



Genomic Analysis of Global *Staphylococcus argenteus* Strains Reveals Distinct Lineages With Differing Virulence and Antibiotic Resistance Gene Content

Cosmika Goswami¹, Stephen Fox¹, Matthew Holden², Alistair Leanord^{1,3} and Thomas J. Evans^{1*}

¹ Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, United Kingdom, ² School of Medicine, University of St Andrews, St Andrews, United Kingdom, ³ Scottish Microbiology Reference Laboratories, Glasgow, United Kingdom

OPEN ACCESS

Edited by:

Iain Sutcliffe,
Northumbria University,
United Kingdom

Reviewed by:

Julianne Valentina Kus,
Public Health Ontario, Canada
Meiji Soe Aung,
Sapporo Medical University, Japan

*Correspondence:

Thomas J. Evans
tom.evans@glasgow.ac.uk

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 14 October 2021

Accepted: 05 November 2021

Published: 02 December 2021

Citation:

Goswami C, Fox S, Holden M,
Leanord A and Evans TJ (2021)
Genomic Analysis of Global
Staphylococcus argenteus Strains
Reveals Distinct Lineages With
Differing Virulence and Antibiotic
Resistance Gene Content.
Front. Microbiol. 12:795173.
doi: 10.3389/fmicb.2021.795173

Infections due to *Staphylococcus argenteus* have been increasingly reported worldwide and the microbe cannot be distinguished from *Staphylococcus aureus* by standard methods. Its complement of virulence determinants and antibiotic resistance genes remain unclear, and how far these are distinct from those produced by *S. aureus* remains undetermined. In order to address these uncertainties, we have collected 132 publicly available sequences from fourteen different countries, including the United Kingdom, between 2005 and 2018 to study the global genetic structure of the population. We have compared the genomes for antibiotic resistance genes, virulence determinants and mobile genetic elements such as phages, pathogenicity islands and presence of plasmid groups between different clades. 20% ($n = 26$) isolates were methicillin resistant harboring a *mecA* gene and 88% were penicillin resistant, harboring the *blaZ* gene. ST2250 was identified as the most frequent strain, but ST1223, which was the second largest group, contained a marginally larger number of virulence genes compared to the other STs. Novel *S. argenteus* pathogenicity islands were identified in our isolates harboring *tsst-1*, *seb*, *sec3*, *ear*, *selk*, *selq* toxin genes, as well as chromosomal clusters of enterotoxin and superantigen-like genes. Strain-specific type I modification systems were widespread which would limit interstrain transfer of genetic material. In addition, ST2250 possessed a CRISPR/Cas system, lacking in most other STs. *S. argenteus* possesses important genetic differences from *S. aureus*, as well as between different STs, with the potential to produce distinct clinical manifestations.

Keywords: *Staphylococcus argenteus* pathogenicity islands, *Staphylococcus argenteus*, antibiotic resistance, virulence genes, CRISPR/Cas

INTRODUCTION

Staphylococcus argenteus is a newly described species of bacteria associated with community-associated infection, which was first identified as a divergent clade within the *Staphylococcus aureus* clonal complex (CC)75 (Holt et al., 2011). It was formally classified as a separate species in 2015 based on its genetic signature (Tong et al., 2015). This species shows many phenotypic

characteristics with *S. aureus*, and is indistinguishable from *S. aureus* by routine clinical diagnostic methods, such as tests for coagulase or clumping factor. MLST (Multi Locus Sequence Typing) and *crtM* gene absence may provide specific identification (Becker et al., 2019). *S. argenteus* was first reported in a community of northern Australia in 2002 as causing a skin and soft tissue infection (Ng et al., 2009) and gradually it was reported in other tropical countries (Chantratita et al., 2016; Chen et al., 2018; Kitagawa et al., 2020) as well as in Europe (Dupieux et al., 2015; Argudín et al., 2016; Hallböck et al., 2018), including the United Kingdom. *S. argenteus* was initially suggested to be less virulent than *S. aureus* (Holt et al., 2011) but later studies suggested that infections with *S. argenteus* carry a significantly higher risk of producing respiratory infection, higher morbidity and mortality than methicillin-susceptible *S. aureus* (MSSA) (Chen et al., 2018; Johansson et al., 2019). In addition to skin and respiratory infections, *S. argenteus* is also reported to cause necrotizing fasciitis (Holt et al., 2011), bone and joint infection (Rigault et al., 2018), bacteremia (Chantratita et al., 2016), and food poisoning (Suzuki et al., 2017). Comparison of the *S. argenteus* dominant ST ST2250 with various *S. aureus* strains isolated in Thailand show distinctive antibiotic profile variation between them. Isolates from Thailand are predominantly ST2250 which is methicillin susceptible but resistant to penicillin (Moradigaravand et al., 2017), whereas sequenced ST2250 isolates from Denmark and United States are often methicillin resistant strains, harboring the Staphylococcal cassette chromosome *mec* type IV (SCCmecIV) (Okuma et al., 2002; Hansen et al., 2017). However, given the non-random selection of strains for sequencing, these patterns may not reflect the overall prevalence of antibiotic resistance within *S. argenteus* in these regions. Another emerging ST of *S. argenteus*, ST1223, which had been reported related to food poisoning (Suzuki et al., 2017; Wakabayashi et al., 2018) was also found to be resistant to methicillin.

Other *S. argenteus* strains have been identified in marginally lower numbers in humans (Thaipadungpanit et al., 2015; Zhang et al., 2016; Aung et al., 2017), a few in livestock (Argudín et al., 2016) and also from wildlife such as the African gorilla and bat (Akobi et al., 2012; Schuster et al., 2017). Virulence genes were mostly carried in mobile genetic elements such as plasmids, pathogenicity islands and bacteriophages (Goerke et al., 2009). Almost all *S. aureus* genomes contain staphylococcal pathogenicity islands (SaPI)-like elements and genomic islands vSa α and vSa β (Baba et al., 2002). Different types of SaPIs and associated virulence genes have been identified (Novick et al., 2010). In *S. argenteus*, a SaPI-like element SargPID7903 was identified in a Danish isolate harboring staphylococcus exotoxin B gene (*seb*) and a penicillin-binding protein gene *ear* (Hansen et al., 2017). Comprehensive analysis of mobile genetic elements in *S. argenteus* and associated virulence determinants has not been performed. In addition, genomic studies of inter lineage variation within *S. argenteus* have been limited to small numbers of isolates.

Staphylococcus argenteus can cause significant human disease, and clearly shares many features in common with the much better understood related species *S. aureus*. Previous studies have

explored genetic similarities and differences between these two species with either small numbers (Long et al., 2020) or from one sequence type and geographical location (Moradigaravand et al., 2017). Whole genome sequencing gives a fine-scale insight into bacterial genomes that enables macroevolutionary relationship of lineages to be established and the variety of virulence determinants, mobile genetic elements and antibiotic resistance genes to be determined. The aim of this study was to use global whole genome sequence data to obtain a greater insight into the genetic differences and similarities between *S. argenteus* and *S. aureus*.

MATERIALS AND METHODS

Data Collection and Sequencing

A total of 132 *S. argenteus* genomes were collected from fourteen different countries deposited between the years 2005 and 2018. Of these, 129 genome sequences were collected from National Centre for Biotechnology Information Collection (NCBI) genome database and three United Kingdom isolates were from Whole Genome Shotgun (WGS) sequencing project database PRJEB12513. Out of the 129 genomes from the NCBI database, Illumina short reads were not available but the assembled genome contigs were used for our analysis. Clinical information available for these samples were the sample collection date, source of collection or host (e.g., Human, Pig, Gorilla, etc.) and country of collection; for some studies, source of infection was also available. From the United Kingdom, three genomes were collected from the Reference Laboratory in Glasgow, Scotland from years 2015–2016 and were sequenced at the Wellcome Sanger Institute, Cambridge using the Illumina HiSeq 2000 platform. The sequencing generated 100-bp paired end reads with an average sequencing depth of 66-fold. The sample accession numbers, assembly accession, assembly size and GC% of all the samples used in this study are shown in **Supplementary Table 1**. Where comparison is made to specific *S. aureus* sequences, their accession numbers are given in the text or figure legend. Publications from the Bioproject studies citing these sequences are as follows: PRJEB9575 (Thailand), (Moradigaravand et al., 2017); PRJEB20633 (Denmark) (Hansen et al., 2017); PRJNA317277 (China) (Zhang et al., 2018); PRJNA305687 (Sweden) (Hallböck et al., 2018).

Genome Assembly and Accessory Genes

To identify the STs from the assembled contigs, we have used MLST v2.16.2¹ with the *S. aureus* database from BIGSdb (Jolley and Maiden, 2010; Maiden et al., 2013). *De novo* assembly of those samples, with available sequences reads, were performed using Shovill v1.0.0² with minimum FASTQ read depth 150, minimum contig coverage 2 and assembler SPAdes for the paired-end reads with k-mer sizes 21, 33, and 55. After filtering small contigs (less than 1000 bp), the

¹<https://github.com/tseemann/mlst>

²<https://github.com/tseemann/shovill>

assemblies were rearranged according to a reference genome *S. argenteus* 58113 (Acc. No. AP018562.1) using Mauve v2015-02-13 genome aligner (Darling et al., 2010) for a complete alignment of the conserved regions among the input genome sequences with initial seed-size $\log_2(\text{average sequence length})$ and minimum recursive gap length as 200. These genomes were further annotated for genes with Prokka v1.2 (Seemann, 2014) with genus *Staphylococcus* and e value $1e-9$. The pan-genome of these sequences were identified with Roary (Page et al., 2015) allowing paralog gene splits for a minimum 95% identity in protein database BLAST and a threshold of 99% of isolates for a gene to be in core. Non-metric multidimensional scaling (NMDS) and gene accumulation curve was performed using the “vegan” package in R³. For NMDS “metaMDS” function was used to calculate the “Jaccard” distance matrix between genomes with $k = 2$ and $\text{trymax} = 100$. The Gene accumulation curve was calculated using “specaccum” function from “vegan” package with classic “random” method for 100 permutations. For pathway identification between different STs, DAVID geneontology (Huang et al., 2008a,b) helped to classify the functional variation between the STs.

SNPs and Phylogenetic Analysis

For phylogenetic analysis, ParSNP v1.2 (Treangen et al., 2014) was used to filter the core genome SNPs from the assembled sequences using 1000 as the maximum distance between two colinear MUMs. The *S. argenteus* 58113 strain (ST2250, Acc No. AP018562.1) was used as the reference for variant calling. Phages were masked from the reference genome after identifying them with PHASTER web-server (Zhou et al., 2011; Arndt et al., 2016). From this reference-based SNP-alignment, recombination regions were identified using Gubbins (Croucher et al., 2015) for a threshold of minimum of three SNPs to identify a recombination and for maximum of 5 iterations for tree convergence under the “recombination” method of convergence. The Maximum Likelihood SNP based phylogeny for 132 samples was inferred using RAxML (Stamatakis, 2015) with the generalised time reversible (GTR) model and a gamma distribution to model site-specific rate variation for 1000 bootstraps. The isolates were also checked for a temporal signal using TempEst v1.5.3 (Rambaut et al., 2016) that gives the linear regression of the root to tip distance of the population. The best fit of the root to tip divergence for each of the isolates gave a correlation coefficient of 0.36 and a very low R^2 of 0.13. Determination of a temporal signal was not possible due to the absence of raw sequence reads of some of the isolates.

Antibiotic Resistance and Virulence Genes

For identification of antibiotic resistance genes, the database ARG-annot2 (Gupta et al., 2014) was used for nucleotide BLAST (Altschul et al., 1990) search against the contigs. This database includes genes mediating resistance against the following antibiotic classes: aminoglycosides, beta-lactamases, fosfomycin, fluoroquinolones, glycopeptides,

macrolide-lincosamide-streptogramin, phenicols, rifampin, sulfonamides, tetracyclines, and trimethoprim. If an alignment of at least 90% nucleotide sequence identity was detected that covered at least 90% of the gene length, the ABR gene was considered present in the genome. Similarly, Plasmidfinder (Carattoli and Hasman, 2019) database was used for plasmid replication gene identification using nucleotide BLAST search with 90% minimum identity and 85% minimum coverage. Virulence gene identification was done by using database VFDB (Chen, 2004) with the same cut off as plasmids. This database curates all sequences that enable a microorganism to establish itself on or within a host of a particular species and enhance its potential to cause disease, including bacterial toxins, cell surface proteins that mediate bacterial attachment, secretion system components, cell surface carbohydrates and proteins that protect a bacterium, and hydrolytic enzymes that may contribute to the pathogenicity of the bacterium. Statistical analysis on virulence gene variation between the STs was performed using R and significant values were calculated by two-sample Mann-Whitney test, adjusted with Bonferroni corrections in R.

SaPI Island Detection and Phages

To identify the SaPI-like elements, SaPI integrases were extracted from whole SaPI islands of 31 fully identified SaPI references (Novick et al., 2010; Sato'o et al., 2013; Suzuki et al., 2015) using Artemis Genome visualization (Carver et al., 2011). These integrases were used to perform nucleotide BLAST against the contigs for >90% identity and checked for the alignment lengths. The whole SaPI islands were also searched using nucleotide BLAST to investigate for the island similarity for >60% coverage threshold. EasyFig (Sullivan et al., 2011) was used to plot the similarity between different SargPI islands. Phage integrases from each integrase group (Goerke et al., 2009) were also extracted using Artemis. We defined phage groupings based on the phylogenetic relationships of their integrases, as described (Goerke et al., 2009). Phage1 included Sa2, Sa4, and Sa6 integrases; Phage2 contained Sa3, Sa8, and Sa9 integrases; Phage3 consisted of Sa1 and Sa5 integrase group and Phage4 had Sa12 and Sa7 integrase groups.

Restriction Modification System

For the analysis of the hsdS amino-acid phylogeny, sequences were aligned using MAFFT (Katoh et al., 2002) and a maximum likelihood phylogeny created using RAxML (Stamatakis, 2006) using automatic selection of best amino-acid substitution model by setting the -m flag to PROTGAMEAUTO and using a rapid bootstrap calculation (-f flag set to a) and 1000 bootstrap calculations. FigTree⁴ was used to visualize the resultant best tree. To show homologies between different hsdS protein sequences, representative aligned sequences were visualized using the function msaplot in the ggtree package (Yu et al., 2017). Aligned sequences were viewed in Jalview (Waterhouse et al., 2009) to determine Blossum 62 overall homology scores across aligned sequences (Henikoff and Henikoff, 1992; Eddy, 2004).

³<https://cran.r-project.org/>

⁴<http://tree.bio.ed.ac.uk/software/figtree/>

CRISPR-Cas Loci Identification

The identification of signature genes for CRISPR-Cas types and subtypes were searched by the CRISPRCasFinder web-server (Couvin et al., 2018) on the assembled genomes. For each CRISPR array size of flanking regions was set for 100 bp; repeats were identified with repeat length threshold (min. 23 and max 55) and for spacers minimum 60% sequence similarity was taken. The Cas genes were detected using Prodigal and MacSyFinder program by HMM search on the Cas protein library (Abby et al., 2014). A total of 1021 CRISPR were identified along with some small CRISPR-like by-products within our genomes. Discarding these false positives, a collection of 175 CRISPRs were obtained (with evidence level > 1) and a pool of 44 unique spacers were extracted from these CRISPRs. These spacers were then grouped for 90% similarity using cd-hit⁵ to give 24 spacer groups named SCP1–SCP24 having an average length of 37 bp (ranging from 33 bp till 58 bp). Nucleotide based Maximum likelihood tree (with 1000 bootstraps using RAxML as described above) was built to identify the phylogenetic relationship between the spacer sequence representatives and highly similar sequence searches were carried out in the National Center for Biotechnology Information (NCBI) nucleotide database⁶, using BLASTn for bacterial genomes, plasmids and bacteriophages.

fos gene analysis. Selected *fos* genes were aligned using SEAVIEW (Gouy et al., 2009) with the CLUSTAL omega (Sievers and Higgins, 2018). A maximum likelihood phylogeny created using RAxML (Stamatakis, 2006) using a generalized time reversible model and using a rapid bootstrap calculation (-f flag set to a) and 1000 bootstrap calculations. FigTree (see foot note 4) was used to visualize the resultant best tree. To show homologies between different hsdS protein sequences, representative aligned sequences were visualized using the function msaplot in the ggtree package (Yu et al., 2017).

RESULTS

Global Isolate Collection

The 132 *S. argenteus* genomes included in this study were from isolates collected from patients with *S. argenteus* infection from fourteen different countries from the years 2005–2018. The isolates were from Thailand (51%, 68), Denmark (18%, 25), Malaysia (6%, 8), China (5%, 7), Australia (3.8%, 5), United Kingdom (3%, 4), United States (3%, 4), Sweden (2.3%, 3), Singapore (1.5%, 2), and also from France, Germany, Gabon, Israel, and Japan. The source of isolates were 129 from humans, along with three from animals: one pig from China, one cow from Malaysia and one gorilla from Gabon. Clinical metadata was not available for all samples, but where present is indicated in **Supplementary Table 1**. Three of the four United Kingdom isolates included in this study were from Scotland and were clinically misidentified as *S. aureus* until the whole genome SNP analysis identified them as outliers and were subsequently identified as *S. argenteus* by MLST (Thomas et al., 2006) and were

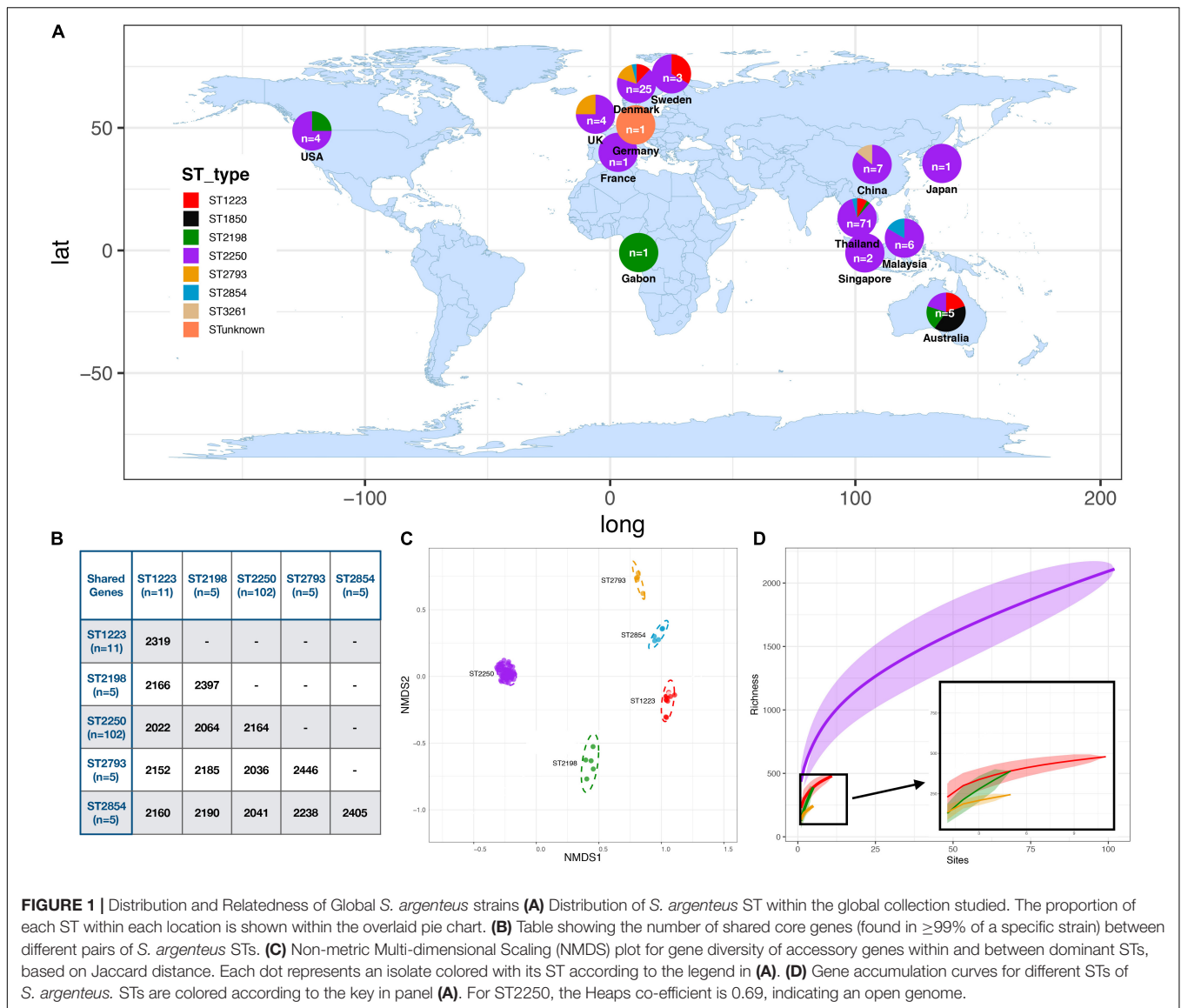
also crtM-negative (Chen et al., 2018). One of the Scottish isolates was from blood cultures of a patient who was severely ill and with a travel history from South-East Asia. Details of the isolates, their genomes, and their accession numbers are in **Supplementary Table 1**. ST assignments of genomes analyzed here were all in agreement with the original designations where published. It is important to note that the range of samples of *S. argenteus* analyzed here is liable to selection bias, since sampling strategies will vary between locations and only whole genomes that have been deposited in public databases can be analyzed.

Core and Accessory Gene Variation

We compared the genomes of these different *S. argenteus* strains to understand the variation in core and accessory gene content. *De novo* assembly of the genomes of these strains gave the genome size from 2.6 to 2.9 Mb and the GC content varied from 31.6 to 32.7% (**Supplementary Table 1**). The MLST scheme identified 7 distinct STs within the isolates, and one sample with indeterminate ST for which sequence quality in one of the alleles used for typing was not sufficient to classify unequivocally within a known ST group. The global collection was dominated by ST2250 (77%, 102) followed by ST1223 (8.2%, 11) along with a diverse collection of STs (ST2854, ST2198, ST2793, ST1850, and ST3261) each with fewer than 10 isolates (**Figure 1A** and **Supplementary Table 2**). The ST ST2250 is an established clade in Thailand (59%, 60/102) and also distributed globally in many eastern and European countries (Moradigaravand et al., 2017). The pig and the cow isolates belong to ST2250 whereas the gorilla isolate was ST2198. The pangenome carried a total of 5,169 genes of which 40% ($n = 2,069$) were present in 99% of our 132 genomes. The accessory genome (contained in > 5% but <99% of isolates) consists of 3,091 genes. These genes were mostly hypothetical genes, only 32% genes having functional annotations, mostly for plasmid genes or pathogenicity island genes but no specific functional pathways were identified. Core genomes of isolates (contained in >99% of isolates with 95% minimum identity) within individual clades ST2250, ST1223 and ST2198 contained 2164, 2319, and 2397 genes, respectively. Comparisons of the genes shared between 99% of the different STs are shown in **Figure 1B**. This shows that hundreds of the core genes of each ST were unique (defined as less than 95% identity), indicating clade-specific gene accumulations (**Figure 1B**). The relationship between each isolate was investigated using the non-metric multidimensional scaling (NMDS) method where pairwise genomic distances were calculated using Jaccard distances between the strains. Formation of distinct clusters for each ST confirmed that accessory genes were conserved within a ST (**Figure 1C**). Gene accumulation curves for the different STs of *S. argenteus* showed a steady accumulation of new genes with successive addition of genomes – for ST 2250 the Heaps law co-efficient was 0.69 (<1) indicating an open genome (Tettelin et al., 2008). This is similar to the open genome of *S. aureus* (John et al., 2019) but in contrast to the closed genome of *S. lugdunensis* (Argemi et al., 2018). A heat map showing gene presence/absence for the pangenome of our samples shows the distribution of the core and accessory genomes across the different STs (**Supplementary Figure 1**).

⁵<https://github.com/weizhongli/cdhit>

⁶<http://www.ncbi.nlm.nih.gov/>



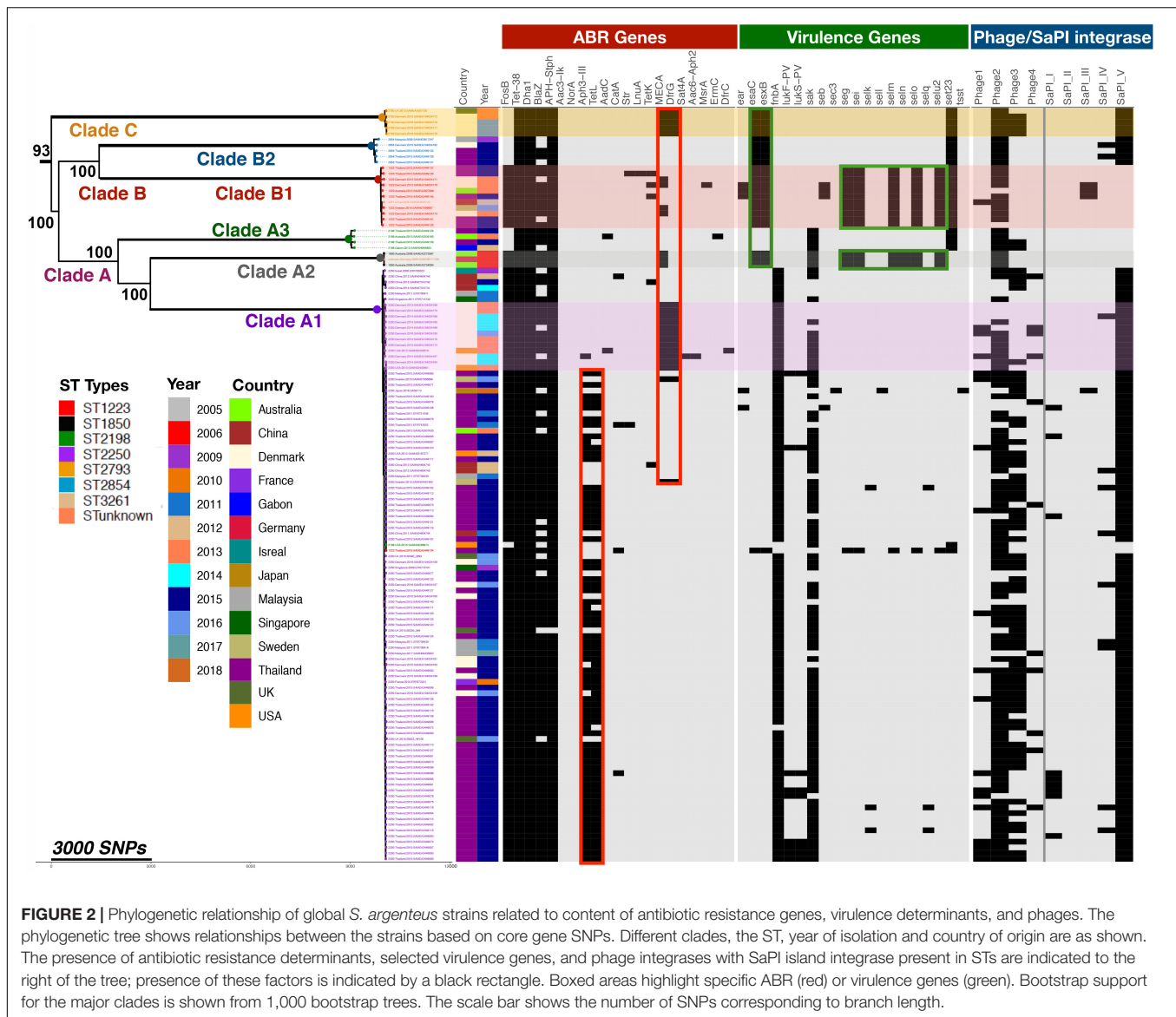
SNP Variation and Evolution of Different Strains

To study in more depth the population structure of *S. argenteus* and its relationship to the distribution of virulence determinants and genes conferring antibiotic resistance, we constructed a phylogenetic tree of this collection of *S. argenteus* strains using core genome SNPs and Maximum-likelihood (ML) phylogenetic analysis. Following exclusion of mobile genetic elements (MGEs) and recombination regions, the core genome was found to have a total of 34,837 SNPs in 132 genomes. This revealed three distinct clades which we have named A, B, and C, with the subdivisions as shown (**Figure 2**). We then overlaid the STs of the different strains on the midpoint rooted phylogenetic tree, which mostly segregated to distinct clades and subclades, but with some exceptions. Clade A1 contained the dominant ST ST2250, but also one sample of ST2198 and one of ST1223. It is not clear why these two samples segregated into this clade. Clade A2 contained

both ST1850 and ST unknown, while clade B1 contained ST1223 and ST3261. We attempted to construct a time-scaled tree of the genomes to establish the evolution of the different clades over time but this was not possible as the raw sequence reads were not available for many of the sequenced strains.

Antibiotic Resistance Genes

We determined the distribution and spread of antibiotic resistance genes within this population to determine if such genes were generally present or clade specific. Compared to *S. aureus*, antimicrobial drug resistance of *S. argenteus* is known to be lower (Chantratita et al., 2016; Aung et al., 2017). In this collection, 21 antimicrobial genes were identified and the distribution of these genes within different clades is shown in **Figure 2**. The *mecA* gene conferring resistance to methicillin was identified in 14% ($n = 14/102$) of ST2250, 45% of ($n = 5/11$) ST1223 and in all five genomes of ST2793. None of



the Thailand isolates have *mecA*, but it has been acquired by isolates from other geographical areas. It was most prevalent in isolates from Denmark (68%, $n = 17/25$), but also found in those from Australia (40%, $n = 2/5$), United States (40%, $n = 2/4$), United Kingdom (25%, $n = 1/4$) and isolates from Sweden and Germany (Figure 2). These isolates all contained a SCCmec IV element. In addition to methicillin resistance, these isolates also contained elements conferring resistant to penicillin (*blaZ*), trimethoprim (*dfrA*), and cadmium (*cadC*) (Moradigaravand et al., 2017). 50% of isolates carry the *tetL* gene mediating resistance to tetracycline, whereas 5% of isolates carry *tetK*; *tetL* was found only in the ST2250 clade (67/103, 65%). About 90% of *S. argenteus* isolates carry the fosfomycin resistance element *fosB*, which is reportedly rare in the *S. aureus* population (Fu et al., 2016b; Xu et al., 2020). The gene was highly conserved within the strains of *S. argenteus* analyzed here, with 99.4% pairwise nucleotide identity over a sequence length of 420

nucleotides. BLAST nucleotide searching in *S. aureus* genomes revealed *fosB* gene sequences with over 98% identity to the *fosB* gene in *S. argenteus* (21,609 records found in 2,199,446 total *S. aureus* nucleotide sequences in the NCBI database (0.98%) - in both organisms where complete assemblies were available, the gene was located within the bacterial chromosome. A range of *fos* genes have been described which share the common feature of inactivating fosfomycin by enzymatic opening of its oxirane ring (Castañeda-García et al., 2013). A number of subtypes with alphabetical suffixes have been described in both Gram-positive and Gram-negative bacteria, including a range of *fosB* genes identified in plasmids of *S. aureus* (Fu et al., 2016a) that differ from the chromosomal *fosB* gene. BLAST searching for these plasmid *fosB* sequences in the NCBI database found only 8 *S. aureus* records containing sequence matches with >95% identity. Representative examples of these different *fos* genes were aligned and these comparisons with a phylogenetic tree of the

sequences are shown in **Supplementary Figure 2**. This highlights the close similarity of the chromosomal *fosB* genes of *S. aureus* and *S. argenteus*, which differ significantly from plasmid borne *fosB* genes described in *S. aureus*.

Virulence Gene Content

Comparative analysis of the 132 *S. argenteus* sequences for a range of virulence factors identified a total of 213 virulence genes that differed between the different strains, as well as differing phage content (**Figure 2** and **Supplementary Table 3**). These include bacterial toxins, cell surface proteins that mediate bacterial attachment, secretion system components, cell surface carbohydrates and proteins that protect a bacterium, and hydrolytic enzymes that may contribute to the pathogenicity of the bacterium. The average number of virulence genes in ST1223 and ST2793 was found to be significantly higher ($p < 6.7 \times 10^{-9}$ and $p < 4.9 \times 10^{-5}$, respectively) compared to ST2250 (**Supplementary Figure 3**). ST1223 has previously been noted to have a larger number of virulence genes (Aung et al., 2021). Substantial differences in virulence factors were found between STs (**Figure 3**). Adherence and secretion system genes were present selectively in all STs. However, toxin genes were more common in ST1223 compared to ST2250. These genes were mostly found in bacteriophages, genomic and pathogenicity islands (**Figures 4, 5**). In most cases, the virulence determinants were similar to those found in *S. aureus*.

In *S. aureus*, important virulence determinants are frequently found in genomic islands. One of these is the genomic island vSa α (Type I and Type II) associated with a cluster of staphylococcal superantigen-like (*ssl*) genes which varies considerably between different strains of *S. aureus* (Lina et al., 2004). vSa α also contains a cluster of genes encoding lipoproteins. This island was found in all the *S. argenteus* strains studied here (**Figure 4A**). As is the case with *S. aureus* (Lindsay and Holden, 2004; Lindsay, 2010), there is some variation within the *ssl* genes and other genes within the island. However, the overall structure was well conserved within the different strains, and there was no significant difference between different STs.

The *S. aureus* enterotoxins *seg*, *sei*, *selm*, *seln*, *selo*, and *selU2* are harbored in a genomic island termed vSa β (Type I) (Omoe et al., 2005); this was also identified in *S. argenteus*, almost exclusively in clades B1 and clade A2 (**Figure 2**). These enterotoxins cluster together in what is termed the enterotoxin gene cluster (*egc*) (Jarraud et al., 2001; Thomas et al., 2006) which is widely distributed within strains of *S. aureus*, although there is considerable allelic variation of this locus between strains (Lindsay and Holden, 2004, 2006). The *egc* locus found in the different strains of *S. argenteus* clades B1 and A2 is shown in **Figure 4B**. Broadly, the *egc* locus in these strains was very similar to the locus found in some strains of *S. aureus*, such as the ST45 strain MCRF184 (Aswani et al., 2019), although lacking the gene encoding a helicase (*rep*), which is found in that strain. Coltery et al. (2009) classified the locus into different types based on their gene content; the predominant type found in the *S. argenteus* strains described here is their *egc2*. However, as with *S. aureus*, there is variation; for example in the strain SAMN4604742, the

seg gene is lost and the orientation of the *selu* and *seln* genes is reversed (**Figure 4B**).

There are other differences between the strains of *S. argenteus*. Genes encoding type VII secretion system extracellular proteins (*esaC*, *esxB*) were missing from clade A1 and A3. The genes encoding the PVL Pantan-Valentine leucocidin (PVL), *lukF-PV*, *lukS-PV*, carried by prophages were identified in nine ST2250 isolates. The staphylokinase encoded by *sak*, which activates plasminogen to plasmin and inactivates bacterial defensins (Bokarewa et al., 2006), was found in clade A in 79% (81/102) of ST2250 and in 100% of ST2198. Fibronectin binding protein gene *fmbA* was only found in clade A1 and absent for all others. Some toxin genes were common across all clades. These were: the enterotoxin gene (*selX*) which plays a significant role in the inhibition of host innate immune system pathways; the staphylococcal enterotoxin Y gene (*secY*); and antitoxin component of type-II toxin-antitoxin (TA) system genes (*mazE*, *YefM*).

Pathogenicity Islands

Staphylococcus aureus contains a wide variety of phage mobilized pathogenicity islands (Novick et al., 2010). We searched for these *S. aureus* Pathogenicity Island (SaPI)-like elements present in *S. argenteus*, as described in the Materials and Methods. New genomic islands were found that possess common features of a SaPI and very similar mosaic structure of its open reading frames (ORFs.) They have direct repeats at the both ends, along with the basic components of a SaPI: *int*, *stl*, *str*, *xis*, *pri*, *rep*, and *terS* genes. The integrases (*int*) showed very high sequence similarity to SaPI integrases (>90%) and the attachment sites were similar to those of previously reported SaPIs (Novick et al., 2010). We referred to these genomic islands as SargPIs using the nomenclature previously defined (Subedi et al., 2007; Novick et al., 2010). Within our 132 *S. argenteus* isolates, 115 (86%) had at least one SargPI and 22 (17%) isolates have two of these islands each. Within these SargPIs, four novel SargPIs were identified: SargPI58113 (*att*-site III), SargPI5 (*att*-site IV), SargPIThai098 (*att*-site I), and SargPICChina744 (*att*-site IV) shown in **Figure 5** and **Supplementary Figures 4–7**. The first pathogenicity island, SargPI58113, was identified in reference genome *S. argenteus* 58113, harboring toxin gene *tst-1* along with *ear*, *sec3* and *sell* genes. Its integrase was 98% similar to SaPI_{mw2} but matches 50% of SaPI_{bov1} (**Supplementary Figure 5**). One of the Thai ST2250 strains was found to have this SargPI58113 island. The second island, SargPI5, contained *selk* and *selq* toxin genes and had the integrase of SaPI5, but identical to only 30% of the island. Three of our genomes have this SargPI5 island (**Supplementary Figure 6**). Another novel pathogenicity island, SargPIThai098, was found in 8 of our isolates whose integrase was 88% similar to SaPI1 (**Figure 5A**). These islands did not harbor any known toxin genes but incorporated bacteriophage vSa1 (Φ 55) (Goerke et al., 2009). The fourth island, SargPICChina744 was identified with 10% island similarity with SaPI5 in one genome without any toxin genes (**Figure 5C**). Previously defined pathogenicity island SargPID703 (Hansen et al., 2017) contained virulence genes *seb* and *ear*, and was present in four genomes at the *att*-site III (**Figure 5B**). This island is seen only in clade

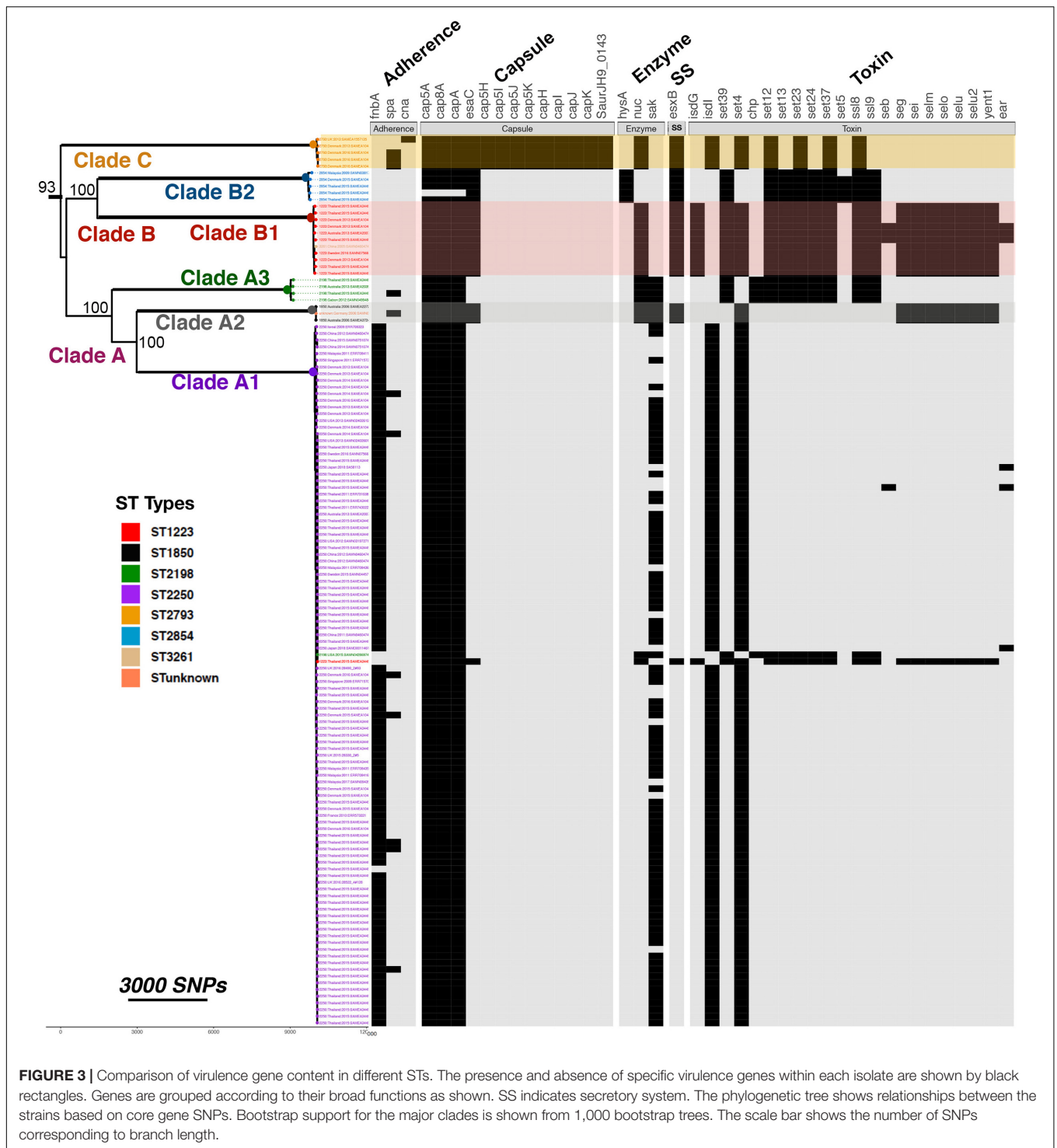


FIGURE 3 | Comparison of virulence gene content in different STs. The presence and absence of specific virulence genes within each isolate are shown by black rectangles. Genes are grouped according to their broad functions as shown. SS indicates secretory system. The phylogenetic tree shows relationships between the strains based on core gene SNPs. Bootstrap support for the major clades is shown from 1,000 bootstrap trees. The scale bar shows the number of SNPs corresponding to branch length.

ST1223. We have also found some SaPIs within our isolates. SaPITokyo12431 carrying *seb* and *ear* toxin genes was present in two isolates of ST2250 (Figure 5A). SargPITokyo11212 was identified in seven genomes of ST2250 isolates, conserving 85% of the SaPITokyo11212 but having lost the enterotoxin genes *seb* and *ear* from it (Figure 5C). The island SargPIhirosaki4 was identified in three isolates of *S. argenteus* at the *att*-site

V, which was 84% similar to SaPIhirosaki4 with missing toxin genes (Figure 5D).

Type I Restriction Modification System

Staphylococcus aureus contains a type I restriction modification system comprising genes encoding DNA modification (*hsdM*), a restriction unit (*hsdR*), and a sequence specificity element (*hsdS*)

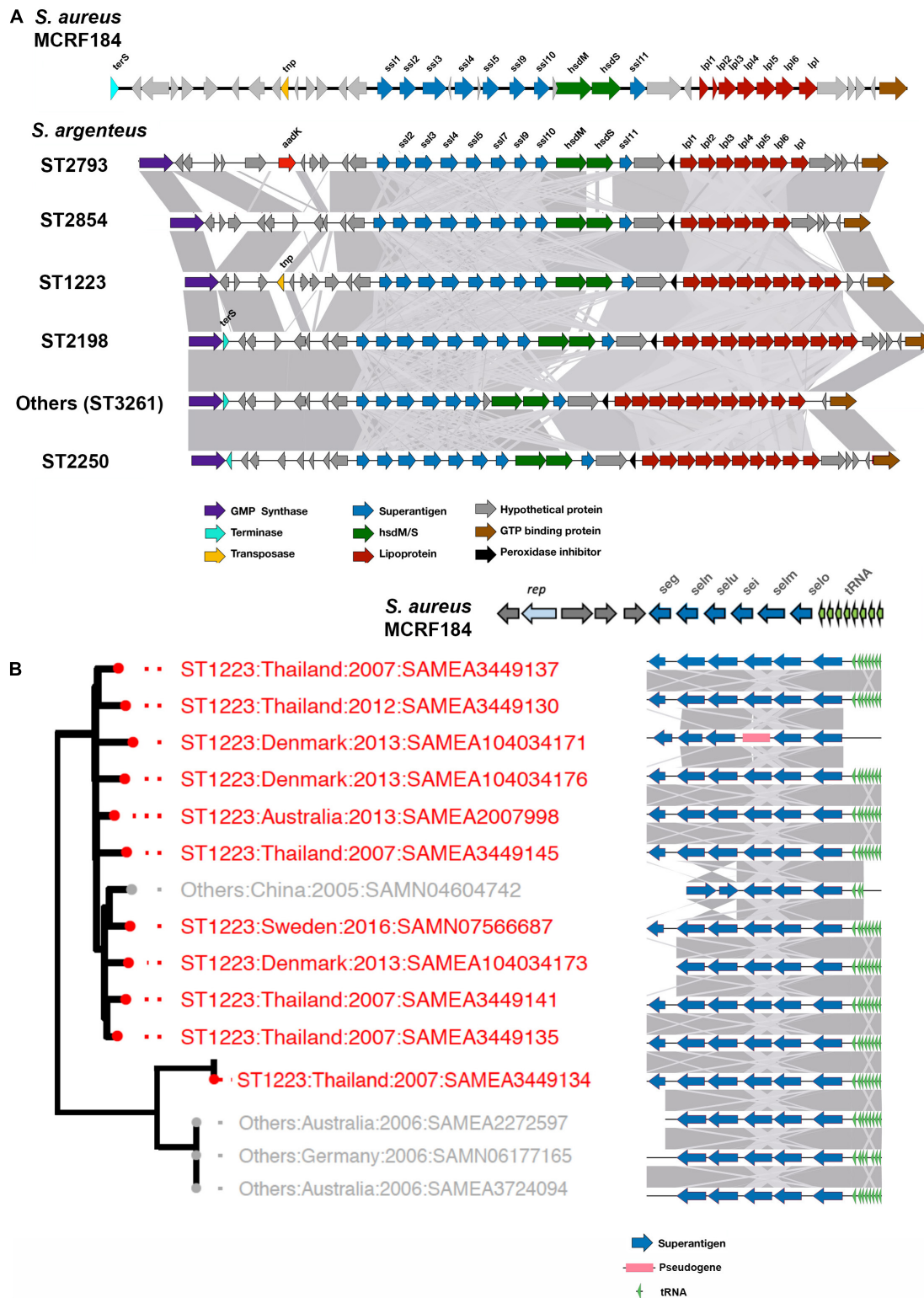
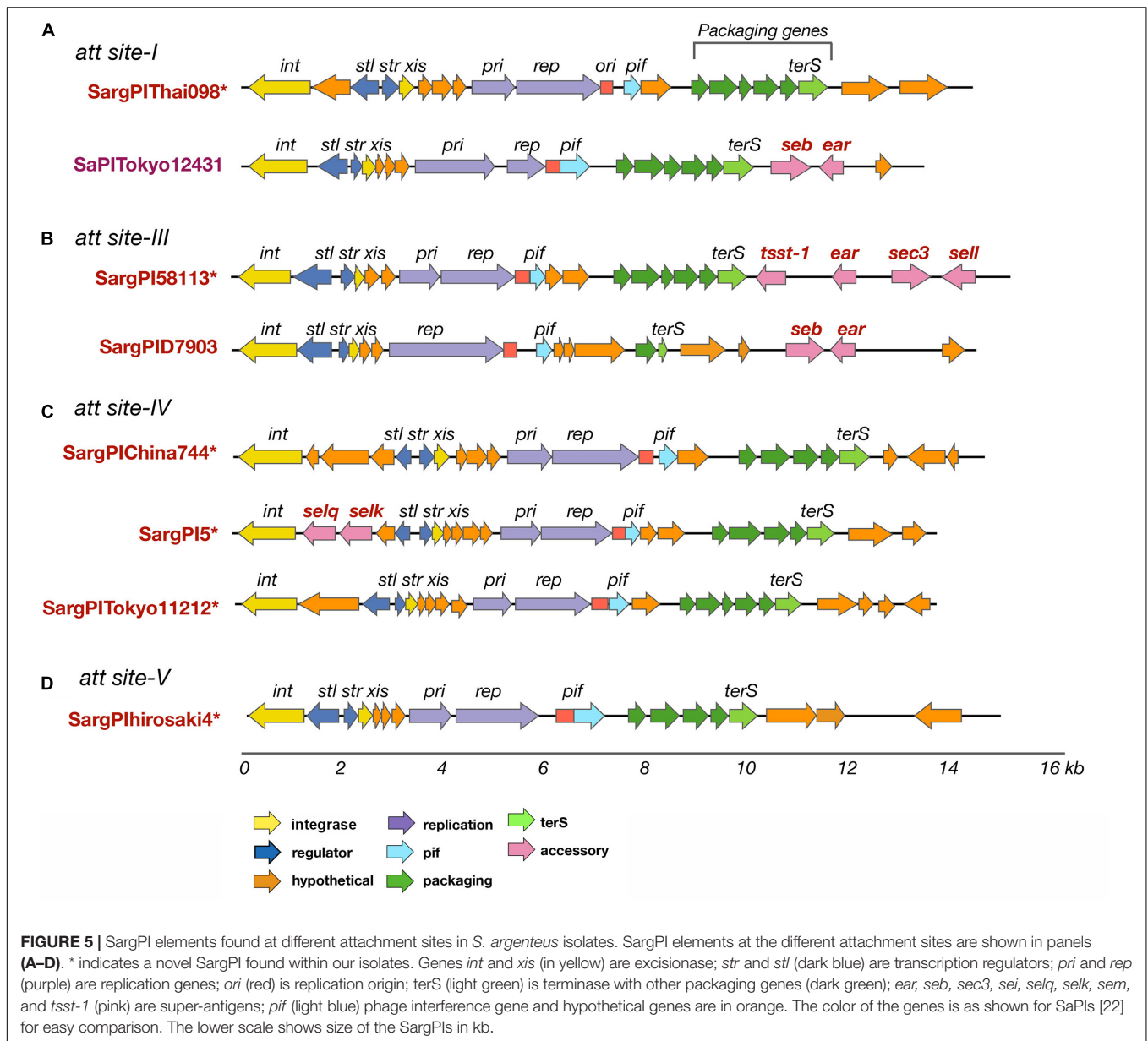


FIGURE 4 | vSA α and vSA β elements in *S. argenteus*. The figure shows the gene composition of vSA α (A) and vSA β (B) elements in the strains of *S. argenteus* indicated. In panel (A), a representative member from each ST is shown. Gene functions are color coded as indicated. The degree of similarity between the different *S. argenteus* strains is shown by the gray shading. In both panels, the arrangement of genes in the cognate elements of the MCRF184 strain of *S. aureus* is shown. *rep* is a bacterial replicase (helicase).



(Waldron and Lindsay, 2006). The *hsdM* and *hsdS* encoded subunits together form a sequence specific methylase – they can also combine with the *hsdR* encoded product to form a DNase that cleaves DNA at unmethylated sequences as specified by the *hsdS* product. In this fashion, the organism protects itself against uptake of foreign DNA, such as harmful bacteriophages. The specificity of the *hsdS* protein subunit is specific to particular sequence types of *S. aureus*, thus limiting horizontal transfer of genes from other strains (Cooper et al., 2017). This sequence type specificity of the *hsdS* protein sequence is not seen in the closely related species *S. epidermidis*, where the type I restriction modification system shows significant variation even within a single lineage (Lee et al., 2019); however, this report found that 38% of the *S. epidermidis* strains studied lacked a type I modification system. In order to understand how a type I

restriction modification system might influence horizontal gene transfer in *S. argenteus* and whether this differed from *S. aureus*, we undertook a detailed analysis of the *hsdS* protein sequences in the sequences studied here. 121 (92%) of the 132 sequences analyzed contained at least one copy of the *hsdS* gene; of these 101 isolates had one copy [ST2250 (91 isolates), ST2854 (5 isolates), ST2198 (5 isolates)], 17 isolates had 2 copies [ST1223 (11 isolates), ST2793 (5 isolates) and ST 3261 (1 isolate)], and 3 isolates had 3 copies [ST1850 (2 isolates), indeterminate ST (1 isolate)]. Those strains without an *hsdS* gene were all members of ST2250.

In order to determine if specific strains of *S. argenteus* had conserved type I restriction modification systems that were not shared by other strains, we carried out sequence comparisons of the *hsdS* proteins. The protein sequences were chosen for comparison rather than nucleotide as it is the variation in protein

sequence that will dictate the specificity of the hsdS subunit and hence the target of the type I restriction modification system. Strains of the same ST had very similar hsdS proteins. Of those STs with only 1 *hsdS* gene, the hsdS protein sequences had 100% average pairwise identity for ST 2854, 99.12% for ST2198, and 99.99% for ST2250. For those strains that had more than one *hsdS* gene, the protein sequences encoded by the additional gene(s) grouped together in a given ST but were distinct from the other copy (Supplementary Figure 8). Note that some of the copies have significant deletions, e.g., copy 3 of SAMEA2272597 and SAMEA3724094. We compared the sequences of representative hsdS proteins from different STs and constructed a phylogenetic tree using RAxML. Even with 1,000 bootstraps, support for some of the branches was poor. However, the sequence comparison showed the variation in the sequences was located within the two

target recognition domains of the proteins, with conservation of sequence in the proximal, central and distal regions (Figure 6). This suggests that these proteins direct distinct methylation patterns within their respective STs, and thus may serve to limit horizontal gene transfer between STs of *S. argenteus*, in much the same fashion as described for *S. aureus*.

CRISPR/Cas Locus

Many bacteria also limit introduction of foreign DNA from phages or plasmids by a system characterized by clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR associated (Cas) systems. These CRISPR/Cas systems allow bacteria to capture sequences from invading phages or plasmids in heritable DNA arrays. These then are used by the Cas protein to recognize identical sequences when they are

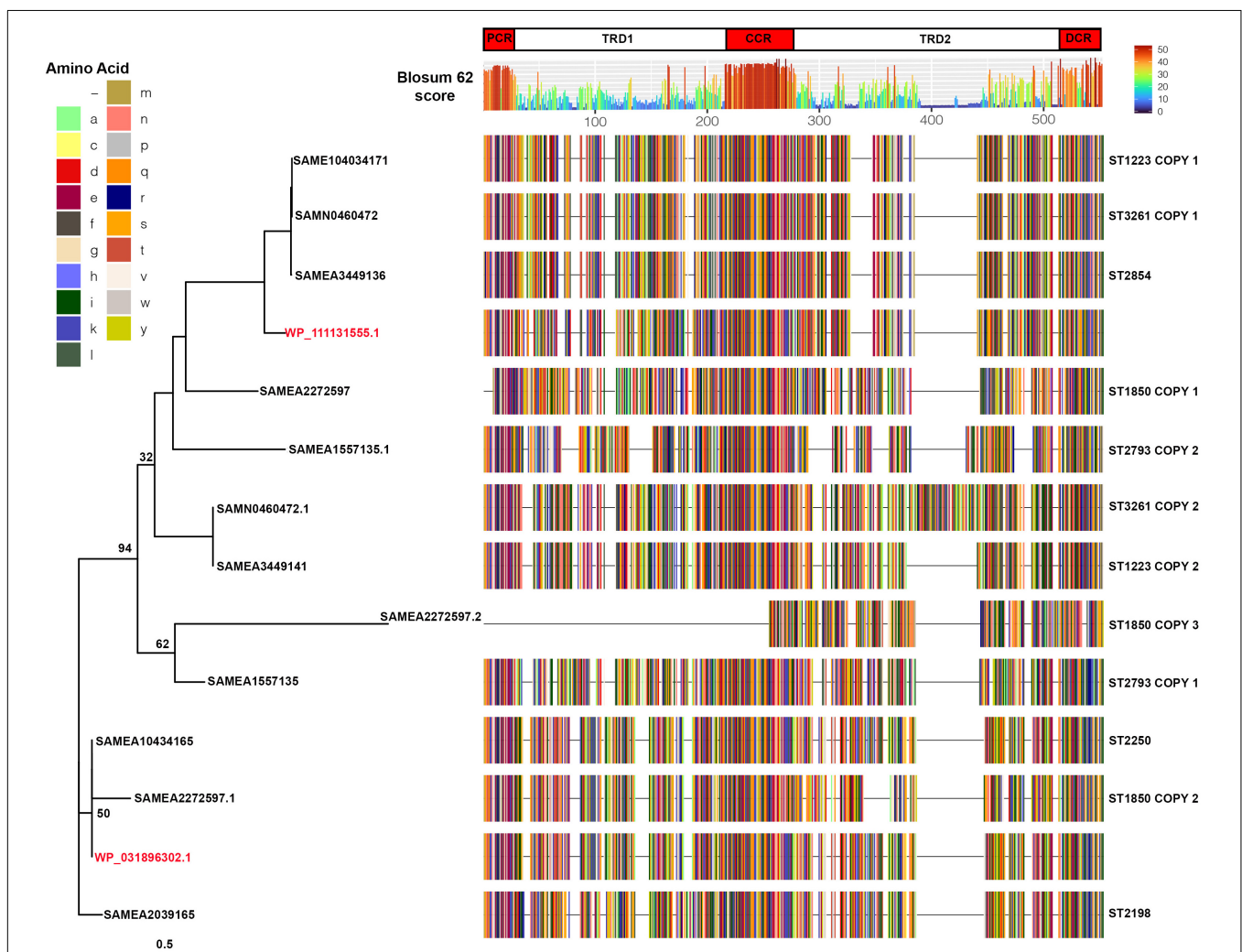


FIGURE 6 | Sequence comparisons of *hsdS* proteins in *S. argenteus* strains. Representative sequences of the aligned *hsdS* proteins from the different *S. argenteus* STs are shown, colored by amino acid residue as shown in the key. Two *S. aureus* *hsdS* protein sequences are shown in red, labeled with their corresponding National Center for Biotechnology Information identifier. Blosum62 homology scores across all the sequences are shown above the alignment, colored as shown in the key. The domains of the *hsdS* protein are shown above the homology score graph – PCR, proximal conserved region; TRD1 and 2, target recognition domains 1 and 2; CCR, central conserved region; DCR, distal conserved region. The phylogenetic tree to the left shows the sequence relationships. Bootstrap support for the major branches is shown (1,000 bootstraps). The scale shows the average number of substitutions per site.

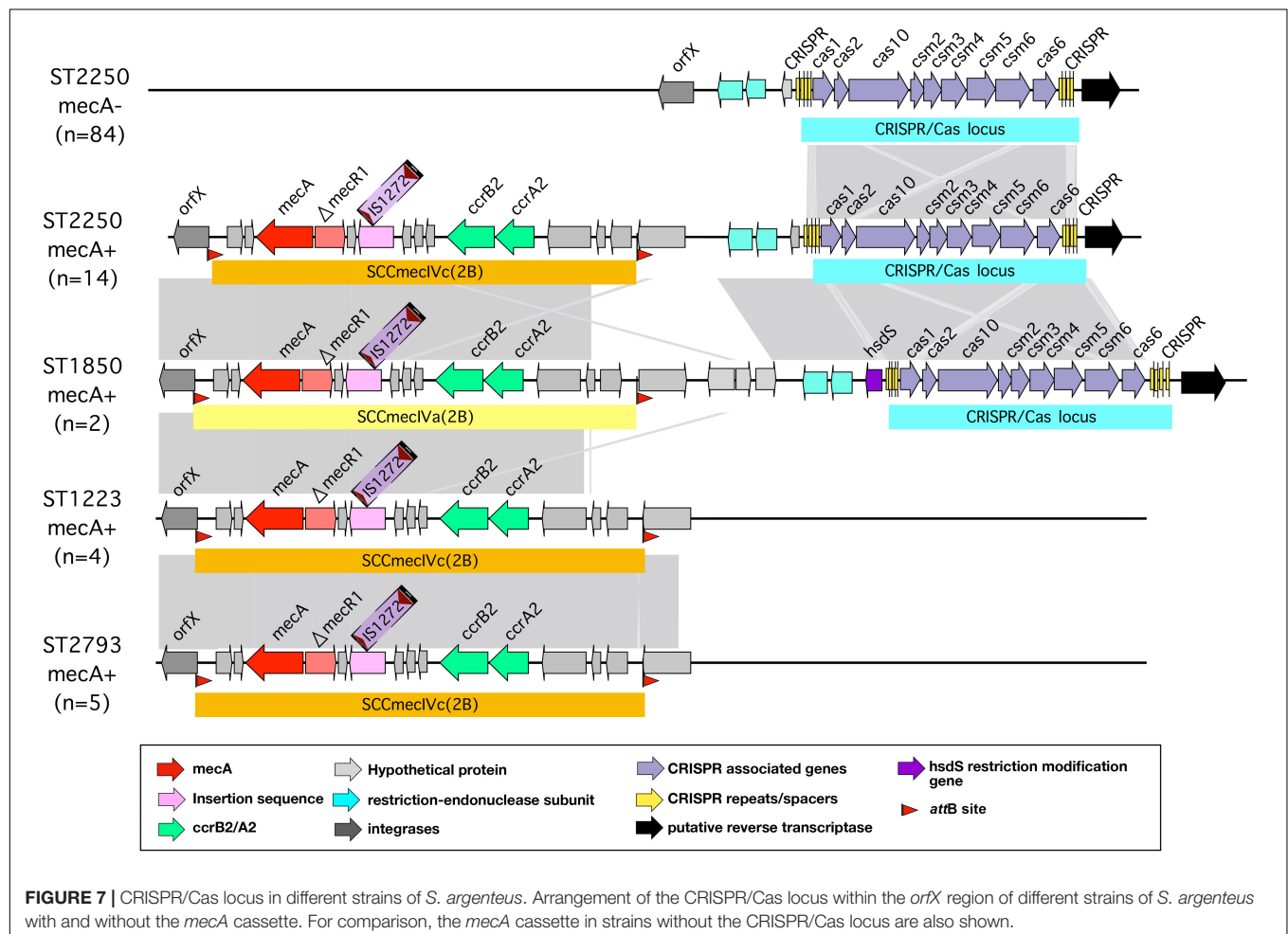
encountered on phage or plasmid invasion and to target them for degradation (Makarova et al., 2011; Rath et al., 2015). In general, few strains of *S. aureus* possess a CRISPR/Cas locus (Cao et al., 2016); previous studies have found this locus closely adjacent to the *mec* locus in MRSA + strains, some of which are prevalent in livestock (Golding et al., 2010; Cao et al., 2016). Analysis of the sequences studied here showed the presence of a type IIIA CRISPR/Cas locus found in almost all (98/102) ST2250 strains, and also in three closely related ST1850 strains. It was not present in the other STs studied. In all cases, the CRISPR/Cas element was found closely adjacent to the chromosomal locus containing the *attB* site at the 3' end of the *rlmH* gene encoding a rRNA methyltransferase (formerly termed *orfX*) - this is the site of insertion of the *mecA* chromosomal cassette where present, as shown in strains with and without this cassette in **Figure 7**. Most strains with a CRISPR/Cas element had two CRISPR regions. The type IIIA CRISPR/Cas locus in these strains is syntenic with the CRISPR/Cas locus found in the livestock-associated *S. aureus* strain 08BA02176 (Golding et al., 2012).

The spacer regions of the CRISPR/Cas arrays frequently contain sequences with high homology to phage and plasmid sequences that would then provide immunity to introduction of these elements into the bacterial cell. Database searching using

the sequences of the spacer elements found in the *S. argenteus* CRISPR/Cas arrays revealed that they contained numerous examples of phage and plasmid derived sequences (**Figure 8** and **Supplementary Table 4**). We grouped overlapping matches and assigned them specific group names. The distribution of the different groups within the *S. argenteus* strains is shown in **Figure 8**. The elements include sequences from a number of different phages known to infect *S. aureus*, as well as plasmid sequences from the pWBG plasmid series first described from remote areas of Western Australia (Shearer et al., 2011). Other sequences from Staphylococcal genomes are from spacer regions of CRISPR/Cas loci of other strains.

DISCUSSION

We have performed a comprehensive analysis of 132 *S. argenteus* genomes from isolates collected from fourteen countries over 13 years. Two sequence types dominate the present global collections - ST 2250 and ST 1223 - with other diverse STs in much smaller numbers. We show that there is significant variation in genetic structure within this organism, even within a specific sequence type, with differing



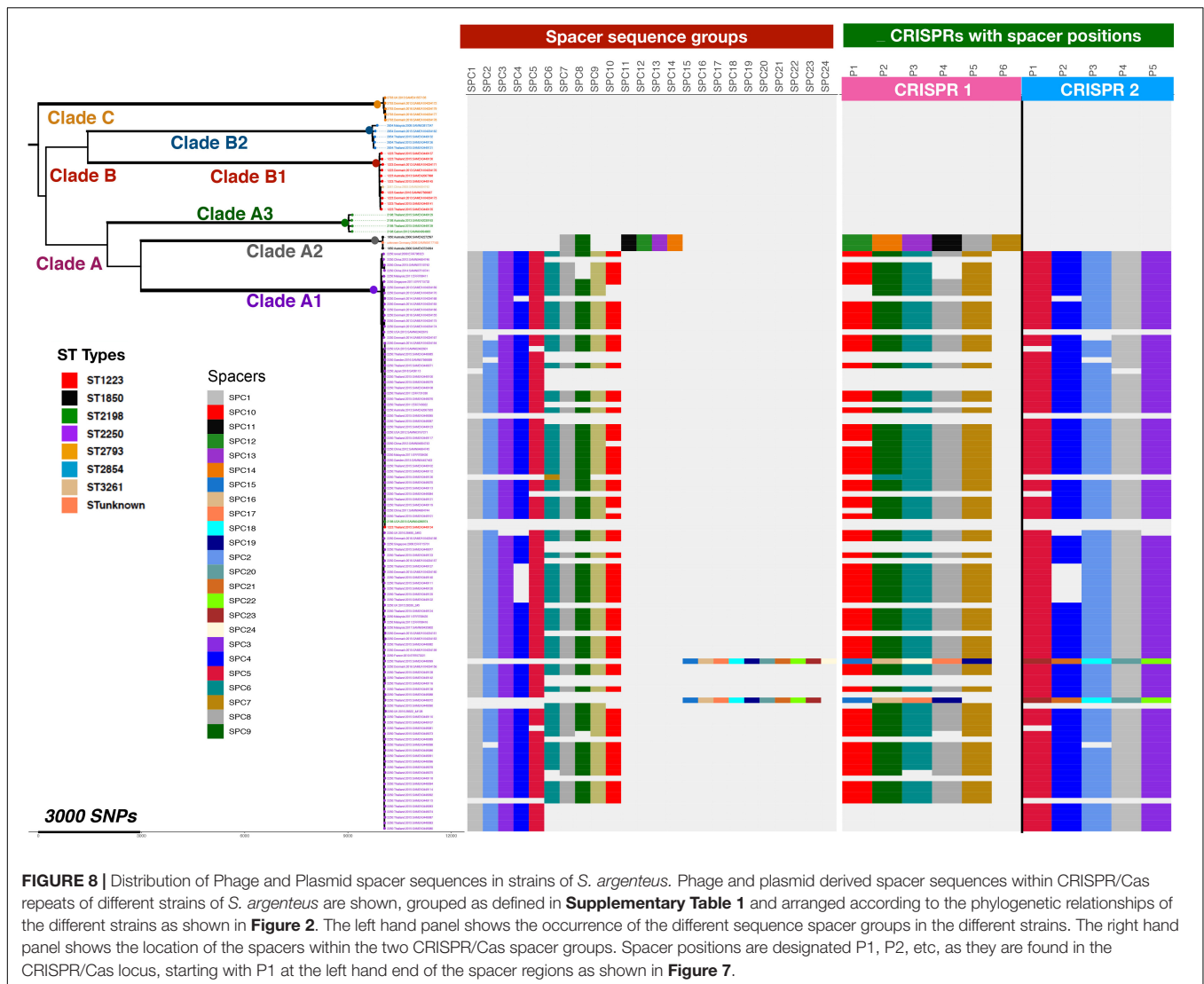


FIGURE 8 | Distribution of Phage and Plasmid spacer sequences in strains of *S. argenteus*. Phage and plasmid derived spacer sequences within CRISPR/Cas repeats of different strains of *S. argenteus* are shown, grouped as defined in **Supplementary Table 1** and arranged according to the phylogenetic relationships of the different strains as shown in **Figure 2**. The left hand panel shows the occurrence of the different sequence spacer groups in the different strains. The right hand panel shows the location of the spacers within the two CRISPR/Cas spacer groups. Spacer positions are designated P1, P2, etc, as they are found in the CRISPR/Cas locus, starting with P1 at the left hand end of the spacer regions as shown in **Figure 7**.

virulence and antibiotic resistance determinants. As with the closely related *S. aureus*, we found the majority of isolates of *S. argenteus* contained at least one mobile *S. aureus* like pathogenicity island, containing a number of different virulence determinants, and identified novel elements. The strains contain strain-specific Type I restriction-modification systems, and almost all ST2250 strains contain a CRISPR/Cas system which is not found in other strains. Taken together, these findings demonstrate that *S. argenteus* has considerable genetic diversity. This may reflect different selection pressures that have allowed it to establish distinct ecological niches occupied by different STs, although neutral mechanisms such as geographical separation and genetic drift may also be important (Sheppard et al., 2018). In addition, it is important to emphasize that the collection of sequences analyzed here is subject to significant bias, since sampling strategies will vary significantly between different locations, particularly where opportunities for surveillance are limited, and resources for whole genome sequencing are limited.

The four most prevalent clades found within these isolates were ST2250, ST1223, ST2198, and ST2854. ST 2250 dominates the current global sequence collection, particularly Thailand, where it has been studied most extensively (Thaipadungpanit et al., 2015; Chantratita et al., 2016; Moradigaravand et al., 2017). However, this may not necessarily reflect its true prevalence in that region because of non-random sampling strategies. Given that *S. argenteus* is not readily differentiated from *S. aureus* by routine microbiological tests, its true global distribution remains unclear. Our study has shown distinct gene differences between the different clades, which may reflect different evolutionary pressures. Differing virulence and antibiotic resistance determinants, as well as distinct mobile genetic elements, demonstrate that this species, although only recently described, has considerable genetic diversity.

Although in general antibiotic resistance in *S. argenteus* appears less widespread than in *S. aureus*, we show here some important resistance determinants. We found that about 90% of

isolates contain the *fosB* gene, which can mediate resistance to the antibiotic fosfomycin (Fu et al., 2016a). However, phenotypic resistance to this antibiotic in *S. argenteus* strains from Thailand was reported as very low (Chantratita et al., 2016). Further studies are required to establish the relationship between possession of the *fosB* gene in *S. argenteus* and phenotypic resistance to fosfomycin. Susceptibility of *S. aureus* to fosfomycin is reported to be greater than 90% in 7 of 9 studies reviewed in Vardakas et al. (2016). Fosfomycin is an old antibiotic, having been discovered in 1969, with antimicrobial activity against a wide spectrum of both Gram-positive and Gram-negative organisms. Its use in humans is limited, although it has significant potential for use as a second-line agent in treating infections due to resistant organisms, particularly in urine (Michalopoulos et al., 2011). It is also used in veterinary medicine, particularly in treating infections in chickens and pigs (Perez et al., 2014). The *fosB* nucleotide sequences were highly similar with 99.4% average pairwise identity across all the *S. argenteus* isolates, suggesting recent acquisition. Complete genome sequences were only available for a few of the isolates to allow unequivocal localization, but where a complete assembly was available, the gene was present in the genome not a plasmid and was distinct from *fosB* genes found in plasmids of *S. aureus*. The widespread presence of the resistance determinant *tetL* in isolates of *S. argenteus* was previously noted in ST2250 isolates from Thailand, apparently carried on a plasmid together with heavy-metal resistance genes and thought to be imported from livestock associated strains of *S. aureus* (Moradigaravand et al., 2017). We did not find the *tetL* gene in other sequence types of *S. argenteus*, and although predominantly present in isolates from Thailand, it was also found in some ST2250 strains from other geographical areas, including Denmark, United States, Malaysia and Singapore. Given the bias toward strains from Thailand in the collection of sequences analyzed here, it is difficult to determine the origin of the *tetL* gene within *S. argenteus*. Geographical limitation of antibiotic resistance is, however, seen with the distribution of the *mecA* element, mediating methicillin resistance. This determinant was not found in ST2250 isolates from Thailand, but was present in ST2250 isolates from Denmark and the United States, as well as in ST2793. The element belonged to the SCCmecIV complex in all the strains where present. MRSA strains are common in pigs in Denmark; the *mec* element in these strains is typically SCCmecVc (Sieber et al., 2018). Whether the *mecA* element found in Danish strains of *S. argenteus* originated in strains of *S. aureus* within pigs will require further study of isolates from both humans and pigs. Taken together, however, the data contained in this report is certainly consistent with gene flow of antibiotic resistance elements from livestock into human strains.

The presence of different pathogenicity islands within specific STs suggests specific survival advantage of these islands within specific STs. Moreover, gene variation analysis shows the accumulation of different virulence factors in different clades over time. The analysis shows that evolution of ST1223 is separate from that of ST2250. For example, ST1223 carries the gene for a siderophore surface protein (*isdB*) which has been

lost during adaptation of clade ST2250. However, ST2250 has acquired prophages harboring PVL genes associated with skin and soft-tissue infections and increased morbidity. Movement of genomic islands in and out of the clades results in success of one clade over another. Novel pathogenicity islands identified in ST2250 and associated multiple phages may be the reason for the clade being successful in its spread compared to ST1223. A number of specific virulence genes encoding enterotoxins are found predominantly in ST1223, such as *sei*, *seg*, *selm*, *selu2*, and *set23*. These genes are typically clustered together in *S. aureus* as a cluster known as the *egc* locus, which has undergone significant rearrangement in different strains (Jarraud et al., 2001; Thomas et al., 2006). In the ST1223 strains analyzed here, we also found clustering of these genes as shown in **Figure 4A**. In their original description of a *S. argenteus* as a divergent CC75 clade, Holt et al. (2011) noted the presence of the vSa β island with an enterotoxin gene cluster similar to that found in non-CC75 *S. aureus* lineages, arguing that this locus is ancestral to the separation of *S. argenteus* from *S. aureus*. The similarity of this locus in ST1223 and related strains we describe here support this view, although it has been largely lost in other lineages, particularly in ST2250 strains. This suggests a specific selection pressure for retention in the ST1223 and related strains that reflects differing ecological niches – the apparent association of ST1223 strains with food poisoning (Wakabayashi et al., 2018) suggests that gastrointestinal spread of this strain may dominate compared to other mechanisms found in non-ST1223 strains. ST1223 has previously been noted to have a larger number of virulence genes (Aung et al., 2021). Further characterization of the disease associations of the different *S. argenteus* strains will clarify these differences.

A type I restriction modification system is widespread in the *S. argenteus* strains studied here. As is the case for *S. aureus*, a given ST of *S. argenteus* has essentially identical hsdS proteins which differ from other STs. This may limit the ability of specific STs to acquire genes by horizontal gene transfer from other STs or indeed strains of *S. aureus*. Despite this potential barrier to horizontal gene flow, *S. aureus* does demonstrate evidence of such transfer (Lindsay, 2010) so it is clearly not an absolute barrier, and the presence of different phages and SAPIs within *S. argenteus* suggest that is also the case in this species. We also found the presence of a CRISPR/Cas locus with captured phage and plasmid sequences found in virtually all of the ST2250 strains. These data strongly suggest that this CRISPR/Cas locus has been active in capturing DNA from relevant phage and plasmid sources and thus providing those strains with immunity from subsequent phage attack or from introduction of plasmids from other staphylococcal strains. The presence of this locus selectively in ST2250 strains suggests it confers a fitness advantage to this ST. Resisting phage lysis is clearly a selective advantage, but it is not clear why the ST2250 strains in particular have acquired the CRISPR/Cas locus and why other STs of *S. argenteus* and most *S. aureus* lack this system. Strains lacking the CRISPR/Cas locus do not have significantly more integrated phages, suggesting this system is not pivotal in resisting phage integration – the type I restriction modification system will also act to limit integration

of phages from different STs. Recent work has highlighted additional functions of the CRISPR/Cas locus in virulence (Louwen et al., 2014; Solbiati et al., 2020); whether such a role has been important in the acquisition of CRISPR/Cas loci in *S. argenteus* strains will require further experimental work.

In summary, the findings reported here show that although *S. argenteus* shares many features with *S. aureus* it also has some distinct features, and significant genomic variability between different STs. Increasingly, *S. argenteus* is recognized as an invasive human pathogen, much as *S. aureus*. Expert microbiological society opinion in Europe has proposed not distinguishing between *S. aureus* and *S. argenteus* for routine reporting purposes, although if laboratories are able to make the distinction, reporting should explicitly state that *S. argenteus* is a member of the *S. aureus* complex (Becker et al., 2019). A recent paper from North America stressed the importance of clinical laboratories differentiating the two species from a research and surveillance perspective, but with the same precautions on specific wording in laboratory reports to capture the close link between *S. argenteus* and *S. aureus* (Eshaghi et al., 2021). The data presented here shows that *S. argenteus* does have some genomic features that differ from *S. aureus*, and better correlations with specific clinical syndromes and outcomes would certainly require more robust differentiation between the two species in clinical practice. Future acquisition of whole genome sequences of *S. argenteus* will add to the analysis conducted here and improve our understanding of this organism.

DATA AVAILABILITY STATEMENT

The authors confirm all supporting data, code and protocols have been provided within the article or through

REFERENCES

- Abby, S. S., Néron, B., Ménager, H., Touchon, M., and Rocha, E. P. C. (2014). MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. *PLoS One* 9:e110726. doi: 10.1371/journal.pone.0110726
- Akobi, B., Aboderin, O., Sasaki, T., and Shittu, A. (2012). Characterization of *Staphylococcus aureus* isolates from faecal samples of the Straw-Coloured Fruit Bat (*Eidolon helvum*) in Obafemi Awolowo University (OAU), Nigeria. *BMC Microbiol.* 12:279. doi: 10.1186/1471-2180-12-279
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Argemi, X., Matelska, D., Ginalski, K., Riegel, P., Hansmann, Y., Bloom, J., et al. (2018). Comparative genomic analysis of *Staphylococcus lugdunensis* shows a closed pan-genome and multiple barriers to horizontal gene transfer. *BMC Genomics* 19:621. doi: 10.1186/s12864-018-4978-1
- Argudín, M. A., Dodémont, M., Vandendriessche, S., Rottiers, S., Tribes, C., Roisin, S., et al. (2016). Low occurrence of the new species *Staphylococcus argenteus* in a *Staphylococcus aureus* collection of human isolates from Belgium. *Europ. J. Clin. Microbiol. Infect. Dis.* 35, 1017–1022. doi: 10.1007/s10096-016-2632-x
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387
- Aswani, V., Najar, F., Pantrangi, M., Mau, B., Schwan, W. R., and Shukla, S. K. (2019). Virulence factor landscape of a *Staphylococcus aureus* sequence type 45 strain, MCRF184. *BMC Genomics* 20:123. doi: 10.1186/s12864-018-5394-2
- Aung, M., San, T., Aye, M., Mya, S., Maw, W., Zan, K., et al. (2017). Prevalence and Genetic Characteristics of *Staphylococcus aureus* and *Staphylococcus argenteus* Isolates Harboring Panton-Valentine Leukocidin, Enterotoxins, and TSST-1 Genes from Food Handlers in Myanmar. *Toxins* 9:241. doi: 10.3390/toxins9080241
- Aung, M. S., Urushibara, N., Kawaguchiya, M., Hirose, M., Ike, M., Ito, M., et al. (2021). Distribution of Virulence Factors and Resistance Determinants in Three Genotypes of *Staphylococcus argenteus* Clinical Isolates in Japan. *Pathogens* 10:163. doi: 10.3390/pathogens10020163
- Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A., et al. (2002). Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* 359, 1819–1827. doi: 10.1016/S0140-6736(02)08713-5
- Becker, K., Schaumburg, F., Kearns, A., Larsen, A. R., Lindsay, J. A., Skov, R. L., et al. (2019). Implications of identifying the recently defined members of the *Staphylococcus aureus* complex *S. argenteus* and *S. schweitzeri*: a position paper of members of the ESCMID Study Group for Staphylococci and Staphylococcal Diseases (ESGS). *Clin. Microbiol. Infect.* 25, 1064–1070. doi: 10.1016/j.cmi.2019.02.028
- Bokarewa, M. I., Jin, T., and Tarkowski, A. (2006). *Staphylococcus aureus*: Staphylokinase. *Int. J. Biochem. Cell Biol.* 38, 504–509. doi: 10.1016/j.biocel.2005.07.005

Supplementary Data Files. Accession numbers for all sequences reported here are contained in **Supplementary Table 1**.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

TE, CG, SF, AL, and MH: conceptualization. TE, CG, SF, and MH: methodology. CG, SF, and TE: analysis. All authors: writing, contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Scottish Executive via the Chief Scientists Office through the provision of a grant to establish the Scottish Healthcare Associated Infection Prevention Institute (SHAIPi). The funders had no role in the study design, collection, analysis or interpretation of data, writing of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.795173/full#supplementary-material>

- Cao, L., Gao, C.-H., Zhu, J., Zhao, L., Wu, Q., Li, M., et al. (2016). Identification and functional study of type III-A CRISPR-Cas systems in clinical isolates of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* 306, 686–696. doi: 10.1016/j.ijmm.2016.08.005
- Carattoli, A., and Hasman, H. (2019). PlasmidFinder and In Silico pMLST: Identification and Typing of Plasmid Replicons in Whole-Genome Sequencing (WGS). *Methods Mole. Biol.* 285–294. doi: 10.1007/978-1-4939-9877-7_20
- Carver, T., Harris, S. R., Berriman, M., Parkhill, J., and McQuillan, J. A. (2011). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28, 464–469. doi: 10.1093/bioinformatics/btr703
- Castañeda-García, A., Blázquez, J., and Rodríguez-Rojas, A. (2013). Molecular Mechanisms and Clinical Impact of Acquired and Intrinsic Fosfomycin Resistance. *Antibiotics* 2, 217–236. doi: 10.3390/antibiotics2020217
- Chantratita, N., Wikraiphat, C., Tandhavanant, S., Wongsuvan, G., Ariyaprasert, P., Suntornsut, P., et al. (2016). Comparison of Community-Onset *Staphylococcus argenteus* and *Staphylococcus aureus* Sepsis in Thailand: A Prospective Multicentre Observational Study. *Clin. Microbiol. Infect.* 22, 458–459 e411. doi: 10.1016/j.cmi.2016.01.008
- Chen, L. (2004). VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* 33, D325–D328. doi: 10.1093/nar/gki008
- Chen, S.-Y., Lee, H., Teng, S.-H., Wang, X.-M., Lee, T.-F., Huang, Y.-C., et al. (2018). Accurate differentiation of novel *Staphylococcus argenteus* from *Staphylococcus aureus* using MALDI-TOF MS. *Fut. Microbiol.* 13, 997–1006. doi: 10.2217/fmb-2018-0015
- Collery, M. M., Smyth, D. S., Tumilty, J. J. G., Twohig, J. M., and Smyth, C. J. (2009). Associations between enterotoxin gene cluster types egc1, egc2 and egc3, agr types, enterotoxin and enterotoxin-like gene profiles, and molecular typing characteristics of human nasal carriage and animal isolates of *Staphylococcus aureus*. *J. Med. Microbiol.* 58, 13–25. doi: 10.1099/jmm.0.005215-0
- Cooper, L. P., Roberts, G. A., White, J. H., Luyten, Y. A., Bower, E. K. M., Morgan, R. D., et al. (2017). DNA target recognition domains in the Type I restriction and modification systems of *Staphylococcus aureus*. *Nucleic Acids Res.* 45, 3395–3406. doi: 10.1093/nar/gkx067
- Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Neron, B., et al. (2018). CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* 46, W246–W251. doi: 10.1093/nar/gky425
- Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 43:e15. doi: 10.1093/nar/gku1196
- Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147
- Dupieux, C., Blondé, R., Bouchiat, C., Meugnier, H., Bes, M., Laurent, S., et al. (2015). Community-acquired infections due to *Staphylococcus argenteus* lineage isolates harbouring the Panton-Valentine leucocidin, France, 2014. *Eurosurveillance* 20:21154. doi: 10.2807/1560-7917.es2015.20.23.21154
- Eddy, S. R. (2004). Where did the BLOSUM62 alignment score matrix come from? *Nat. Biotechnol.* 22, 1035–1036. doi: 10.1038/nbt0804-1035
- Eshaghi, A., Bommersbach, C., Zittermann, S., Burnham, C.-A. D., Patel, R., Schuetz, A. N., et al. (2021). Phenotypic and Genomic Profiling of *Staphylococcus argenteus* in Canada and the United States and Recommendations for Clinical Result Reporting. *J. Clin. Microbiol.* 59, 2470–2420.e. doi: 10.1128/JCM.02470-20
- Fu, Z., Liu, Y., Chen, C., Guo, Y., Ma, Y., Yang, Y., et al. (2016a). Characterization of Fosfomycin Resistance Gene, fosB, in Methicillin-Resistant *Staphylococcus aureus* Isolates. *PLoS One* 11:e0154829. doi: 10.1371/journal.pone.0154829
- Fu, Z., Ma, Y., Chen, C., Guo, Y., Hu, F., Liu, Y., et al. (2016b). Prevalence of Fosfomycin Resistance and Mutations in murA, glpT, and uhpT in Methicillin-Resistant *Staphylococcus aureus* Strains Isolated from Blood and Cerebrospinal Fluid Samples. *Front. Microbiol.* 6:1544. doi: 10.3389/fmicb.2015.01544
- Goerke, C., Pantucek, R., Holtfreter, S., Schulte, B., Zink, M., Grumann, D., et al. (2009). Diversity of prophages in dominant *Staphylococcus aureus* clonal lineages. *J. Bacteriol.* 191, 3462–3468. doi: 10.1128/JB.01804-08
- Golding, G. R., Bryden, L., Levett, P. N., McDonald, R. R., Wong, A., Graham, M. R., et al. (2012). whole-genome sequence of livestock-associated st398 methicillin-resistant staphylococcus aureus Isolated from Humans in Canada. *J. Bacteriol.* 194, 6627–6628. doi: 10.1128/jb.01680-12
- Golding, G. R., Bryden, L., Levett, P. N., McDonald, R. R., Wong, A., Wylie, J., et al. (2010). Livestock-associated methicillin-resistant *Staphylococcus aureus* sequence type 398 in humans, Canada. *Emerg. Infect. Dis.* 16, 587–594. doi: 10.3201/eid1604.091435
- Gouy, M., Guindon, S., and Gascuel, O. (2009). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mole. Biol. Evol.* 27, 221–224. doi: 10.1093/molbev/msp259
- Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., et al. (2014). ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* 58, 212–220. doi: 10.1128/AAC.01310-13
- Hallböck, E. T., Karami, N., Adlerberth, I., Cardew, S., Ohlén, M., Engström Jakobsson, H., et al. (2018). Methicillin-resistant *Staphylococcus argenteus* misidentified as methicillin-resistant *Staphylococcus aureus* emerging in western Sweden. *J. Med. Microbiol.* 67, 968–971. doi: 10.1099/jmm.0.000760
- Hansen, T. A., Bartels, M. D., Høgh, S. V., Dons, L. E., Pedersen, M., Jensen, T. G., et al. (2017). Whole Genome Sequencing of Danish *Staphylococcus argenteus* Reveals a Genetically Diverse Collection with Clear Separation from *Staphylococcus aureus*. *Front. Microbiol.* 8:1512. doi: 10.3389/fmicb.2017.01512
- Henikoff, S., and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U.S.A.* 89:10915. doi: 10.1073/pnas.89.22.10915
- Holt, D. C., Holden, M. T. G., Tong, S. Y. C., Castillo-Ramirez, S., Clarke, L., Quail, M. A., et al. (2011). A Very Early-Branching *Staphylococcus aureus* Lineage Lacking the Carotenoid Pigment Staphyloxanthin. *Genome Biol. Evol.* 3, 881–895. doi: 10.1093/gbe/evr078
- Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2008a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi: 10.1093/nar/gkn923
- Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2008b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* 4, 44–57. doi: 10.1038/nprot.2008.211
- Jarraud, S., Peyrat, M. A., Lim, A., Tristan, A., Bes, M., Mougél, C., et al. (2001). egc, a highly prevalent operon of enterotoxin gene, forms a putative nursery of superantigens in *Staphylococcus aureus*. *J. Immunol.* 166, 669–677. doi: 10.4049/jimmunol.166.1.669
- Johansson, C., Rautelin, H., and Kaden, R. (2019). *Staphylococcus argenteus* and *Staphylococcus schweitzeri* are cytotoxic to human cells in vitro due to high expression of alpha-hemolysin Hla. *Virulence* 10, 502–510. doi: 10.1080/21505594.2019.1620062
- John, J., George, S., Nori, S. R. C., and Nelson-Sathi, S. (2019). Phylogenomic Analysis Reveals the Evolutionary Route of Resistant Genes in *Staphylococcus aureus*. *Genome Biol. Evol.* 11, 2917–2926. doi: 10.1093/gbe/evz213
- Jolley, K. A., and Maiden, M. C. (2010). BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinform.* 11:595. doi: 10.1186/1471-2105-11-595
- Katoh, K., Misawa, K., Kuma, K. I., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436
- Kitagawa, H., Ohge, H., Hisatsune, J., Masuda, K., Aziz, F., Hara, T., et al. (2020). Low incidence of *Staphylococcus argenteus* bacteremia in Hiroshima, Japan. *J. Infect. Chemother.* 26, 140–143. doi: 10.1016/j.jiac.2019.07.011
- Lee, J. Y. H., Carter, G. P., Pidot, S. J., Guérillot, R., Seemann, T., Gonçalves, et al. (2019). Mining the Methylome Reveals Extensive Diversity in *Staphylococcus epidermidis* Restriction Modification. *mBio* 10, e00277–11. doi: 10.1128/mBio.02451-19
- Lina, G., Bohach, G. A., Nair, S. P., Hiramatsu, K., Jouvin-Marche, E., and Mariuzza, R. (2004). Standard nomenclature for the superantigens expressed by *Staphylococcus aureus*. *J. Infect. Dis.* 189, 2334–2336. doi: 10.1086/420852
- Lindsay, J. A. (2010). Genomic variation and evolution of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* 300, 98–103. doi: 10.1016/j.ijmm.2009.08.013
- Lindsay, J. A., and Holden, M. T. (2004). *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol.* 12, 378–385. doi: 10.1016/j.tim.2004.06.004
- Lindsay, J. A., and Holden, M. T. (2006). Understanding the rise of the superbug: investigation of the evolution and genomic variation of *Staphylococcus aureus*. *Funct. Integr. Genomics* 6, 186–201. doi: 10.1007/s10142-005-0019-7

- Long, Q. X., Tang, X. J., Shi, Q. L., Li, Q., Deng, H. J., Yuan, J., et al. (2020). Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* 26, 1200–1204. doi: 10.1038/s41591-020-0965-6
- Louwen, R., Staals, R. H., Endtz, H. P., Van Baaren, P., and Van Der Oost, J. (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 78, 74–88. doi: 10.1128/MMBR.00039-13
- Maiden, M. C. J., Van Rensburg, M. J. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A., et al. (2013). MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat. Rev. Microbiol.* 11, 728–736. doi: 10.1038/nrmicro3093
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. doi: 10.1038/nrmicro2577
- Michalopoulos, A. S., Livaditis, I. G., and Gougoutas, V. (2011). The revival of fosfomycin. *Int. J. Infect. Dis.* 15, e732–e739. doi: 10.1016/j.ijid.2011.07.007
- Moradigaravand, D., Jamroz, D., Mostowy, R., Anderson, A., Nickerson, E. K., Thaipadungpanit, J., et al. (2017). Evolution of the *Staphylococcus argenteus* ST2250 Clone in Northeastern Thailand Is Linked with the Acquisition of Livestock-Associated Staphylococcal Genes. *mBio* 8, e00802–e00817. doi: 10.1128/mBio.00802-17
- Ng, J. W. S., Holt, D. C., Lilliebridge, R. A., Stephens, A. J., Huygens, F., Tong, S. Y. C., et al. (2009). Phylogenetically Distinct *Staphylococcus aureus* Lineage Prevalent among Indigenous Communities in Northern Australia. *J. Clin. Microbiol.* 47, 2295–2300. doi: 10.1128/jcm.00122-09
- Novick, R. P., Christie, G. E., and Penadés, J. R. (2010). The phage-related chromosomal islands of Gram-positive bacteria. *Nat. Rev. Microbiol.* 8, 541–551. doi: 10.1038/nrmicro2393
- Okuma, K., Iwakawa, K., Turnidge, J. D., Grubb, W. B., Bell, J. M., O'Brien, F. G., et al. (2002). Dissemination of New Methicillin-Resistant *Staphylococcus aureus* Clones in the Community. *J. Clin. Microbiol.* 40, 4289–4294. doi: 10.1128/jcm.40.11.4289-4294.2002
- Omoe, K., Hu, D.-L., Takahashi-Omoe, H., Nakane, A., and Shinagawa, K. (2005). Comprehensive analysis of classical and newly described staphylococcal superantigenic toxin genes in *Staphylococcus aureus* isolates. *FEMS Microbiol. Lett.* 246, 191–198. doi: 10.1016/j.femsle.2005.04.007
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421
- Perez, D. S., Tapia, M. O., and Soraci, A. L. (2014). Fosfomycin: Uses and potentialities in veterinary medicine. *Open Vet. J.* 4, 26–43.
- Rambaut, A., Lam, T. T., Carvalho, L. M., and Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* 2:vev007. doi: 10.1093/ve/vev007
- Rath, D., Amlinger, L., Rath, A., and Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie* 117, 119–128. doi: 10.1016/j.biochi.2015.03.025
- Rigall, J., Grattard, F., Grange, S., Forest, F., Haddad, E., Carricajo, A., et al. (2018). Community-Acquired *Staphylococcus argenteus* Sequence Type 2250 Bone and Joint Infection, France, 2017. *Emerg. Infect. Dis.* 24, 1958–1961. doi: 10.3201/eid2410.180727
- Sato'o, Y., Omoe, K., Ono, H. K., Nakane, A., and Hu, D. L. (2013). A novel comprehensive analysis method for *Staphylococcus aureus* pathogenicity islands. *Microbiol. Immunol.* 57, 91–99. doi: 10.1111/1348-0421.12007
- Schuster, D., Rickmeyer, J., Gajdiss, M., Thye, T., Lorenzen, S., Reif, M., et al. (2017). Differentiation of *Staphylococcus argenteus* (formerly: *Staphylococcus aureus* clonal complex 75) by mass spectrometry from *S. aureus* using the first strain isolated from a wild African great ape. *Int. J. Med. Microbiol.* 307, 57–63. doi: 10.1016/j.ijmm.2016.11.003
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shearer, J. E. S., Wireman, J., Hostetler, J., Forberger, H., Borman, J., Gill, J., et al. (2011). Major families of multiresistant plasmids from geographically and epidemiologically diverse staphylococci. *G3* 1, 581–591. doi: 10.1534/g3.111.000760
- Sheppard, S. K., Guttman, D. S., and Fitzgerald, J. R. (2018). Population genomics of bacterial host adaptation. *Nat. Rev. Genetics* 19, 549–565. doi: 10.1038/s41576-018-0032-z
- Sieber, R. N., Skov, R. L., Nielsen, J., Schulz, J., Price, L. B., Aarestrup, F. M., et al. (2018). Drivers and Dynamics of Methicillin-Resistant Livestock-Associated *Staphylococcus aureus* CC398 in Pigs and Humans in Denmark. *mBio* 9, e02142–18. doi: 10.1128/mBio.02142-18
- Sievers, F., and Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Prot. Sci.* 27, 135–145. doi: 10.1002/pro.3290
- Solbiati, J., Duran-Pinedo, A., Godoy Rocha, F., Gibson, F. C. III, and Frias-Lopez, J. (2020). Virulence of the Pathogen *Porphyromonas gingivalis* Is Controlled by the CRISPR-Cas Protein Cas3. *mSystems* 5, e00852–20. doi: 10.1128/mSystems.00852-20
- Stamatakis, A. (2006). RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446
- Stamatakis, A. (2015). Using RAXML to Infer Phylogenies. *Curr. Prot. Bioinform.* 51, 6.14.11–16.14.14. doi: 10.1002/0471250953.bi0614s51
- Subedi, A., Ubeda, C., Adhikari, R. P., Penadés, J. R., and Novick, R. P. (2007). Sequence analysis reveals genetic exchanges and intraspecific spread of SaPI2, a pathogenicity island involved in menstrual toxic shock. *Microbiology* 153, 3235–3245. doi: 10.1099/mic.0.2007/006932-0
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- Suzuki, Y., Kubota, H., Ono, H. K., Kobayashi, M., Murauchi, K., Kato, R., et al. (2017). Food poisoning outbreak in Tokyo, Japan caused by *Staphylococcus argenteus*. *Int. J. Food Microbiol.* 262, 31–37. doi: 10.1016/j.ijfoodmicro.2017.09.005
- Suzuki, Y., Kubota, H., Sato'o, Y., Ono, H. K., Kato, R., Sadamasu, K., et al. (2015). Identification and characterization of novel *Staphylococcus aureus* pathogenicity islands encoding staphylococcal enterotoxins originating from staphylococcal food poisoning isolates. *J. Appl. Microbiol.* 118, 1507–1520. doi: 10.1111/jam.12786
- Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008). Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477. doi: 10.1016/j.mib.2008.09.006
- Thaipadungpanit, J., Amornchai, P., Nickerson, E. K., Wongsuvan, G., Wuthiekanun, V., Limmathurotsakul, D., et al. (2015). Clinical and Molecular Epidemiology of *Staphylococcus argenteus* Infections in Thailand. *J. Clin. Microbiol.* 53, 1005–1008. doi: 10.1128/JCM.03049-14
- Thomas, D. Y., Jarraud, S., Lemerrier, B., Cozon, G., Echasserieu, K., Etienne, J., et al. (2006). Staphylococcal enterotoxin-like toxins U2 and V, two new staphylococcal superantigens arising from recombination within the enterotoxin gene cluster. *Infect. Immun.* 74, 4724–4734. doi: 10.1128/IAI.00132-06
- Tong, S. Y. C., Schaumburg, F., Ellington, M. J., Corander, J., Pichon, B., Leendertz, F., et al. (2015). Novel staphylococcal species that form part of a *Staphylococcus aureus*-related complex: the non-pigmented *Staphylococcus argenteus* sp. nov. and the non-human primate-associated *Staphylococcus schweitzeri* sp. nov. *Int. J. Systemat. Evol. Microbiol.* 65, 15–22. doi: 10.1099/ijso.062752-0
- Treangen, T. J., Ondov, B. D., Koren, S., and Phillippy, A. M. (2014). The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol.* 15, 1–15. doi: 10.1186/s13059-014-0524-x
- Vardakas, K. Z., Legakis, N. J., Triarides, N., and Falagas, M. E. (2016). Susceptibility of contemporary isolates to fosfomycin: a systematic review of the literature. *Int. J. Antimicrob. Agents* 47, 269–285. doi: 10.1016/j.ijantimicag.2016.02.001
- Wakabayashi, Y., Umeda, K., Yonogi, S., Nakamura, H., Yamamoto, K., Kumeda, Y., et al. (2018). Staphylococcal food poisoning caused by *Staphylococcus argenteus* harboring staphylococcal enterotoxin genes. *Int. J. Food Microbiol.* 265, 23–29. doi: 10.1016/j.ijfoodmicro.2017.10.022
- Waldron, D. E., and Lindsay, J. A. (2006). Saul: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into *Staphylococcus aureus* and between *S. aureus* isolates of different lineages. *J. Bacteriol.* 188, 5578–5585. doi: 10.1128/JB.00418-06
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
- Xu, W., Chen, T., Wang, H., Zeng, W., Wu, Q., Yu, K., et al. (2020). Molecular Mechanisms and Epidemiology of Fosfomycin Resistance in *Staphylococcus*

- aureus Isolated From Patients at a Teaching Hospital in China. *Front. Microbiol.* 11:290. doi: 10.3389/fmicb.2020.01290
- Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36. doi: 10.1111/2041-210X.12628
- Zhang, D.-F., Xu, X., Song, Q., Bai, Y., Zhang, Y., Song, M., et al. (2016). Identification of *Staphylococcus argenteus* in Eastern China based on a nonribosomal peptide synthetase (NRPS) gene. *Fut. Microbiol.* 11, 1113–1121. doi: 10.2217/fmb-2016-0017
- Zhang, J., Suo, Y., Zhang, D., Jin, F., Zhao, H., and Shi, C. (2018). Genetic and Virulent Difference Between Pigmented and Non-pigmented *Staphylococcus aureus*. *Front. Microbiol.* 9:598. doi: 10.3389/fmicb.2018.00598
- Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011). PHAST: A Fast Phage Search Tool. *Nucleic Acids Res.* 39, W347–W352. doi: 10.1093/nar/gkr485
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2021 Goswami, Fox, Holden, Leanord and Evans. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.