Sequence characterisation and novel insights into bovine mastitis-associated *Streptococcus uberis* in dairy herds

Ben Vezina¹,³,⁸, Hulayyil Al-harbi², Hena R. Ramay⁴, Martin Soust⁵, Robert J. Moore⁶, Timothy W. J. Olchowy¹,⁷ & John I. Alawneh¹,²,⁸

*Streptococcus uberis* is one of the most frequent mastitis-causing pathogens isolated from dairy cows. Further understanding of *S. uberis* genetics may help elucidate the disease pathogenesis. We compared the genomes of *S. uberis* isolates cultured from dairy cows located in distinctly different geographic regions of Australia. All isolates had novel multi locus sequence types (MLST) indicating a highly diverse population of *S. uberis*. Global clonal complexes (GCC) were more conserved. GCC ST86 and GCC ST143 represented 30% of the total isolates (n = 27) and were clustered within different geographic regions. Core genome phylogeny revealed low phylogenetic clustering by region, isolation source, and MLST. Identification of putative sortase (srtA) substrates and generation of a custom putative virulence factor database revealed genes which may explain the affinity of *S. uberis* for mammary tissue, evasion of antimicrobial efforts and disease pathogenesis. Of 27 isolates, four contained antibiotic resistance genes including an antimicrobial resistance cluster containing mel/mef(A), mrsE, vatD, InuD, and transposon-mediated InuC was also identified. These are novel genes for *S. uberis*, which suggests interspecies lateral gene transfer. The presence of resistance genes across the two geographic regions tested within one country supports the need for a careful, tailored, implementation and monitoring of antimicrobial stewardship.

Mastitis poses a risk to public health, animal welfare, and farm profitability worldwide¹⁻⁴. Increased costs related to clinical mastitis are the result of reduced milk production and price, discarded milk, animal culling, mortality, labour and herd veterinary costs⁵. Both contagious and environmental pathogens have been implicated in intramammary infections and bovine mastitis. *Streptococcus uberis* is one of the most frequently identified environmental pathogens responsible for mastitis in dairy herds, a trend which appears to be increasing worldwide⁶⁻⁸. Mastitis control programs have been effective in decreasing the prevalence of contagious pathogens in well managed dairy herds⁹, but *S. uberis* remains responsible for a significant proportion of both clinical and subclinical intramammary infections⁵. In Australia, *S. uberis* associated mastitis accounts for one of every three cases of clinical mastitis⁹. This pathogen can be found in a variety of sites in the dairy cow environment such as bedding materials and milking equipment, as well as on other items within the dairy cows’ daily living spaces⁴. However, *S. uberis* is not only an environmental risk, this pathogen has also been associated with cow-to-cow transmission consistent with a contagious pathogen role for *S.uberis*¹⁰⁻¹² and indicative of the complexity of the bacteria, environment and dairy cow relationship.

A number of different DNA-based techniques have been used to investigate the genetic diversity of *S. uberis* isolated from cases of bovine mastitis, including MLST¹¹⁻¹⁴. While MLST has provided insights into the population diversity of *S. uberis*¹¹⁻¹⁴, it is limited in its ability to delve deeply into genetic variation of isolates. The recent availability of low-cost Whole Genome Sequencing (WGS) technology and advances in computational

¹Good Clinical Practice Research Group (GCPRG), The University of Queensland, School of Veterinary Science, Gatton, QLD 4343, Australia. ²The University of Queensland, School of Veterinary Science, Gatton, QLD 4343, Australia. ³Centre for Cell Factories and Biopolymers, Griffith Institute for Drug Discovery, Griffith University, Nathan, Australia. ⁴International Microbiome Centre, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada. ⁵Terragen Biotech Pty Ltd., Coolum Beach, QLD 4573, Australia. ⁶School of Science, RMIT University, Bundoora, Melbourne 3083, Australia. ⁷Faculty of Veterinary Medicine, University of Calgary, Calgary, AB T3R 1J3, Canada. ⁸These authors contributed equally: Ben Vezina and John I. Alawneh. *email: j.alawneh@uq.edu.au*
Biology has made it possible to compare the entirety of the *S. uberis* genome and completely characterise the levels of heterogeneity between strains. WGS analysis of the genomic sequence of *S. uberis* strains has revealed both a capacity for nutritional flexibility and a large variety of metabolic capabilities that promote bacterial survival in different environments, potentially responsible for the high degree of genetic diversity of *S. uberis* within the dairy herd environments. Variations in genetic profile among *S. uberis* strains were not only observed in isolates originating from different dairy herds, but also in the pathogens present within the same herd. This inherently high genetic diversity renders the attribution of a specific set of virulence factors (and associated pathogenesis) to a given *S. uberis* strain/s challenging. The development of strategies to control this important mastitis-causing organism necessitates monitoring of the genetic diversity and determination of the significance of any detected diversity to the disease pathogenesis. Therefore, the main objectives of this study were to compare *S. uberis* isolates cultured from dairy cows with clinical mastitis in distinctly different geographic and climatic regions of Australia. We then identified antibiotic resistance genes, virulence factors, and mobile genetic elements which could then inform antimicrobial stewardship programs. A secondary objective was to examine the diversity of the isolates, in particular their specific antimicrobial resistance profiles from these different regions to identify geographic or climatic associations that may exist as a potential basis for development of more effective and sustainable mastitis treatments for dairy cows.

**Results**

**S. uberis mastitis isolates, sequence types and phylogenetic relationships.** Eleven isolates from 10 locations in Victoria (VIC) and 16 isolates from 8 sites in Queensland (QLD) for a total of 27 isolates were collected and analysed in this study (Fig. 1). The 27 *S. uberis* genomes ranged in size from 1,832,555 to 2,019,389 bp.
There were 21 MLST sequence types (STs) represented within this collection. All sequence types were considered novel as they matched no known sequence types in the S. uberis MLST database (pubmlst.org/suberis). The closest ST matches were ST ~ 1223 (4/27) followed by ST ~ 222 (3/27). These were only present in QLD farms. Analysis of GCC STs showed the most abundant STs to be GCC ST86 (8/27) and GCC ST143 (8/27) with GCC ST5 representing 4/27 isolates. All GCC ST86 isolates were from herds with high bulk milk somatic cell counts (270–400 × 10^3 cells/mL). GCC ST143 isolates were predominantly found in Victoria (75%) while GCC ST86 isolates were predominantly found in Queensland (88%). Seven isolates (27%) did not belong to a defined GCC. Based on the goeBURST analysis, a possible founder ST for those 7 isolates was ST251. The relationship between the STs of this study and those reported globally in the S. uberis MLST database (pubmlst.org/suberis) is depicted in Fig. 2.

Pangenome analysis. Pangenome analysis showed a total gene pool of 3764 genes (including RNA encoding), with the core genome making up 1509 genes (Table 2). The soft-core genome, found in ≥ 26/27 isolates, was made up of 1542 genes. Of these, 1380 genes (89.49%) could be assigned a Gene Ontology (GO) identity and descriptors (Data S5). The most abundant class of genes, based on predicted localisation of the encoded proteins was as ‘integral member of the membrane’ followed by ‘cytoplasm location’. Most genes appeared to be housekeeping. In addition, 64 encoded proteins predicted to be associated with proteolysis, 5 associated with virulence, 4 with adhesion, 1 with capsule polysaccharide biosynthetic process, and 1 was pilus-associated. Of the 5 virulence-associated genes, 3 encoded proteins that were predicted to be secreted: pauA, LytR, and the YSIRK-type signal peptide-containing protein. Several proteins associated with nutrients commonly found in abundance within the mammary tissue of dairy cows were also identified including fibrinogen binding proteins (WP_046389274.1, WP_012658173.1) and lactose catabolism proteins (lacA, WP_000215993.1; lacB, WP_000686149.1).

Virulence factors. To address the lack of traditional S. uberis virulence factors detected from the Virulence Factor Database (VFDB), a custom S. uberis virulence database was constructed, named the S. uberis Putative Virulence Database (SuPVDB; Data S2). This was built based on three main sources. Firstly, identification of translated homologues from the VFDB using relaxed alignment statistics (> 70% query cover and > 60% similarity)

| Isolate | Genome size (bp) | Source | State | MLST ST | GCC ST |
|---------|------------------|--------|-------|---------|--------|
| S. uberis 22A | 1,938,913 | QLD1 | Queensland | ~ 222 | 86 |
| S. uberis 31A | 1,900,760 | QLD2 | Queensland | ~ 97 | – |
| S. uberis 36A | 1,914,697 | QLD3 | Queensland | ~ 222 | 86 |
| S. uberis 38B | 1,958,229 | QLD3 | Queensland | ~ 253 | 5 |
| S. uberis 42 | 1,883,527 | VIC1 | Victoria | ~ 271 | 143 |
| S. uberis 43 | 1,925,541 | VIC2 | Victoria | ~ 402 | 5 |
| S. uberis 44 | 1,879,665 | VIC3 | Victoria | ~ 89 | 143 |
| S. uberis 45 | 1,961,274 | VIC4 | Victoria | ~ 190 | – |
| S. uberis 46 | 1,899,378 | VIC5 | Victoria | ~ 132 | 143 |
| S. uberis 47 | 1,971,757 | QLD4 | Queensland | ~ 280 | 143 |
| S. uberis 48 | 1,927,709 | QLD5 | Queensland | ~ 222 | 86 |
| S. uberis 49 | 2,019,389 | QLD6 | Queensland | ~ 145 | 143 |
| S. uberis 50 | 1,938,127 | QLD7 | Queensland | ~ 425 | 86 |
| S. uberis 51 | 1,984,097 | QLD8 | Queensland | ~ 56 | 86 |
| S. uberis 56A | 1,906,743 | VIC6 | Victoria | ~ 261 | 143 |
| S. uberis 60A | 1,967,409 | VIC6 | Victoria | ~ 412 | 86 |
| S. uberis 69B | 1,926,363 | VIC7 | Victoria | ~ 271 | – |
| S. uberis 70A | 1,919,457 | VIC8 | Victoria | ~ 256 | 143 |
| S. uberis 75B | 1,844,929 | VIC9 | Victoria | ~ 156 | 143 |
| S. uberis 86A | 1,977,432 | VIC10 | Victoria | ~ 90 | – |
| S. uberis 93A | 1,856,167 | QLD9 | Queensland | ~ 67 | 5 |
| S. uberis 106A | 1,832,555 | QLD6 | Queensland | ~ 1223 | 86 |
| S. uberis 107A | 1,832,592 | QLD6 | Queensland | ~ 1223 | – |
| S. uberis 114A | 1,927,632 | QLD7 | Queensland | ~ 1223 | – |
| S. uberis 118A | 1,894,094 | QLD8 | Queensland | ~ 850 | 5 |
| S. uberis 122A | 1,886,974 | QLD8 | Queensland | ~ 80 | – |
| S. uberis 124A | 1,909,330 | QLD8 | Queensland | ~ 1223 | 86 |

Table 1. Summary of S. uberis strains identified in this study, isolated from cows affected by bovine mastitis. MLST STs are also shown. ~ indicates closest ST. All STs were considered novel. The isolation sources have been given codes for privacy reasons.
(Data S3). Secondly, experimentally-confirmed virulence factors identified from a number of publications\(^{26–41}\). Finally, inclusion of genes with virulence-associated GO descriptors (Data S5). Despite HasA\(^{42}\), PauA\(^{43}\), and PauB\(^{39}\) being non-essential for bovine mastitis, these three were included in the search for the sake of completeness. MtsB and regulator ScaR were also included as potential virulence factors because they were associated with the MtuA operon. To identify homologous virulence factors in \(S.\) \(uberis\), the \(Streptococcal\) VFDB amino acid sequences were screened against the translated pangenome. Any sequences with > 70% query cover and > 60% similarity were considered putative homologues. Twenty-eight previously undescribed putative virulence factors were identified in \(S.\) \(uberis\) based on homology to known virulence factors in other \(Streptococcal\) species (Data S3). This custom database consisted of 53 total sequences (Data S2), which were screened against each isolate, as well as all 40 \(S.\) \(uberis\) isolates present within GenBank (Data S6).

The GO virulence-associated genes encoded an enolase/phosphopyruvate hydratase (WP_012658173.1), LytR family regulatory protein (SQG46876.1), YSIRK-type signal peptide-containing protein (WP_154591130.1), PauA a plasminogen activator A (WP_046388868.1), and a HU family DNA-binding protein, HlpA (WP_012658741.1). 90 soft core proteins were predicted to be secreted including PauA, LytR, and the YSIRK-type signal peptide-containing protein were all predicted to be secreted, along with pneumococcal-type histidine triad protein (Pht).

**Figure 2.** Multilocus sequence type (MLST)-based minimal spanning tree of 1220 \(Streptococcus\) \(uberis\) isolated from bovine mastitis milk as per NCBI database\(^{103}\) (as of 8 June 2020) to date. Coloured circles are Australian isolates from bovine origin \((n = 240)\). The tree was calculated and generated using the goeBURST version 1.2.1 (http://www.phylolviz.net/goeburst/\(^{113}\)) full MST algorithm in Phyloviz version 2.0 (www.phylolviz.net\(^{112}\)). Node sizes reflect the number of isolates with specific MLST profile. Numbers within the nodes indicate the ST. Node colours refer to types of Global Clonal Complexes (GCC): GCC Strain Type (ST) 86 (green; left side of the plot), GCC ST143 (orange, right-bottom side of the plot), GCC ST5 (black, right-top side of the plot), isolates from Queensland are in red colour and Victorian isolates are in blue. Dashed rectangle show the location of the seven isolates to their possible ST founder ST251 (asterisk).
The average number of virulence factors per isolate was 27.7 factors (± Standard deviation [σ] = 2.6) from the current study. When the publicly available S. uberis genomes were screened, an average of 27.86 ± 2.14 σ putative virulence factors were identified from isolates associated with bovine mastitis (Data S6). The 3 human mastitis isolates also contained a similar amount, 27.67 ± 4.04 σ. The aquatic S. uberis isolate CAIM 1894 (SAMN03093229) contained a total of 3 putative virulence factors.

In the current study, the number of virulence factors per isolate ranged from 35 in S. uberis 36A to 24 in S. uberis 124A. Twenty-two putative virulence factors were found in every isolate. These factors included biofilm putative glycosyltransferase 2, cpsB, cpsC, cpsD/cps4D, cyIA, fabG 1/cyLG, fbpS, gtaB/hasC, hasC homologue gpsA, lmb, mga, mtaA, oppF, pauA, putative surface-anchored protein, rqcH/fbp54, scaR, scpA, sua, tagU 3/cps4A and srtA (Fig. 4, Data S4). Moreover, there was a tendency for virulence factors to cluster by state (p = 0.09). No clustering was observed between isolation source and virulence factors. The accessory virulome was made up of putative virulence factors not present in every isolate. They consisted of 24 sequences (Fig. 4), with individual isolates containing the majority of them. S. uberis 56A was the only isolate to contain the putative virulence factors cpsL and neuABCD (epsM/neuD, legI/neuB, neuA_1, neuC).

### Putative sortase substrates, antimicrobial resistance and mobile genetic elements.

**Streptococcal** sortase srtA has been shown to be essential for bovine mastitis and a previous study identified 9 proteins attached to the cell wall of S. uberis which were likely virulence-associated genes [45]. This enzyme cleaves and processes proteins containing particular motifs to allow external cell wall-anchoring. To confirm the importance of cell wall-anchored proteins, isolates were screened for the srtA gene, which had two variants with 85.9% similarity between them. 22/27 strains contained one srtA variant (WP_046390884.1) while 5/27 strains contained another (WP_037592200.1). This was the result of the 90% identity cut-off criteria used in the pangenome analysis. Within this constraint, sortase srtA was still considered a core gene.

To analyse SrtA substrates, core protein-coding sequences were screened for several motifs including LPXTG [44, 45], LPXXXD [44], LPXTA [46, 47], QVPTGV and LPSTGE [48]. 12 protein-coding sequences were identified when excluding the more variable LPXXXD motif, and 102 including it. As only core genes were screened, there was an increased likelihood of finding essential virulence factors. Some putative sortase substrates of note included a transmembrane protein containing a xanthine permease domain (WP_046392124.1), 7 coding sequences (CDS) associated with the bacterial global response to DNA damage (SOS), peroxiredoxin (WP_015912043.1), and transmembrane branched-chain amino acid transport protein (WP_046393325.1) (Data S5).

Overall, antibiotic resistance genes were not commonly found within the genomes. Only S. uberis 45, 47, 48 and 51 contained resistance genes and most of these were predicted to encode resistance to lincosamides, due to presence of lnuC or lnuD (Table 3). Antibiotic resistance genes were found in both Victorian and Queensland isolates and in a number of STs. Screening NCBI for LnuC revealed it was not identified in any S. uberis isolates in the database. A multidrug antibiotic resistance cluster was identified in both S. uberis 48 and 51. The cluster consisted of mefA, mrsE, vatD and lnuD. Despite S. uberis 47 containing lnuD, it did not contain the cluster. The two isolates containing the cluster were found in different isolation sources within Queensland. They both belonged to GCC ST86. The gene cluster is associated with bacterial resistance to lincosamides, macrolides, oxazolidinones, phenicols, pleuromutilins, streptogramins and tetracyclines.

One-third (9/27) of the isolates contained putatively intact bacteriophages (Table 4). S. uberis 38B contained two bacteriophages. The Streptococcus prophage 315.2 was found in isolates S. uberis 38B and S. uberis 47, both Queensland isolates. There were no virulence genes associated with any bacteriophages. Transposons and insertion sequences were identified in 21/27 isolates (77.78%) (Table 5). None were associated with virulence factors. No plasmids were found in the isolates.

### Discussion

Twenty-nine herds located in three climatically distinct geographic regions of Australia were screened and from them 27 independent S. uberis isolates were obtained and their genetic relatedness was assessed. The diverse range of MLSTs observed in this study is consistent with previous surveys of bovine mastitis-associated S. uberis in that there was a large variety of STs identified [49]. In one study, 33 MLST STs were identified in their group of 46 isolates, the most common STs being MLST ST60 and ST155 [49]. In the current study, none of the MLST STs in the isolates were identical, although some were similar such as MLST ST ~ 1223. However, GCC STs were more consistently represented (Table 1) with three GCC STs (ST86, ST143, GCC ST5) representing 20 of the 27 isolates. This finding is mostly consistent with the published literature that has identified GCC ST143 and GCC ST5 as the most common GCCs found in S. uberis isolates obtained from Australian cases of clinical mastitis [11, 16, 17, 23].

| Isolate/resistance gene | Lincosamides | Streptogramins | Macrolides, oxazolidinones, phenicols, pleuromutlins, streptogramins, tetracyclines |
|-------------------------|--------------|----------------|-----------------------------------------------------------------------------|
| S. uberis 45            | lnuD         |                |                                                                              |
| S. uberis 47            | lnuC         |                |                                                                              |
| S. uberis 48            | lnuD, mel, mrsE, vatD | mel, mrsE       |
| S. uberis 51            | lnuD, mel, mrsE, vatD | mel, mrsE       |

Table 3. Table showing antibiotic resistance profiles of S. uberis isolates in this study. The antibiotic class is shown along the x-axis, and the isolate is shown along the y-axis. The gene responsible for resistance is shown.
Queensland clade including 51, 124A, 114A, 106A and 107A, which almost all shared the same MLST.

Sanger sequencing reactions required for MLST and GCC ST analysis, it would be appropriate to rely more on whole genome sequencing and greater power for genomic and phylogenetic analysis compared with the multiple S. uberis isolates identified in this study. Given the reducing cost of between GCC STs (Fig. 3) meant that these GCCs did not appear to accurately capture the assumed evolutionary relationships.

The observation that GCCs clustered by region was notable, given that GCC ST86 isolates were more prevalent in QLD compared to Victoria, while GCC ST143 isolates were more prevalent in Victoria. This study provides the relative success rate of treatment protocols and management practices and the apparent ‘success’ and ‘failure’ rates in different regions or on different properties. At present, attention to possible environmental sources and disease potential of the particular strains circulating in the cow’s environment could also impact on transmission in the development and risk of mastitis also depend on geographical factors, and the transmissibility of infection, good standards of milking hygiene, and proper treatment of S. uberis mastitis remain important measures in attempts to control transmission of S. uberis.

A maximum likelihood phylogeny of the core genome demonstrated the clustering of MLST STs and indicated that the scheme was, at times, suitable (Fig. 3) for S. uberis. For example, MLST ST ~ 1223 clustered well, however MLST ST ~ 222 did not. Given the amount of novel STs identified (100% of isolates), the usefulness of the ST was of limited value to describing the evolutionary lineage of this highly recombinant species. While the GCCs clustered with other bovine mastitis-associated S. uberis (Fig. 2), the inconsistent phylogenetic clustering between GCC STs (Fig. 3) meant that these GCCs did not appear to accurately capture the assumed evolutionary lineages of the bovine mastitis-associated S. uberis isolates identified in this study. Given the reducing cost of whole genome sequencing and greater power for genomic and phylogenetic analysis compared with the multiple Sanger sequencing reactions required for MLST and GCC ST analysis, it would be appropriate to rely more on higher resolution WGS in future diagnostics. There appeared to be two clades containing state-only isolates: a Queensland clade including S. uberis 51, 124A, 114A, 106A and 107A, which almost all shared the same MLST.

### Table 4

| Isolate | Region length (kb) | Closest bacteriophage hit | Accession of closest hit |
|---------|-------------------|---------------------------|--------------------------|
| S. uberis 38B | 36.8 | Streptococcus phage SMP | NC_008721 |
| S. uberis 38B | 42.2 | Streptococcus prophage 315.2 | NC_005855 |
| S. uberis 47 | 43.9 | Streptococcus prophage 315.2 | NC_005855 |
| S. uberis 49 | 40.8 | Enterococcus phage phiFL4A | NC_013644 |
| S. uberis 50 | 55.5 | Streptococcus phage 5093 | NC_012753 |
| S. uberis 60A | 59.5 | Streptococcus phage PH10 | NC_012756 |
| S. uberis 70A | 36.4 | Streptococcus phage T12 | NC_028700 |
| S. uberis 93A | 36.7 | Streptococcus prophage 315.3 | NC_005856 |
| S. uberis 118A | 42.8 | Streptococcus prophage P9 | NC_009819 |
| S. uberis 122A | 41.9 | Streptococcus phage T12 | NC_028700 |

### Table 5

| IS family | Isolates | Total |
|-----------|----------|-------|
| IS256     | S. uberis 46 | 1     |
| IS3       | S. uberis 106A; S. uberis 107A; S. uberis 114A; S. uberis 122A; S. uberis 124A; S. uberis 125A; S. uberis 388; S. uberis 43; S. uberis 44; S. uberis 47; S. uberis 49; S. uberis 51; S. uberis 56A; S. uberis 70A; S. uberis 75B; S. uberis 86A | 19 |
| IS30      | S. uberis 118A; S. uberis 42; S. uberis 44; S. uberis 56A; S. uberis 69B; S. uberis 70A; S. uberis 75B | 7 |
| IS6       | S. uberis 31A | 1     |
| ISL3      | S. uberis 86A | 1     |
| Novel     | S. uberis 22A | 1     |

The observation that GCCs clustered by region was notable, given that GCC ST86 isolates were more prevalent in QLD compared to Victoria, while GCC ST143 isolates were more prevalent in Victoria. This study provides observational evidence that S. uberis GCC ST86 isolates, which had been previously thought to lack pathogenicity required to cause clinical mastitis, can contribute to disease. Further investigations such as experimental animal trials will be required to confirm this observation. If GCC ST86 isolates were indeed responsible for clinical mastitis, it would also be consistent with the lack of strong phylogenetic clustering of GCC ST86 (Fig. 3); the putative virulence factor profiles which showed consistent differences between GCCs (Fig. 4). Isolation of S. uberis GCC ST86 strains from cows afflicted with bovine mastitis may be indicative of a shift in pathogenic factors promoting invasion of host tissue, survival in the host environment, evasion of the host immune response, and internalization in the mammary gland cells.

These observations indicate that the relative importance of an environmental source and possible cow-to-cow transmission in the development and risk of mastitis also depend on geographical factors, and the transmissibility and disease potential of the particular strains circulating in the cow’s environment. It could also impact on the relative success rate of treatment protocols and management practices and the apparent ‘success’ and ‘failure’ rates in different regions or on different properties. At present, attention to possible environmental sources of infection, good standards of milking hygiene, and proper treatment of S. uberis mastitis remain important measures in attempts to control transmission of S. uberis.
ST (~ 1223) and GCC ST (86); and a Victorian clade containing S. uberis 75B, 56A, and 70A, in which almost all sharing the same GCC ST (143). Within these two specific clades, it is likely that the isolates diverged from a common lineage within their respective states. Every other identified clade appeared to contain a mixture of Victorian and Queensland isolates. Phylogenetic clustering within a source point was not observed in many cases, such as with QLD7, QLD8, QLD9 and VIC6 isolates. This observed divergence was much greater than originally expected. The most likely explanation of this finding is environmental acquisition of novel S. uberis bovine mastitis-associated isolates23, 41, 50, rather than an acquisition of monophyletic descendants such as would be expected with a contagious pathogen. Given this observation, further analysis of environmental S. uberis reservoirs is recommended.

The pan and core genome genes were consistent with previous pangenome analyses, showing a core genome of 1550 genes using 13 clinical and sub-clinical bovine mastitis isolates27 and ~ 1530 genes using 21 bovine mastitis isolates52. Core genes associated with nutrients found in abundance within the mammary tissue of cows, including two fibrinogen binding proteins, were identified (WP_046389274.1, WP_012658173.1). This is a notable finding given the fibrinogen concentration in the milk of cows affected by mastitis is significantly higher than that of healthy cows53. Two genes associated with lactose catabolism were identified, lacA (WP_000215993.1) and lacB (WP_000686149.1), which is also notable given the abundance of lactose in bovine milk. This may explain the presence or adaptation of S. uberis to mammary tissue.

As strain specific pathogenicity has been observed with S. uberis54, this indicates bacterial pathogenic factors are involved, not host-specific factors. Notably, the isolates from this study shared similar numbers of putative virulence factors as the publicly available bovine mastitis isolates, as well as human mastitis isolates (Fig. 4, Data S4, Data S6). This indicates a common set of virulence determinants responsible for cross-host mastitis pathogenesis.

The core virulence factors identified in every isolate (Fig. 4, Data S4) may partially explain the pathogenic behaviour of S. uberis. There were several core genes of note including cpsBCD and cpsA which encode products involved in capsular polysaccharide formation39; as-well-as, two putative glycosyltransferases involved with biofilm formation40, cylAG were identified, which contribute to hemolytic activity38, corresponding to an ABC transporter and an enoyl reductase involved in fatty acid biosynthesis56. This analysis confirmed hasA as a non-essential virulence factor, as it was only found in 24/27 isolates, and also confirmed mtuA as an essential virulence factor33, as it was found in all isolates. rqcH/fbp54, a fibronectin-binding protein and putative fibronectin-binding protein fbpS were also found in all isolates. This has been shown to be important for adhesion and a key virulence factor in Streptococcus pyogenes57, as-well-as, in Streptococcus suis, but not essential for virulence58. Also found in every isolate was scpA, a C5a peptidase, which has been shown to delay accumulation of leukocytes attracted to S. pyogenes by cleaving the serum chemotaxin C5a, indicating it is an essential virulence factor59. hasC was also found in all isolates; it is essential for hyaluronic acid capsule production50. Upon loss of this gene, S. uberis
becomes susceptible to phagocytosis by bovine neutrophils, indicating it is an important virulence factor\(^29\). \textit{Lmb}, encoding a laminin binding protein important for bacterial colonisation\(^37\) was also found in every isolate. \textit{Mga}, a virulence regulator has been shown to be important for a number of phenotypes including biofilm formation, growth in whole human blood and soft tissue and phagocytosis resistance\(^60, 61\). This was also identified in all isolates of \textit{S. uberis} in this study. \textit{sua}, an adhesin, has been experimentally verified in \textit{S. uberis} isolates from dairy cows and identified as a conserved gene\(^62\). Our study confirms this as it is conserved in all isolates. \textit{oppF}, an oligopeptide permease, shown to be important for auxotrophic amino acid acquisition from bovine milk\(^34\), was also found in every isolate. \textit{scaR}, another virulence regulator which upregulates during low levels of Mn\(^{2+}\)\(^63\) was also found in every isolate. Given that manganese is present in cow’s milk at 0.02–0.05 µg/mL\(^64\), it is likely \textit{scaR} upregulates virulence-associated genes in response to mastitis infection. Finally, a putative surface anchored

Figure 4. Heatmap showing the presence/absence of virulence and putative virulence genes (x-axis) within the \textit{S. uberis} isolates identified in this study (y-axis). Presence of virulence genes shown in red, absence in light pink. Core virulence genes are found in every isolate (27/27), while accessory virulence genes are found in < 27 isolates. Victorian isolates are coloured in yellow, Queensland in purple. GCCs are coloured, GCC ST86 coloured red, GCC ST5 coloured blue and GCC ST143 coloured grey. Data generated using ABRIcage version 1.0.1 (https://github.com/tseemann/abricage\(^\text{121}\)) with the VFDB\(^122\) and the custom PVDB (Data S2). Heatmap produced using ClustVis (https://biit.cs.ut.ee/clustvis/\(^\text{124}\)) and Inkscape version 0.92 (https://www.inkscape.org).
protein was also found in all isolates. This was experimentally identified as a sortase substrate, anchored to the surface in *S. uberis*. It contains a G5 domain, which is involved with N-acetylglucosamine binding and is found in a number of different enzymes, ranging from biofilm proteins to IgA-cleaving peptidases. The *neuABCD* locus is involved in sialic acid synthesis and their presence in *S. uberis* S6A indicates it may produce a capsule to assist with immune system evasion.

Other core virulence genes included enolase/phosphophryvuate hydratase (WP_012658173.1), involved in plasminogen and adhesion in *Streptococcal* species and LytR family regulatory protein (SQG46876.1) essential for attaching capsules and teichoic acids to cell wall. Also a YSIRK-type signal peptide-containing protein was then screened via NCBI for *mel/mef(A)*, *mrsE*, *vatD* and *lnuD*. The presence of an antimicrobial resistance gene cluster not associated with any known insertion sequences or mobilizable elements indicates this may have been acquired through uptake and integration of free DNA. The *lnuC* gene has been previously reported in *S. pyogenes* 90 but not in *S. uberis*. *lnuC* is a transposon-mediated resistance gene, this may represent the first observed interspecies jump into bovine mastitis-associated *S. uberis*. It is possible the genes highlighted here demonstrate the Red Queen hypothesis between host and pathogen, which can be summarised as follows: organisms living within a dynamic environment (bovine host) require matching rates of evolution to maintain colonisation. Co-evolution between host and parasite is a widely recognised consequence of natural selection and continually develops as genetic arms-race.

There were a limited number of antibiotic resistance genes found in this study. This finding is consistent with some previous dairy herd surveys, but inconsistent with others that detected a broad and diverse group of antibiotic resistance genes. Four *S. uberis* isolates found across Victoria and Queensland herds contained a combination of *lnuC* and a multidrug resistance cluster including *mel/mef(A)*, *mrsE*, *vatD* and *lnuD*. The presence of an antimicrobial resistance gene cluster not associated with any known insertion sequences or mobilizable elements indicates this may have been acquired through uptake and integration of free DNA. The two isolates that contain this cluster, *S. uberis* 48 and 51 are not directly phylogenetically related in the context of other isolates, which may indicate this chromosomal integration occurred independently. Further evidence of this can be seen in the variable genomic context of this cluster, which is not consistently placed within the genome (data not shown). The *mel/mef(A)* gene has been reported in *S. pneumoniae* 80 and *S. pyogenes*, but not in *S. uberis*. *S. uberis* was then screened via NCBI for *mel/mef(A)*. Only one protein sequence was identified (accession: WP_046391446.1) from *S. uberis* 6780 (accession: JATD00000000), also isolated from milk from a *S. agalactiae* infection not associated with known mobilizable elements. The *mel/mef(A)* gene encodes resistance to lincosamides, macrolides, oxazolidinones, phenicols, pleuromutilins, streptogramins and tetracyclines. The *msr family ABC-F type ribosomal protection gene, which is homologous to the coding sequence of *MsrE* (99% query coverage, 79% similarity), encodes resistance to streptogramins, phenicols, pleuromutilins, lincosamides, oxazolidinones, tetracyclines, macrolides and other antibiotics which target the peptidyl-transferase region of the ribosomal subunit. The next gene in the cluster encodes a hypothetical protein with no known function or domains. This is followed by *vatD*, a streptogramin A O-acetyltransferase resistance gene. The final gene in the cluster is *lnuD*, encoding a nucleotidyltransferase, which has been shown to provide lincosamides resistance. The *lnuD* and non-cluster *lnuC* genes encode O-nucleotidyltransferases and inactivate lincosamides by adenylylation. The *lnuD* has been previously reported in *S. uberis* isolated from clinical bovine mastitis, but the *lnuC* gene has only been identified in human *S. agalactiae* isolates, suggesting recent acquisition by bovine mastitis-associated *S. uberis*. As *lnuC* is a transposon-mediated resistance gene, this may represent the first observed interspecies jump into bovine mastitis-associated *S. uberis*. Lincosamide (lincomycin, clindamycin, and pirlimycin) usage and resistance should be carefully monitored now this mobilizable gene is within bovine mastitis-associated *S. uberis* populations. It would be of value to determine if a similar pattern exists in other streptococcal species present in the dairy farm environment. Further work is needed to verify these findings and determine antimicrobial gene
reservoirs and the mechanism of transfer. A limitation of this study was the small number of herds and limited number of isolates cultured from clinical mastitis cases. However, herd (farm) type, breed of dairy cow, and herd size were representative of Australian herds. Nevertheless, the presence of these genes in some of the isolates indicates that antibiotic usage should be carefully tracked and monitored to produce and maintain an up-to-date understanding of the level of antimicrobial resistance within dairy populations in the country.

Materials and methods

The study and all experimental procedures and protocols were approved by the Animal Ethics Committee at the University of Queensland (animal ethics approval numbers: SVS/ANRFA/540/18 and SVS/043/18/TERRAGEN), and all methods were performed in accordance with the relevant institutional guidelines and regulations. The methods described in the current study and reported results were compliant with ARRIVE guidelines for reporting animal research.

Sample collection. This study was part of a larger project conducted between March and June 2019 based on 430 milk samples collected from 29 dairy herds located in Queensland (QLD) (North Queensland, n = 9 herds; Southeast QLD, n = 9 herds), and Victoria (VIC, n = 11 herds) (Fig. 1; map figure was generated using ozmap package version 0.3.6 [https://CRAN.R-project.org/package=ozmaps] implemented within R version 4.0.2 [https://www.R-project.org/]). A 2-week rolling average of herd bulk somatic cell counts (BMTSCCs) was used to arbitrarily divide the herds into three categories (≤ 150 × 1000 cells/mL, > 150 to 300 × 1000 cells/mL, and > 300 × 1000 cells/mL). BMTSCCs categories were used as an a priori assumption for the risk of mastitis in the herd. Herd selection was based on ease of access to the farm location and the cooperation of the dairy farm owners and their associated veterinary practices.

Detailed methodology is described elsewhere. Briefly, milk samples were collected from eligible dairy cows with a new case of clinical mastitis. Chronic mastitis cases (apparently healthy cow with lumps palpable in the udder, and mild changes to milk) and subclinical mastitis cases were not eligible for study enrolment. An enrolment eligible clinical mastitis case was defined as a previously apparently healthy lactating dairy cow of any age, breed, or stage of lactation that was experiencing a new case of clinical mastitis, defined as either the first occurrence of a mastitis event in the current lactation or a mastitis event occurring at least 21 days following a previous mastitis event that has clinically resolved or achieved a clinical cure. Eligible cases must have not received systemic or intramammary antimicrobials, anti-inflammatory medications, or topical treatments in the 2 weeks prior to becoming a new case, nor immediately prior to sample collection. Milk samples were collected aseptically from individual quarters into separate sterile tubes, immediately capped and placed in a − 20 °C freezer. Collected samples were transported and delivered frozen to the Veterinary Laboratory Services of the University of Queensland for bacterial culture.

Streptococcus uberis isolation. At the laboratory, milk samples were mixed thoroughly, 100 µL streaked onto Sheep Blood Agar (SBA, P2133 Sheep Blood Columbia Agar Plates, Thermofisher), and the SBA plate incubated aerobically at 37 °C for 18–24 h. S. uberis isolates were initially identified using conventional microbiology laboratory tests (Gram stain appearance and catalase production). Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS; Bruker Daltonik, Bremen, Germany) was used to confirm the identification. Individual colonies were sub-cultured on SBA plates and incubated aerobically at 37 °C for 18–24 h. Pure isolates were then incubated in 2 mL of Brain Heart Infusion (BHI) broth, subsequently mixed with 20% glycerol, and stored at − 80 °C.

DNA extraction. Stored S. uberis isolates were batch thawed and cultured in BHI broth (37 °C, orbital shaker at 300 rpm, 18–24 h). Genomic DNA was extracted using DNeasy PowerFood Microbial Kit (QIAGEN Chadstone, Victoria, Australia) with minor modifications. Eight mL of culture liquid was centrifuged (15 min, 4 °C, 20,000×g) to pellet the bacteria. The pellet was then resuspended in 450 µL of lysis buffer and incubated for 10 min at 65 °C. Proteinase K (Proteinase K, QIAGEN Chadstone, Victoria, Australia) (25 µL) was added and incubation continued for an additional 20 min at 65 °C. This Proteinase K modification to the DNeasy protocol increased the final quantity of extracted DNA (data not shown). Thereafter, the whole component was transferred to a Powerbead tube, secured horizontally to a vortex adapter, and vortexed at maximum speed for 10 min. After washing to remove protein and other inhibitors, purified DNA was eluted and the concentration and purity of the isolated genomic DNA determined using a NanoDrop ND-1000 spectrophotometer. A sample of DNA was considered acceptable if the A260/280 ratio was ~ 1.8.

Whole genome sequencing. The New England Biolabs NEBNext Ultra II FS DNA Library Prep Kit for Illumina for multiplexing was used according to the manufacturer’s instructions (New England Biolabs, Notting Hill, Victoria, Australia) to construct whole genome sequencing libraries. Libraries were sequenced (Illumina NextSeq 500 instrument, 300 cycle mid-output kit; 2 × 150 bp paired end) at the Centre of Health Genomics and Informatics, University of Calgary, Canada, to obtain an average coverage depth of > 100 and a read retention of > 98%. The draft genome sequences of the 27 S. uberis isolates have been deposited with GenBank under BioProject PRJNA669385.

Whole genome assembly and annotation. Primer sequences were removed and reads were quality trimmed using cutadapt. The Nullarbor pipeline was used to process the samples. SKESA assembler (version 2.3.0) was used for de-novo assembly. Assemblies were annotated using prokka (version 1.14.5).
Details about the parameters used in different packages can be found at https://github.com/bananabenana/genomics_methods_S.uberis/blob/main/commands_run.

**Multi-locus sequence typing.** Assemblies were taxonomically assigned, analysed for MLST and were possible assigned a GCC, sequence types (STs)\(^{108,109}\), and scored for genome distance using Type (Strain) Genome Server (TYGS) (https://tygs.dsmz.de)\(^{110}\) (date accessed: 19 May 2020). MLSTs were confirmed using mlst_check version 2.1.1706216\(^{111}\) and MLST databases accessed on 5 June 2020. Phyloviz version 2.0 (www.phyloviz.net)\(^{112}\) and the goeBURST algorithm\(^{113}\) were used to visualise the data. Pairwise comparison of genome sequences among the set of genomes were conducted by calculating precise distances using the Genome BLAST Distance Phylogeny approach (GBDP) under the algorithm 'coverage' and distance formula d5\(^{114}\). These distances were used to determine the genome similarities for each of the genome pairs. 100 distance replicates were calculated each. Digital DDH values and confidence intervals were calculated using the recommended settings of the GGDC 2.1\(^{114}\).

**Pangenome analysis.** Pangenome analysis was performed using roary (version 3.13.0)\(^{115}\), splitting paralogues and setting the identity threshold at 90%. This list was then annotated and characterised via screening sequences against NCBI and Gene Ontology\(^{116}\) using the PANNZER2 server\(^{117}\). Signal sequences were obtained from the isolates genomes reported in this current study, 40 additional publically available genomes (Data S6)\(^{27}\), S. iniae YST01-82 (accession number: GCF_000831485.1) and in-group S. uberis 0140J (accession number: AM946015) ST1. RAxML (raxmlHPC-PTHREADS-SSE3 version 8.2.10)\(^{120}\) was used for maximum likelihood and rapid bootstrap analysis with 1000 replicates. The bipartitions output file was visualised in FigTree version 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and bootstraps visualised. Inkscape version 0.92 (https://www.inkscape.org) was used to prepare the figure.

**Identification of putative virulence factors.** ABRicate version 1.0.1 (https://github.com/tseemann/abricate)\(^{121}\) was used with the VFDB\(^{122}\) and a custom gene list, the sPVDB (Data S2). The custom gene list was based on sequences identified in the literature described to be S. uberis virulence factors with experimental or strong computational evidence\(^{26-41}\), as well identifying homologues of Streptococcus virulence factors. In addition to the isolates genomes reported in this current study, 40 additional publically available genomes (Data S6)\(^{27}\),\(^{123}\) were retrieved from GenBank and screened for putative virulence factors. To identify homologous virulence factors in S. uberis, the Streptococcus VFDB amino acid sequences were screened against the translated pangenome. Any sequences with > 70% query cover and > 60% similarity were considered putative virulence factors. The output was then visualised in ClustVis (https://biit.cs.ut.ee/clustvis/)\(^{123}\). No scaling was performed on the data and clustering was performed on rows and columns using the 'Euclidean' measure. To identify core, putative sortase A substrates, grep version 3.1 (https://git.savannah.gnu.org/cgit/grep.git/tree/AUTHORS) was used.

**Antibiotic resistance determination.** ABRicate version 1.0.1 (https://github.com/tseemann/abricate)\(^{121}\) was used with the CARD\(^{125}\), database accessed 4/5/2020.

**Bacteriocin identification.** Bacteriocins were identified by downloading the BAGEL4 database\(^{26}\) and converting into a multifasta file. The pangenome was then aligned against this using blastp version 2.9.0+\(^{127}\),\(^{128}\) and results were filtered for ≥ 80% coverage and ≥ 50% similarity/positives (to account for different signal sequences).

**Plasmid identification and analysis.** Mob-suite version 3.0.0\(^{129}\) was used to identify plasmids using the mob recon function, then visualised in ClustVis (https://biit.cs.ut.ee/clustvis/)\(^{124}\). No scaling was performed on the data and clustering was performed on rows and columns using the 'correlation' distance measure. Putative mobilization of plasmids was determined using the mob_type function.

**Bacteriophage and prophages identification.** The PHASTER (https://phaster.ca/)\(^{130}\) web server was used to identify bacteriophage and prophage regions using the API tool. 'Questionable' phages (score of 70–90) were grouped with 'incomplete'.

**Transposon/Insertion sequence identification.** ISESCAN version 1.7.2 (https://github.com/xiezhq/ISEScan)\(^{131}\) was used to identify transposons and insertion sequences.

**Statistical analysis.** Fisher's exact test was performed to identify significance in the proportion of strain types, virulence genes, and plasmids from isolates (unit of analysis) recovered from NQLD, QLD and VICT. A two sided p-value obtained by Monte-Carlo simulation (n = 2000) of at least 0.05 was considered to be significant. Statistical analysis was conducted using stats package implement in R.
References

1. Rollin, E., Dhuyvetter, K. C. & Overton, M. W. The cost of clinical mastitis in the first 30 days of lactation: An economic modeling tool. Prev. Vet. Med. 122, 257–264. https://doi.org/10.1016/j.prevetmed.2015.11.006 (2015).

2. Huijps, K., Lam, T. J. G. M. & Hogeveen, H. Costs of mastitis: Facts and perception. J. Dairy Res. 75, 113–120. https://doi.org/10.1017/S0022029907002392 (2008).

3. Hogeveen, H., Huijps, K. & Lam, T. J. G. M. Economic aspects of mastitis: New developments. N. Z. Vet. J. 59, 16–23 (2011).

4. Shum, L. W., McConnel, C. S., Grunn, A. A. & House, J. K. Environmental mastitis in intensive high-producing dairy herds in New South Wales. Aust. Vet. J. 87, 469–475. https://doi.org/10.1111/j.1751-0867.2009.00223.x (2009).

5. Ruegg, P. A 100-Year Review: Mastitis detection, management, and prevention. J. Dairy Sci. 100, 10381–10397. https://doi.org/10.3168/jds.2017-13023 (2017).

6. Verbeke, J., Piepers, S., Supre, K. & De Veiregh, S. Pathogen-specific incidence rate of clinical mastitis in Flemish dairy herds, severity, and association with herd hygiene. J. Dairy Sci. 97, 6926–6934. https://doi.org/10.3168/jds.2014-8173 (2014).

7. Phuektes, P. et al. Molecular epidemiology of Streptococcus uberis isolates from dairy cows with mastitis. J. Clin. Microbiol. 39, 1460–1466 (2001).

8. Bradley, A., Leach, K., Breen, J., Green, L. & Green, M. Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales. Vet. Rec. 160, 253–258 (2007).

9. Charman, N. R., Dyson, R., Hodge, A., Robertson, N. & Chaplin, S. A Survey of Mastitis Pathogens in the South Eastern Australian Dairy Industry (Dairy Focus, Moncton, 2012).

10. Zadoks, R. N. A 100-Year Review: Mastitis detection, management, and prevention. J. Dairy Sci. 100, 10381–10397. https://doi.org/10.3168/jds.2017-13023 (2017).

11. Tomazi, T. et al. Genotyping and antimicrobial resistance of Streptococcus uberis isolated from bovine clinical mastitis. PLoS One 14, e0223719 (2019).

12. Hassan, A., Khan, I., Abdulmawjood, A. & Lämmler, C. Evaluation of PCR methods for rapid identification and Differentiation of Streptococcus uberis and Streptococcus parvulus. J. Clin. Microbiol. 39, 1618–1621 (2001).

13. Lundberg, A., Nyman, A., Unerstad, H. E. & Waller, K. P. Prevalence of bacterial genotypes and outcome of bovine clinical mastitis due to Streptococcus dysgalactiae and Streptococcus uberis. Acta Vet. Scand. https://doi.org/10.1186/s13028-014-0080-0 (2014).

14. Coffey, T. J. et al. First insights into the evolution of Streptococcus uberis: A multilocus sequence typing scheme that enables investigation of its population biology. Appl. Environ. Microbiol. 72, 1420–1428. https://doi.org/10.1128/AEM.72.2.1420-1428.2006 (2006).

15. Douglas, V. L., Fenwick, S. G., Pfeiffer, D. U., Williamson, N. B. & Holmes, C. W. Genomic typing of Streptococcus uberis isolates from cases of mastitis, in New Zealand dairy cows, using pulsed-field gel electrophoresis. Vet. Microbiol. 75, 27–41. https://doi.org/10.1016/S0378-1135(00)00184-X (2000).

16. Leelahapongsa,thon, K., Schukken, Y. H., Strithanuwan, A. & Suryasathapon, W. Molecular epidemiology of Streptococcus uberis intramammary infections: Persistent and transient patterns of infection in a dairy herd. J. Dairy Sci. 103, 3565–3576. https://doi.org/10.3168/jds.2019-17281 (2020).

17. Werner, C., Sauerwald, C., Sundrum, A., El-Sayed, A. & Zschöck, M. Genotyping of Streptococcus uberis isolates in healing process of bovine clinical mastitis. Vet. Res. 50, 1–18 (2019).

18. Ward, P. N. et al. Evidence for niche adaptation in the genome of the bovine pathogen Streptococcus uberis. BMC Genom. 10, 54 (2009).

19. Pullinger, G. D. et al. Application of Streptococcus uberis multilocus sequence typing: Analysis of the population structure detected among environmental and bovine isolates from New Zealand and the United Kingdom. Appl. Environ. Microbiol. 72, 1429–1436. https://doi.org/10.1128/AEM.72.2.1429-1436.2006 (2006).

20. Lundberg, A. et al. Udder infections with Staphylococcus aureus, Streptococcus dysgalactiae, and Streptococcus uberis at calving in dairy herds with suboptimal udder health. J. Dairy Sci. 99, 2102–2117 (2016).

21. Reinoso, E. B., Lasagno, M. C. & Odierno, L. M. Genetic patterns of Streptococcus uberis isolated from bovine mastitis. Rev. Argent. Microbiol. 47, 108–111 (2015).

22. McDougall, S., Parkinson, T. J., Leyland, M., Anniss, F. M. & Fenwick, S. G. Duration of infection and strain variation in Streptococcus uberis isolated from cows’ milk. J. Dairy Sci. 87, 2062–2072. https://doi.org/10.3168/jds.2007-0022 (2004).

23. Abureeema, S., Smooker, P., Malmo, J. & Deighton, M. Molecular epidemiology of recurrent clinical mastitis due to Streptococcus uberis: Evidence of both an environmental source and recurring infection with the same strain. J. Dairy Sci. 97, 285–290. https://doi.org/10.3168/jds.2013-7074 (2014).

24. Werner, C., Sauerwald, C., Sundrum, A., El-Sayed, A. & Zschöck, M. Genotyping of Streptococcus uberis isolates in healing process of bovine clinical mastitis. Vet. Res. 50, 1–18 (2019).

25. Jolley, K., Bray, J. & Maiden, M. Open-access bacterial population genomics: BIGSdb software, the PubMedstrg website and their applications [version 1; peer review: 2 approved]. Wellcome Open Res. https://doi.org/10.12688/wellcomeopenres.14826.1 (2018).

26. Leish, J. A., Egan, S. A., Ward, P. N., Field, T. R. & Coffey, T. J. Sortase anchored proteins of Streptococcus uberis play major roles in the pathogenesis of bovine mastitis in dairy cattle. Vet. Res. 41, 63 (2010).

27. Novak, M. et al. Virulence related sequences; insights provided by comparative genomics of Streptococcus uberis of differing virulence. BMC Genom. 16, 334. https://doi.org/10.1186/s12864-015-1512-6 (2015).

28. Egan, S. A., Ward, P. N., Watson, M., Field, T. R. & Leish, J. A. Vru (Sub0144) controls expression of proven and putative virulence determinants and alters the ability of Streptococcus uberis to cause disease in dairy cattle. Microbiology 158, 1581–1592. https://doi.org/10.1099/mic.0.055863-0 (2012).

29. Ward, P. N., Field, T. R., Ditcham, W. G. F., Maquin, E. & Leish, J. A. Identification and disruption of two discrete loci encoding Hyaluronic acid capsule biosynthesis genes hasA, hasB, and hasC in Streptococcus uberis. Infect. Immun. 69, 392–399. https://doi.org/10.1128/IAI.69.1.392-399.2001 (2001).

30. Oliver, S. P., Almeida, R. A. & Calvinho, L. F. Virulence factors of Streptococcus uberis isolated from cows with mastitis. J. Vet. Med. B 45, 461–471. https://doi.org/10.1111/j.1439-0450.1998.tb08178.x (1998).

31. Skalka, B. & Smola, J. Lethal effect of CAMP-factor and UBERIS-factor—a new finding about diffusible exosubstances of Streptococcus agalactiae and Streptococcus uberis. Zentralbl. Bakteriol. Mikrobiol. Hyg. A 249, 190–194. https://doi.org/10.1007/S00174-001-0073-X (1981).

32. Jiang, M., Babiu, I. A. & Potter, A. A. Cloning, sequencing and expression of the CAMP factor gene of Streptococcus uberis. Microb. Pathog. 20, 297–307. https://doi.org/10.1006/mpat.1996.0028 (1996).
33. Smith, A. J. et al. MutA, a lipoprotein receptor antigen from Streptococcus uberis, is responsible for acquisition of manganese during growth in milk and essential for the function of the lactating bovine mammary gland. Infect. Immun. 71, 4842–4849. https://doi.org/10.1128/IAI.71.13.4842-4849.2003 (2003).

34. Smith, A. J., Kitt, A. J., Ward, P. N. & Leigh, J. A. Isolation and characterization of a mutant strain of Streptococcus uberis, which fails to utilize a plasmin derived β-casein peptide for the acquisition of methionine. J. Appl. Microbiol. 93, 631–639. https://doi.org/10.1046/j.1365-2672.2002.01723.x (2002).

35. Tian, X. Y. et al. Antimicrobial resistance and virulence genes of Streptococcus isolated from dairy cows with mastitis in China. Microb. Pathog. 131, 33–39. https://doi.org/10.1016/j.micpath.2019.03.035 (2019).

36. Lakshmi, R. Understanding the regulation of Group B Streptococcal virulence factors. Future Microbiol. 4, 201–221. https://doi.org/10.2217/17460913.4.2.201 (2009).

37. Spellerberg, B. et al. Lmb, a protein with similarities to the LraI adhesin family, mediates attachment of Streptococcus agalactiae to human laminin. Infect. Immun. 67, 871–878. https://doi.org/10.1128/IAI.67.3.871-878.1999 (1999).

38. Czabanska, A. Plasmin induced degradation of β-casein and β-casein-derived peptides in milk. PLoS One 9, e109883. https://doi.org/10.1371/journal.pone.0109883 (2014).

39. Barnett, T. C., Patel, A. R. & Scott, J. R. A novel Sortase A (SrtA) substrates from Streptococcus pyogenes anchors a surface protein containing a QVPTGV Motif to the cell wall. J. Bacteriol. 186, 5865–5875. https://doi.org/10.1128/jb.186.17.5865-5875.2004 (2004).

40. Tomita, T. et al. Identification of Streptococcus uberis multilocus sequence types highly associated with mastitis. Appl. Environ. Microbiol. 74, 114–124. https://doi.org/10.1128/aem.01373-07 (2008).

41. Lalioui, L. et al. The SrtA sortase of Streptococcus agalactiae is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for colonization of the mouse intestine. Infect. Immun. 73, 3342. https://doi.org/10.1128/IAI.73.8.3342-3350.2005 (2005).

42. Lang, P. et al. Virulence gene content differences across strains of Streptococcus uberis identified using oligonucleotide microarray comparative genomic hybridization. Infect. Genet. Evol. 9, 179–188. https://doi.org/10.1016/j.meegid.2008.10.015 (2009).

43. Tabrizi, A. D., Batavani, R. A., Rezaei, S. A. & Ahmadi, M. Fibrinogen and ceruloplasmin in plasma and milk from dairy cows with subclinical and clinical mastitis. Pak. J. Biol. Sci. 11, 571–576 (2008).

44. Spellerberg, B. et al. Identification of genetic determinants for the hemolytic activity of Streptococcus suis serotype 2 capsular polysaccharide. J. Bacteriol. 189, 631–639. https://doi.org/10.1128/jb.189.3.631-639.2007 (2007).

45. Cho, K. H. & Caparon, M. G. Patterns of virulence gene expression differ between biofilm and tissue communities of Streptococcus gordonii. Science 309, 140–153. https://doi.org/10.1126/science.1110755 (2005).

46. Le Bertron, Y. et al. Genome-wide discovery of novel MT1 group A streptococcal determinants important for fitness and virulence during soft-tissue infection. PLoS Pathog. 13, e1006584. https://doi.org/10.1371/journal.ppat.1006584 (2017).

47. Cho, K. H. & Caparon, M. G. Patterns of virulence gene expression differ between biofilm and tissue communities of Streptococcus gordonii. Microbes. Mol. Biol. Rev. 75, 1455–1556. https://doi.org/10.1128/mbr.00033-05 (2011).

48. Lalioui, L., Tabrizi, A. D. & Lang, P. Plasticity of virulence by the alteration of bacterial surface architecture. Future Microbiol. 12, 75–86. https://doi.org/10.2241/fm.2017.2988.2005.02579.x (2011).

49. Li, Q. et al. Fibrinogen-/fibrinolysin-binding protein (FPBS) is not a critical virulence factor for the Streptococcus suis serotype 2 strain ZY05719. Vet. Microbiol. 208, 38–46. https://doi.org/10.1016/j.vetmic.2017.07.010 (2017).

50. O’Connor, S. P. & Cleary, P. Purpose of Streptococcus pyogenes Csa peptidase activity: Analysis by transposon- and nitroso- guanidine-induced mutants. J. Infect. Dis. 156, 495–504. https://doi.org/10.1086/363883 (1987).

51. Bateman, A., Holden, M. T. G. & Yeats, C. The G5 domain: A potential N-acetylglucosamine recognition domain involved in biofilm formation. Bioinformatics 21, 1301–1303. https://doi.org/10.1093/bioinformatics/bti206 (2004).

52. Feng, Y. et al. Site-specific chemoenzymatic labeling of aerolysin enables the identification of new aerolysin receptors. J. Bacteriol. 189, 631–639. https://doi.org/10.1128/jb.189.3.631-639.2007 (2007).

53. Auger, J.-P., Dolbec, D., Roy, D., Segura, M. & Gottschalk, M. Role of the Streptococcus suis type 2 capsule polysaccharide in the interactions with dendritic cells is strain-dependent but remains critical for virulence. PLoS One 13, e0200453. https://doi.org/10.1371/journal.pone.0200453 (2018).
68. Fulde, M. & Bergmann, S. Moonlighting Proteins: Novel Virulence Factors in Bacterial Infections 245–268 (Wiley, New York, 2017).

69. Ye, W. et al. Pneumococcal LytR protein is required for the surface attachment of Streptococcus pneumoniae to respiratory epithelial cells. Infect. Immun. 82, 1683–1691. https://doi.org/10.1128/iai.00699-13 (2014).

70. Kamashvili, D. & Rouviere-Yaniv, J. The histone-like protein HU binds specifically to DNA recombination and repair intermediates. EMBO J. 19, 6527–6535. https://doi.org/10.1093/emboj/19.23.6527 (2000).

71. Oberto, J., Nabi, S., Jooste, V., Mignot, H. & Rouviere-Yaniv, J. The HU Regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. PLoS One 4, e3467. https://doi.org/10.1371/journal.pone.0003467 (2009).

72. Stucky, K. et al. Cloning and characterization of brnQ, a gene encoding a low-affinity, branched-chain amino acid carrier in Lactobacillus delbrückii subsp. lactis DSM7290. Mol. Genet. Genom. 249, 682–690. https://doi.org/10.1007/BF00418035 (1998).

73. Wohlt, J. E., Clark, J. H., Derrig, R. G. & Davis, C. L. Valine, leucine, and isoleucine metabolism by lactating bovine mammary glands. J. Dairy Sci. 60, 1875–1882. https://doi.org/10.3362/0303-0207(77)91188-0 (1977).

74. Grispoldi, L. et al. The relationship between S. aureus and branched-chain amino acids content in composite cow milk. Animals 9, 20 (2019).

75. Tiemeyer, W., Stohrer, M. & Giesecke, D. Metabolites of nucleic acids in bovine milk. Anim. Feed Sci. Technol. 43, 619–626. https://doi.org/10.1016/0377-8401(93)90053-5 (1993).

76. Harrison, R. Milk xanthine oxidase: Properties and physiological roles. Int. Dairy J. 16, 546–554. https://doi.org/10.1016/j.idairyj.2005.08.016 (2006).

77. Stevens, C. R. et al. Antibacterial resistance of bovine xanthine oxidase in human milk. Lancet 356, 829–830. https://doi.org/10.1016/S0140-6736(00)02660-X (2000).

78. Stockmann, C. et al. Clinical and epidemiological evidence of the Red Queen hypothesis in pneumococcal serotype dynamics. Clin. Infect. Dis. 63, 1611–1616. https://doi.org/10.1093/cid/ciw083 (2016).

79. MacAogáin, M. et al. Metagenomics reveals a core macrolide resistome related to microbiota in chronic respiratory disease. Am. J. Respir. Crit. Care Med. 202, 433–447. https://doi.org/10.1164/rccm.201911-2202OC (2020).

80. Su, W. et al. Ribosome protection by antibiotic resistance ATP-binding cassette protein. Proc. Natl. Acad. Sci. 115, 5157. https://doi.org/10.1073/pnas.1803313115 (2018).

81. Kallio, A. et al. Role of Pho proteins in attachment of Streptococcus pneumoniae to respiratory epithelial cells. Infect. Immun. 82, 1683–1691. https://doi.org/10.1128/iai.00699-13 (2014).

82. Kallio, A. et al. Role of Pho proteins in attachment of Streptococcus pneumoniae to respiratory epithelial cells. Infect. Immun. 82, 1683–1691. https://doi.org/10.1128/iai.00699-13 (2014).

83. Fiscione, A. et al. Survey and sequence characterisation of bovine mastitis-associated Streptococcus agalactiae strains in UK dairy herds. Proc. Natl. Acad. Sci. 115, 626–630. https://doi.org/10.1073/pnas.1217962115 (2018).

84. Hammervold, A. M., Flannagan, S. E., Clewell, D. B. & Jensen, L. B. Indication of transposition of a mobile DNA element containing the rat(D) and erm(B) genes in Enterococcus faecium. Antimicrob. Agents Chemother. 45, 3223–3229. https://doi.org/10.1128/AAC.45.11.3223-3229.2001 (2001).

85. Bozdogan, B. et al. New resistance gene, linB, conferring resistance to lincomidines by nucleotideidylation in Enterococcus faecium HM1025. Antimicrob. Agents Chemother. 43, 925–929. https://doi.org/10.1128/AAC.43.2.925-929.1999 (1999).

86. Kibler, J. D. et al. Molecular cloning and functional characterization of a novel macrolide resistance determinant, mefA, from Enterococcus faecalis. Antimicrob. Agents Chemother. 41, 470–476. https://doi.org/10.1128/AAC.41.2.470-476.1997 (1997).

87. Deboer, T. et al. The EnteroBase user’s guide, with case studies on antimicrobial resistance. Nucleic Acids Res. 44, D7–19. https://doi.org/10.1093/nar/gkz1290 (2016).

88. Martin, M. Catadapt removes adapter sequences from high-throughput sequencing reads. EMBO J. 17(3), 2011. https://doi.org/10.1038/s41598-021-82357-3 (2018).

89. Zhou, Z. et al. The EnterOveruse guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny and Escherichia coli genomic diversity. Genome Res. 30, 138–152. https://doi.org/10.1101/gr.251678.119 (2019).
Acknowledgements

We would like to acknowledge Joanne Mollinger and Ameh James for their contributions to this work.

Author contributions

Conception and experimental design: J.I.A., T.O., M.S. Acquisition of data: H.A., H.R., J.I.A. Analysis and interpretation of data: B.V., J.I.A., R.M. Drafted the manuscript: B.V., J.I.A., H.A. All authors contributed to and revised the manuscript.

Competing interests

This work was funded by Science with Impact Scheme, the University of Queensland, Sub-tropical Dairy Board, Dairy Australia, and Terragen Biotech Ltd, Australia. The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-82357-3.

Correspondence and requests for materials should be addressed to J.I.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
