Genome Comparison and Phylogenetic Analysis of Mastitis-Related Staphylococci with a Focus on Adhesion, Biofilm, and Regulatory-Related Genes

Lucas José Luduverio Pizauro (lucaspizauro@gmail.com)
Sao Paulo State University
Camila Chioda Almeida
Sao Paulo State University
Saura Rodrigues Silva
Sao Paulo State University
Janet I. MacInnes
University of Guelph
Andrew M. Kropinski
University of Guelph
Luiz Francisco Zafalon
Brazilian Agricultural Research Corporation (EMBRAPA), Embrapa Southeast Livestock
Fernando Antônio Avila
Sao Paulo State University
Alessandro Mello Varani
Sao Paulo State University

Research Article

Keywords: Adhesins, biofilm, comparative genomics, intramammary infection, virulence factors

Posted Date: April 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-433702/v1

Abstract

Bovine mastitis is the costliest diseases on dairy farms and is caused by different *Staphylococcus* species. However, staphylococci associated with clinical mastitis infections are different from subclinical ones, indicating a complex mechanism related to bovine mastitis pathogenesis. Here, we performed genomic analyses to determine the prevalence of adhesion, biofilm, and regulatory genes in 478 staphylococcal spp. associated with clinical and subclinical mastitis deposited in public databases. The most prevalent adhesin genes were the *ebpS*, *atl*, *pls*, *sasH* and *sasF* genes found in both clinical and subclinical isolates. However, the *ebpS* gene is absent in subclinical isolates of *Staphylococcus arlettae, S. succinus, S. sciuri, S. equorun, S. galinarum*, and *S. saprophyticus*. In constrast, the *coa*, *eap*, *emp, efb*, and *vWbp* genes were present more frequently in clinical mastitis isolates and highly correlated with the presence of the *icaABCD* and *icaR* biofilm genes. We also revealed that many adhesins, biofilm, and associated regulatory genes were potentially horizontally disseminated between clinical and subclinical isolates. Taken together, our results indicate that several adhesins, biofilm, and regulatory-related genes have been overlooked in previous studies and that these virulence factors may arise in staphylococcal species not generally associated with clinical mastitis by horizontal gene transfer.

Introduction

Bovine mastitis is one of the costliest diseases seen on dairy farms, with an estimate of US$19.7 to US$32 billion loss due to reduced milk production and withdrawal periods related to antibiotic usage [1]. Mastitis may also cause death directly or lead to the slaughter of chronically infected animals [1]. *S. aureus* is generally considered the most important cause of both clinical and subclinical mastitis, while coagulase-negative staphylococci or non-aureus *Staphylococcus* spp. are thought to be of lesser importance or are opportunistic pathogens [2]. A higher prevalence of subclinical versus clinical infections has been reported [3]. Moreover, the prevalence of subclinical mastitis is likely to be underestimated due to the lack of obvious signs except for changes in milk quantity and quality (which can only be detected by specific tests such as the California Mastitis Test and by somatic cell counting) [1]. It is generally believed that the *Staphylococcus* spp. strains associated with chronic infections are different from those that cause acute infections and are more likely to be transmitted and persist in the herd due to better host adaptation and the absence of clinical signs [1,2].

There are many distinct virulence factors related to the disease establishment. In *Staphylococcus* spp. factors such as the fibronectin-binding proteins (*fnbA* and *fnbB*), elastin binding proteins (*ebpS*), clumping factors (*clfA* and *clfB*), and collagen-binding protein (*cna*) play important roles in binding to host cells, colonization, and invasion [2]. In addition, the *ica* genes, which are associated with the synthesis of polysaccharide intercellular adhesin (PIA), are thought to play a crucial role in biofilm development in these bacteria. The autolytic protein (*alt*) degrades the peptidoglycan cell wall layer and plays a key role in bacterial cell wall metabolism[4].

Furthermore, the *pls* gene encodes the plasmin-sensitive protein that also has a role in bacterial adherence [5]. The surface proteins encoded by *sasH* and *sasF*, play an import role in virulence because they can bind to host extracellular matrix and plasma components, and only recently, they have been reported as prevalent adhesins in a genome comparison study of *Staphylococcus* spp. isolates from bovine [6].

Therefore, there is still much to learn about *Staphylococcus* spp. pathogenesis, especially about attachment factors and their regulation. Molecular epidemiology-based methods such as specific PCR assays, MLST, and PFGE have been used to analyze the genetic diversity and virulence factors and to track the dissemination of *Staphylococcus* spp. infections, but they have their limitations [3]. Accordingly, this work aimed to ascertain the prevalence of adhesion and biofilm genes by investigating whole genome sequences of *Staphylococcus* spp. from clinical and subclinical mastitis cases and determine the phylogenetic relationship of these isolates and if any adhesin or biofilm genes might be associated with acute bovine/bubaline mastitis.

Results

Assessment of Clinical and subclinical mastitis Staphylococcus spp. isolates

Staphylococcus chromogenes (28.7%), *S. simulans* (20.0%). *S. aureus* (18.7%) and the *S. sciuri* (10.0%) were the most prevalent clinical mastitis species deposited in the NCBI GenBank database. The remaining 22.5% of the clinical mastitis strains were *S. epidermidis* (5.00%), *S. haemolyticus* (5.00%), *S. agnetis* (2.50%), *S. xylosus* (2.50%), *S. arlettae* (1.25%), *S. capitis* (1.25%), *S. cohnii* (1.25%), *S. devriesei* (1.25%), *S. gallinarum* (1.25%) and *S. hominis* (1.25%). The most frequent staphylococcal species associated with subclinical mastitis were *S. chromogenes* (15.6%), *S. simulans* (6.8%), *S. xylosus* (6.5%), *S. haemolyticus* (6.3%), *S. cohnii* (5.8%), *S. epidermidis* (5.5%), *S. capitis* (5.3%), *S. sciuri* (5.3%), *S. gallinarum* (5.0%), *S. warneri* (4.8%), *S. equorum* (4.5%), *S. saprophyticus* (4.0%), *S. succinus* (3.8%), *S. arlettae* (3.5%), *S. agnetis* (3.3%), *S. aureus* (3.3%) and *S. hominis* (3.0%). The remaining 7.50% of subclinical isolates included S. *devriesei* (1.76%), *S. pasteuri* (1.51%), *S. vitulinus* (1.51%), *S. auricularis* (0.50%), *S. caprae* (0.50%), *S. fleurettii* (0.50%), *S. hyicus* (0.50%), *S. nepalensis* (0.50%) and *S. kloosii* (0.25%) (Fig. 1).

Distribution of Adhesin, biofilm and regulatory genes across clinical and subclinical mastitis Staphylococcus spp. isolates

In the mastitis related staphylococci genomes analyzed (n = 478) the most prevalent genes associated with adhesion and biofilm formation were: *ebpS* (71.3%), *atl* (70.9%), *sasF* (70.7%), *sasH* (53.3%), *araC*

(52.1%), *tcaR* (52.1%), *sarA* (52.1%), *sigB* (52.1%) *pls* (44.6%), *sasA* (37.2%) and *sasC* (30.8%) (Fig. 2). The *icaC* (17.3%), *icaR* (14.0%), *sasD* (13.8%), *sdrE* (13.4%), *icaA* (11.5%), *icaB* (11.5%), *icaD* (11.5%), *sdrC* (11.0%), *clfA* (10.6%), *fnbA* (9.62%), *spa* (9.21%), *vWbp* (8.79%), *fnbB* (6.9%), *efb* (6.07%), *coa* (5.86%), *eap* (5.86%), *emp* (5.65%), *clfB* (5.44%), *aap* (5.23%), *cna* (5.02%), *sasG* (3.97%), *sasK* (3.77%) and *sdrD* (3.14%) genes were detected less frequently. The *sasI* gene was absent in all isolates (Fig. 2)

In strains associated with clinical mastitis, the *ebpS* (83.8%), *atl* (83.8%), *sasF* (83.8%), *sasH* (77.5%), *atl* (56.3%), *rbf* (56.3%), *tcaR* (56.3%), *sarA* (56.3%), *sigB* (56.3%), *pls* (55.0%), *sasA* (47.5%), *pls* (37.5%) and *sasC* (30.0%) genes were most frequently detected while s*drE* (22.5%), *fnbA* (21.2%), *spa* (20.0%), *clfA* (22.5%) *vWbp* (18.7%) *icaC* (17.5%), s*drC* (17.5%), *efb* (17.5%), *coa* (17.5%), *eap* (17.5%), *emp* (17.5%) *icaR* (16.2%) *icaA* (16.2%) *icaB* (16.2%) *icaD* (16.2%) *clfB* (16.2%) *sasD* (15.0%) *fnbB* (11.2%) *sasG* (10.0%) *sasK* (8.75%) *sdrD* (7.50%) *cna* (6.25%) and *aap* (2.50%) were present less often.

The carriage of adhesin/biofilm related genes in isolates associated with subclinical mastitis was less frequent (e.g., *ebpS* (68.8%), *atl* (68.3%), *sasF* (68.1%), *atl* (51.3%), *rbf* (51.3%), *tcaR* (51.3%), *sarA* (51.3%), *sigB* (51.3%) *sasH* (48.5%) *pls* (42.5%), *sasA* (35.2%) and *sasC* (30.9%). Also, a lower frequency of the following genes was also observed: *icaC* (17.3%) *icaR* (13.5%) *sasD* (13.5%) *sdrE* (11.5%) *icaA* (10.5%) *icaB* (10.5%) *icaD* (10.5%) *sdrC* (9.80%) *clfA* (9.05%) *fnbA* (7.29%) *spa* (7.04%) *vWbp* (6.78%) *fnbB* (6.03%) *aap* (5.78%) *cna* (4.77%) *efb* (3.77%) *coa* (3.52%) *eap* (3.52%) *emp* (3.27%) *clfB* (3.27%) *sasG* (2.76%) *sasK* (2.76%) and *sdrD* (2.26%). The *sasI* gene was absent in all subclinical isolates.

Most of the subclinical isolates of *S. aureus, S. capitis, S. chromogenes, S. cohnii, S. epidermidis, S. haemolyticus, S. warneri* and *S. simulans* had the *ebpS, atl, pls, sasH, sasC* and *sasF* adhesion-associated genes with the *clfB* and *emp* genes found only in *S. aureus* strains.

The PIA production operon *icaADBC* and its regulator *icaR* was only present in *S. aureus*, *S. chromogenes*, *S. capitis*, and *S. epidermidis*; most *S. cohnii* and *S. saprophyticus* carried the *icaC* gene. In the clinical isolates, the *icaADBC* operon and *icaR* gene were present only in *S. aureus* isolates. Other biofilm regulatory genes i.e., *rbf*, *tcaR*, *sarA* and *sigB* were found in subclinical isolates of *S. aureus*, *S. chromogenes*, *S. epidermidis* and *S. haemolyticus*, but not in *S. simulans* isolates.

Phylogenetic analyses reveal no clear relationship between clinical and subclinical isolates showing an uneven distribution of adhesin, biofilm and regulatory genes

Analysis of the 16S RNA genes from the genome sequences of the *Staphylococcus* spp. from bovine and buffalo mastitis cases revealed that the clinical and subclinical isolates (n = 478) are present in a wide variety of clades and do not show any clear relationship (Supplementary Fig. 1). The 16S RNA gene phylogeny also indicated that the mastitis related *S. aureus, S. epidermidis*, and *S. capitis* have a close phylogenetic relationship. These species also possess many adhesion genes (avg. no. = 26, 11, and 17

respectively), followed by *S. chromogenes* and *S. warneri* (avg. no.= 9 and 12, respectively). *S. capitis* has a close phylogenetic relationship to the species that are mainly associated with clinical mastitis (*S. aureus* and *S. epidermidis*). *S. chromogenes*, which was implicated in cases of clinical (n = 23/80) and subclinical mastitis (n = 61/398) is most closely related to *S. agnetis* and *S. hyicus* species that were only associated with subclinical mastitis. “Subclinical species” *S. saprophyticus, S. xylosus, S. gallinarum* and *S. arlettae* formed a distinct node with few strains involved in clinical mastitis and with most of these strains not carrying known adhesion/biofilm related genes. The “subclinical species” *S. warneri* and *S. pasteuri* were also phylogenetically related and carried biofilm/adhesion associated genes (n = 35; avg. no. of genes = 12 and 10, respectively). No specific pattern was observed between clinical and subclinical strains *ebpS, rbf, sarA, sasH, sigB*, and *tcaR* gene phylogeny (Supplementary Figs. 2–7, respectively). Overall, the clinical and subclinical strains of most species were in the same clade.

The co-phylogenetic analysis suggests the occurrence of different events horizontal gene transfer (HGT) between virulence genes. For instance, the *ebpS* gene between clinical and subclinical isolates of *S. simulans, S chromogenes* and *S. aureus* (Supplementary Figure S8); the *pls*, gene from clinical and subclinical isolates of *S. haemolyticus, S. chromogenes* and *S. simulans* (Supplementary Figure S9); the *rbp* gene among clinical and subclinical isolates of *S. chromogenes, S. aureus* and *S. haemolyticus* (Supplementary Figure S10), and the *sarA* gene, between clinical and subclinical isolates of *S. chromogenes* and *S. aureus*, respectively (Supplementary Figure S11). Additional evidence of potential HGT were also observed for the *sasH* gene between *S. aureus*, *S. simulans* and *S. chromogenes* (Supplementary Figure S12), and the *tcaR* gene among the *S. aureus*, and *S. chromogenes* clinical and subclinical isolates (Supplementary Figure S13)

Data analysis indicates adhesion and biofilm genes exclusively related to clinical isolates

Hierarchical clustering analysis based on the presence/absence of the adhesin, biofilm, and regulatory genes revealed 20 different clusters (Supplementary Table 1). One hundred and twenty-seven (26.5%) strains (13 clinical and 114 sub-clinical) of the 478 genomes evaluated lacked the 35 adhesion and biofilm-associated genes identified by the RAST annotation tool. The staphylococcal species lacking these genes included: *S. arlettae, S. equorum S. gallinarum S. sciuri, S. succinus, S. vitulinus* and *S. xylosus*. In species heatmaps (Fig. 3), the pattern of adhesion/biofilm genes in clinical isolates differs from that of sub-clinical isolates. The presence of the *clfA, clfB, fnbA, spa, sdrC, coa, eap, emp, vWbP, sasD, icaA, icaB, icaC, icaD* and *icaR* genes is highly correlated in clinical isolates, while in subclinical isolates, no specific gene correlations were observed (Spearman coefficient > 0.8). Based on hierarchical matrix clustering, clusters 9 and 10; 19 and 18 and 4 and 5 (Supplementary Table 1) contained most of the strains that harbored a typical pattern of nine genes (*rbf, pls, sasF, sarA, atl, sasH, sigB, tcaR* and *ebpS*) in both clinical and subclinical isolates. This pattern is also demonstrated in the heatmap of the gene frequency (Fig. 4).

Discussion

Staphylococcus spp. are the most common etiologic agents of mastitis, with *S. aureus* considered the most important, while coagulase-negative staphylococci and non-aureus staphylococci considered less significant [7]. Based on 16S RNA identification of the 478 available genome sequences, *S. chromogenes* (28.7%) and *S. simulans* (20.0%) were the species most frequently associated with mastitis. *S. aureus* was the next most prevalent staphylococci (18.7%) associated with clinical mastitis and was rarely (3.3%) associated with subclinical mastitis deposited in GenBank. Most subclinical cases were associated with *S. chromogenes* (15.6%) followed by *S. simulans* (6.8%), *S. xylosus* (6.5%), *S. haemolyticus* (6.3%), *S. cohnii* (5.8%), *S. epidermidis* (5.5%), *S. capitis* (5.3%), *S. sciuri* (5.3%), *S. gallinarum* (5.0%), *S. warneri* (4.8%), *S. equorum* (4.5%), *S. saprophyticus* (4.0%), *S. succinus* (3.8%), *S. arlettae* (3.5%), and *S. agnetis* (3.3%). These findings are consistent with a growing number of reports that coagulase-negative staphylococci are emerging pathogens associated with mastitis and persistence of intramammary infection in bovine worldwide [8]. In a recent study, the five most common species found in Canadian mastitis cases were *S. chromogenes, S. simulans, S. xylosus, S. haemolyticus* and *S. epidermidis*[9]. These were also the most common species found in subclinical strains that were sequenced in the database.

Adherence is considered the first step of staphylococcal infection, and the presence of biofilm aids it. Accordingly, adhesion-related genes are thought to be key virulence factors [9]. In the current study, the most frequently observed adherence and biofilm-forming genes were e*bpS, atl, pls, sasH, sasF, rbf, tcar, sarA* and *sigB* in both clinical and subclinical isolates. The elastin binding protein gene (*ebpS)* was absent in the *S. arlettae, S. succinus, S. sciuri, S. equorun, S. galinarum* and *S. saprophyticus* genomes analyzed. These isolates also lacked most of the adherence and biofilm genes, which could indicatehat these species are more likely to be contaminants associated with the milk microbiota [10,11] rather than subclinical mastitis agents. Nevertheless, the *S. equorum. S. sciuri, S. galinarum*, and *S. succinus* isolates have been associated with skin infections and urinary tract in humans and mice [12–14,] which indicates that they can potentially harbor virulence genes.

The *atl* gene was the second most frequently detected gene. It encodes an autolytic protein that can cause the lysis of other bacterial that compete with *Staphylococcus* spp. for nutrients in the milk. It acts in the peptidoglycan cell wall layer's degradation is thought to play a key role in bacterial cell wall metabolism, daughter-cell separation, and antibiotic-mediated cell lysis [4]. This gene is also associated with bacterial internalization and secretion of proteins in *S. aureus*[15]. This gene's presence in most mastitis isolates could be attributed to the fact that it is implicated in diverse functions such as bacterial attachment to surface, lysis mediated biofilm formation and secretion of the cytoplasmic proteins from the staphylococcal cell wall. The *atl* gene has been implicated in adherence to fibronectin, heparin, and gelatin[16,] which could confer an advantage during infection as heparin is released by mast cells and basophils at the site upon tissue damage [19]. The same could be noted about the *pls* gene, which encodes the plasmin-sensitive protein that has a role in adherence and is an important virulence factor in

mouse septic arthritis model [5]. To date, these genes have been underrated and have not yet been highlighted even in recent genomic comparison studies of S. *aureus* isolates from bovine mastitis [17,18].

The surface proteins encoded by *sasH* and *sasF*, play an essential role in virulence because they can bind to host extracellular matrix and plasma components. They have been recently reported as the most prevalent adhesins in a genome comparison study of 24 bovine-associated staphylococcus isolates, with all isolates positive for the two genes [6]. The *sasH* gene is significantly associated with invasive disease isolates due to its ability to inhibit the oxidative burst and promote *S. aureus* survival in neutrophils[19,] thus allowing the organism to avoid the bovine immune response and colonize the mammary gland. In the current study, the *sasH* gene was not only present in *S. aureus*, but it was also present in *S. chromogenes, S. haemolyticus, S. simulans, S. agnetis, S. capitis*, and *S. warneri*. The *sasF* gene is also of interest, but little is known about its role in *S. aureus* virulence [20]. It is believed that it may have an important role in thromboembolic lesions[20] and play a role in advanced states in mastitis when capillary damage occurs[21]

The *coa*, *eap*, *emp efb* and *vWbp* genes were most frequently present in clinical mastitis isolates and their presence was highly correlated with the presence of the *ica*ABCD and R genes in these isolates. This correlation was not observed in subclinical isolates. Manual examination of the *S. aureus* SAMN02603524 (NC_021670.1) genome reveals that the *emp* gene is located 300 nucleotides downstream from the *vWbp* gene (Supplementary Fig. 14) no other close spatial relationships were observed with the other genes. Therefore, it is possible that the *emp* gene would work with the *vWbp* gene and a have possible role in *S. aureus* immune response evasion [22]. Nevertheless, a high correlation indicates that a clinical mastitis *Staphylococcus* spp. strain that have either *coa*, *eap*, *emp efb* and *vWbp* is more likely to carry the *icaABCDR* genes. It is known that the *coa* gene generate coagulase genotypes that were proven to be more resistant to neutrophil activities [23]. Also, the *vWbp* is another know coagulase in *Staphylococcus* that would present a similar effect [24]. The eap gene productas recently been shown to be able to suppress the formation of “neutrophil extracellular traps” (NETs), which are thought to function as a neutrophil-mediated extracellular trapping mechanism [25].

The *icaADBC* locus contains the most studied *Staphylococcus* biofilm forming genes and [26] it is most frequently reported in mastitis isolates. In the current study, the *icaC* gene was the most prevalent in subclinical isolates, in contrast to previous report of which *icaA* and *icaD* are the most prevalent [27]. The most prevalent biofilm regulatory genes were *rbf*, *tcaR*, *sarA* and *sigB*. The *rbf* gene is an important biofilm regulatory gene since its inactivation results in a biofilm negative phenotype [28]. Also, it has been recently shown that *rbf* mutants exhibit significantly increased pathogenicity compared to the wild type *S. aureus* strains [29,] thus highlighting its importance for host adaptation. Also, the *rbf* gene product negatively regulates hemolytic activity by repressing the *hla* and *psmA* genes’ expression. It also upregulates *sarX*, which, in turn, activates the *icaADBC* locus leading to biofilm production [30].

The *tcaR* gene increases the production of polysaccharide intracellular adhesin (PIA) by regulating the expression of the *icaADBC* operon and the *spa*, *sasF* and *sarS* genes [30]. Given the high frequency of the *sasF* observed in this study, a high frequency of detection of its transcriptional regulators would be expected. The *sarA* family of transcriptions regulators proteins are responsible for controlling many target genes involved in virulence, with the *sarA* been responsible for regulating the *agr* loci, which is a pivotal regulator for virulence in *S. aureus*[31]. This gene's importance in mastitis is highlighted with a recent study demonstrating that it was present in 100% of the 84 *S. aureus* isolates from mastitis in Xinjiang, China [31]. The rRNA polymerase sigma factor (*sigB*) has a central role in stress homeostasis. The *sigB* contributes to several virulence determinants defining staphylococcal pathogenesis, including the transcriptional activation of many surface proteins (such as *clf*A and *fnb*A) while downregulating the production of secreted toxins and proteases (such as Aur, SspA, SspB) [32].

The phylogenetic analysis of the 16S RNA gene indicated that *S. aureus, S. epidermidis*, and *S. capitis* have a close relationship. These species possess a high number of adhesion genes. Although phylogenetic analysis of *Staphylococcus* related to mastitis has been done before, they mainly focus on fewer isolates, and in *S. aureus* species, a broad-scale analysis as performed in this study has not been made before. Nevertheless, as observed with *S. aureus*[33], strains that had different origins were clustered together. The authors suggested that dairy cows can be natural carriers by being subclinically infected with *S. aureus* subtypes that can cause clinical mastitis if the right conditions are present. In addition, in this study, it was possible to observe that clinical and subclinical strains were clustered together by the 16S RNA but have peculiar, different gene content regarding biofilm and coagulase genes. Therefore, it is safe to assume that besides the suitable condition, a genomic content will also be present for the *S. aureus*, *S. epidermidis* and *S. capitis* to cause clinical mastitis. Also, this study shows that *S. chromogenes* that was implicated in cases of clinical and subclinical mastitis were closer related with *S. agnetes* and *S. hyicus*. Besides, evidence of potential HGT between clinical and subclinical isolates in *S. chromogenes, S. aureus* and *S. simulans* were observed for the *ebpS, rbp, sarA, tcaR, pls* and, specifically for *sigB* that only presented potential HGT for the *S. aureus* species.

The phylogenetic relationship of the adhesion and biofilm genes (*ebpS, sasH, atl, sar*A, *rbf* and *tca*R*)* are not similar to the 16S RNA gene phylogenetic distribution. This finding and the evidence of potential HGT among clinical and subclinical isolates strongly suggest that virulence factors may arise in staphylococcal species not generally associated with clinical mastitis.

Methods

Genomic data

The genomes of *Staphylococcus* spp. from clinical (n = 80) and subclinical (n = 398) mastitis cases worldwide were downloaded from the National Center for Biotechnology Information (NCBI). Only complete genomes identified as mastitis isolates from *Bos taurus* or *Bubalus bubalis* in publications or

their BioSample descriptions were used (Supplementary Table 1). It was assumed that mastitis states were classified according to the clinical presentation and standard triage test described by Radostits et al. [3]. Genomes in this study were from bacteria isolated in Brazil[34], Canada[35] India, Netherlands, and United States (information regarding isolates BioSample available on Supplementary Table 1).

Genome annotation and adhesion-related gene identification

Genomes were annotated using Rapid Annotation using Subsystem Technology (RAST) [36,37]. The sequences of the genes classified as adhesion/adhesins or implicated in biofilm formation and their respective regulatory genes were downloaded and analyzed manually. 16S rRNA gene sequences were obtained from the complete genomes using the Basic Rapid Ribosomal RNA Predictor (Barrnap) v 0.9 (https://github.com/tseemann/barrnap).

Data analysis

The presence or absence of selected genes was used in hierarchical clustering analysis with PAST software v4.03 [38]. Clusters of the isolates were created based on the most and least frequent genes. The Spearman test was used to analyze the correlation between the presence/absence of adhesin and biofilm genes in both clinical and subclinical mastitis isolates (A coefficient close to 1.0 indicates a high correlation). Gene profiling by frequency heatmaps was calculated using Numpy [39]. Graphs were made using Matplotlib [40] and, when needed, with R-software [41]. The statistical significance of gene presence and mastitis state was obtained by logistic regression with R software.

Phylogenetic analysis and tree construction

The phylogenetic correlation of the 16S RNA, *araC*, *pls*, *sasF*, *sarA*, *atl*, *sasH*, *sigB tcaR*, and *ebpS* genes was determined and phylogenetic trees were generated with the aid of the MAFFT v7[42]. Tree visualization was done with iTOL v5 [43]. The co-phylogenetic tree construction was done using phytools in R-software [41]. The *Escherichia coli* strains (accession numbers): 2014C-3057 (NZ_CP027387.1); 2015C-4944 (NZ_CP027390.1); 2013C-4538 (NZ_CP027582.1); E2865 (NZ_AP018808.1); 97-3250 (NZ_JHEW00000000.1); FORC_028 (NZ_CP012693.1); 2013C-4225 (NZ_CP027577.1); 2014C-3050 (NZ_CP027472.1); 2012C-4606 (NZ_CP027352.1) and CFSAN027343 (NZ_CP037943.1) were used as an outlier group.

Declarations

Author contributions statement

L. J. L. P and A. M. V. Conceived the experiment. L J L P, C. C. A and S. R. S conducted the experiment, J. I. M, A. M. K, L. F. Z and F. A. A helped elaborating the draft. All authors contributed to analyze the results and reviewed the manuscript.

Additional information

The authors declare that they have no conflict of interest

References

1. Rocha, L. S. *et al.* Comparative genomics of Staphylococcus aureus associated with subclinical and clinical bovine mastitis. *PLoS ONE* **14**, 1–19 (2019).
2. Schukken, Y. H. *et al.* Host-response patterns of intramammary infections in dairy cows. *Veterinary Immunology and Immunopathology* **144**, 270–289 (2011).
3. Radostits, O. M., Gay, C. C., Hinchcliff, K. W. & Constable, P. D. Diseases of the mammary gland. in *Veterinary Medicine – A Textbook of the Diseases of Cattle, Horses, Sheep, Pigs, and Goats.* 673–763 (2007).
4. Singh, V. K. High level expression and purification of atl, the major autolytic protein of staphylococcus aureus. *International Journal of Microbiology* **2014**, (2014).
5. Efthimiou, G., Tsiamis, G., Typas, M. A. & Pappas, K. M. Transcriptomic Adjustments of Staphylococcus aureus COL (MRSA) Forming Biofilms Under Acidic and Alkaline Conditions. *Frontiers in Microbiology* **10**, (2019).
6. Åvall-Jääskeläinen, S. *et al.* Comparative genome analysis of 24 bovine-associated *Staphylococcus* isolates with special focus on the putative virulence genes. *PeerJ* **6**, e4560 (2018).
7. Wald, R., Hess, C., Urbantke, V., Wittek, T. & Baumgartner, M. Characterization of staphylococcus species isolated from bovine quarter milk samples. *Animals* **9**, 1–16 (2019).
8. Corbeil, A. *et al.* Short communication: Search for superantigen genes in coagulase-negative staphylococci isolated from bovine milk in Canada. *Journal of Dairy Science* **102**, 2008–2010 (2019).
9. Zuniga, E. *et al.* Occurrence of genes coding for MSCRAMM and biofilm-associated protein Bap in Staphylococcus spp. isolated from bovine subclinical mastitis and relationship with somatic cell counts. *Microbial Pathogenesis* **89**, 1–6 (2015).
10. Oikonomou, G. *et al.* Milk Microbiota: What Are We Exactly Talking About? *Frontiers in Microbiology* **11**, 1–15 (2020).
11. Taponen, S. *et al.* Bovine milk microbiome: A more complex issue than expected. *Veterinary Research* **50**, 1–15 (2019).
12. Nováková, D. *et al.* Staphylococcus equorum and Staphylococcus succinus isolated from human clinical specimens. *Journal of Medical Microbiology* **55**, 523–528 (2006).
13. Kengkoom, K. & Ampawong, S. Staphylococcus sciuri associated to subcutaneous abscess and dermatitis in ICR mouse. *Arquivo Brasileiro de Medicina Veterinaria e Zootecnia* **69**, 117–122 (2017).
14. Latham, R. H., Running, K. & Stamm, W. E. Urinary Tract Infections in Young Adult Women Caused by Staphylococcus saprophyticus. *JAMA: The Journal of the American Medical Association* **250**, 3063–

3066 (1983).

15. Josse, J., Laurent, F. & Diot, A. Staphylococcal adhesion and host cell invasion: Fibronectin-binding and other mechanisms. *Frontiers in Microbiology* **8**, 1–8 (2017).
16. Porayath, C. *et al.* Autolysin mediated adherence of Staphylococcus aureus with Fibronectin, Gelatin and Heparin. *International Journal of Biological Macromolecules* **110**, 179–184 (2018).
17. Fursova, K. *et al.* Virulence Factors and Phylogeny of Staphylococcus aureus Associated With Bovine Mastitis in Russia Based on Genome Sequences. *Frontiers in Veterinary Science* **7**, 1–10 (2020).
18. Hoekstra, J. *et al.* Genomic analysis of European bovine Staphylococcus aureus from clinical versus subclinical mastitis. *Scientific Reports* **10**, 1–11 (2020).
19. Thammavongsa, V., Missiakas, D. M. & Schneewind, O. Staphylococcus aureus degrades neutrophil extracellular traps to promote immune cell death. *Bone* **23**, 1–7 (2013).
20. Jenkins, A. *et al.* Differential Expression and Roles of Staphylococcus aureus Virulence Determinants during Colonization and Disease. *mBio* **6**, 1–10 (2015).
21. Akers, R. M. & Nickerson, S. C. Mastitis and its impact on structure and function in the ruminant mammary gland. *Journal of Mammary Gland Biology and Neoplasia* **16**, 275–289 (2011).
22. McCarthy, A. J. & Lindsay, J. A. Genetic variation in staphylococcus aureus surface and immune evasion genes is lineage associated: Implications for vaccine design and host-pathogen interactions. *BMC Microbiology* **10**, (2010).
23. Javid, F. *et al.* Molecular typing of Staphylococcus aureus based on coagulase gene. *Veterinary World* vol. 11 423–430 (2018).
24. Trivedi, U. *et al.* Staphylococcus aureus coagulases are exploitable yet stable public goods in clinically relevant conditions. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E11771–E11779 (2018).
25. Eisenbeis, J. *et al.* The staphylococcus aureus extracellular adherence protein Eap Is a DNA binding protein capable of blocking neutrophil extracellular trap formation. *Frontiers in Cellular and Infection Microbiology* vol. 8 (2018).
26. Azmi, K., Qrei, W. & Abdeen, Z. Screening of genes encoding adhesion factors and biofilm production in methicillin resistant strains of Staphylococcus aureus isolated from Palestinian patients. *BMC Genomics* **20**, 1–12 (2019).
27. A. Salina , F.F. Guimarães, V.B. Richini Pereira, B.D Menozzi, V.L.M. Rall, H. L. Detection of icaA, icaD, and bap genes and biofilm production in Staphylococcus aureus and non-aureus staphylococci isolated from subclinical and clinical bovine mastitis. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia* **72**, 1034–1038 (2020).
28. Rowe, S. E. *et al.* AraC-type regulator Rbf controls the Staphylococcus epidermidis biofilm phenotype by negatively regulating the icaADBC repressor SarR. *Journal of Bacteriology* **198**, 2914–2924 (2016).

29. Fang, B., Liu, B. & Sun, B. Transcriptional regulation of virulence factors Hla and phenol-soluble modulins α by AraC-type regulator Rbf in Staphylococcus aureus. *International Journal of Medical Microbiology* **310**, 151436 (2020).
30. Hoang, T. M. *et al.* Transcriptional regulation of icaADBC by both IcaR and TcaR in staphylococcus epidermidis. *Journal of Bacteriology* **201**, 1–17 (2019).
31. Ballal, A. & Manna, A. C. Expression of the sarA family of genes in different strains of Staphylococcus aureus. *Microbiology (Reading, England)* **155**, 2342–52 (2009).
32. Marbach, H. *et al.* Within-host evolution of bovine Staphylococcus aureus selects for a SigB-deficient pathotype characterized by reduced virulence but enhanced proteolytic activity and biofilm formation. *Scientific Reports* **9**, 1–12 (2019).
33. Ronco, T. *et al.* Genomic investigation of Staphylococcus aureus isolates from bulk tank milk and dairy cows with clinical mastitis. *Veterinary Microbiology* **215**, 35–42 (2018).
34. Pizauro, L. J. L. *et al.* Complete Genome Sequences of 11 Staphylococcus sp . Strains Isolated from Buffalo Milk and Milkers ' Hands. *Microbiology Resource Announcements* 11–13 (2019).
35. Naushad, S. *et al.* Comprehensive phylogenetic analysis of bovine non-aureus staphylococci species based on whole-genome sequencing. *Frontiers in Microbiology* **7**, (2016).
36. Overbeek, R. *et al.* The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research* **42**, 206–214 (2014).
37. Aziz, R. K. *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**, 75 (2008).
38. Hammer, Øyvind, David A.T. Harper, and P. D. R. PAST: PALEONTOLOGICAL STATISTICS SOFTWARE PACKAGE FOR EDUCATION AND DATA ANALYSIS. *Palaeontoligia Electronica* **4**, 1–9 (2001).
39. Van Der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Computing in Science and Engineering* **13**, 22–30 (2011).
40. Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science and Engineering* **9**, 99–104 (2007).
41. R Core Team. R: A language and environment for statistical computing. (2014).
42. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772–780 (2013).
43. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic acids research* **44**, W242–W245 (2016).

Figures

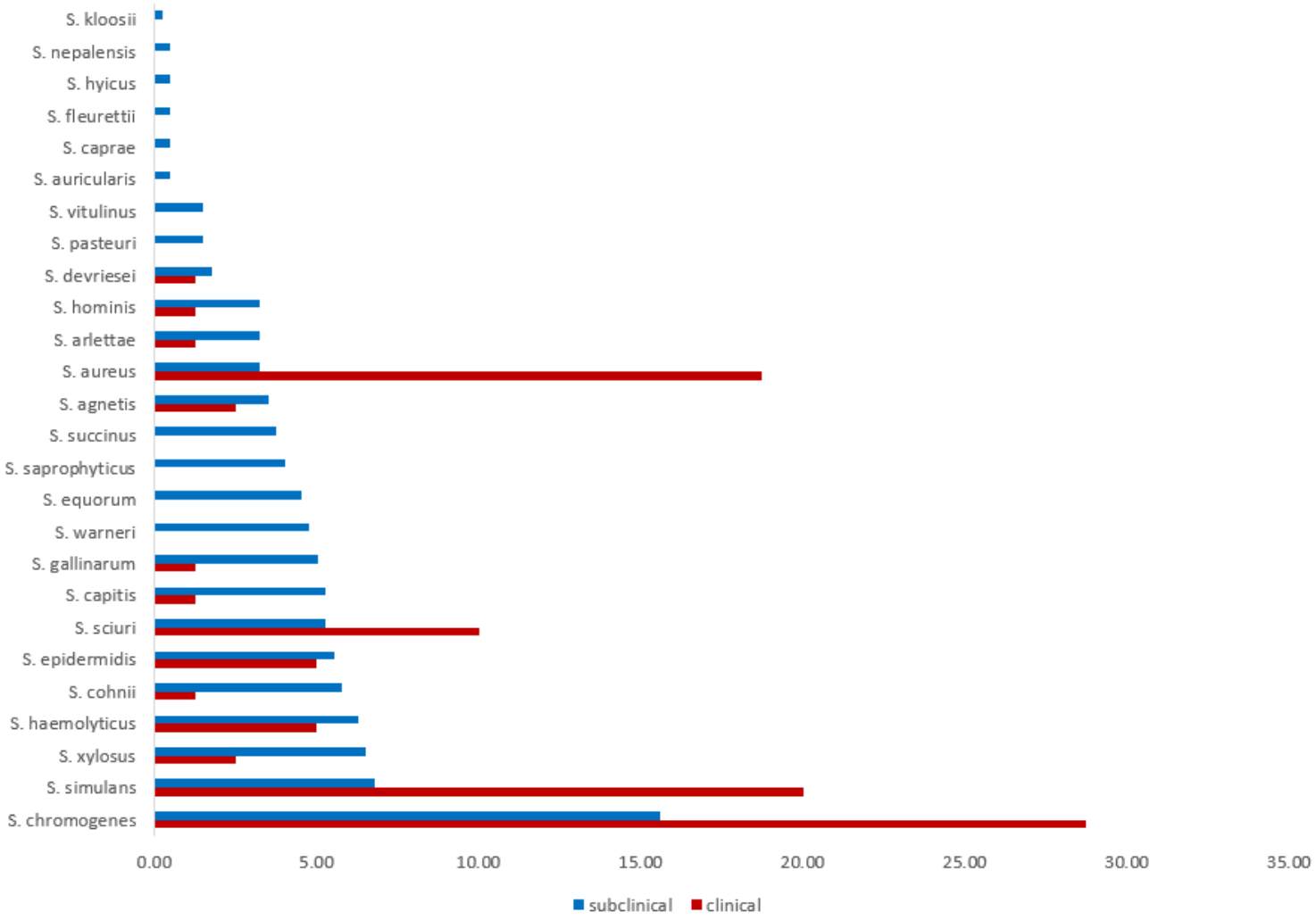

Figure 1

Frequency of staphylococcal species associated with clinical and subclinical mastitis

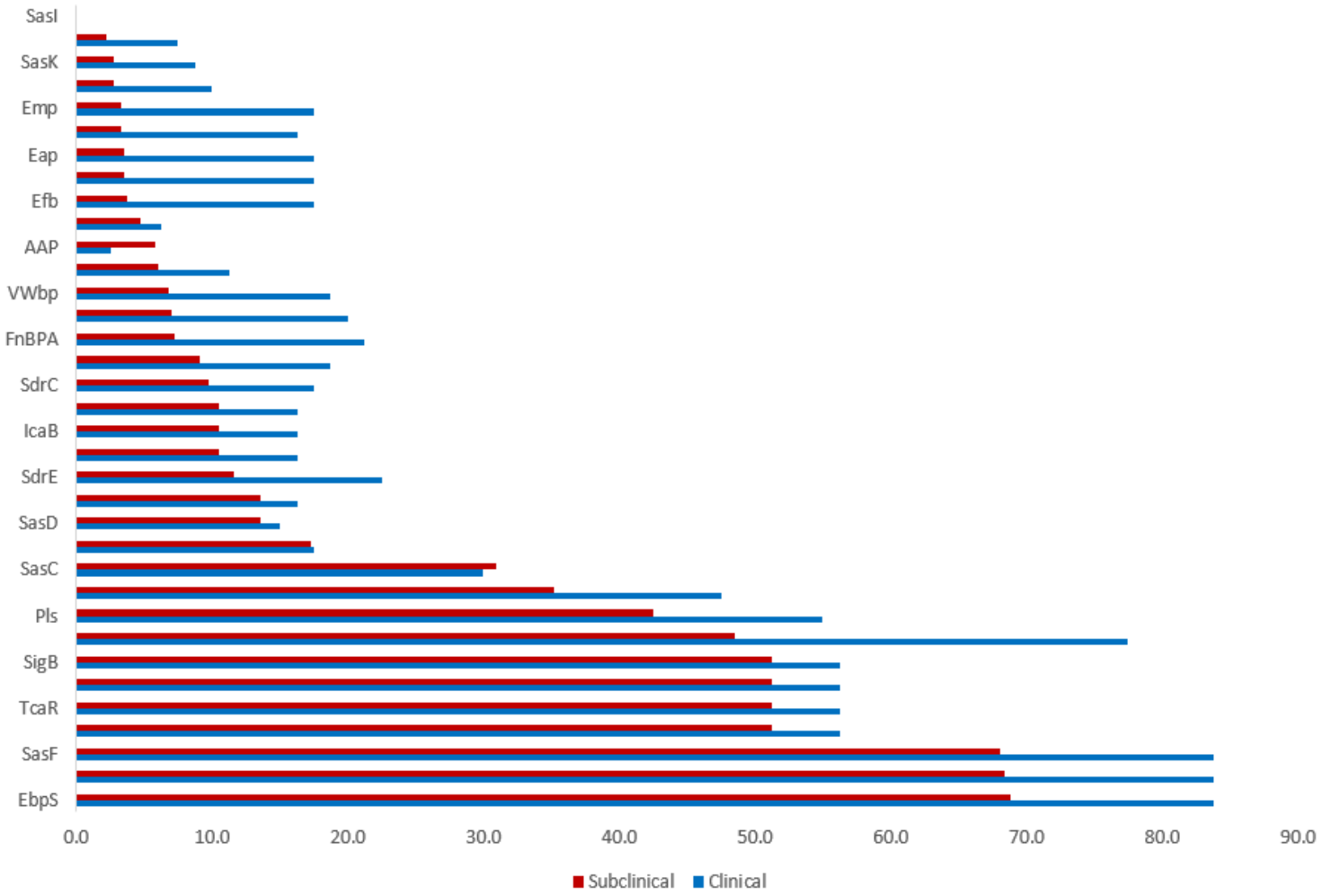

Figure 2

Frequency of adhesin and biofilm-related genes in mastitis-associated staphylococcal isolates.

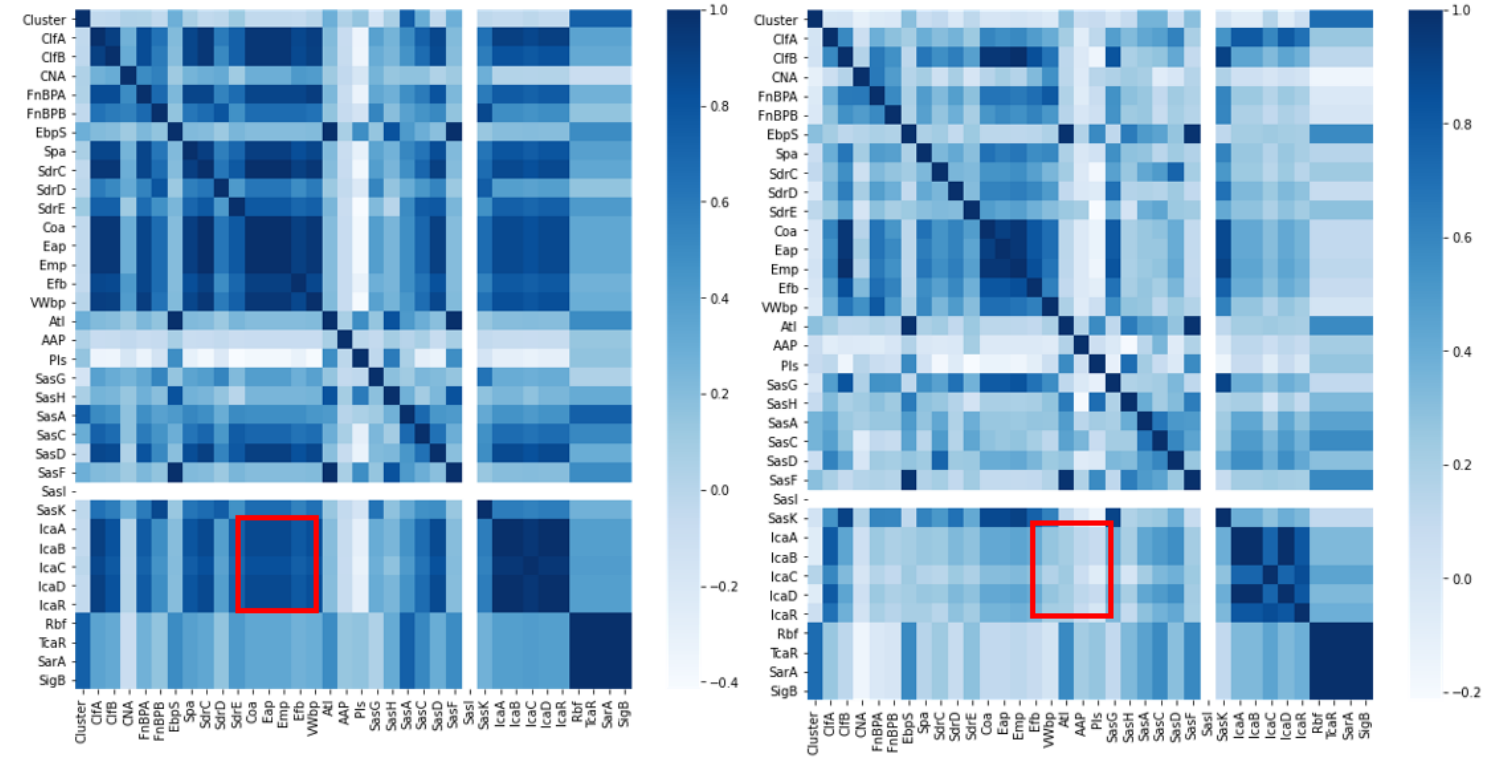

Figure 3

Heat map of adhesion and biofilm genes in clinical and subclinical staphylococcal isolates (Left – Clinical mastitis; Right – Subclinical mastitis). A higher correlation between the presence of the icaABCDR and the coa, eap,emp,efb and vWbp was found in clinical (vs. subclinical) isolates (red box)

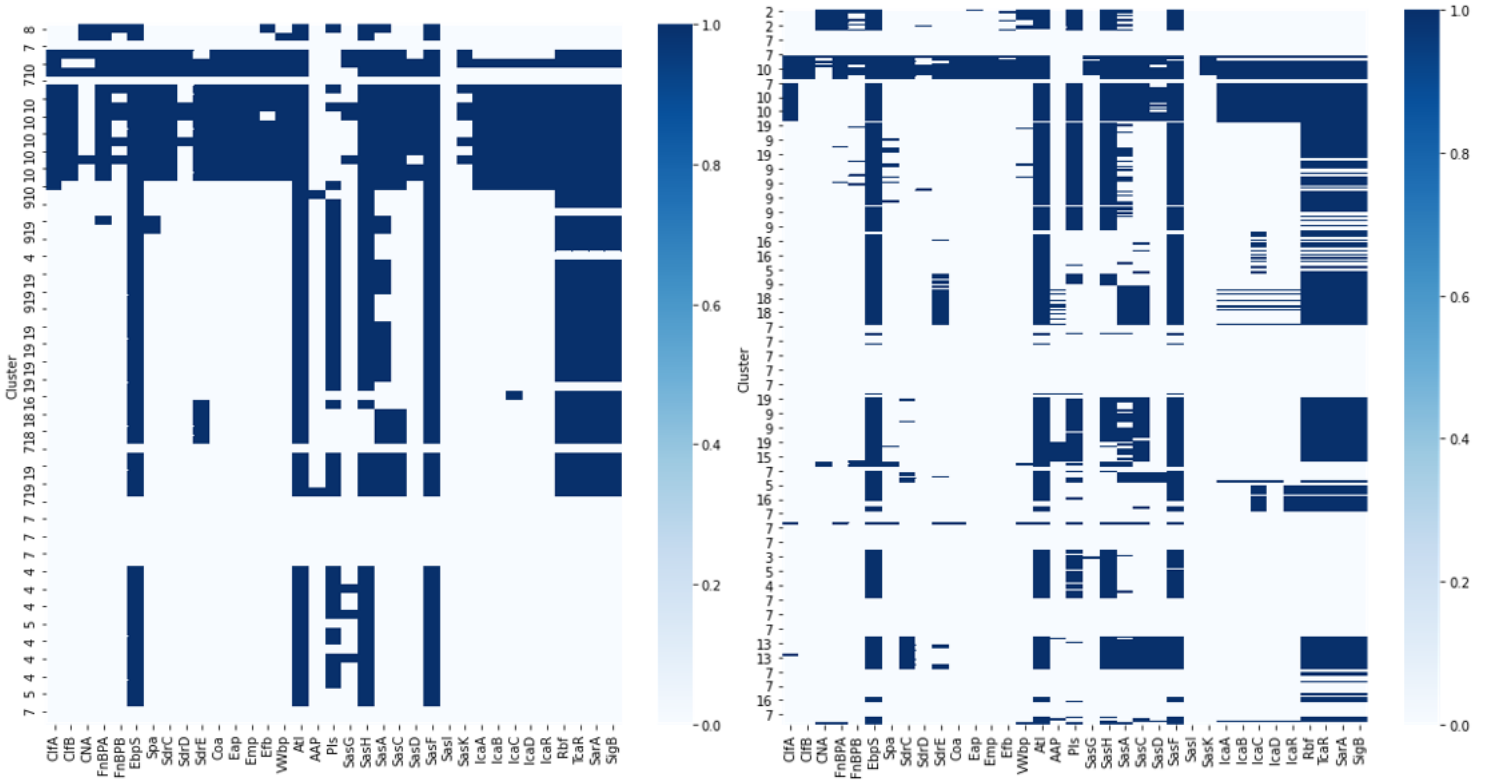

Figure 4

Heat map of the frequency of adhesin and biofilm genes found in Staphylococcus spp. clusters associated with clinical and subclinical mastitis (Left – Clinical mastitis; Right – Subclinical mastitis). The ebpS, atl, pls,sasH, sasF, tcaR, sarA and sigB were most frequently observed

Supplementary Files

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

- supplementaryfigure1legend.pdf
- supplementaryfigure2legend.pdf
- supplementaryfigure3legend.pdf
- supplementaryfigure4legend.pdf
- supplementaryfigure5legend.pdf
- supplementaryfigure6legend.pdf
- supplementaryfigure7legend.pdf
- supplementaryfigure8legend.pdf
- supplementaryfigure9legend.pdf
- supplementaryfigure10legend.pdf
- supplementaryfigure11legend.pdf
- supplementaryfigure12legend.pdf
- supplementaryfigure13legend.pdf
- supplementaryfigure14legend.pdf
- supplementarytable.xls