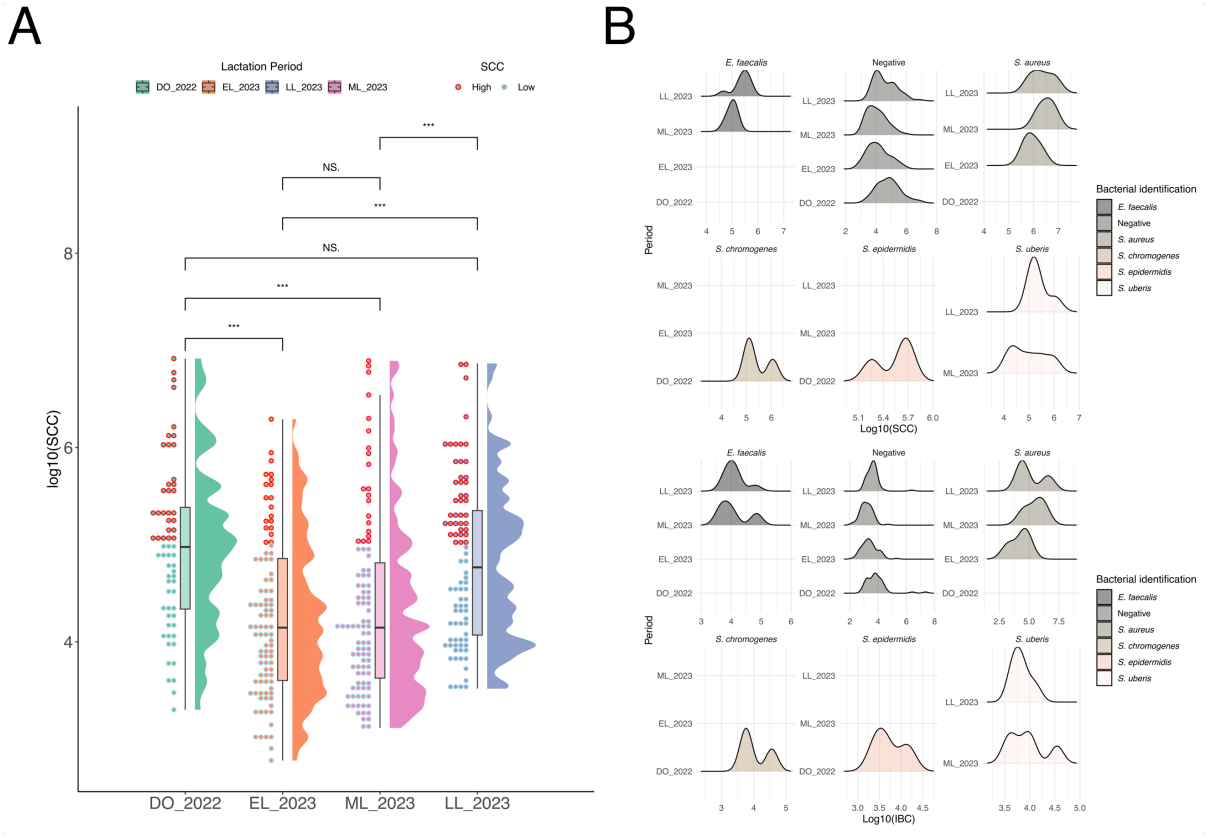


Figure 2

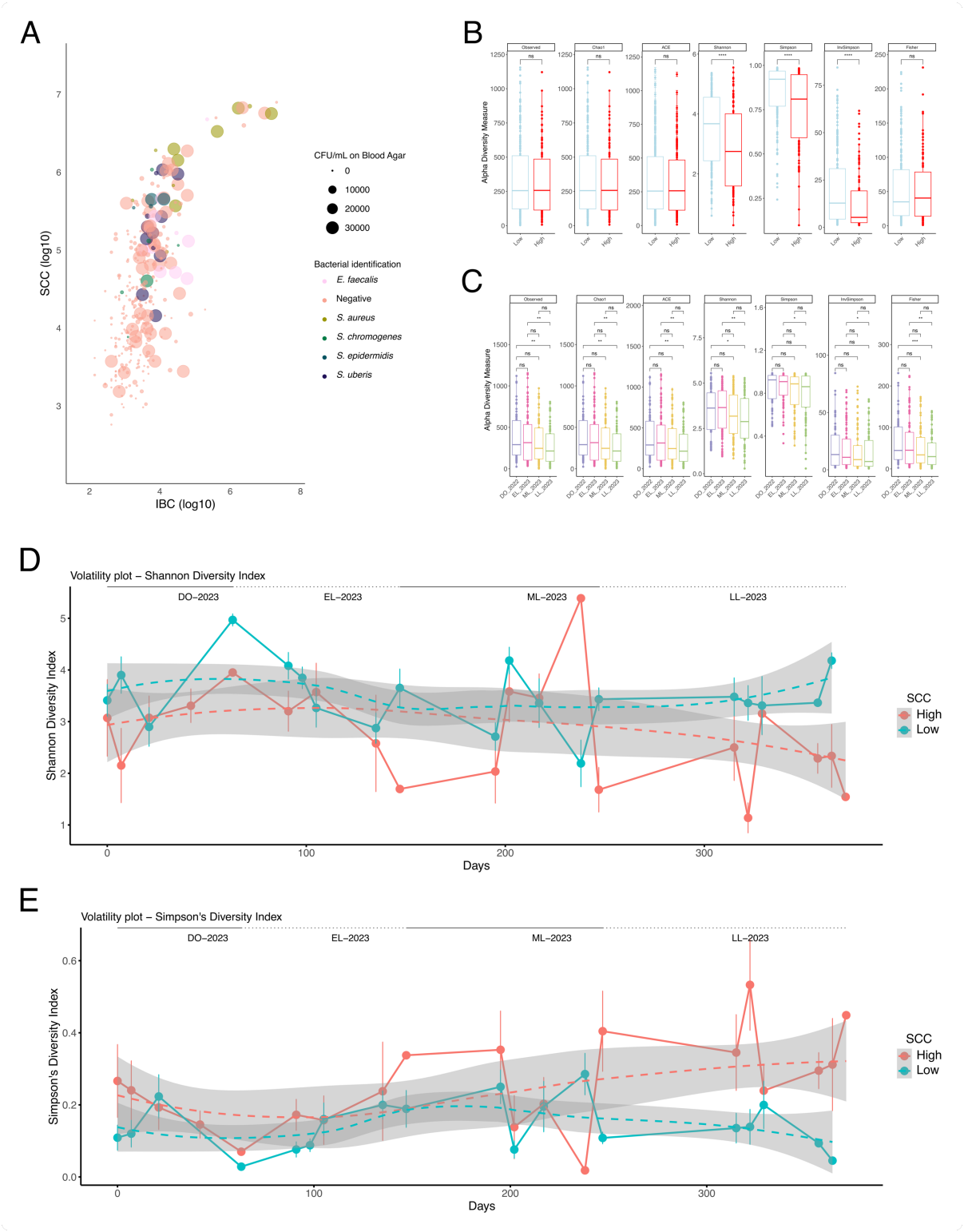

A
SCC (log10)
IBC (log10)
CFU/mL on Blood Agar
0
10000
20000
30000
Bacterial identification
E. faecalis
Negative
S. aureus
S. chromogenes
S. epidermidis
S. uberis
B
Alpha Diversity Measure
Observed
Chao1
ACE
Shannon
Simpson
InvSimpson
Fisher
Low
High
C
Alpha Diversity Measure
D
Volatility plot – Shannon Diversity Index
DO–2023
EL–2023
ML–2023
LL–2023
Shannon Diversity Index
Days
SCC
High
Low
E
Volatility plot – Simpson's Diversity Index
DO–2023
EL–2023
ML–2023
LL–2023
Simpson's Diversity Index
Days
SCC
High
Low

Figure 3

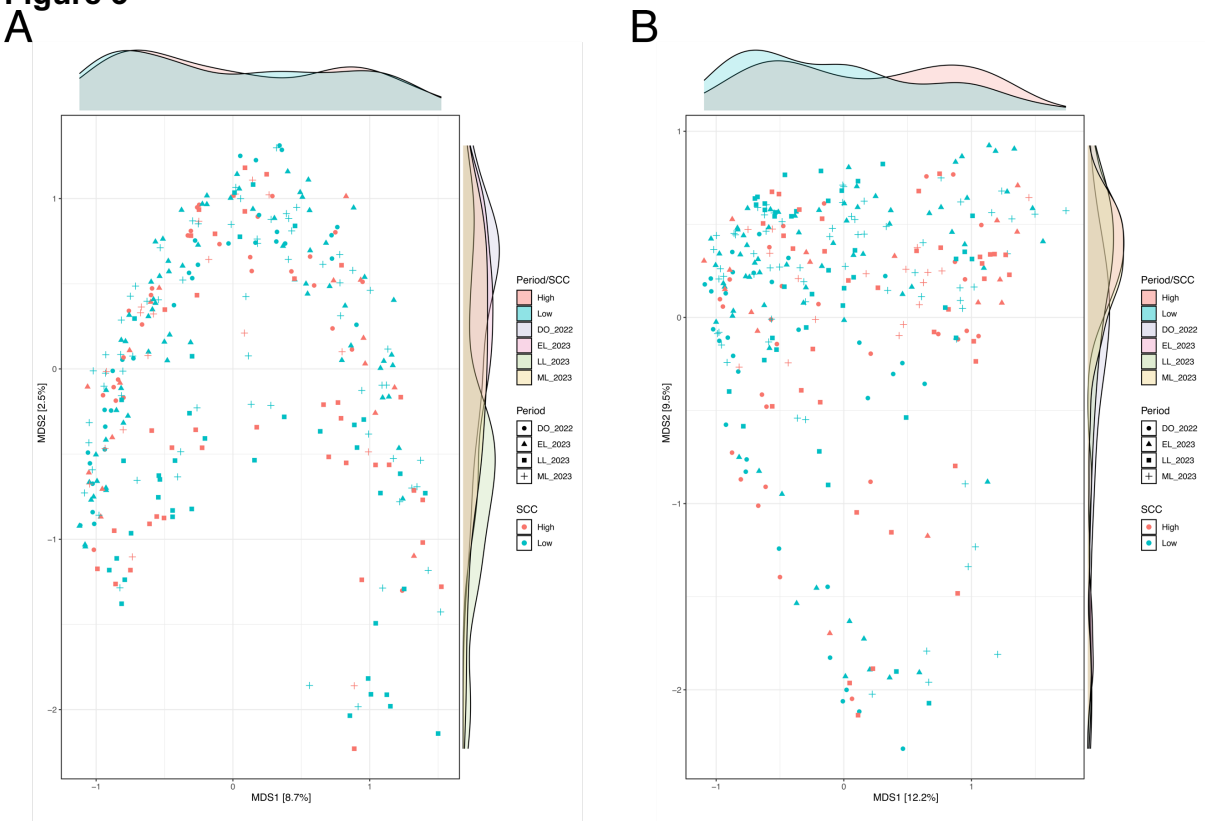

A
B
MDS1 [8.7%]
MDS2 [2.5%]
MDS1 [12.2%]
MDS2 [9.5%]
Period/SCC
High
Low
DO_2022
EL_2023
LL_2023
ML_2023
Period
DO_2022
EL_2023
LL_2023
ML_2023
SCC
High
Low

Figure 4

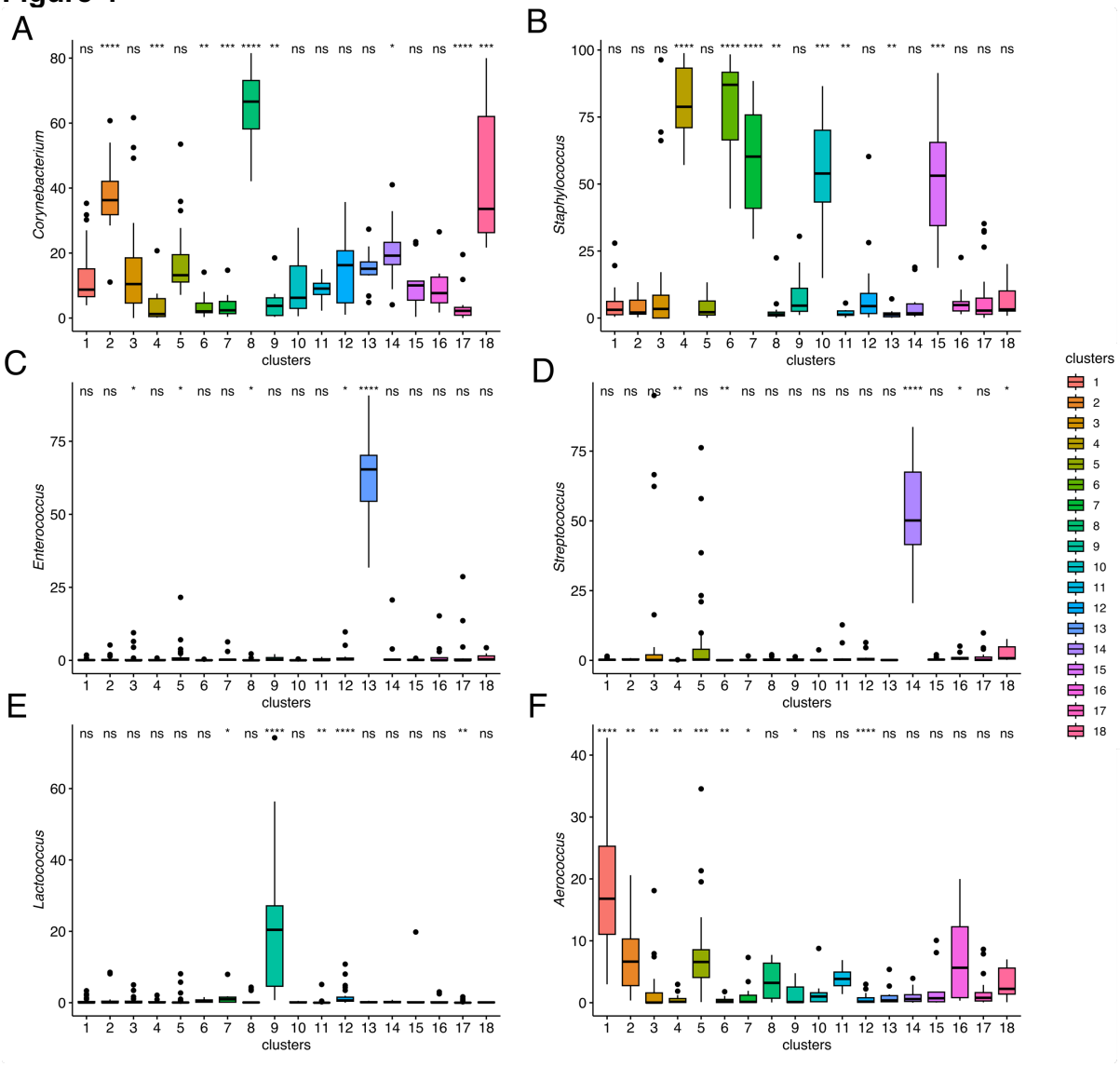

A
Corynebacterium
clusters
B
Staphylococcus
clusters
C
Enterococcus
clusters
D
Streptococcus
clusters
E
Lactococcus
clusters
F
Aerococcus
clusters
clusters
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

Figure 5

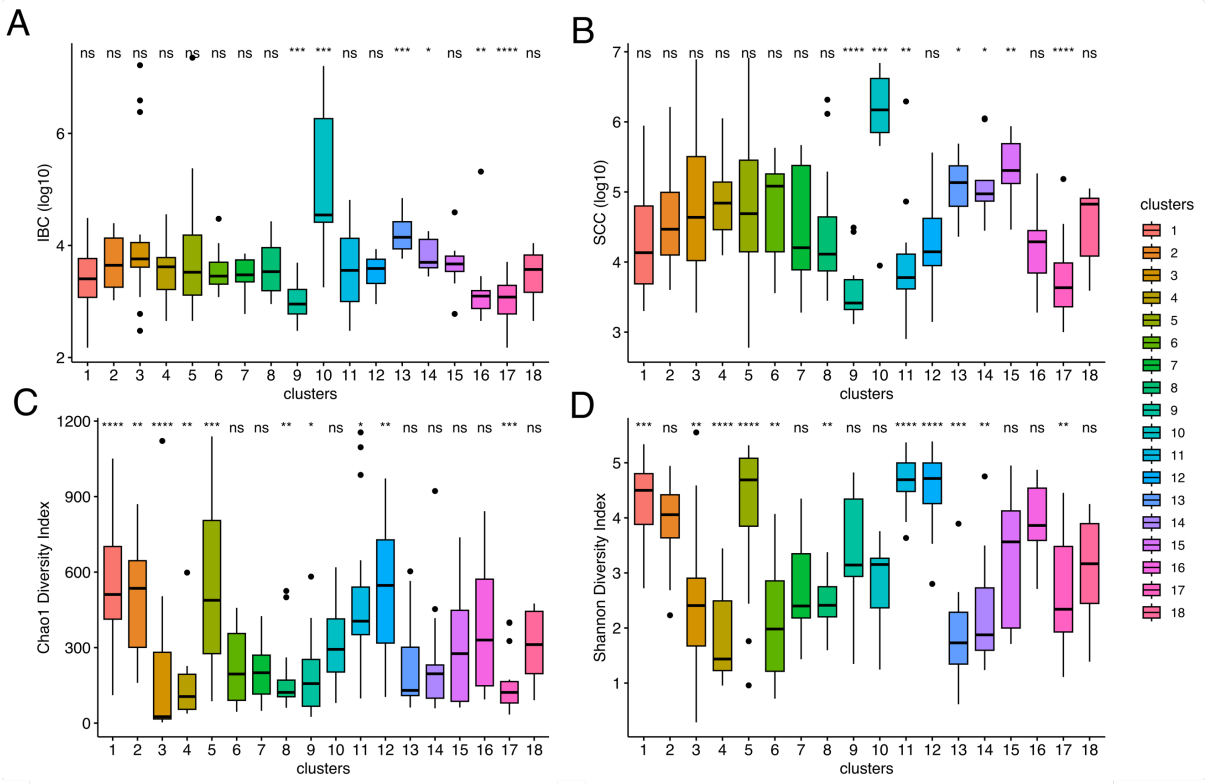
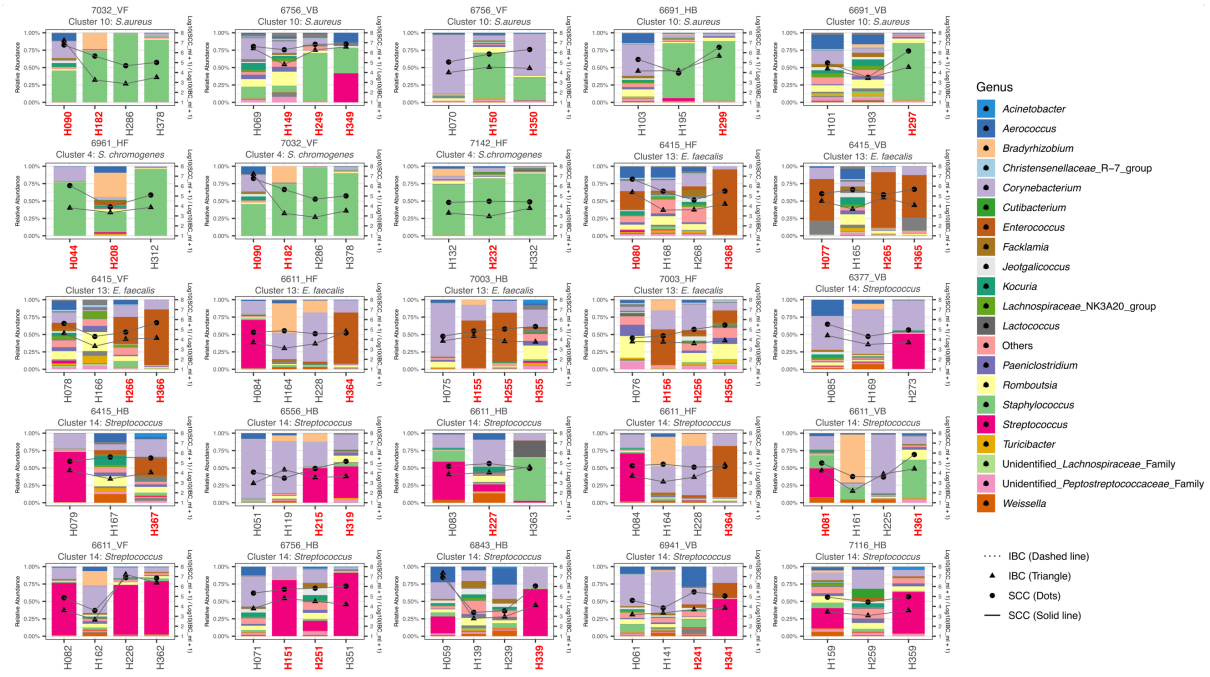

A
B
C
D
IBC (log10)
SCC (log10)
Chao1 Diversity Index
Shannon Diversity Index
clusters
clusters 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Figure 6

Cluster 4 – Staphylococcus
Cluster 6 – Staphylococcus
Cluster 7 – Staphylococcus
Cluster 10 – Staphylococcus
Cluster 15 – Staphylococcus
Cluster 2 – Corynebacterium
Cluster 8 – Corynebacterium
Cluster 18 – Corynebacterium
Cluster 3 – Streptococcus
Cluster 14 – Streptococcus
Relative Abundance
ASV

2075 **Figure 7**

2076



2077
2078
2079
2080

Figure 8

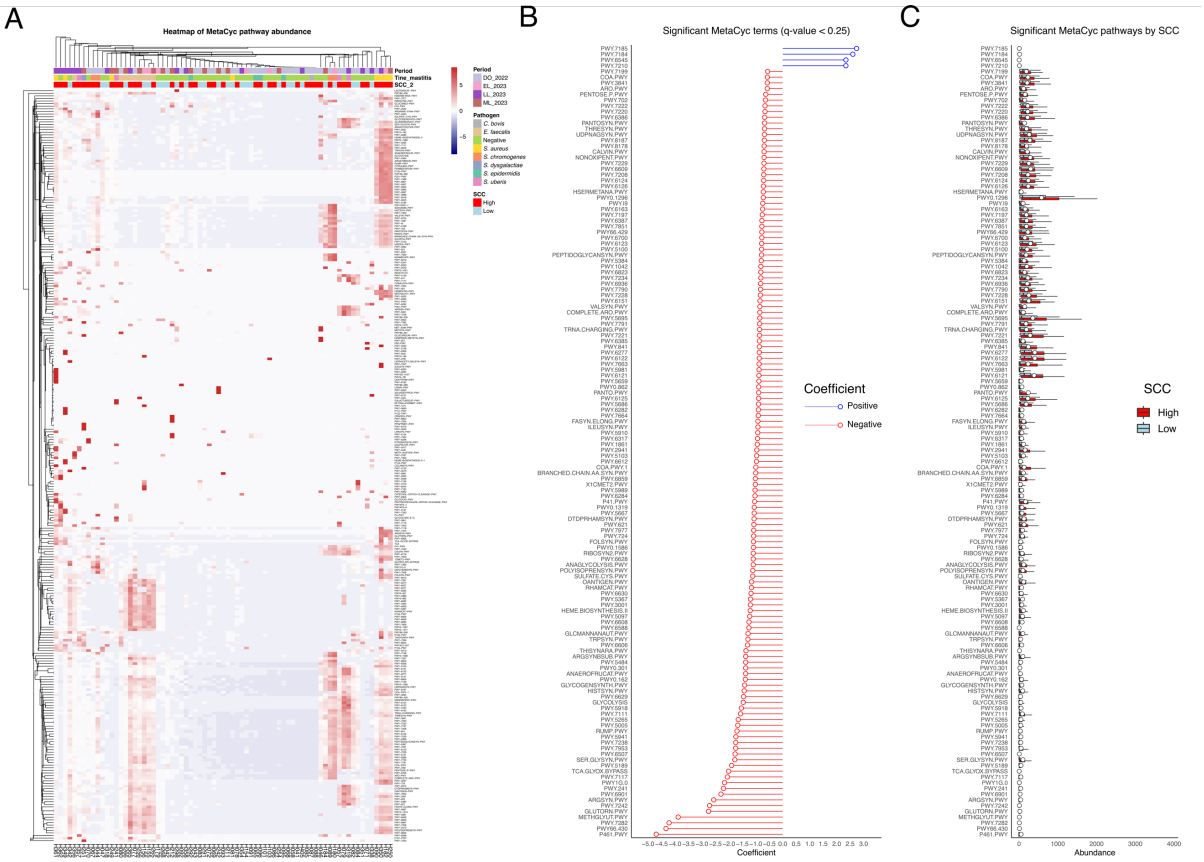
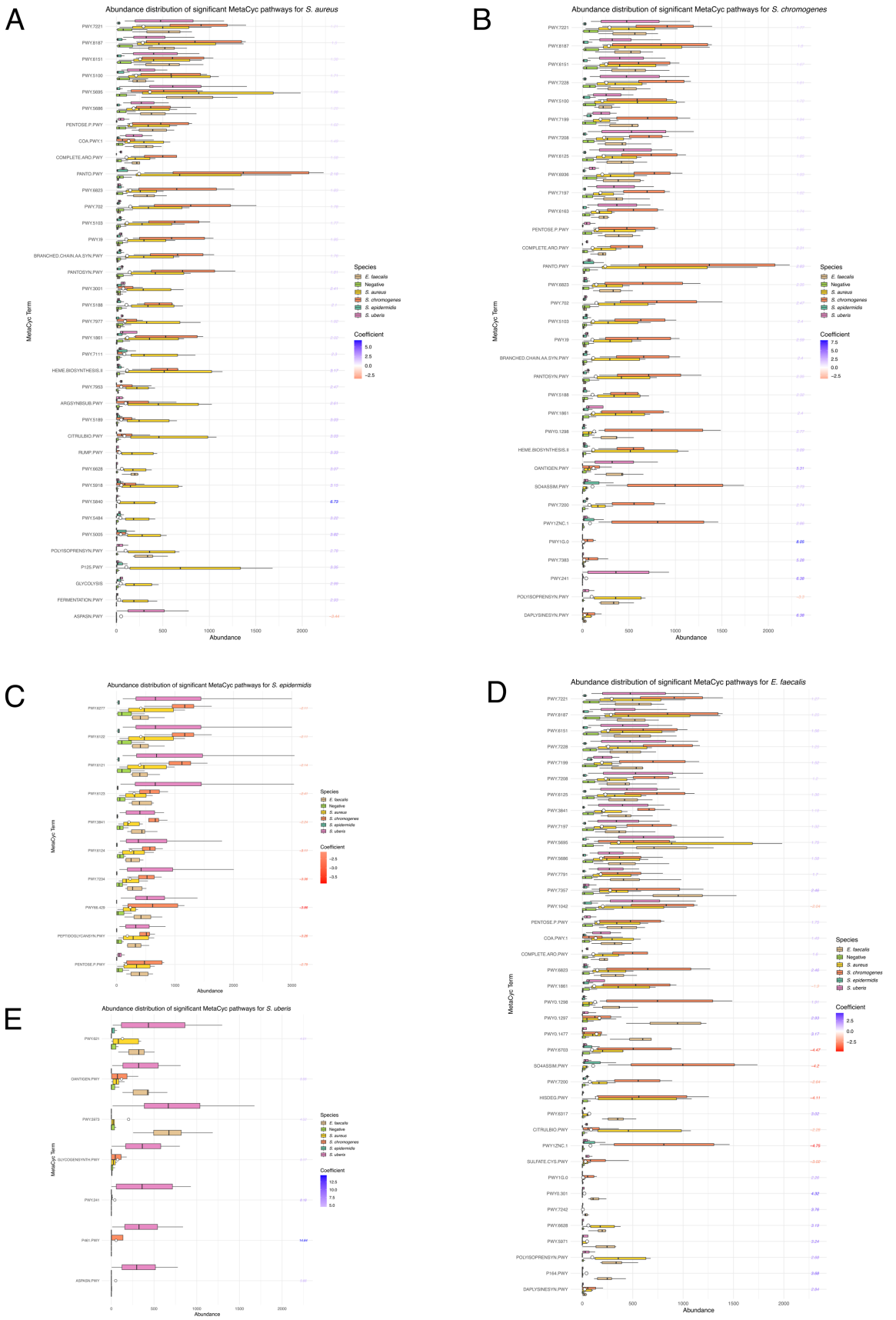
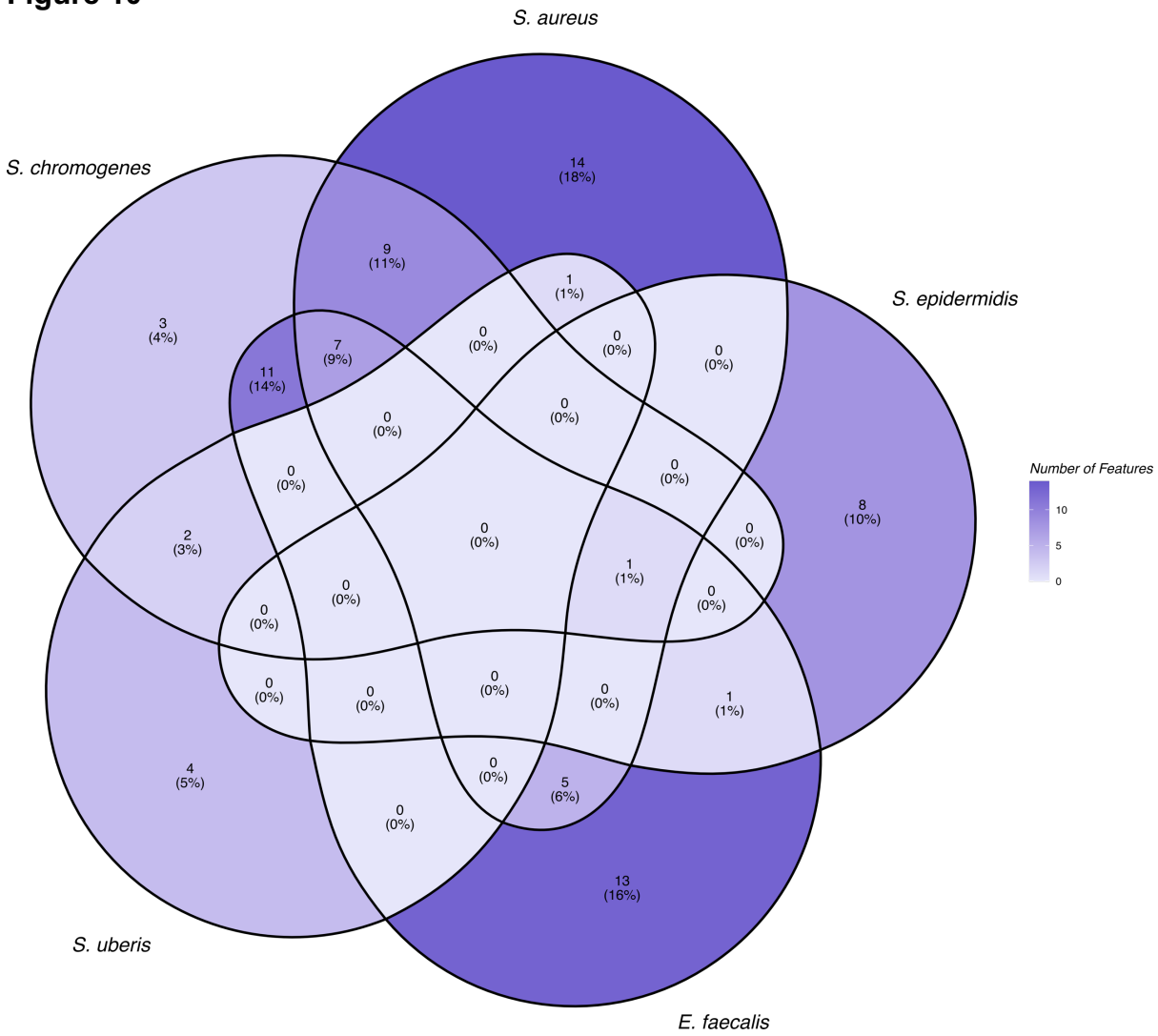

A
Heatmap of MetaCyc pathway abundance
Period
Time_mastitis
SCC_2
Period
DO_2022
EL_2023
LL_2023
ML_2023
Pathogen
C. bovis
E. faecalis
Negative
S. aureus
S. chromogenes
S. dysgalactiae
S. epidermidis
S. uberis
SCC
High
Low
B
Significant MetaCyc terms (q-value < 0.25)
Coefficient
Positive
Negative
Coefficient
C
Significant MetaCyc pathways by SCC
SCC
High
Low
Abundance

Figure 9

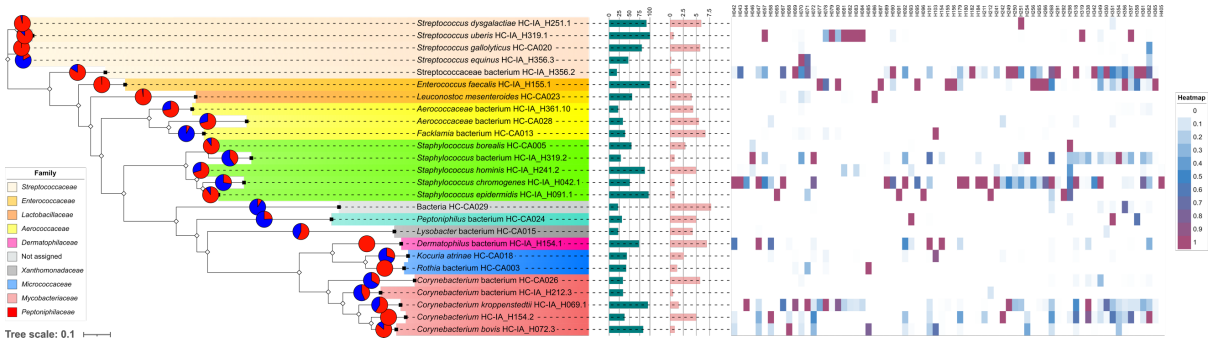

A
Abundance distribution of significant MetaCyc pathways for S. aureus
B
Abundance distribution of significant MetaCyc pathways for S. chromogenes
C
Abundance distribution of significant MetaCyc pathways for S. epidermidis
D
Abundance distribution of significant MetaCyc pathways for E. faecalis
E
Abundance distribution of significant MetaCyc pathways for S. uberis
MetaCyc Term
Abundance
Species
E. faecalis
Negative
S. aureus
S. chromogenes
S. epidermidis
S. uberis
Coefficient

2088 **Figure 10**

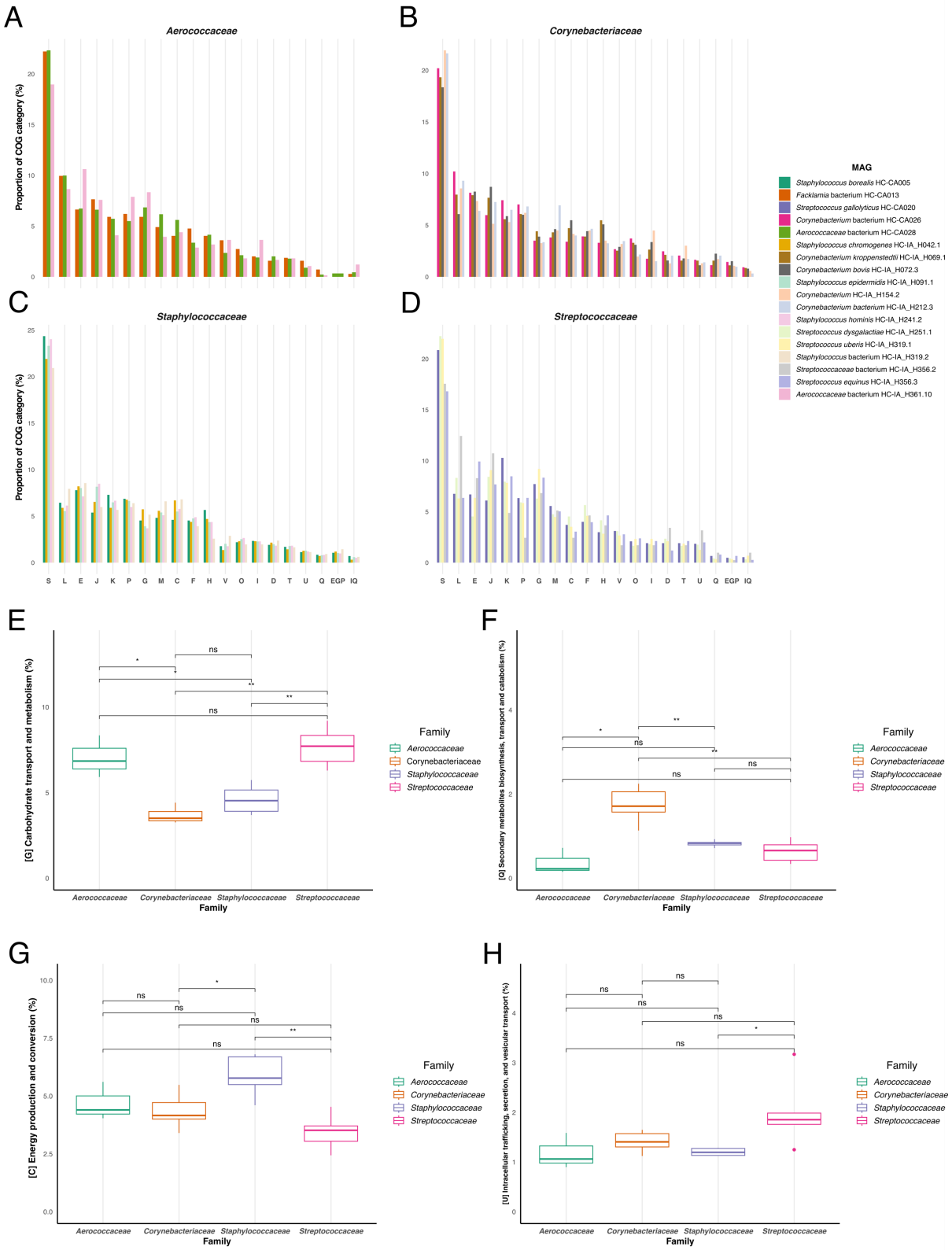


2089
2090
2091

Figure 11


Streptococcus dysgalactiae HC-IA_H251.1
Streptococcus uberis HC-IA_H319.1
Streptococcus gallolyticus HC-CA020
Streptococcus equinus HC-IA_H356.3
Streptococcaceae bacterium HC-IA_H356.2
Enterococcus faecalis HC-IA_H155.1
Leuconostoc mesenteroides HC-CA023
Aerococcaceae bacterium HC-IA_H361.10
Aerococcaceae bacterium HC-CA028
Facklamia bacterium HC-CA013
Staphylococcus borealis HC-CA005
Staphylococcus bacterium HC-IA_H319.2
Staphylococcus hominis HC-IA_H241.2
Staphylococcus chromogenes HC-IA_H042.1
Staphylococcus epidermidis HC-IA_H091.1
Bacteria HC-CA029
Peptoniphilus bacterium HC-CA024
Lysobacter bacterium HC-CA015
Dermatophilus bacterium HC-IA_H154.1
Kocuria atrinae HC-CA018
Rothia bacterium HC-CA003
Corynebacterium bacterium HC-CA026
Corynebacterium bacterium HC-IA_H212.3
Corynebacterium kroppenstedtii HC-IA_H069.1
Corynebacterium HC-IA_H154.2
Corynebacterium bovis HC-IA_H072.3
Family
Streptococcaceae
Enterococcaceae
Lactobacillaceae
Aerococcaceae
Dermatophilaceae
Not assigned
Xanthomonadaceae
Micrococcaceae
Mycobacteriaceae
Peptoniphilaceae
Tree scale: 0.1
Heatmap

2098 **Figure 12**



2099
2100
2101

Figure 13

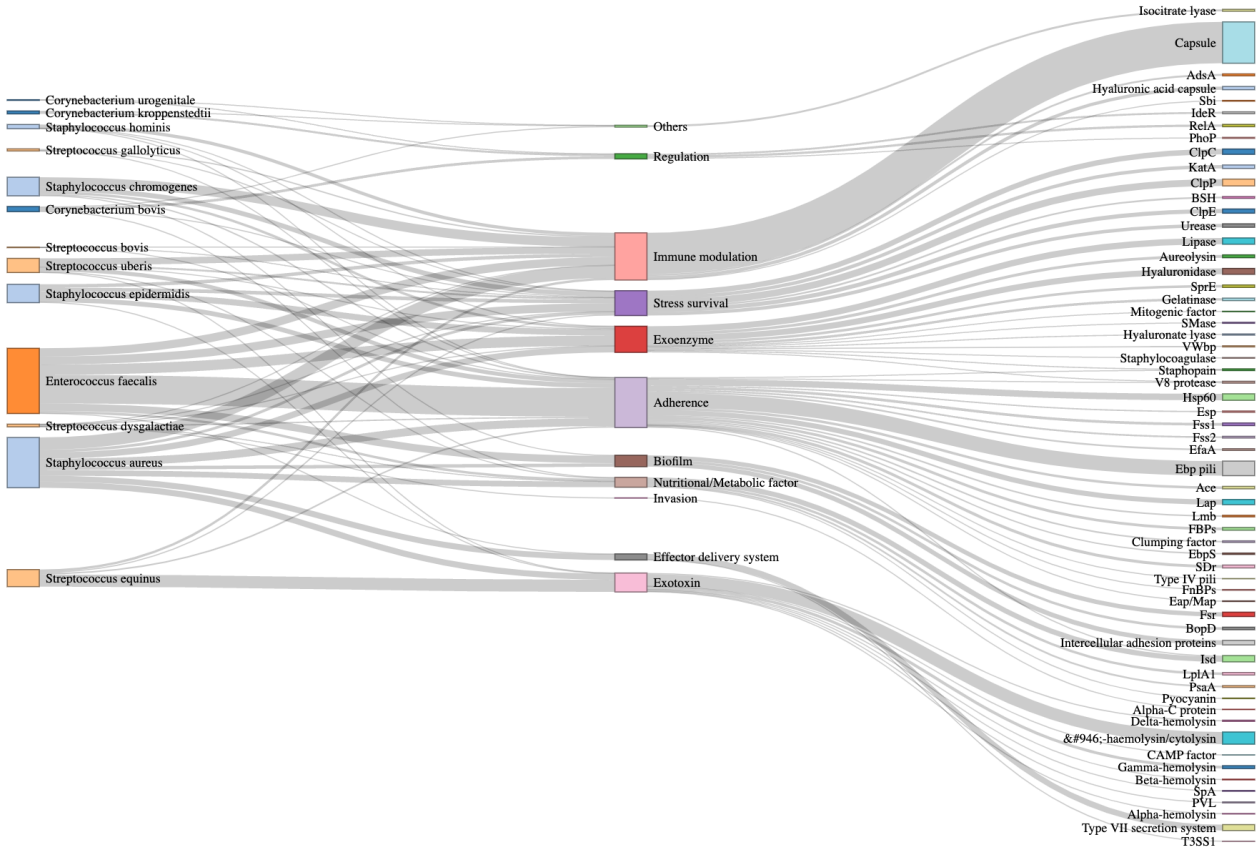

Corynebacterium urogenitale
Corynebacterium kroppenstedtii
Staphylococcus hominis
Streptococcus gallolyticus
Staphylococcus chromogenes
Corynebacterium bovis
Streptococcus bovis
Streptococcus uberis
Staphylococcus epidermidis
Enterococcus faecalis
Streptococcus dysgalactiae
Staphylococcus aureus
Streptococcus equinus
Others
Regulation
Immune modulation
Stress survival
Exoenzyme
Adherence
Biofilm
Nutritional/Metabolic factor
Invasion
Effector delivery system
Exotoxin
Isocitrate lyase
Capsule
AdsA
Hyaluronic acid capsule
Sbi
IdeR
RelA
PhoP
ClpC
KatA
ClpP
BSH
ClpE
Urease
Lipase
Aureolysin
Hyaluronidase
SprE
Gelatinase
Mitogenic factor
SMase
Hyaluronate lyase
VWbp
Staphylocoagulase
Staphopain
V8 protease
Hsp60
Esp
Fss1
Fss2
EfaA
Ebp pili
Ace
Lap
Lmb
FBPs
Clumping factor
EbpS
SDr
Type IV pili
FnBPs
Eap/Map
Fsr
BopD
Intercellular adhesion proteins
Isd
LplA1
PsaA
Pyocyanin
Alpha-C protein
Delta-hemolysin
β-haemolysin/cytolysin
CAMP factor
Gamma-hemolysin
Beta-hemolysin
SpA
PVL
Alpha-hemolysin
Type VII secretion system
T3SS1

Figure 14

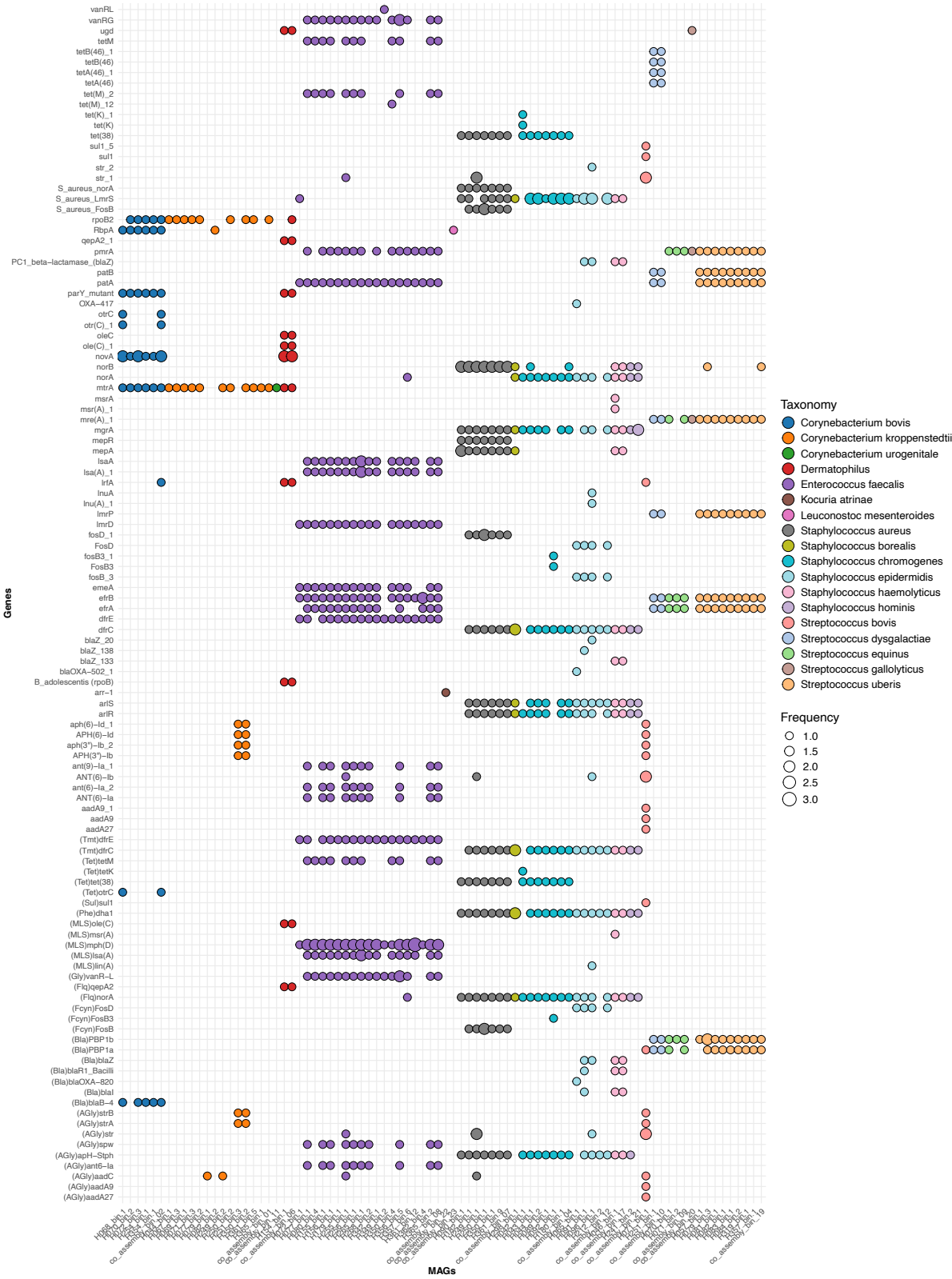
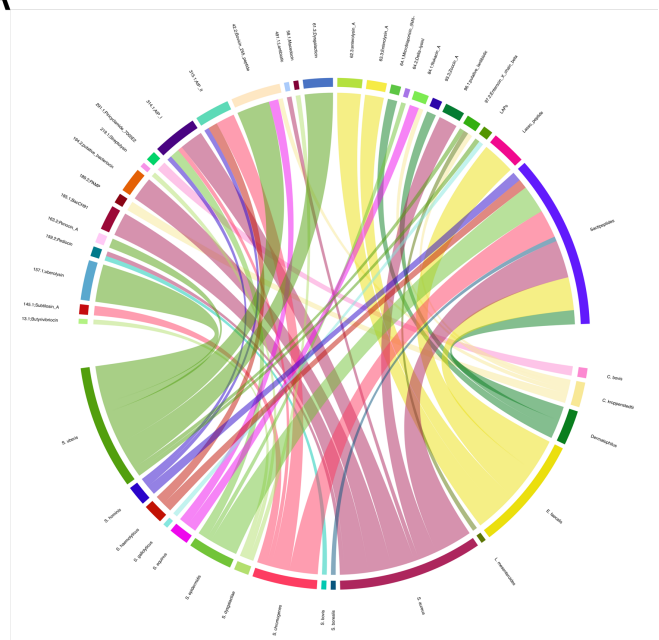
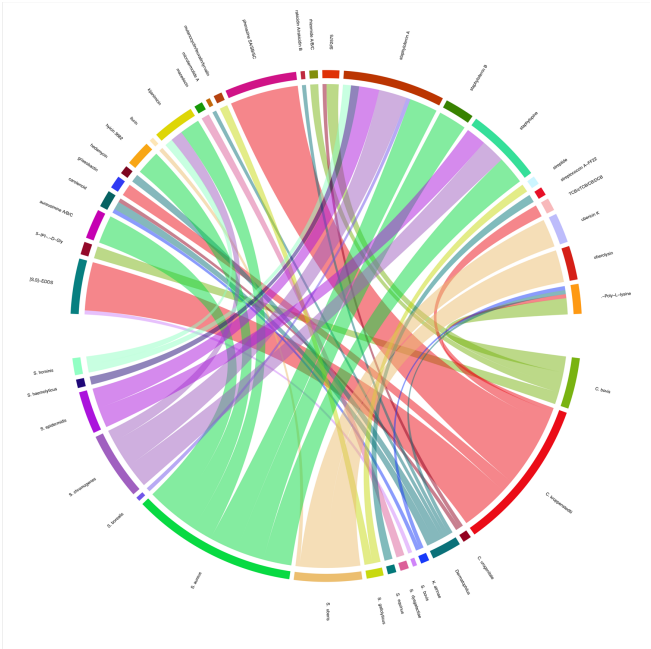

Genes
MAGs
Taxonomy
Corynebacterium bovis
Corynebacterium kroppenstedtii
Corynebacterium urogenitale
Dermatophilus
Enterococcus faecalis
Kocuria atrinae
Leuconostoc mesenteroides
Staphylococcus aureus
Staphylococcus borealis
Staphylococcus chromogenes
Staphylococcus epidermidis
Staphylococcus haemolyticus
Staphylococcus hominis
Streptococcus bovis
Streptococcus dysgalactiae
Streptococcus equinus
Streptococcus gallolyticus
Streptococcus uberis
Frequency
1.0
1.5
2.0
2.5
3.0

Figure 15

A



B