

Whole-Genome Characterization of *Mammaliicoccus sciuri* Sequence Type 605 from Subclinical Bubaline Mastitis Reveals a Distinct Lineage and Intrinsic Antimicrobial-Resistance Repertoire

Md. Mahfujur Rahman

Sylhet Agricultural University

Md. Shahidur Rahman Chowdhury

Sylhet Agricultural University

Sabbir Khan

Gazipur Agricultural University

Hemayet Hossain

Teesta University

M. Nazmul Hoque

nazmul90@gau.edu.bd

Gazipur Agricultural University

Research Article

Keywords: *Mammaliicoccus sciuri*, Buffalo mastitis, Whole-genome sequencing, Resistome, Pangenome, Antimicrobial resistance, One Health

Posted Date: October 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11270530/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Whole-Genome Characterization of *Mammaliicoccus sciuri* Sequence Type 605 from Subclinical Bubaline Mastitis Reveals a Distinct Lineage and Intrinsic Antimicrobial-Resistance Repertoire

Authors

Md. Mahfujur Rahman^{1#*}, Md. Shahidur Rahman Chowdhury^{1#}, Sabbir Khan^{2#}, Hemayet Hossain³, M. Nazmul Hoque^{2*}

Affiliations

¹ Department of Medicine, Sylhet Agricultural University, Sylhet 3100, Bangladesh

² Molecular Biology and Bioinformatics Laboratory, Department of Gynecology, Obstetrics and Reproductive Health, Gazipur Agricultural University, Gazipur 1706, Bangladesh

³ Department of Veterinary Science and Animal Husbandry, Teesta University, Rangpur 5404, Bangladesh

Equal contribution: These authors contributed equally as the first author.

*** Correspondence to:** Md. Mahfujur Rahman, Email: mahfuj.vetmed@sau.ac.bd and M. Nazmul Hoque, Email: nazmul90@gau.edu.bd

Running title: *Mammaliicoccus sciuri* ST605 from Dairy Buffalo

Abstract

Background: *Mammaliicoccus sciuri*, a non-aureus staphylococcus associated with intramammary infection, carries the chromosomal *mecA1* gene, the evolutionary precursor of the methicillin-resistance determinant *mecA*. It is an emerging reservoir of transferable antimicrobial resistance (AMR). No *M. sciuri* genome from South Asian dairy livestock has yet been reported.

Objective: To characterize the genome of *M. sciuri* 30RR from a swamp buffalo (*Bubalus bubalis*) with subclinical mastitis, focusing on its resistome architecture, phylogenomics and pangenome structure.

Methods: *M. sciuri* 30RR was isolated from a buffalo with subclinical mastitis (CMT-positive) in Sylhet, Bangladesh, and sequenced on an Illumina MiSeq. Assembly used SPAdes and annotation used NCBI PGAP. Resistance genes were screened against CARD, ResFinder, MEGARES and AMRFinderPlus and compared across 22 genomes through the Roary pangenome. Phylogenomics combined TYGS, FastANI and a bootstrapped IQ-TREE core-genome phylogeny.

Results: Subclinical mastitis was detected in 67.9% of quarters (1,046/1,540) and 80.8% of buffaloes (311/385). Non-aureus staphylococci including *M. sciuri* 30RR was detected in 16.3% of SCM-positive quarter samples. *M. sciuri* 30RR was phenotypically resistant to ampicillin, amoxicillin–clavulanate, cefoxitin, amikacin and tetracycline. The draft genome spans 2,545,126 bp (40 contigs; GC 32.35%; N50 116,238 bp; 100% BUSCO completeness; 2,565 CDS). MLST assigned the novel sequence type ST605. Strain 30RR carries five intrinsic resistance genes (*mecA1*, *blaZ*, *sal(A)*, *aadK* and *qacA*) but no acquired or mobile antimicrobial-resistance genes (ARGs), in contrast to food-animal and human comparator strains that harbour an acquired SCCmec element, *tet(M)/tet(S)*, *cfr*, *optrA* and other cargo. The pangenome was open, with cloud

gene clusters making up 50.9% of the pan-genome. *M. sciuri* 30RR is a genomically distinct, early-diverging lineage (maximum average nucleotide identity, (ANI) 96.7%). Phenotypic beta-lactam resistance was concordant with *mecA1* and *blaZ*. Core-genome MLST confirmed this distinctiveness, with 30RR differing from any other genome at $\geq 1,923$ of 2,142 core loci, and no intact *icaADBC* exopolysaccharide or capsule (*cap*) locus was detected.

Conclusions: *M. sciuri* ST605 is the first genome of this species from dairy buffalo in Bangladesh and from South Asian dairy livestock. Its purely intrinsic resistome and low-virulence profile contrast with multidrug-resistant poultry- and human-associated lineages, indicating a distinct regional sub-lineage. This genome provides a reference for One Health surveillance at the buffalo–human interface.

Keywords: *Mammaliicoccus sciuri*; Buffalo mastitis; Whole-genome sequencing; Resistome; Pangenome; Antimicrobial resistance; One Health

1. Background

Mastitis remains one of the most economically damaging infectious diseases of dairy livestock worldwide, driving losses through reduced milk yield, elevated somatic cell counts, treatment costs, and premature culling [1-3]. The etiology of mastitis in dairy animals are phylogenetically diverse [4, 5]. Of these, the non-aureus staphylococci (NAS) have long been regarded as environmental contaminants and are now routinely recovered from subclinical mastitis quarters in well managed herds, and widely recognized as true pathogens of intramammary infections in swamp buffalo as well as dairy cattle [6-8]. One member of this group, *M. sciuri* reclassified from *S. sciuri* following phylogenomic analysis of the *Staphylococcaceae* [9], is both an opportunistic pathogen and a reservoir of transferable antimicrobial resistance (AMR) determinants. Although

predominantly a commensal of mammalian skin and mucous membranes, it causes mastitis, wound infections, dermatitis, and, in severe cases, fatal systemic disease in companion animals and humans [10, 11]. Isolates have been recovered from dairy cattle, dairy buffalo, food products, seabirds, wildlife, and clinical human specimens across Europe, Asia, and the Americas, reflecting a host and geographic range that spans both food-production systems and human clinical medicine [11-15].

The AMR significance of *M. sciuri* centers on the chromosomally encoded *mecA1* gene, present in virtually all strains and identified by phylogenomic analyses as the evolutionary precursor of *mecA*, the determinant that confers methicillin resistance in methicillin-resistant *Staphylococcus aureus* (MRSA) [12, 13]. The *mecA1* gene acquired from a mobile SCCmec element has been found to be associated with regulatory sequences that are essential to confer high-level resistance to beta-lactam antibiotics, the acquisition of which is also directly associated with the evolution of MRSA from *M. sciuri* [15, 16].

Whole-genome sequencing (WGS) studies from European, Asian, and South American dairy and veterinary settings have documented multidrug-resistant (MDR) *M. sciuri* strains in which *mecA1* co-occurs with *sal(A)*, alongside aminoglycoside, tetracycline, and phenicol resistance genes [10, 17]. The capacity of *M. sciuri* to donate *mecA* to *S. aureus*, together with the risk of onward transfer of multidrug-resistance cassettes, establishes it as a genuine One Health concern [14, 15]. Despite this accumulating global genomic evidence, South Asian dairy livestock remain entirely absent from published *M. sciuri* WGS studies. Bangladesh has a substantial dairy buffalo (*Bubalus bubalis*) sector, reflecting the broader importance of buffalo production in Asia, which harbors more than 97% of the global buffalo population and contributes approximately 15% of worldwide milk production [18, 19]. Cross-sectional surveys spanning nine districts in Bangladesh have

identified NAS, including *M. sciuri*, as the dominant IMI-associated bacteria in Bangladeshi buffalo herds, recovered from 32% of culture-positive quarter samples and present on 72% of surveyed farms [20]. However, no whole genome sequences of *M. sciuri* from South Asian dairy livestock have been deposited in any public genome database. Only clinical hospital isolates of *S. sciuri* with a novel *ccrA7* allotype have been previously molecularly characterized in Bangladesh and no genome level analysis of any dairy or livestock strains [21]. Here, we report the genomic characterization of *M. sciuri* strain 30RR, dairy buffalo milk isolate from Sylhet, Bangladesh, that includes resistome profiling, a comparative resistome across the species, phylogenomic placement of the strain among global lineages, functional annotation and pangenome analysis. To the best of our knowledge, it is the first *M. sciuri* genome from swamp buffalo and from South Asian dairy livestock.

2. Materials and methods

2.1. Ethics approval and consent to participate

This study was approved by the Animal Experimentation and Ethics Committee (AEEC) of Sylhet Agricultural University, Bangladesh (Approval No. AUP2023001). All animals were privately owned by smallholder farmers, and informed consent was obtained from the owners before sampling. Milk samples were collected non-invasively by qualified personnel in accordance with institutional ethical guidelines, and animal welfare was maintained throughout the study.

2.2 Study population and sampling

We conducted a cross-sectional study between February 2023 and June 2024 in five Upazilas of the Sylhet district of Bangladesh (Jaintapur, Gowainghat, Kanaighat, Balaganj, and Fenchuganj), areas with a predominantly buffalo-rearing population. These Upazilas lie between approximately 24°36' and 25°11' N latitude and 91°38' and 92°30' E longitude. The study design, study area, and

sampling methodology have been described in detail in our previous report [22]. The present investigation was conducted on 1,540 quarter milk samples collected from 385 lactating swamp buffaloes (*Bubalus bubalis*). All animals were privately owned, smallholder-reared buffaloes sampled directly on their farms/households; no animals were purchased, relocated, or experimentally housed for this study. A multistage random cluster sampling design was employed. Within each randomly selected cluster, individual animals were enrolled by simple random selection, irrespective of age, parity, or stage of lactation. The milk samples were collected aseptically from each quarter of apparently healthy buffaloes, following the guidelines set by the National Mastitis Council [23]. The California Mastitis Test (CMT) [2] was utilized as the initial screening methods for detecting subclinical mastitis (SCM). Before sample collection, quarters were soaked with 70.0% ethanol and dried off using tissue paper. Initially, one to two drops of milk were discarded and then 10–15 mL of milk from each of the four quarters (left front; LF, left hind; LH, right front; RF, and right hind; RH) was collected aseptically in sterile plastic tubes (15 mL Falcon Tube®) during morning milking (08:00–10:00 h) [24]. Immediately after collection, milk samples were subjected to CMT to diagnose quarters with SCM. After on farm screening, the SCM-positive samples (n = 1,046) were kept in icebox (4–8°C) and transported to the laboratory for further bacteriological examination and molecular analysis.

2.3 Bacterial isolation and identification

The SCM positive milk samples (n = 1,046) were subjected to pre-enrichment in nutrient broth (Biolife, Italy) at 37°C under aerobic conditions. After enrichment, 2 µL of broth was streaked onto nutrient agar (HiMedia, India) and incubated at 37°C for 24 h. The enriched broth was also serially diluted to 10⁻³ and plated onto mannitol salt agar (MSA; Oxoid, UK) [8, 17, 25]. Colonies with staphylococcal morphology (circular, creamy-white to yellow) were sub-cultured to purity

and identified by Gram staining and biochemical testing including catalase, coagulase, and sugar fermentation [8, 26]. The species-level identification *M. sciuri* 30RR isolate was confirmed with the VITEK-2 system v9.01(bioMerieux, France) [27], with definitive confirmation by whole-genome sequencing.

2.4 Antimicrobial susceptibility testing

The *in vitro* antibiogram profile of *M. sciuri* 30RR isolate was determined using the disk diffusion method following the Clinical Laboratory Standards Institute (CLSI) M100 34th edition guidelines [28]. Antimicrobial susceptibility testing (AST) was performed on Mueller–Hinton agar plates (HiMedia Laboratories Pvt. Ltd., Mumbai, India). A panel of 13 antibiotics representing eight distinct antimicrobial classes was evaluated: penicillins (ampicillin 10 µg, amoxicillin–clavulanic acid 20/10 µg), tetracyclines (tetracycline 30 µg), macrolides (azithromycin 15 µg), quinolones (ciprofloxacin 5 µg, nalidixic acid 30 µg), folate pathway antagonists (trimethoprim–sulfamethoxazole 1.25/23.75 µg), phenicols (chloramphenicol 30 µg), cephalosporins (cefoxitin 30 µg, ceftriaxone 30 µg), and aminoglycosides (gentamicin 10 µg, amikacin 30 µg, streptomycin 10 µg). Overnight cultures at approximately 10⁹ CFU/mL were swabbed uniformly onto Mueller–Hinton Agar (MHA; HiMedia) plates under aseptic conditions. Antibiotic-impregnated discs were applied, and inhibition zone diameters were recorded after 24 h of incubation at 37°C. All tests were performed in triplicate, with results categorized as resistant (R), intermediate (I), or susceptible (S) per CLSI M100 (2023) breakpoints [28]. Cefoxitin disc diffusion served as a surrogate marker for methicillin resistance and was interpreted using CLSI criteria for staphylococci [29, 30]. All tests were performed in triplicate to ensure consistency and reproducibility. *M. sciuri* 30RR isolate was resistant to five of the thirteen antibiotics tested, and was classified as MDR.

2.5 Whole-genome sequencing and assembly

Genomic DNA from the *M. sciuri* 30RR isolate was extracted from a pure colony of strain 30RR using the QIAamp DNA Mini Kit (QIAGEN, Germany) following the manufacturer's Gram-positive protocol. Libraries were prepared from 1 ng of DNA using the Nextera DNA Flex kit (Illumina, USA) and sequenced on an Illumina MiSeq platform (2 × 250 bp paired-end). Reads were quality-assessed with FastQC v0.12.1 [31] and trimmed with Trimmomatic v0.40 [32] with parameters LEADING:20, TRAILING:20, SLIDINGWINDOW:4:20, MINLEN:50. Assembly used SPAdes v4.2.0 [33] and annotation used the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.10 [34]. Genome completeness was assessed with CheckM v1.2.4 [35] using the *Mammaliicoccus*-specific marker set and with BUSCO v6.1.0 [36] using the bacteria_odb10 lineage dataset.

2.6 Multilocus sequence typing and core-genome MLST

Multilocus sequence typing (MLST) was performed using mlst v2.32.2 (<https://github.com/tseemann/mlst>) with the seven-locus *M. sciuri* MLST scheme and the PubMLST sequence-definition database. Four loci matched existing alleles (*ack-30*, *aroE-55*, *fisZ-2*, and *tpiA-27*), whereas three loci (*gmk*, *ptaI*, and *glpK*) contained novel alleles that were submitted to and accepted by PubMLST, receiving allele numbers 61, 129, and 120, respectively (BIGSdb submission IDs: 20260312192623, 20260312190228, and 20260312191502). The complete allelic profile was subsequently submitted to PubMLST (BIGSdb submission ID: BIGSdb_20260314200052_2880295_67847) and assigned the novel sequence type ST605. Core-genome MLST (cgMLST) was performed with chewBBACA v3.4.0 [37]. A de novo gene-by-gene schema was created from all 23 genomes (CreateSchema), followed by allele calling (AlleleCall) and extraction of the core-genome allele matrix with ExtractCgMLST, retaining loci

present in at least 95% of genomes (2,142 core loci; 1,887 loci were present in all 23 genomes). Pairwise allelic-difference profiles were used to construct a minimum-spanning tree (MST) with the MSTreeV2 algorithm in GrapeTree v2.2 [38], which was also used for visualization; nodes were collapsed by sequence type, with node size proportional to the number of genomes.

2.7 Resistome, virulome and mobile genetic element analysis

Antimicrobial resistance genes (ARGs) in the *M. sciuri* 30RR genome were identified by screening the assembled genome against the Comprehensive Antibiotic Resistance Database (CARD) v6.0.5 2023 [39]. Virulence factor genes (VFGs) were detected using the Virulence Factor Database (VFDB) v6.0 [40]. Plasmid replicons, staphylococcal cassette chromosome mec (SCCmec) elements, and prophage regions were identified using PlasmidFinder v2.0.1 [41], SCCmecFinder v1.2 [42], and PHASTER [43], respectively. CRISPR-Cas systems were identified using CRISPRCasFinder v2.0.3 [44], and the pathogenic potential of the isolate was assessed using PathogenFinder2 [45]. To place the *M. sciuri* 30RR resistome in species context, the Roary gene presence/absence matrix (Section 2.9) was mined for resistance-associated clusters across all 23 genomes; candidate clusters were verified by BLASTp against reference proteins and by eggNOG-mapper annotation, and were classified as intrinsic (core/near-core) or acquired (accessory, mobile-element-associated). Insertion-sequence (IS) elements were identified with ISEScan [46]. Genomic islands were predicted with IslandViewer 4 [47], with the 30RR genome ordered against the *S. sciuri* subsp. *sciuri* NCTC12103 reference chromosome.

2.8 Phylogenetic analysis and average nucleotide identity

Phylogenetic placement was assessed relative to 22 publicly available *M. sciuri* genomes (**Table S1**). Species identity was confirmed with the Type Strain Genome Server (TYGS; <https://tygs.dsmz.de/>) and digital DNA-DNA hybridisation (dDDH). Pairwise average nucleotide

identity (ANI) was computed with FastANI [48]. A core-genome single-nucleotide alignment (152,856 bp) produced by Roary was used to infer a maximum-likelihood phylogeny with IQ-TREE v3.1.3 [49] under the best-fit model (GTR+F+I+R4) selected by ModelFinder, with branch support from 1,000 SH-aLRT replicates and 1,000 ultrafast bootstrap replicates; the tree was midpoint-rooted and visualized with ggtree [50]. In addition, a reference-based core-genome single-nucleotide-polymorphism (SNP) phylogeny was inferred with Parsnp v2.1.3 [51], using *M. sciuri* NCTC12103 as the reference genome (-r) and forcing inclusion of all 23 genomes in the core alignment (-c); the resulting tree was annotated with country of origin, source/host group and core-genome SNP distance from 30RR using ggtree [50].

2.9 Pangenome analysis

Pangenome analysis used Roary v3.13.0 [52] with a 95% BLASTp identity threshold across all 23 genomes. Gene clusters were classified as core (present in 99-100% of genomes), soft-core (95-99%), shell (15-95%) and cloud (<15%). Pangenome openness was modelled with Heap's law [53] by fitting a power law ($N = \kappa n^\gamma$; where N is the pan-genome size, n is the number of genomes and γ is the scaling exponent) to the pan-genome rarefaction curve and an exponential decay to the core-genome curve in R v4.5.0 [54]; the pangenome was considered open when $\gamma > 0$ (equivalently, Tettelin alpha, $\alpha = 1 - \gamma$; where α is the Tettelin decay parameter and γ is the scaling exponent).

2.10 Functional annotation

Subsystem content was characterized with RAST v2.0 [55]. Clusters of Orthologous Groups (COG) functional categories were assigned with eggNOG-mapper v5.0 [56], and KEGG pathway mapping was performed with KAAS [57] using the bidirectional best-hit method. Carbohydrate-active enzymes were identified with the dbCAN3 meta-server [58] using HMMER searches

against CAZy HMM profiles (E-value $< 1 \times 10^{-15}$; coverage > 0.35) and classified according to the CAZy database [59].

2.11 Secondary metabolite and bacteriocin analysis

Secondary metabolite biosynthetic gene clusters (BGCs) were predicted with antiSMASH v8.0.4 [60] and primary metabolic gene clusters with gutSMASH v2.0.1 [61]. Bacteriocin-encoding clusters were identified with BAGEL4 (<http://bagel4.molgenrug.nl/>), and candidate domains were confirmed by BLASTp against the NCBI non-redundant protein database. The genomic neighbourhoods of the *mecA1* and *sal(A)* loci were extracted from the PGAP annotation and visualised as gene-arrow maps.

2.12. Prediction of exopolysaccharide and capsule biosynthesis gene clusters

Exopolysaccharide (EPS) and capsular polysaccharide biosynthesis capacity was assessed in all 23 genomes. Reference protein sequences for staphylococcal capsule biosynthesis (CapA–CapP of the *S. aureus* *cap5/cap8* operons [62]), PNAG/PIA exopolysaccharide biosynthesis and regulation (IcaADBC and IcaR [63]) and previously described *M. sciuri* EPS-associated proteins were retrieved from UniProt [64] and used as queries in BLASTP [65] searches against the Prokka-annotated proteome of 30RR and the NCBI PGAP-annotated proteomes of the 22 public genomes (e-value $\leq 1 \times 10^{-5}$, $\geq 30\%$ amino-acid identity and $\geq 40\%$ reference-protein coverage). Hits were mapped to gene neighbourhoods from the Prokka annotations, and biosynthetic clusters were defined as two or more hits within 10 kb on the same contig. Genes were categorised by predicted function.

2.13. Epidemiological context

To provide background context for the study region, antimicrobial-resistance-gene surveillance data from mastitic milk quarters collected across five places of Sylhet (Balaganj, Fenchuganj,

Gowainghat, Jaintapur and Kanaighat) during the study period were summarised as the proportion of quarters positive for each resistance-gene class.

3. Results

3.1. Prevalence of *M. sciuri* in bubaline SCM

The prevalence of SCM, based on the CMT, was 67.9% (1046/1540; 95% CI: 65.6–70.3%) at the quarter level, and 80.8% (311/385; 95% CI: 76.5–84.6%) at the animal level. Among the SCM-positive samples, 171 NAS isolates were identified, including *M. sciuri* 30RR. This isolate was further tested for AST and whole-genome sequencing.

3.2. Antimicrobial susceptibility and genotype-phenotype correlation

M. sciuri 30RR was resistant to five of the 13 antibiotics tested (38.5%), including the β -lactams ampicillin, amoxicillin–clavulanic acid, and cefoxitin, as well as amikacin and tetracycline. It showed intermediate susceptibility to ceftriaxone, gentamicin, streptomycin, and nalidixic acid, while azithromycin, chloramphenicol, ciprofloxacin, and trimethoprim–sulfamethoxazole remained active (**Table 1**). Cefoxitin resistance phenotypically confirmed methicillin resistance, consistent with the *mecA1*-encoded PBP2a.

Correlating phenotype with genotype (**Table 1**), the beta-lactam resistance (ampicillin, amoxicillin-clavulanic acid, cefoxitin) is fully accounted for by *mecA1* (PBP2a) together with the *blaZ* beta-lactamase. The streptomycin intermediate phenotype is consistent with the *aadK* ANT(6) adenylyltransferase. Two phenotypes remain incompletely resolved at the genotype level: amikacin resistance, because ANT(6) enzymes characteristically modify streptomycin rather than amikacin, and tetracycline resistance, for which the only candidate is a conserved, divergent Tet(X)-like flavin-dependent monooxygenase (~30% identity to characterised Tet(X)) that is present in all 23 genomes. The susceptible phenotypes (azithromycin, chloramphenicol,

ciprofloxacin, trimethoprim-sulfamethoxazole) are concordant with the absence of cognate resistance genes. The complete antibiotic-by-antibiotic genotype–phenotype assessment is provided in **Table S1**.

3.3. Genome assembly and general features of the *M. sciuri* 30RR

The WGS of the *M. sciuri* 30RR isolate yielded 996,952 paired-end raw read pairs, of which 97.98% (976,858 pairs) passed quality trimming. The resulting draft genome was 2,545,126 bp in size and assembled into 40 contigs, with a GC content of 32.35%, N50 of 116,238 bp, L50 of 6, and mean coverage of 99.66×. Genome completeness was estimated at 100% by CheckM and BUSCO. NCBI PGAP annotation identified 2,565 coding sequences (CDS), 4 rRNAs, 30 tRNAs, and 1 tmRNA (**Table 2**). Compared with the 22 publicly available reference *M. sciuri* genomes, the 30RR assembly was among the most contiguous. 19 of the 22 reference genomes are highly fragmented (median ~2,750 contigs; N50 ~1.1 kb) and only Colony295, Colony296 and WHA07 are near-complete, whereas 30RR comprises 40 contigs with an N50 of 116 kb (**Table S2**).

3.4. Functional annotation

dbCAN3 identified 50 CAZyme-encoding genes across 26 families, with glycoside hydrolases (GH; 23) and glycosyltransferases (GT; 21) predominant. eggNOG-mapper assigned COG categories to 2,357 of 2,596 proteins (90.8%); amino-acid transport and metabolism (COG-E; 249) and transcription (COG-K; 213) were the largest assigned categories, and 23.0% of annotated proteins were of unknown function (COG-S) (**Fig. S1; Table S3**). KEGG mapping returned 1,510 unique KO terms, with ABC transporters the most abundantly mapped pathway. RAST assigned 972 functional roles across 227 subsystems.

3.5. Taxonomic identity and sequence type of *M. sciuri* 30RR

TYGS whole-genome phylogenetic analysis (average bootstrap support 91.2%) placed strain 30RR within the *M. sciuri* species cluster (**Fig. S2**). The highest digital DNA–DNA hybridization (dDDH) d4 value was 75.4% (95% CI: 72.4–78.2%) against *S. sciuri* subsp. *carnaticus* CCUG 39509, exceeding the 70% species threshold. Against the *S. sciuri* type strain DSM 20345, d4 was 69.9% (95% CI: 66.9–72.7%), marginally below the threshold but with a confidence interval extending above it. MLST assigned four alleles matching existing PubMLST entries (ack-30, aroE-55, ftsZ-2, tpiA-27); the remaining three loci (glpK, gmk, and pta1) harbored novel sequences, assigned allele numbers glpK-120, gmk-61, and pta1-129 following curator review in March 2026. The complete profile was assigned sequence type ST605 (BIGSdb_20260314200052), a novel sequence type in the *M. sciuri* PubMLST database; the isolate has been deposited as PubMLST ID 1043. Full dDDH values are given in **Table S4**, and the seven-locus allelic profiles of all 23 genomes in **Table S5**.

Core-genome MLST across the 2,142 core loci resolved the 23 genomes into 18 sequence types, and the minimum-spanning tree was fully concordant with seven-locus MLST (**Fig. 1; Table S6**). Genomes sharing an ST were nearly identical at the core-genome level (the three ST219 lake-sediment isolates differed at 0 loci, and the two ST82 human isolates at 11 loci), whereas 30RR was highly divergent from every comparator: it differed from its closest relative, the camel isolate IVB6216, at 1,923 of 2,142 core loci (89.8%) and from the type strain NCTC12103 at 2,076 loci, occupying an isolated singleton branch in the MST (**Fig. 1**). This genome-scale distinctiveness independently supports the novel ST605 designation.

3.6. Phylogenomics and average nucleotide identity

In the midpoint-rooted core-genome maximum-likelihood phylogeny (**Fig. 2**), 30RR occupied an early-diverging position: the camel isolate IVB6216 branched basally, and 30RR formed the sister lineage to the main *M. sciuri* radiation, with the split separating 30RR and IVB6216 from all other genomes receiving maximal statistical support (100% bootstrap). This distinct placement is consistent with the novel ST605 designation and the below-average ANI. An independent reference-based core-genome SNP phylogeny (Parsnp; **Fig. S3**) was topologically concordant: 30RR again occupied an early-diverging position together with IVB6216, and the remaining genomes clustered in agreement with sequence type, host group and country of origin. Pairwise FastANI across all 23 genomes confirmed species-level cohesion: ANI ranged from 95.19% to 100% (mean 98.58%), and all pairwise comparisons (253/253) exceeded the 95% species threshold (**Fig. 3**). Strain 30RR shared no more than 96.7% ANI with any reference genome (closest: MSS74, 96.66%; PT0981, 96.66%; Dog031, 96.64%), below the dataset mean, indicating that 30RR has no close characterized relative. The most divergent pair overall was CAB16–IVB6216 (95.19%), and the most divergent comparator for 30RR was the camel isolate IVB6216 (96.30%). The full pairwise ANI matrix is provided in **Table S7**.

3.7. Pangenome structure

Pangenome analysis across 23 genomes yielded 5,745 gene clusters: 1,853 core (32.3%), 201 soft-core (3.5%), 768 shell (13.4%) and 2,923 cloud (50.9%) (**Fig. 4; Fig. S4; Table S8**). The pangenome rarefaction curve had not plateaued at $n = 23$ and followed Heap's law ($N = 2423n^{0.27}$; $R^2 = 0.99$); the exponent gamma = 0.27 (Tettelin alpha = $0.73 < 1$) confirms an open pangenome, while the core genome decayed to a stable estimate of ~1,846 genes. 30RR carried 211 unique genes (**Table S9**), of which 155 (73.5%) encode hypothetical proteins; the annotated remainder is

enriched for mobile elements and defence functions, including two *cas9* copies, IS3-family transposases and an SPbeta-prophage-derived endonuclease.

3.8. Resistome analysis and species-wide comparison

Screening across four AMR platforms (CARD, ResFinder, MEGARES and AMRFinderPlus) concordantly identified two ARGs in 30RR: *mecA1* and *sal(A)*. Mining the 23-genome pangenome matrix (**Table S10**), verified by BLASTp and eggNOG, resolved three further intrinsic resistance determinants that the strict-identity databases had missed: a class A beta-lactamase (*blaZ*; 68% protein identity to *S. aureus* BlaZ), an aminoglycoside 6-adenylyltransferase (*aadK*; ANT(6) family) and a putative MFS efflux transporter (*qacA*). *M. sciuri* 30RR therefore carries five intrinsic resistance genes (*mecA1*, *blaZ*, *sal(A)*, *aadK*, *qacA*), all of which are core or near-core across the species (present in 21–23 of 23 genomes).

The comparative resistome was the most striking finding (**Fig. 5**). *M. sciuri* 30RR carries no acquired or mobile ARGs. By contrast, an acquired SCCmec element (*mecA* together with the *mecI* and *mecR1* regulators and IS431mec) was detected in only three genomes, the duck isolate GDK8D6P, the chicken isolate WHA07 and the human isolate MSS74 which also carry the acquired cargo *tet(M)*, *tet(S)*, *cfr*, *optrA*, *fexA*, *erm(A)*, *linA* and *qacC* (acquired-ARG burden: GDK8D6P 6, WHA07 5, MSS74 3, 30RR 0). The intrinsic *mecA1* (22/23 genomes) is distinct from the SCCmec-acquired *mecA* (3/23), consistent with *mecA1* being the ancestral, vertically inherited homologue. PlasmidFinder detected no plasmid replicons, SCCmecFinder found no SCCmec element in 30RR, and VFDB returned no virulence-factor hits. The distribution of all 16 curated resistance genes across the 23 genomes is given in **Table S11**.

mecA1 (AY820253; 1,998 bp) is located on contig_10 (positions 14,882–16,882, forward strand) at 96.85% nucleotide identity and 100% coverage; ResFinder predicted resistance to 14 beta-

lactam agents, including penicillins, beta-lactam/inhibitor combinations, extended-spectrum cephalosporins, cephamycins and the three carbapenems tested. *sal(A)* (KC693025; 1,626 bp) is on contig_2 (positions 251,802–253,427, forward strand) at 96.13% identity and 100% coverage and encodes an ABC-F ribosomal-protection protein with predicted resistance to lincosamides, streptogramin A compounds and pleuromutilins. The two loci occupy contrasting genomic neighbourhoods (**Fig. 6**): *mecA1* lies within a cargo-rich region of contig_10 that also carries an NI-siderophore biosynthetic cluster (*sbnE/sbnH* with the *yfiY/yfiZ/yfhA* uptake genes), a mevalonate pathway (*mvaS/thlA/mvaA*) and efflux and metal-resistance determinants, whereas *sal(A)* sits within a conserved housekeeping block on contig_2 flanked by tRNA-ligase, stringent-response and DNA-repair genes.

3.9. Secondary metabolites, CRISPR-Cas, prophage remnants and mobile genetic elements

antiSMASH predicted four BGCs: a cyclic-lactone autoinducer cluster on NODE_6, a terpene-precursor cluster on NODE_9, a merged T3PKS/NI-siderophore region on NODE_10 (co-localized with *mecA1*) and a terpene cluster on NODE_25. BAGEL4 identified one sactipeptide locus on contig_6. CRISPRCasFinder identified five single-spacer CRISPR arrays (all Evidence Level 1) and a Cas9 gene on contig_13 (Type II-U), together forming a partial adaptive-immune locus. PHASTER detected two incomplete prophage remnants (contig_4 and contig_7). PathogenFinder2 classified 30RR as a predicted non-pathogen (ensemble score 0.439 ± 0.138). gutSMASH v2.0.1 predicted six primary metabolic gene clusters, encoding a putative 2-oxoglutarate ferredoxin oxidoreductase, a fumarate-to-succinate pathway, a nitrate reductase, a TPP-dependent amino-acid metabolism cluster, a caffeate-respiration/fatty-acid cluster and a PAPS-to-sulfide (sulfate-reduction) pathway. Insertion-sequence analysis identified 17 IS

elements (8 complete and 9 partial) spanning seven families: IS3 (n = 8), IS256 (n = 3), IS6 (n = 2) and single copies of IS21, IS110, IS200/IS605 and IS1182 distributed across 15 contigs. IslandViewer 4 predicted four genomic islands (9.2, 17.3, 37.0 and 47.6 kb; 104 genes in total), all detected by the IslandPath-DIMOB method, with the genome ordered against the *S. sciuri* subsp. *sciuri* NCTC12103 reference chromosome (**Fig. 7**). The islands harbour IS6-family transposases (IS431R), a cadmium-resistance transcriptional regulator (*cadC*), efflux proteins and an SPbeta-prophage-derived endonuclease (*yokF*); no curated virulence-factor or antimicrobial-resistance genes were annotated within the predicted islands. Notably, although 30RR carries IS431-family elements, the insertion sequences associated with SCCmec in the *mecA*-positive members of this species, no SCCmec element was detected, consistent with the SCCmecFinder and PlasmidFinder results.

3.10. Exopolysaccharide and capsule biosynthesis potential

Screening of the 23 proteomes against the curated staphylococcal capsule and EPS reference proteins identified 276 homologues (9-18 per genome; **Fig. 8; Table S12**). A complete *icaADBC* operon, encoding the PNAG/PIA exopolysaccharide, was found in only three genomes (WHA07, MSS74 and CAB16), and extended capsule (*cap*) loci carrying *capB* and *capD* together with seven or more *cap* genes were restricted to four genomes (Dog009_1, Dog108, GDK8D6P and IVB6216). In contrast, *M. sciuri* 30RR carried only 10 homologues on five contigs: the *ica* regulator *icaR* (34.4% identity) and a divergent *icaC* (30.4%) without *icaA*, *icaD* or *icaB*, and *cap* homologues limited to *cap5O* (*wecC*; 76.0%), *capO* (*wecC* paralog; 75.5%), *cap8P* (*mnaA*; 62.4%) and *capC* (*ywqE*; 36.6%); apart from the *tagG*–*wecC*–*icaC* region conserved across all 23 genomes, no contiguous *cap* biosynthetic cluster was detected. Thus, unlike several food-, poultry-

and dog-associated comparators, 30RR lacks the genetic capacity for PNAG/PIA exopolysaccharide and capsular polysaccharide production.

3.11. Epidemiological context

Across 1,540 complete milk quarter records from five Sylhet district locations, resistance-gene positivity was highest for gentamicin (32.3%) and tetracycline (30.3%), followed by sulphonamide (20.2%) and streptomycin (12.0%) (**Fig. S5; Table S13**). Positivity was markedly higher in Gowainghat and Jaintapur (~62–68% for gentamicin and tetracycline) than in Balaganj, Fenchuganj and Kanaighat (~16%), indicating substantial background AMR pressure in the region from which 30RR was isolated.

4. Discussion

This study provides the first whole-genome characterization of *M. sciuri* from dairy livestock in South Asia, isolated from a buffalo with SCM in Sylhet, Bangladesh. The high prevalence of subclinical mastitis in the study herds (67.9% of quarters; 80.8% of animals) is consistent with the well-documented burden of non-aureus staphylococci in buffalo livestock of Bangladesh. Subsequent antimicrobial susceptibility testing identified 30RR as multidrug-resistant, with resistance to five of thirteen antibiotics, prompting whole-genome sequencing to resolve its resistome and phylogenomic position.

One outstanding resistome result is that no acquired or mobile ARGs are observed in 30RR and that it possesses only five intrinsic ARGs, conserved across the species (*mecA1*, *blaZ*, *sal(A)*, *aadK*, and *qacA*). This is a stark contrast to the food-animal (duck GDK8D6P, chicken WHA07) and human (MSS74) comparators, which possess an acquired SCCmec element and a set of mobile resistance genes (*tet(M)*, *tet(S)*, *cfr*, *optrA*, *fexA*, *erm(A)*, *linA*, *qacC*). This makes 30RR more similar to the ancestral, wild-type resistome of *M. sciuri*, whilst the acquired cargo is confined to

428 lineages exposed to greater antibiotic selection pressures. As found in this species-wide analysis,
429 the intrinsic *mecA1* (22/23 genomes) is structurally and phylogenetically different from the
430 SCCmec-acquired *mecA* (3/23) and is accepted to be the evolutionary progenitor of the *mecA* gene
431 that defines MRSA [12, 13]. The IS-element complement reinforces this picture: although 30RR
432 carries IS431-family transposases, which flank SCCmec in the *mecA*-carrying members of the
433 species, no SCCmec element or other mobile ARG cargo was detected, consistent with a purely
434 intrinsic resistome.

435 The genotype-phenotype correlation helps explain apparent conflicts in the data. Although no
436 SCCmec cassette was found and the typical SCCmec regulatory elements (*mecR1* and *mecI*) are
437 lacking, 30RR was phenotypically resistant to the methicillin surrogate cefoxitin, suggesting that
438 the intrinsic PBP2a is expressed and functional in this strain. This, along with *blaZ*, accounts for
439 the beta-lactam resistance phenotype. By contrast, there is no definitive genotypic correlate for the
440 amikacin and tetracycline resistance phenotypes: the *aadK* ANT(6) enzyme is characteristically
441 streptomycin-specific, and the only candidate for tetracycline is a divergent Tet(X)-like
442 monooxygenase found across the species. These discrepancies leave a compelling case for
443 considering phenotypic susceptibility testing in addition to resistome analysis from WGS, both of
444 which are targets for further functional testing and MIC verification.

445 The result of the MLST provides geographic context. The novel sequence type ST605 contains
446 three novel alleles not found in the *M. sciuri* PubMLST database prior to this submission. 30RR
447 is genomically distinct from all characterized references (maximum ANI 96.7%, less than the
448 dataset mean of 98.6%) and occupies an early-diverging branch in the core-genome phylogeny.
449 We observe a difference in closest neighbour between the two methods: the core-genome tree
450 places 30RR as sister to IVB6216 (camel), while whole-genome ANI gives slightly higher values

451 to the human, sewage and dog isolates MSS74, PT0981 and Dog031. To explain this difference,
452 ANI averages over the whole genome of an isolate, whereas the core-genome tree reflects the
453 clonal frame; together with the long branches involved, recombination via a homologous process
454 best suggests that 30RR is a distinct lineage with no close relative, consistent with long-term
455 geographic isolation of a South Asian sub-lineage.

456 This all takes place in the background of the open pangenome (5,745 clusters, 50.9% cloud),
457 indicating a high level of continuous gene flow that may allow the species to take up, retain and
458 spread resistance determinants [11, 15]. A particularly interesting co-localization is that an NI-
459 siderophore BGC shares contig_10 with *mecA1*; a link to iron acquisition in the bacteriostatic
460 (lactoferrin-mediated) milk environment is speculative and requires experimental evidence. The
461 30RR draft assembly (40 contigs, N50 116 kb) is less contiguous than the 22 complete or near-
462 complete reference *M. sciuri* genomes (complete or chromosome level); however, its 100%
463 BUSCO/CheckM completeness indicates that gene models and hence the pangenome and
464 resistome comparisons are essentially unaffected by the draft-level fragmentation.

465 The functional genome is consistent with a commensal, host-adapted organism rather than a
466 primary pathogen: ABC transporters were the most abundantly mapped KEGG pathway, VFDB
467 returned no virulence factors, and PathogenFinder2 predicted a non-pathogenic phenotype. In line
468 with published clinical data, *M. sciuri* does not appear to be a primary pathogen in healthy animals,
469 but rather is responsible for the development of mastitis and possible sepsis in companion animals,
470 as well as sporadic infections in humans with reduced immune status [11, 13]. The regional
471 surveillance context is that a third of mastitic quarters were single or multiply positive for
472 resistance genes to gentamicin and tetracycline, which together indicates the selective environment
473 to which such lineages are added. This commensal profile extends to surface polysaccharides:

30RR lacks both the *icaADBC* exopolysaccharide operon and an intact capsule (*cap*) locus (**Fig. 8**), in contrast to several food-, poultry- and dog-associated comparators, suggesting limited biofilm and immune-evasion capacity consistent with its low-virulence genomic signature.

There are several limitations that need to be taken into account. Phenotypic susceptibility was determined by disc diffusion, but a broth-microdilution MIC determination was not performed, and therefore WGS-predicted resistance categories should be confirmed by MIC determination, especially for the unresolved resistance phenotypes of amikacin and tetracycline. The draft assembly (40 contigs) currently fails to resolve some of the flanking regions of the chromosomal loci. The *blaZ*, *aadK* and *qacA* assignments and the Tet(X)-like candidate are annotation and homology-based, and are supported by eggNOG and BLAST, but have not been functionally validated. Lastly, it is important to consider that one isolate cannot provide any insight into the distribution of this sub-lineage, and more circulation-based surveillance in South Asian buffalo herds is required to make inferences.

5. Conclusions

Strain 30RR is the first genome of an *M. sciuri* from dairy buffalo (swamp type) and South Asian dairy livestock. Unlike the multidrug-resistant food-animal/human lineages, it was isolated from buffalo milk in Sylhet, Bangladesh, and harbours an intrinsically acquired resistome, including the genes *mecA1*, *blaZ*, *sal(A)*, *aadK*, and *qacA*, without a detectable SCCmec cassette or an acquired resistance gene. Its early divergent phylogenomic position and pangenome suggests it belongs to a species with a global distribution, which is genetically dynamic, offering a genomic reference to be used for surveillance at the dairy buffalo-human interface in Bangladesh. The systematic

496 monitoring of non-aureus staphylococci for food-producing animals throughout the region should
497 be the next step in WGS-based monitoring.

498

499

500

501

502

503

504 **Abbreviations**

505 AMR: Antimicrobial resistance

506 ANI: Average nucleotide identity

507 ARGs: Antimicrobial resistance genes

508 AST: Antimicrobial susceptibility testing

509 BGCs: Biosynthetic gene clusters

510 CDS: Coding sequences

511 cgMLST: Core-genome multilocus sequence typing

512 CLSI: Clinical and Laboratory Standards Institute

513 CMT: California Mastitis Test

514 COG: Clusters of Orthologous Groups

515 dDDH: Digital DNA–DNA hybridization

516 EPS: Exopolysaccharide

517 GIs: Genomic islands

518 IMI: Intramammary infection

519 IS: Insertion sequence

520 MDR: Multidrug-resistant

521 MLST: Multilocus sequence typing

522 MST: Minimum-spanning tree

523 PBP2a: Penicillin-binding protein 2a

524 PGAP: Prokaryotic Genome Annotation Pipeline

525 QAC: Quaternary ammonium compound

526 SCM: Subclinical mastitis

527 SNP: Single-nucleotide polymorphism

528 ST: Sequence type

529 TYGS: Type Strain Genome Server

530 VFDB: Virulence Factor Database

531 VFGs: Virulence factor genes

532 WGS: Whole-genome sequencing

533

534 Tables

535 **Table 1.** In vitro antimicrobial susceptibility of *Mammaliicoccus sciuri* 30RR (Kirby-Bauer disc

536 diffusion, CLSI M100 2023) correlated with the WGS-predicted resistome.

Antibiotic	Abbreviation	Drug class	Phenotype (disc diffusion)	Candidate genotype	Genotype-phenotype assessment
Ampicillin	AMP	beta-lactam	R	blaZ + mecA1	Concordant
Amoxicillin-clavulanic acid	AMC	beta-lactam/inhibitor	R	mecA1 (PBP2a; clavulanate-insensitive)	Concordant
Cefoxitin	FOX	Cephamicin	R	mecA1 (PBP2a)	Concordant (methicillin resistance)
Ceftriaxone	CTR	3rd-gen cephalosporin	I	mecA1	Consistent
Amikacin	AK	Aminoglycoside	R	aadK (putative; ANT (6) typically streptomycin-specific)	Putative / unresolved

Gentamicin	CN	Aminoglycoside	I	aadK (putative)	Putative
Streptomycin	STR	Aminoglycoside	I	aadK (ANT (6) adenylyltransferase)	Consistent
Tetracycline	TE	Tetracycline	R	tet(X)-like FMO (divergent, ~30% id; candidate)	Putative / unresolved
Azithromycin	AZM	Macrolide	S	none detected (no erm)	Concordant (susceptible)
Chloramphenicol	C	Phenicol	S	none detected (no fexA/cfr)	Concordant (susceptible)
Ciprofloxacin	CIP	Fluoroquinolone	S	none detected	Concordant (susceptible)
Nalidixic acid	NA	Quinolone	I	none detected	Borderline
Trimethoprim-sulfamethoxazole	SXT	Folate-pathway inhibitor	S	none detected (no dfr/sul)	Concordant (susceptible)

Here: R, resistant; I, intermediate; S, susceptible.

545

546

547

548

549

550

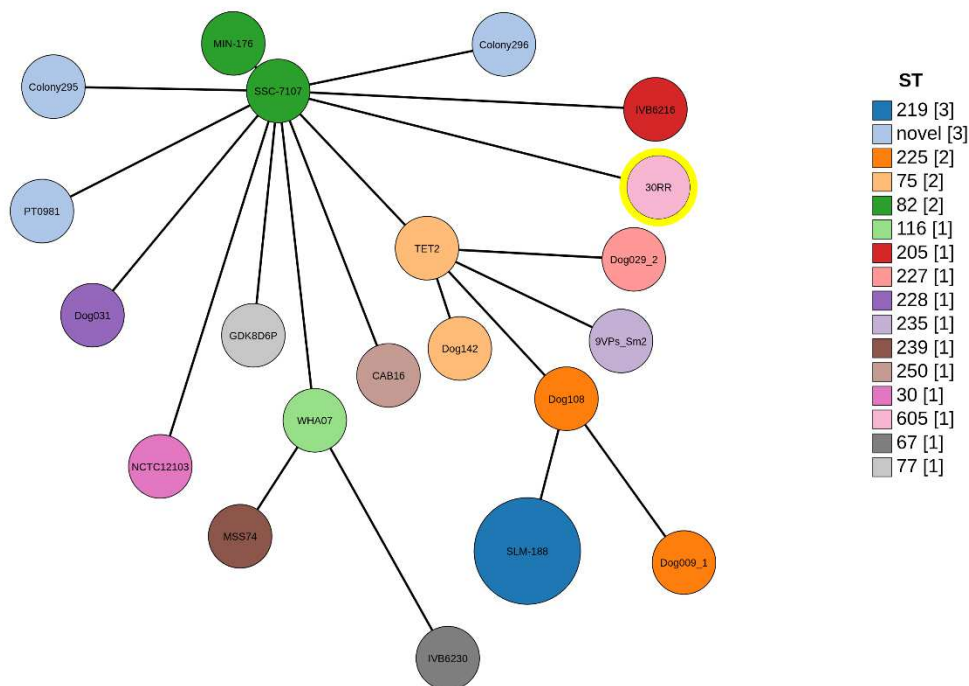
551 **Table 2.** Genomic features of *Mammaliicoccus sciuri* strain 30RR.

Feature	Value
Assembled genome size (bp)	2,545,126
GC content (%)	32.35
Contigs (≥ 500 bp)	40
N50 (bp)	116,238
L50	6
Largest contig (bp)	289,680
Mean sequencing depth (x)	99.66
BUSCO completeness (%)	100
Coding sequences (CDS)	2,565
rRNAs / tRNAs / tmRNA	4 / 30 / 1
MLST sequence type	ST605 (novel; PubMLST ID 1043)
Antimicrobial-resistance genes (intrinsic)	mecA1, blaZ, sal(A), aadK, qacA
Acquired / mobile ARGs	None detected
Virulence factor genes (VFDB)	None detected
Biosynthetic gene clusters (antiSMASH)	4 (cyclic-lactone autoinducer, 2x terpene, T3PKS/NI-siderophore)
CRISPR-Cas	Partial Type II-U (5 arrays; Cas9 on contig_13)
Prophage regions (PHASTER)	2 (incomplete)
Plasmid replicons (PlasmidFinder)	None detected
SCCmec element (SCCmecFinder)	Absent (methicillin-susceptible background)

552

PathogenFinder2 score	0.439 ± 0.138 (predicted non-pathogen)
-----------------------	--

553 **Figure legends**



554

555 **Fig. 1.** Core-genome MLST minimum-spanning tree of *Mammaliicoccus sciuri* (n = 23 genomes).

556 The tree was constructed and visualised with the MSTreeV2 algorithm in GrapeTree from pairwise

557 allelic differences across 2,142 core loci (chewBBACA). Nodes represent unique core-genome

558 allelic profiles, labelled by a representative genome; node size is proportional to the number of

559 genomes sharing each profile (counts in the legend). Strain 30RR (ST605) is highlighted with a

560 yellow ring. Edge lengths are proportional to the number of allelic differences.

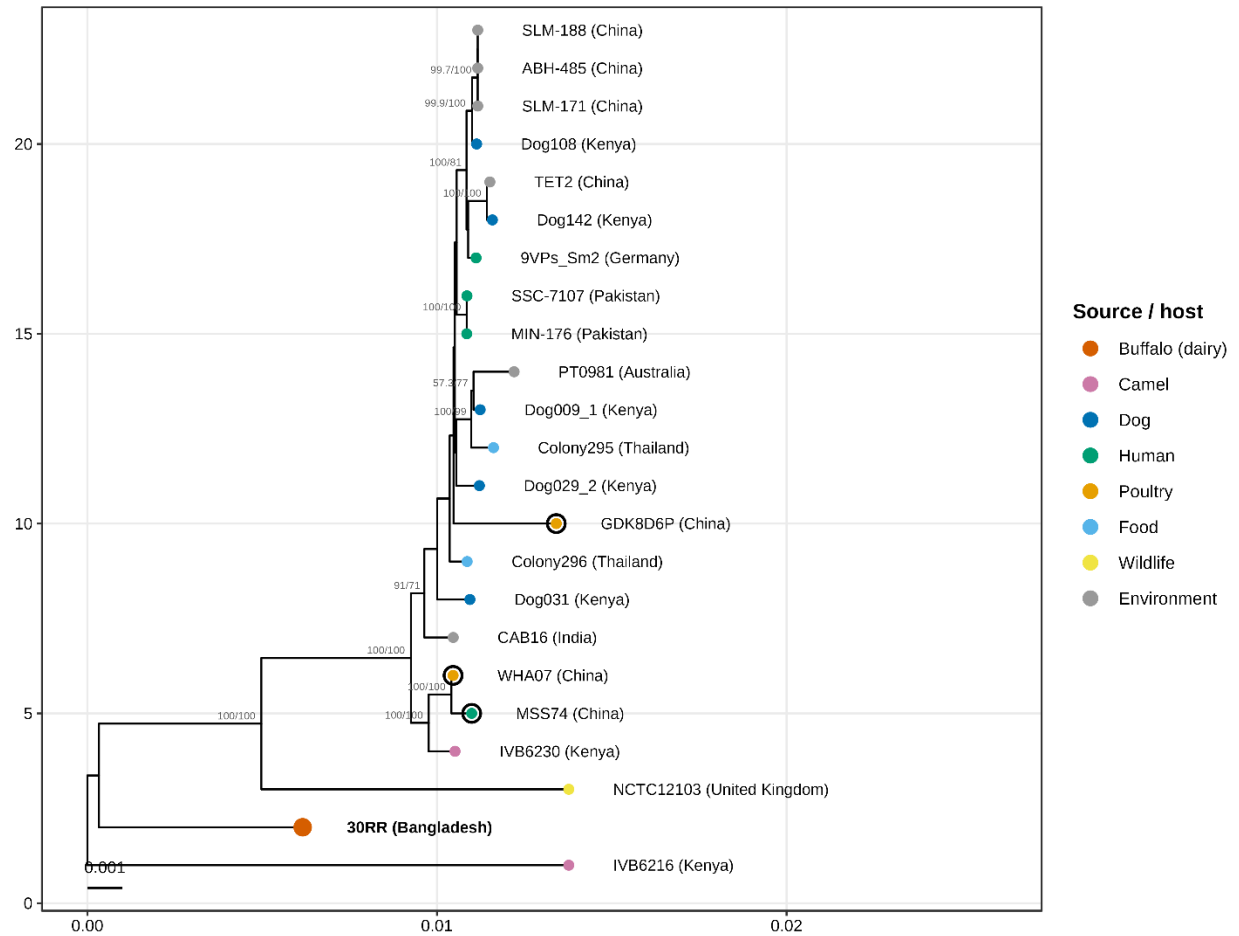


Fig. 2. Core-genome maximum-likelihood phylogeny of *Mammaliicoccus sciuri* 30RR and 22 reference genomes. The tree was inferred with IQ-TREE from a 152,856 bp core-genome alignment under the GTR+F+I+R4 model and midpoint-rooted. Node labels show SH-aLRT / ultrafast-bootstrap support (%). Tip points are coloured by source/host; black rings mark genomes carrying an acquired SCCmec element. Strain 30RR (bold) occupies an early-diverging position.

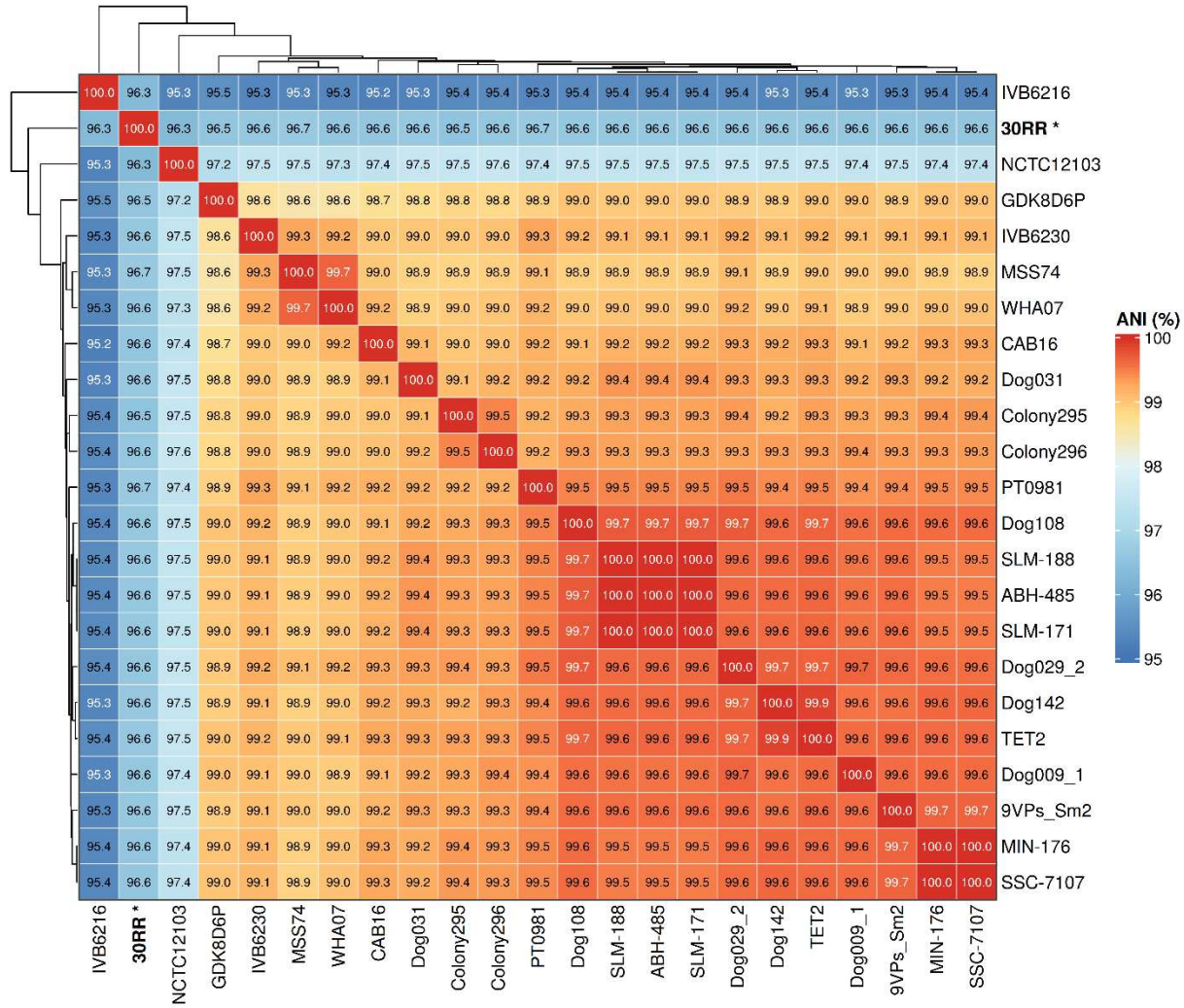


Fig. 3. Average nucleotide identity (ANI) among *Mammaliicoccus sciuri* genomes (n = 23). Pairwise FastANI values (%) are coloured by nucleotide identity; genomes are ordered by UPGMA hierarchical clustering of ANI distances. Strain 30RR (bold, marked with an asterisk) shares no more than 96.7% ANI with any reference genome.

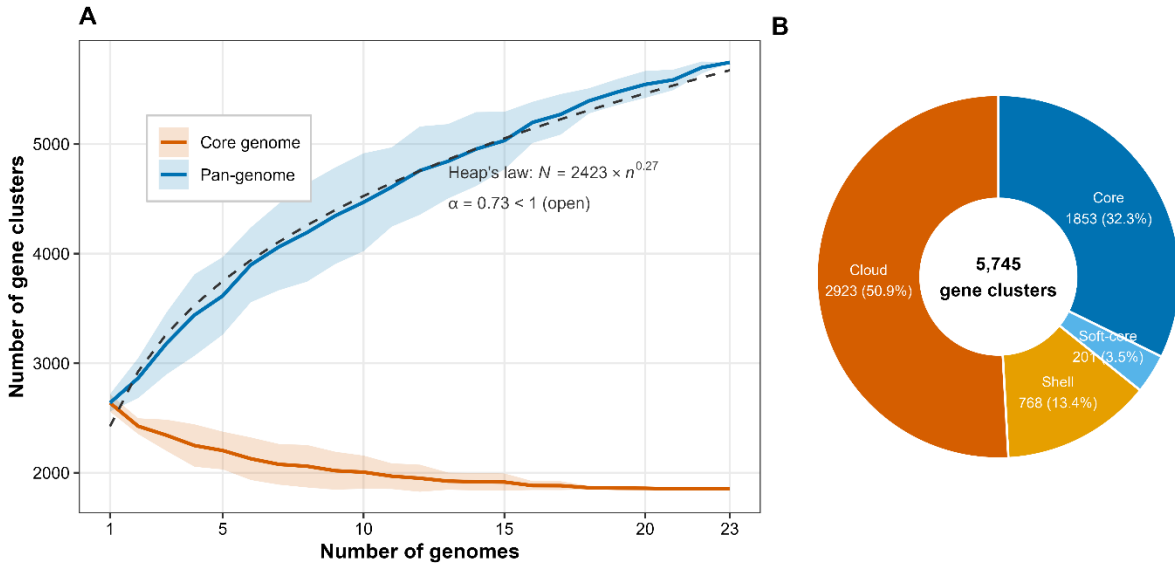


Fig. 4. Pangenome structure of *Mammaliicoccus sciuri* (n = 23 genomes). (A) Pan-genome and core-genome rarefaction curves (mean $\pm$ sd across Roary permutations); the dashed line is the Heap's-law power-law fit to the pan-genome (gamma = 0.27; Tettelin alpha = 0.73 < 1, open pangenome). (B) Pangenome composition: core, soft-core, shell and cloud gene clusters.

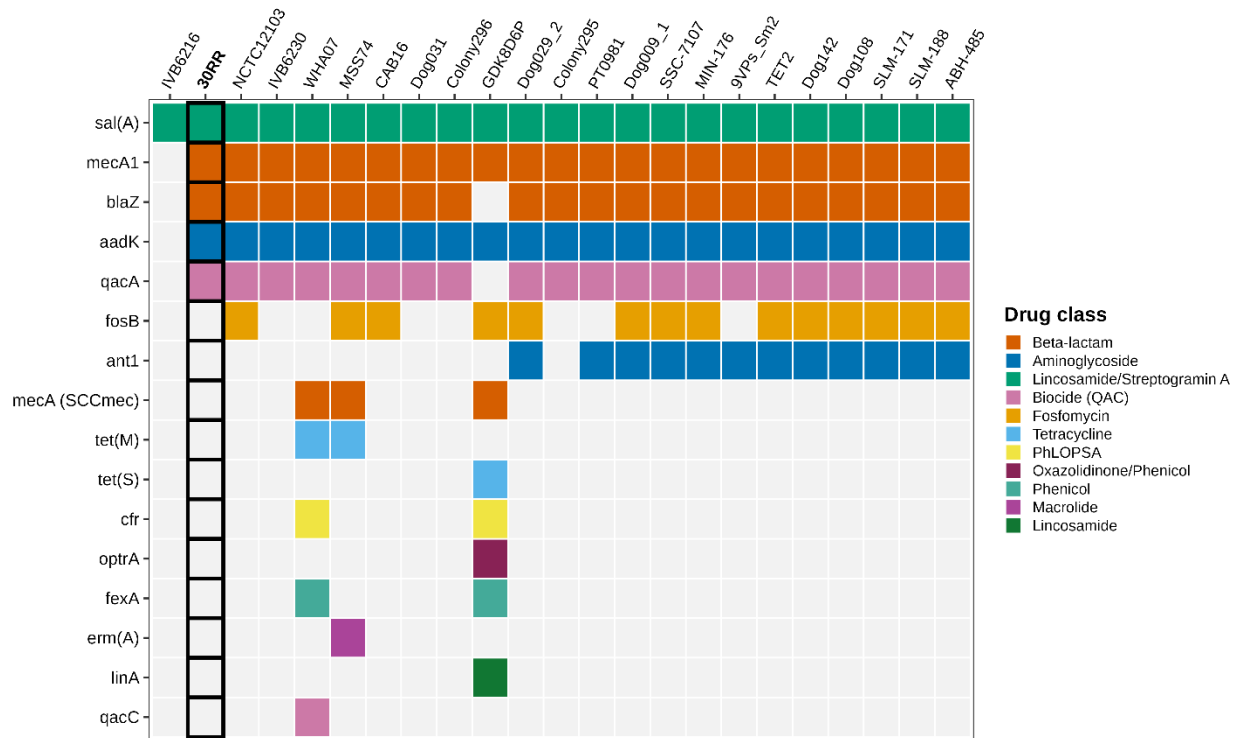
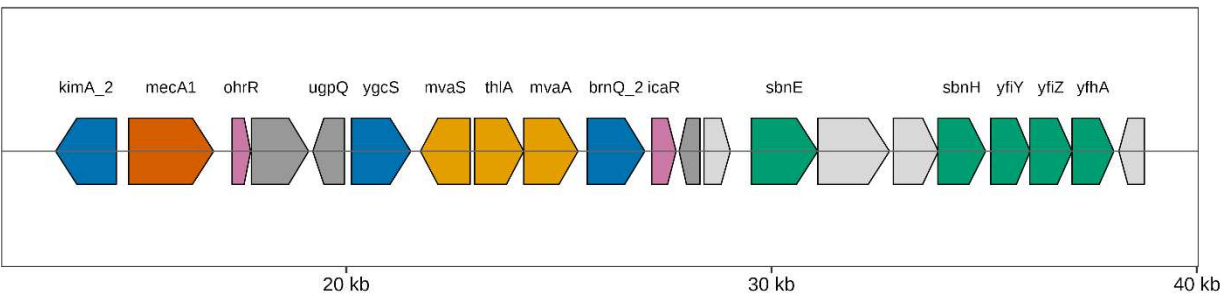


Fig. 5. Comparative resistome of *Mammaliicoccus sciuri*. Presence (coloured by drug class) and absence (light grey) of 16 curated antimicrobial-resistance genes across the 23 genomes (columns, in phylogenetic order). 30RR (boxed) carries only the five intrinsic genes (*sal(A)*, *mecA1*, *blaZ*, *aadK*, *qacA*) and no acquired resistance genes, whereas acquired cargo (SCCmec-*mecA*, *tet(M)/tet(S)*, *cfr*, *optrA*, *fexA*, *erm(A)*, *linA*, *qacC*) is concentrated in the duck (GDK8D6P), chicken (WHA07) and human (MSS74) isolates. The accessory determinants *fosB* and *ant1* are present in 14 and 12 of the 23 genomes, respectively, but absent from 30RR.

A *mecA1* locus (contig NODE_10)



B *sal(A)* locus (contig NODE_2)

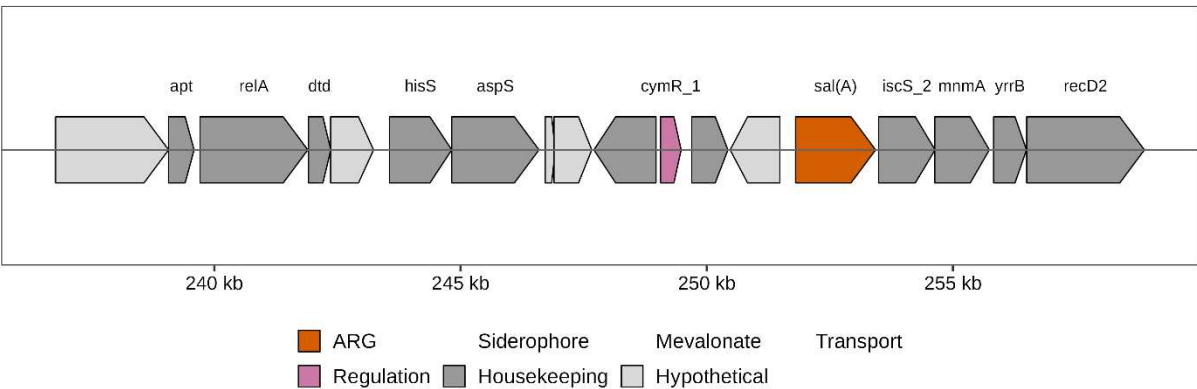


Fig. 6. Genomic neighbourhoods of the *mecA1* and *sal(A)* loci in *Mammaliicoccus sciuri* 30RR. (A) The *mecA1* locus on contig NODE_10 lies within a cargo-rich region containing an NI-siderophore biosynthetic cluster (*sbnE/sbnH* and the *yfiY/yfiZ/yfhA* uptake genes), a mevalonate pathway and efflux/metal-resistance determinants. (B) The *sal(A)* locus on contig NODE_2 lies within a conserved housekeeping block. Arrows denote coding sequences (direction = strand), coloured by functional category; the resistance genes are highlighted.

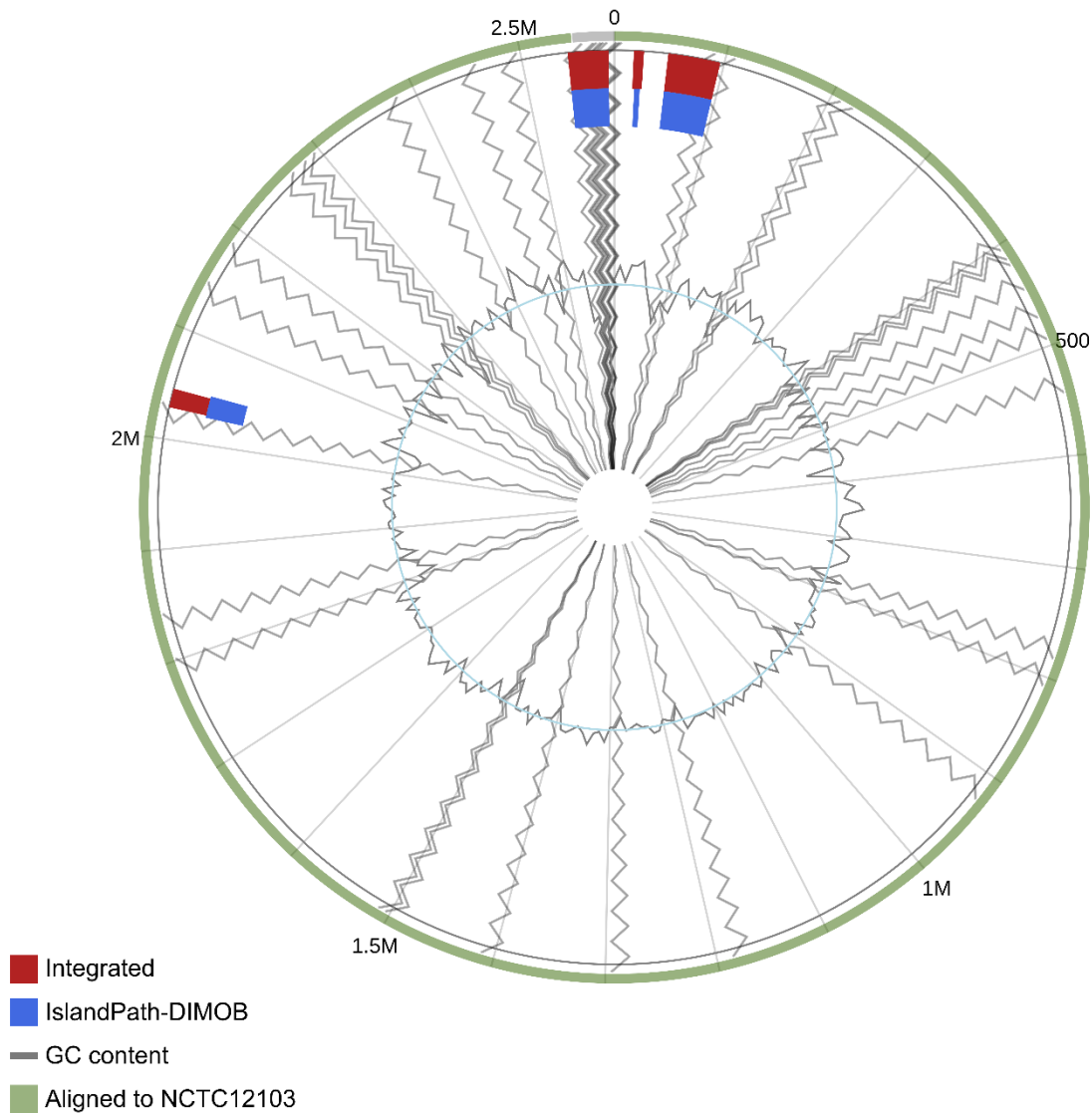


Fig. 7. Overview of predicted genomic islands (GIs) in *Mammaliicoccus sciuri* 30RR. Circular genome map generated with IslandViewer 4, with the genome ordered against the *S. sciuri* subsp. *sciuri* NCTC12103 reference chromosome (green outer ring). Islands predicted by the integrated method are shown in red and IslandPath-DIMOB predictions in blue; the inner grey trace shows GC content. Four islands (9.2–47.6 kb) were predicted; no curated virulence-factor or antimicrobial-resistance gene annotations were detected within them.

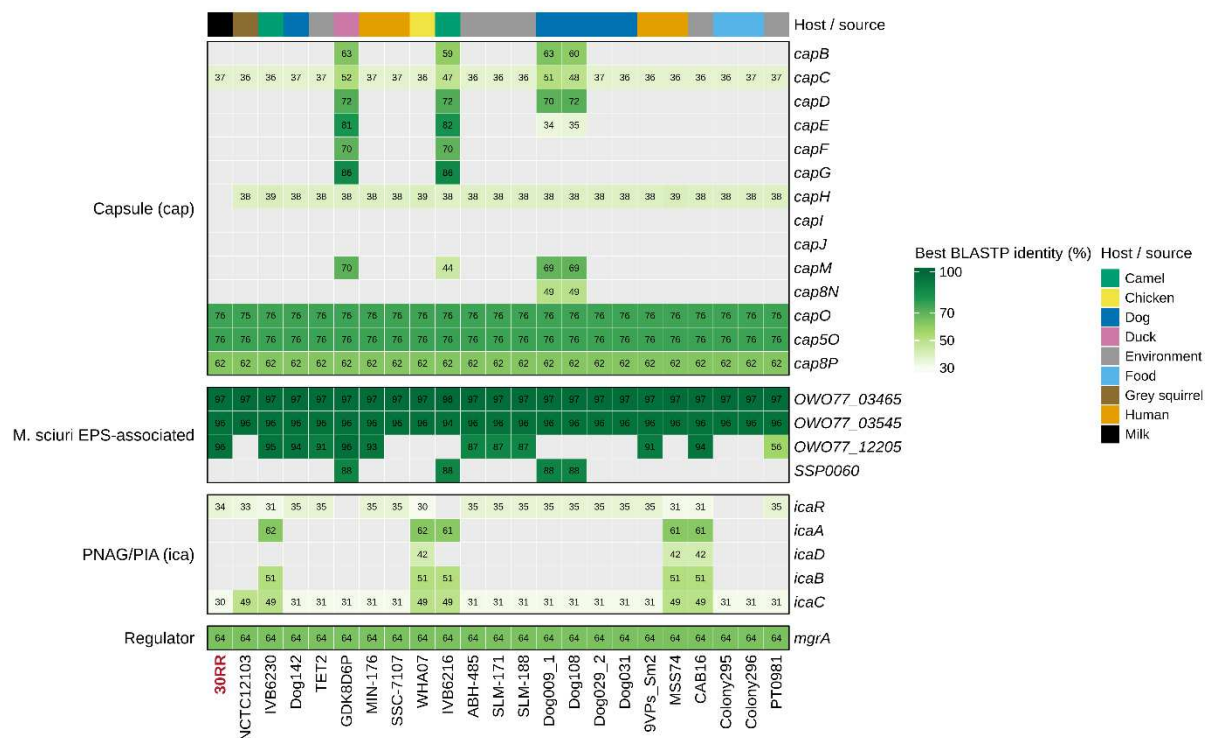


Fig. 8. Distribution of exopolysaccharide and capsule biosynthesis gene homologues across *Mammaliicoccus sciuri* genomes (n = 23). Tiles show the best BLASTP amino-acid identity (%) for each reference gene per genome (e-value $\leq 1 \times 10^{-5}$, $\geq 30\%$ identity, $\geq 40\%$ coverage); grey tiles indicate absence. Rows are grouped by functional category and columns are ordered by sequence type, with 30RR (red) first. A complete icaADBC operon occurs only in WHA07, MSS74 and CAB16, and extended cap loci only in Dog009_1, Dog108, GDK8D6P and IVB6216; 30RR carries only scattered, divergent homologues.

Acknowledgements

The authors thank the dairy buffalo farmers of Sylhet, Bangladesh, for their cooperation and consent, the Sylhet Agricultural University Research System (SAURES) for facilitating part of the study, and Dr. Michael Szostak (Institute of Microbiology, University of Veterinary Medicine Vienna, Austria) for curation of the novel MLST alleles in PubMLST.

Author Contributions Statement

MMR: Conceptualization, methodology, validation, visualization, supervision, and writing—original draft preparation; provided guidance in developing and refining the research concept and reviewed the research project and manuscript. MSRC: Investigation, sample collection, laboratory investigation, data curation, and writing—original draft preparation. SK: Data curation, software, formal analysis, and writing—original draft preparation. HH: Data curation, software, formal analysis, and writing—original draft preparation. MNH: Validation, visualization, supervision, and writing—review and editing; provided critical intellectual input and reviewed and revised the manuscript. All authors read and approved the final manuscript.

Funding

This study was partially funded by the University Grants Commission (UGC) of Bangladesh (Project Code: UGC-2021-22), awarded to Md. Mahfujur Rahman. The funding body had no role in study design, data collection, analysis, interpretation, or the decision to submit for publication.

Data availability statement

This Whole Genome Shotgun project has been deposited in GenBank under accession no. JBWGSY000000000 (version JBWGSY010000000.1). Raw sequencing reads are available under SRA accession no. SRR37536882. The BioProject and BioSample accession numbers are

642 PRJNA1434438 and SAMN56400945, respectively. Novel MLST alleles and the ST605 profile
643 are deposited in PubMLST (isolate ID 1043).

644 **Consent for publication**

645 Not applicable

646 **Competing interests**

647 The authors declare no competing interests.

648 **Consent to participate**

649 Not applicable

650

651

652

653

654

655

656

657

658

659

660

661

662 **References**

- 663 1. Kateete, D.P., et al., *Prevalence and Antimicrobial Susceptibility Patterns of Bacteria*
664 *from Milkmen and Cows with Clinical Mastitis in and around Kampala, Uganda*. PLoS
665 ONE, 2013. **8**(5): p. e63413.
- 666 2. Hoque, M.N., et al., *Different screening tests and milk somatic cell count for the*
667 *prevalence of subclinical bovine mastitis in Bangladesh*. Tropical animal health and
668 production, 2015. **47**(1): p. 79-86.
- 669 3. Hoque, M., et al., *Molecular characterization of Staphylococcus aureus strains in bovine*
670 *mastitis milk in Bangladesh*. International journal of veterinary science and medicine,
671 2018. **6**(1): p. 53-60.
- 672 4. Hoque, M.N., et al., *Metagenomic deep sequencing reveals association of microbiome*
673 *signature with functional biases in bovine mastitis*. Scientific reports, 2019. **9**(1): p.
674 13536.
- 675 5. Hoque, M.N., et al., *Insights into the resistome of bovine clinical mastitis microbiome, a*
676 *key factor in disease complication*. Frontiers in Microbiology, 2020. **11**: p. 860.
- 677 6. Kurban, D., et al., *Diagnosing Intramammary Infection: Meta-Analysis and Mapping*
678 *Review on Frequency and Udder Health Relevance of Microorganism Species Isolated*
679 *from Bovine Milk Samples*. Animals, 2022. **12**(23): p. 3288.
- 680 7. Addis, M.F., et al., *Non-aureus staphylococci and mammaliicocci isolated from bovine*
681 *milk in Italian dairy farms: a retrospective investigation*. Veterinary Research
682 Communications, 2024. **48**(1): p. 547-554.
- 683 8. Hoque, M.N., et al., *Genomic features and pathophysiological impact of a multidrug-*
684 *resistant Staphylococcus warneri variant in murine mastitis*. Microbes and Infection,
685 2024. **26**(3): p. 105285.
- 686 9. Madhaiyan, M., J.S. Wirth, and V.S. Saravanan, *Phylogenomic analyses of the*
687 *Staphylococcaceae family suggest the reclassification of five species within the genus*
688 *Staphylococcus as heterotypic synonyms, the promotion of five subspecies to novel*
689 *species, the taxonomic reassignment of five Staphylococcus species to Mammaliicoccus*
690 *gen. nov., and the formal assignment of Nosocomiicoccus to the family*
691 *Staphylococcaceae*. International Journal of Systematic and Evolutionary Microbiology,
692 2020. **70**(11): p. 5926-5936.
- 693 10. Sacramento, A.G., et al., *Genomic features of mecA-positive methicillin-resistant*
694 *Mammaliicoccus sciuri causing fatal infections in pets admitted to a veterinary intensive*
695 *care unit*. Microbial pathogenesis, 2022. **171**: p. 105733.
- 696 11. Van Der Veken, D., et al., *Genome-Based Characterization of a Plasmid-Associated*
697 *Micrococцин P1 Biosynthetic Gene Cluster and Virulence Factors in Mammaliicoccus*
698 *sciuri IMDO-S72*. Applied and Environmental Microbiology, 2022. **88**(4): p. e02088-21.
- 699 12. Bonnard, F.C., et al., *Frequency, Distribution, and Antimicrobial Resistance of*
700 *Methicillin-Resistant Staphylococci and Mammaliicoccus sciuri Isolated from Dogs and*
701 *Their Owners in Rio de Janeiro*. Antibiotics, 2025. **14**(4): p. 409.
- 702 13. Sacramento, A.G., et al., *Genomic features of mecA-positive methicillin-resistant*
703 *Mammaliicoccus sciuri causing fatal infections in pets admitted to a veterinary intensive*
704 *care unit*. Microbial Pathogenesis, 2022. **171**: p. 105733.
- 705 14. Megat Zakaria, P.N., et al., *Genomic Insights into Multidrug-resistant Mammaliicoccus*
706 *sciuri Carrying Novel Staphylococcal Cassette Chromosome mec (SCCmec) Elements*
707 *and Plasmids from Farm Ruminants on the East Coast of Peninsular Malaysia*. Current
708 Microbiology, 2026. **83**(1): p. 72.

15. Lienen, T., et al., *Mammaliococcus spp. from German Dairy Farms Exhibit a Wide Range of Antimicrobial Resistance Genes and Non-Wildtype Phenotypes to Several Antibiotic Classes*. Biology, 2022. **11**(2): p. 152.
16. Paterson, G.K., *Genomic epidemiology of methicillin-resistant Staphylococcus sciuri carrying a SCCmec-mecC hybrid element*. Infection, Genetics and Evolution, 2020. **79**: p. 104148.
17. Chowdhury, M.S.R., et al., *Detection and prevalence of antimicrobial resistance genes in multidrug-resistant and extensively drug-resistant Staphylococcus and Streptococcus species isolated from raw buffalo milk in subclinical mastitis*. Plos one, 2025. **20**(6): p. e0324920.
18. Singha, S., et al., *Factors influencing somatic cell counts and bacterial contamination in unpasteurized milk obtained from water buffalo in Bangladesh*. Tropical Animal Health and Production, 2023. **55**(4): p. 242.
19. Chowdhury, M.S.R., et al., *Emergence of highly virulent multidrug and extensively drug resistant Escherichia coli and Klebsiella pneumoniae in buffalo subclinical mastitis cases*. Scientific reports, 2025. **15**(1): p. 11704.
20. Singha, S., et al., *Pathogen group-specific risk factors for intramammary infection in water buffalo*. PLOS ONE, 2024. **19**(4): p. e0299929.
21. Urushibara, N., et al., *Analysis of Staphylococcal Cassette Chromosome mec in Staphylococcus haemolyticus and Staphylococcus sciuri : Identification of a Novel ccr Gene Complex with a Newly Identified ccrA Allotype (ccrA7)*. Microbial Drug Resistance, 2011. **17**(2): p. 291-297.
22. Chowdhury, M.S.R., et al., *Biofilm production and virulence traits among extensively drug-resistant and methicillin-resistant Staphylococcus aureus from buffalo subclinical mastitis in Bangladesh*. Scientific Reports, 2025. **15**(1): p. 34425.
23. Middleton, J.R., et al., *The National Mastitis Council: a global organization for mastitis control and milk quality, 50 years and beyond*. Journal of mammary gland biology and neoplasia, 2014. **19**(3): p. 241-251.
24. Sultana, T., et al., *Draft genome sequencing of multidrug-resistant Pseudomonas asiatica strains isolated from dairy cows with clinical mastitis and their farm environment*. Microbiology Resource Announcements, 2024. **13**(11): p. e00907-24.
25. Hoque, M.N., et al., *Whole-Genome Sequence of Multidrug-Resistant Escherichia coli Strains Isolated from Mice with Mastitis*. Microbiol Resour Announc, 2023. **12**(7): p. e0032023.
26. Hoque, M.N., et al., *Molecular characterization of Staphylococcus aureus strains in bovine mastitis milk in Bangladesh*. International Journal of Veterinary Science and Medicine, 2018. **6**(1): p. 53-60.
27. Kumaran, D., C. Laflamme, and S. Ramirez-Arcos, *A multiphasic approach to solve misidentification of Cutibacterium acnes as Atopobium vaginae during routine bacterial screening of platelet concentrates using the VITEK 2 system*. Access Microbiology, 2023. **5**(6).
28. Rai, S., D. Dash, and N. Agarwal, *Introducing the new face of CLSI M100 in 2023: An explanatory review*. Indian Journal of Medical Microbiology, 2023. **46**: p. 100432.
29. Brown, D.F., et al., *Guidelines for the laboratory diagnosis and susceptibility testing of methicillin-resistant Staphylococcus aureus (MRSA)*. J Antimicrob Chemother, 2005. **56**(6): p. 1000-18.
30. Cockerill, F.R., III, et al., *Performance Standards for Antimicrobial Disk Susceptibility Tests; Approved Standard—Eleventh Edition*, in CLSI document M02-A11. 2012, Clinical and Laboratory Standards Institute: Wayne, PA.
31. Andrews, S., *FastQC: a quality control tool for high throughput sequence data*. 2010, Babraham Bioinformatics: Cambridge, United Kingdom.

32. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data*. Bioinformatics, 2014. **30**(15): p. 2114-2120.
33. Bankevich, A., et al., *SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing*. Journal of Computational Biology: A Journal of Computational Molecular Cell Biology, 2012. **19**(5): p. 455-477.
34. Tatusova, T., et al., *NCBI prokaryotic genome annotation pipeline*. Nucleic Acids Research, 2016. **44**(14): p. 6614-6624.
35. Parks, D.H., et al., *CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes*. Genome Research, 2015. **25**(7): p. 1043-1055.
36. Manni, M., et al., *BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes*. Molecular Biology and Evolution, 2021. **38**(10): p. 4647-4654.
37. Silva, M., et al., *chewBBACA: A complete suite for gene-by-gene schema creation and strain identification*. Microb Genom, 2018. **4**(3).
38. Zhou, Z., et al., *GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens*. Genome Res, 2018. **28**(9): p. 1395-1404.
39. Alcock, B.P., et al., *CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database*. Nucleic acids research, 2020. **48**(D1): p. D517-D525.
40. Chen, L., et al., *VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on*. Nucleic Acids Research, 2016. **44**(D1): p. D694-D697.
41. Carattoli, A. and H. Hasman, *PlasmidFinder and in silico pMLST: identification and typing of plasmid replicons in whole-genome sequencing (WGS)*, in *Horizontal gene transfer: methods and protocols*. 2019, Springer. p. 285-294.
42. Kaya, H., et al., *SCC mec Finder, a web-based tool for typing of staphylococcal cassette chromosome mec in Staphylococcus aureus using whole-genome sequence data*. Msphere, 2018. **3**(1): p. 10.1128/msphere. 00612-17.
43. Arndt, D., et al., *PHASTER: a better, faster version of the PHAST phage search tool*. Nucleic acids research, 2016. **44**(W1): p. W16-W21.
44. Couvin, D., et al., *CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins*. Nucleic acids research, 2018. **46**(W1): p. W246-W251.
45. Cosentino, S., et al., *PathogenFinder-distinguishing friend from foe using bacterial whole genome sequence data*. PloS one, 2013. **8**(10): p. e77302.
46. Xie, Z. and H. Tang, *ISEScan: automated identification of insertion sequence elements in prokaryotic genomes*. Bioinformatics, 2017. **33**(21): p. 3340-3347.
47. Bertelli, C., et al., *IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets*. Nucleic Acids Res, 2017. **45**(W1): p. W30-W35.
48. Jain, C., et al., *High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries*. Nature Communications, 2018. **9**(1): p. 5114.
49. Wong, T.K.F., et al., *IQ-TREE 3: phylogenomic inference software using complex evolutionary models*. Mol Biol Evol, 2026. **43**(5).
50. Yu, G., et al., *Two Methods for Mapping and Visualizing Associated Data on Phylogeny Using Ggtree*. Mol Biol Evol, 2018. **35**(12): p. 3041-3043.
51. Kille, B., et al., *Parsnp 2.0: scalable core-genome alignment for massive microbial datasets*. Bioinformatics, 2024. **40**(5).
52. Page, A.J., et al., *Roary: rapid large-scale prokaryote pan genome analysis*. Bioinformatics, 2015. **31**(22): p. 3691-3693.
53. Tettelin, H., et al., *Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome"*. Proc Natl Acad Sci U S A, 2005. **102**(39): p. 13950-5.

54. Team, R.C., *R: a language and environment for statistical computing*. 2025, R Foundation for Statistical Computing: Vienna, Austria.
55. Aziz, R.K., et al., *The RAST Server: Rapid Annotations using Subsystems Technology*. BMC Genomics, 2008. **9**(1): p. 75.
56. 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 Research, 2019. **47**(D1): p. D309-D314.
57. Moriya, Y., et al., *KAAS: an automatic genome annotation and pathway reconstruction server*. Nucleic Acids Research, 2007. **35**(Web Server): p. W182-W185.
58. Zheng, J., et al., *dbCAN3: automated carbohydrate-active enzyme and substrate annotation*. Nucleic Acids Research, 2023. **51**(W1): p. W115-W121.
59. Lombard, V., et al., *The carbohydrate-active enzymes database (CAZy) in 2013*. Nucleic Acids Research, 2014. **42**(D1): p. D490-D495.
60. Blin, K., et al., *antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation*. Nucleic acids research, 2025. **53**(W1): p. W32-W38.
61. Zhu, Y., et al., *gutSMASH 2.0: Extended Identification of Primary Metabolic Gene Clusters From the Human Gut Microbiota*. J Mol Biol, 2026. **438**(18): p. 169567.
62. O'Riordan, K. and J.C. Lee, *Staphylococcus aureus capsular polysaccharides*. Clin Microbiol Rev, 2004. **17**(1): p. 218-34.
63. Cramton, S.E., et al., *The intercellular adhesion (ica) locus is present in Staphylococcus aureus and is required for biofilm formation*. Infect Immun, 1999. **67**(10): p. 5427-33.
64. UniProt, C., *UniProt: the Universal Protein Knowledgebase in 2025*. Nucleic Acids Res, 2025. **53**(D1): p. D609-D617.
65. Camacho, C., et al., *BLAST+: architecture and applications*. BMC Bioinformatics, 2009. **10**: p. 421.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.xlsx](#)
- [SupplementaryFigures.docx](#)