Genome-wide comparison of four MRSA clinical isolates from Germany and Hungary

Romen Singh Naorem[1], Jochen Blom[2] and Csaba Fekete[1]

[1] Department of General and Environmental Microbiology, University of Pécs, Pécs, Hungary
[2] Bioinformatics & Systems Biology, Justus-Liebig-Universität Gießen, Gießen, Germany

ABSTRACT

Staphylococcus aureus is a drug-resistant pathogen, capable of colonizing diverse ecological niches and causing a broad spectrum of infections related to a community and healthcare. In this study, we choose four methicillin-resistant *S. aureus* (MRSA) clinical isolates from Germany and Hungary based on our previous polyphasic characterization finding. We assumed that the selected strains have a different genetic background in terms of the presence of resistance and virulence genes, prophages, plasmids, and secondary metabolite biosynthesis genes that may play a crucial role in niche adaptation and pathogenesis. To clarify these assumptions, we performed a comparative genome analysis of these strains and observed many differences in their genomic compositions. The Hungarian isolates (SA H27 and SA H32) with ST22-*SCCmec* type IVa have fewer genes for multiple-drug resistance, virulence, and prophages reported in Germany isolates. Germany isolate, SA G6 acquires aminoglycoside (*ant(6)-Ia and aph(3')-III*) and nucleoside (*sat-4*) resistance genes via phage transduction and may determine its pathogenic potential. The comparative genome study allowed the segregation of isolates of geographical origin and differentiation of the clinical isolates from the commensal isolates. This study suggested that Germany and Hungarian isolates are genetically diverse and showing variation among them due to the gain or loss of mobile genetic elements (MGEs). An interesting finding is the addition of SA G6 genome responsible for the drastic decline of the core/pan-genome ratio curve and causing the pan-genome to open wider. Functional characterizations revealed that *S. aureus* isolates survival are maintained by the amino acids catabolism and favor adaptation to growing in a protein-rich medium. The dispersible and singleton genes content of *S. aureus* genomes allows us to understand the genetic variation among the CC5 and CC22 groups. The strains with the same genetic background were clustered together, which suggests that these strains are highly alike; however, comparative genome analysis exposed that the acquisition of phage elements, and plasmids through the events of MGEs transfer contribute to differences in their phenotypic characters. This comparative genome analysis would improve the knowledge about the pathogenic *S. aureus* strain's characterization, and responsible for clinically important phenotypic differences among the *S. aureus* strains.

INTRODUCTION

Staphylococcus aureus is a notorious nosocomial, and community-acquired pathogen (*Chambers & Deleo, 2009*). It has the capability of colonizing diverse ecological niches within its human host, including the skin, blood, respiratory tract, and nasal passages (*Deleo, Diep & Otto, 2009*) and causing diverse ranges of the hospital and community-acquired infections such as skin and soft tissue infections (SSI) for example, carbuncles, abscesses, styes, and impetigo and life-threatening infections such as bacteremia, necrotizing pneumonia, osteomyelitis, endocarditis, and sepsis (*Götz, Bannerman & Schleifer, 2006*; *Mottola et al., 2016*). Methicillin-resistant *S. aureus* (MRSA) acquired a mobile genetic element called Staphylococcal chromosomal cassette *mec* (*SCCmec)* accompanied by methicillin resistance gene (*mecA*) (*Zhang, Mcclure & Conly, 2012*). The β-lactam insensitive protein, penicillin-binding protein (PBP2a) encoded by *mecA* gene reduces affinity to penicillin and β-lactam antibiotics including methicillin, oxacillin, cefoxitin, *etc.,* and develop resistance toward the β-lactam antibiotics (*Jansen et al., 2006*; *Mistry et al., 2016*). MRSA acquires an arsenal of antibiotic resistance genes (ARGs) and virulence factor encoding genes (VFGs) through horizontal gene transfer (HGT) and recombination (*Chan, Beiko & Ragan, 2011*; *Hughes & Friedman, 2005*).

MRSA can anchor and colonize on epithelial surfaces and produce biofilm (*Goudarzi et al., 2018*). The biofilm produced by MRSA strains encase its cells in the exopolysaccharide matrix reduces the activity of antibacterial agents and immune defense. The dispersal of bacterial cells from the biofilm can result in secondary site infections and leading infections worsen and difficult to eradicate (*Lister & Horswill, 2014*). Biofilm formation is a complex process that consists of an extracellular polymeric matrix (ECM) formation involving polysaccharide intercellular adhesin (PIA), protein-protein interactions, and the incorporation of extracellular DNA (eDNA) (*O'Gara, 2007*; *Payne & Boles, 2015*). The biofilm formation is determined by the *icaADBC* gene cluster, responsible for PIA and capsular polysaccharide/adhesion synthesis (*Chaieb, Mahdouani & Bakhrouf, 2005*). MRSA possesses adhesive matrix molecules that are encoded by elastin (*ebps*), laminin (*eno*), clumping factors A and B (*clfA* and *clfB*), fibronectin A and B (*fnbA* and *fnbB*), collagen (*cna*), fibrinogen (*fib*), bone sialoprotein (*bbp*), *etc* (*Lindsay et al., 2006*). These molecules are exported to the bacterial cell surface to enable adherence with host tissues, leading to play a role in pathogenesis (*Mazmanian, 1999*).

S. aureus acquires an arsenal of ARGs and VFGs that are subjected to HGT and recombination (*Chan, Beiko & Ragan, 2011*; *Hughes & Friedman, 2005*). Hospital-associated MRSA (HA-MRSA) is often associated with metastatic infections and significant morbidity and mortality (*Gould, 2005*). However, Community-associated MRSA (CA-MRSA) infections have seen a high increase in prevalence, posing a greater threat to the public (*Morens & Fauci, 2013*). The genomic plasticity of *S. aureus* has facilitated the development of hypervirulent and drug-resistant strains, result in challenging issues to antibiotic treatment and health concern.

The classical techniques such as antibiotic susceptibility test (AST) patterns and molecular typing methods such as *SCCmec*-typing, Pulse-Field Gel Electrophoresis (PFGE),

Multi-Locus Sequence Typing (MLST), Multi-locus variable-number tandem-repeat (VNTR) analysis (MLVA), *S. aureus* protein A (*spa*)-typing, accessory gene regulator (*agr*)-typing are widely used to detect and differentiate several MRSA strains, and helpful for identifying the risk factors associated with MRSA infection which support the establishment of adequate infection control programs (*Zhang, Mcclure & Conly, 2012*; *Mistry et al., 2016*). However, these methods are expensive and time-consuming, and have limitations in infection control and investigating the nosocomial transmission due to low resolution (*Du et al., 2011*). In this modern era, whole-genome sequence-based typing offers an excellent resolution in global and local epidemiologic investigations of pathogen outbreaks and offers further data mining activities essentially for ARGs and VFGs profiling (*Köser et al., 2012*). So, the Next Generation Sequencer (NGS) based-genome sequencing technique has become an essential tool in the clinical microbiology arenas for comparative genomic analysis of several other species of the *Staphylococcus* genus in terms of the niche adaptation, combat antibiotics, and emergence of new virulent strains in real-time.

In our preliminary study, the polyphasic characterization of 35 *S. aureus* strains originated from Germany, and Hungary was performed. This characterization included antibiotic resistance test (ART), biochemical tests, biofilm-forming assay, and PCR based typing techniques involving the amplification of *mecA*, *pvl*, *SCCmec*-type, *spa* type, *coa-HaeIII*-RFLP, and biofilm-associated genes. Principal component analysis from polyphasic characterization data showed that the strains originated from the same geographical region were found in the close group while SA G8, Germany strain was grouped with other Hungarian strains (*Naorem et al., 2020*). The Hungarian strains (SA H27 and SA H32) belonged to the same Clonal Complex (ST22/*SCCmec*-IV) were clustered in the same group; however, these strains were isolated from the different sites of infections (nostrils and trachea) and showed different antibiotic resistance patterns and biofilm-forming abilities. Similarly, the strains collected from Germany *viz.,* SA G6, and SA G8 belonged to the same Clonal Complex (ST228/*SCCmec*-I and ST225/ *SCCmec*-II) having similar antibiotic resistance pattern, and biofilm-forming profiles, but these strains were isolated from the different site of infections (skin and other body sites) and not clustered in the same group (*Naorem et al., 2020*). Based on this information, these four *S. aureus* strains were chosen for in-depth comparative genome levels study to better understand the genomic differences among the strains. We assumed that the selected strains have a different genetic background in terms of the presence of ARGs, VFGs, prophages, plasmids, and secondary metabolite biosynthesis genes that may play a crucial role in niche adaptation and pathogenesis. To clarify these assumptions, we performed a comparative genome analysis of these four strains and observed many differences in their genomic compositions.

MATERIALS & METHODS

Bacterial strains used in this study

In this study, four *S. aureus* isolates collected from Germany (SA G6, and SA G8) and Hungarian (SA H27, and SA H32) were used. Hungarian isolate, SA H27 was reported as a strong biofilm producer among them (*Naorem et al., 2020*).

pH tolerance assay

S. aureus strains were cultured overnight at 37 °C in tryptic soy broth (TSB) (DB, Germany). The cell density (colony forming units, CFU) was adjusted to a final concentration of ∼10^6 CFU/ml in pH 4.5 TSB and pH 9.5 TSB. Cell suspension (200 µl) were loaded into the 96-well flat-bottomed polystyrene microtiter plate (Costar 3599; Corning; USA). The plates were incubated at 37 °C for 24 h without shaking, then the growth was measured at 492 nm wavelength using a Multiskan Ex microtiter plate reader (Thermo Electron Corporation, USA). The experiments were performed in triplicate and analyzed using GraphPad Prism 6 software package (Graphpad Software Inc, San Diego, CA, USA).

Genomic DNA isolation and sequencing

The genomic DNA was extracted using the GenElute^{TM} Bacterial Genomic DNA Kit (Sigma, USA) following the manufacturer instructions. The concentration and purity of genomic DNA was measured using dsDNA HS (High Sensitivity) Assay Kit in Qubit 3.0 fluorometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) and subsequently DNA quality was visualized by agarose gel electrophoresis.

Genomic libraries were prepared by using the NEB Next Fast DNA Fragmentation and Library Preparation Kit, developed for Ion Torrent (New England Biolabs) and used according to 200 bp protocol. After chemical fragmentation, DNA size selection was performed on precast 2% E-Gel Size Select Gel (Thermo Fisher Scientific Inc., Waltham, MA, USA). The quality of the libraries was verified using Agilent high sensitivity DNA assay kit (Agilent Technologies Inc., Santa Clara, CA, USA) in Agilent 2100 Bioanalyzer System (Agilent Technologies Inc., Santa Clara, CA, USA). For the template preparation, Ion PGM Hi-Q View OT2 Kit was used (Thermo Fisher Scientific Inc., Waltham, MA, USA). The template positive beads were loaded on Ion 316v2 Chip and sequenced using Ion PGM Hi-Q View Sequencing Kit on Ion Torrent PGM sequencer (Thermo Fisher Scientific Inc., Waltham, MA, USA).

Genome assembly and annotation

In-silico trimming of adapter and barcode sequences and data analysis were performed using Torrent Suite 5.4.0 (Thermo Fisher Scientific Inc., Waltham, MA, USA) and the trimmed paired-end reads were assembled by de novo assembler SPAdes 3.7.1 software with 21, 33, 55, 77, 99, 127 k-mer values (*Nurk et al., 2013*). The assembly-stats and quality of genome completeness for each strain were estimated using the web platform QUEST (*Gurevich et al., 2013*). For identifying the closely related strains, the genome assemblies were analyzed by the kmerFinder 3.1 web platform (*Larsen et al., 2014*). The genome assembly was aligned against the reference genome for the contigs rearrangement using the 'Move Contigs' algorithm in Mauve 2.4.0 (*Darling, Mau & Perna, 2010*) and further, scaffolds were generated with reference genome/ genome of closely related strains predicted by kmerFinder 2.0 as a guide for alignment using the reference-based scaffolder MeDuSa (*Bosi et al., 2015*). Gene annotation of the genome assemblies was performed via the fully automated RAST (Rapid Annotation using Subsystem Technology) (*Aziz et al., 2008*) and PATRIC 3.5.7 (Pathosystems Resource Integration Center) (*Wattam et al., 2013*) pipelines using the reference genome.

In-silico characterization of genome assemblies

In-silico epidemiologic characterization of genome assemblies was performed using SCCmecFinder-1.2 for the identification of *SCCmec* types (*Kaya et al., 2018*), *spa* Typer 1.0 (*Bartels et al., 2014*) for *spa* type, and MLST 1.8 (*Larsen et al., 2012*) for Multilocus Sequence Type in a web-based server provided by the Center for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/). In-silico *arg* (accessory gene regulator)-typing was performed using the primers described by *Shopsin et al. (2003)* in in-silico PCR amplification tools (*Bikandi et al., 2004*).

The genome assemblies were screened for plasmid replicon (*rep*) genes using PlasmidFinder 2.1 (*Carattoli et al., 2014*) with default parameters. The identified nonaligned contig or scaffold associated with plasmid sequences were extracted and used for the identification of full-length plasmid regions using PLSDB (Plasmid Database) version-2020-03-04 (*Galata et al., 2018*) with search strategy Mash screen, and the default values were a maximum *P*-value of 0.1 and a minimum identity of 0.99 (https://ccb-microbe.cs.uni-saarland.de/plsdb/). Identified plasmids were compared with the closest reference plasmids using Easyfig version 2.2.3 (*Sullivan, Petty & Beatson, 2011*). The identification and annotation of prophage sequences were performed by screening the genome assemblies using PHASTER (PHAge Search Tool Enhanced Release) (*Arndt et al., 2016*), and identified template phages were classified for their lifestyles using PHACTS (Phage Classification Tool Set) (*McNair, Bailey & Edwards, 2012*).

In-silico mining of candidate ARGs and VFGs were performed using CARD (Comprehensive Antibiotic Resistance Database) version 3.0.8 in RGI (Resistance Gene Identifier) version 5.1.0 platform (https://card.mcmaster.ca/analyze/rgi) (*Alcock et al., 2019*), and a comprehensive set of *S. aureus* VFGs was analyzed using VFDB (Virulence Factor Database) in VFanalyzer (*Liu et al., 2018*) and the PATRIC tool version 3.6.3 (https://www.patricbrc.org/) (*Wattam et al., 2013*). Further, heatmap and hierarchical clustering were generated to visualize the presence and absence of VFGS and ARGs in *S. aureus* strains using a web-based application, Morpheus, (https://software.broadinstitute.org/morpheus). Secondary metabolite biosynthesis gene clusters and the detection of genes encoding bacteriocins were analyzed using antiSMASH 5.0 (*Blin et al., 2019*) and BAGEL4 (*Van Heel et al., 2018*). The prediction of chromosomal genomic islands was predicted by using IslandViewer 4 (*Bertelli et al., 2017*).

Comparative genomic analysis

The ANI was determined based on BLAST+ using the JSpeciesWS webserver (*Richter et al., 2015*). The pairwise comparisons between the genomes of *S. aureus* isolates and their nearest reference genomes were conducted using GBDP (Genome BLAST Distance Phylogeny) under the algorithm trimming and distance formula d5, and calculated each distance with 100 replicates (*Meier-Kolthoff et al., 2013*). dDDH (Digital DNA-DNA Hybridization) values and confidence intervals were calculated using the recommended settings of the GGDC 2.1 (*Meier-Kolthoff et al., 2013*).

Genomes of *S. aureus* isolates and their reference strains were compared with CGViewer (Circular Genome Viewer) server (*Grant & Stothard, 2008*). The functional annotation was

performed using EggNOG (Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups) mapper 5.0 database (*Huerta-Cepas et al., 2018*) and RAST server-based SEED viewer (*Overbeek et al., 2013*).

The pan-genome, core-genome, and singletons were calculated using four study genomes of *S. aureus* isolates in EDGAR version 2.0 software framework (*Blom et al., 2016*). This pan-genome analysis was extended using four study genomes coupled with three reference genomes of *S. aureus* strains. The core-genome was analyzed in the genomes set using reciprocal best BLAST hits of all CDS using EDGAR version 2.0 software framework (*Blom et al., 2016*). The singletons were calculated for the contig of a strain by comparing to the CDS of a set of contigs in EDGAR. The CDS that has no match with SRV (Score Ratio Value Plots) higher or equal the master cut-off in any of the contigs were considered as singletons. The development of pan-genome and core-genome sizes was analyzed using the core/pan development feature and as well, the pan *vs.* core development plot was generated in EDGAR. Heap's Law function was applied to calculate whether the pan-genome open or closed using the equation $n = k * N^{(-\alpha)}$ where n = expected a number of genes; N = number of genomes; k and α ($\alpha = 1 - \gamma$) are proportionality constant and exponent, respectively (*Tettelin et al., 2008*). Heap's law predicted that closed pan-genome (when α > 1 (γ < 0)), and open pan-genome (when α < 1 (0 < γ < 1)). According to *Tettelin et al. (2008)*, core-genome and singletons developments were calculated by the least-square fitting of exponential decay functions.

The Rcp (ratio of core-genome to that of pan-genome) was calculated (*Ghatak et al., 2016*). Then, genomic subsets, including the number of core-genome and singletons in the gene pool, were extracted, and flowerplot was drawn using *in-house* R scripts.

Phylogenetic analysis

The genome assemblies of the isolates were used for a whole genome-based phylogeny analysis using TYGS (Type/Strain Genome Server) (*Meier-Kolthoff & Göker, 2019*) engaging with genomes of closely related strains of *S. aureus*. The phylogenomic trees were reconstructed using FastME 2.1.6.1 (*Lefort, Desper & Gascuel, 2015*) from the GBDP (Genome BLAST Distance Phylogeny) distances calculated from genome sequences under the algorithm 'coverage' and distance formula d5 (*Meier-Kolthoff et al., 2013*). The trees were rooted at the midpoint (*Farris, 1972*); branch supports were inferred from 100 pseudo-bootstrap replicates and visualized with Interative Tool Of Life v4 (iTOL) (*Letunic & Bork, 2019*). The core SNPs of genome sequences were extracted using Panseq (*Laing et al., 2010*) and the phylogenetic tree was constructed using PhyML+SMS module in NGPhylogeny.fr (*Lemoine et al., 2019*) to select the best evolutionary model, further the tree was annotated in Interative Tool Of Life v4 (iTOL) (*Letunic & Bork, 2019*).

RESULTS

The *S. aureus* isolates could survive at pH 4.5 through pH 9.5 conditions with a survival rate of ∼45%–84%. SA G8 isolate showed the highest cell survival rate of 84.4% at acidic pH but its cell survival rate drops down by 7% when subjected to alkaline pH conditions (Table S1).

General genomic features of *S. aureus* isolates

The genomic DNA of *S. aureus* isolates was successfully sequenced in the IonTorrent PGM sequencing platform. The average raw reads obtained from the genome sequencing of SA G6, SA G8, SA H27, and SA H32 are ∼88.9, 69.6, 128.3, and 92.7 million bases (Mb) for genomes of SA G6, SA G8, SA H27, and SA H32 strains respectively. The closely related strains identified by kmerFinder 2.0 were *S. aureus subsp. aureus* ST228 (HE579073), *S. aureus subsp. aureus* JH9 (CP000703) for SA G6 and SA G8 strains, respectively. Also, *S. aureus subsp. aureus* HO 5096 0412 (HE681097.1) was identified closely related strains for SA H27 and SA H32 strains. Among the *S. aureus* isolates, SA G8 has the largest genome size (28633393 bp) with high % GC content (32.81%). The numbers of protein-coding sequences (CDSs) in the *S. aureus* strains varied from 2630 (SA H27) to 2743 (SA G8). The comparison of draft genome assemblies, genome annotation, molecular typing, plasmid, and prophage features for *S. aureus* genomes were summarized in Table 1.

Genes encoding plasmids

The putative plasmids were detected in nonaligned contigs or scaffolds that exhibited an unexpected high coverage level after the genome assemblies. A putative plasmid (p1G6) of 13331bp length was identified at Scaffold 4 of the SA G6 genome consisting of the replication gene (*repA*). The p1G6 plasmid has 30.97% sequence coverage with plasmids pTW20_1 (FN433597.1) (Fig. 1A). The sequence coverage region of p1G6 with pTW20_1 constitutes the genes that encode for proteins such as IS6 family transposase, replication-associated protein (Rep), cadmium resistance transporter (CadD), cadmium efflux system accessory protein (CadX), replication initiation protein A (RepA), quaternary ammonium compound efflux MFS transporter (QacA), multidrug-binding transcriptional regulator (QacR), DUF536 domain-containing protein (mP), AAA family ATPase (Abp), hypothetical proteins, HAD hydrolase family protein, and IS257 family transposase. The SA H32 genome also consists of a putative plasmid (p2H32) having a length of 2530 bp located at Scaffold 3 and showed 71.32% sequence coverage with plasmids AR_0472 (NZ_CP029648.1). It consists of a replication gene (*repL*) and carried an erythromycin resistance gene (*emrC*) (Fig. 1B). The identified plasmids of *S. aureus* encode no other factors for their transfer, such plasmids may transfer via phage transduction (*McCarthy & Lindsay, 2012*). The linear graphical map of plasmid comparison was represented in Fig. 1.

Characteristic of prophages-like elements

The genomes of *S. aureus* isolates have several prophages and phage-like element regions and these prophages were belonged to the *Siphoviridae* family and having temperate lifestyles. The highest number of prophage regions was found in the genome of SA G8 isolate including three intact prophages (phiG8.2, phiG8.3, and phiG8.4), a questionable (phiG8.1), and an incomplete (phiG8.5) prophages. Four prophage regions were found in the genome of SA G6 isolate including an intact prophage (phiG6.3), two questionable prophages (phiG6.1 and phiG6.4), and an incomplete prophage (phiG6.2). The genome of SA H27 isolate harbor three intact prophages (phiH27.1, phiH27.2, and phiH27.3) while

Table 1 General genomic features of *S. aureus* genomes in this study.

Strains	SA G6	SA G8	SA H27	SA H32
Size (bp)	2856214	2857863	2783185	2786627
Contigs	103	83	44	63
Scaffolds	22	15	1	3
N50 (bp)	125160	263953	328241	208577
GC%	32.79	32.81	32.73	32.72
CDS	2734	2743	2630	2657
Genes assigned to SEED	2101	2169	2014	2036
rRNA	9	10	8	9
tRNA	61	60	57	60
Prophage Regions	3	5	3	1
[a]Plasmids	p1G6	-	-	p2H32
SCCmec type	I	II	IVa	IVa
MLST	ST228	ST225	ST22	ST22
Spa type	t535	t003	t379	t1258
agr-type	II	II	I	I
Accession no.	RAHA00000000	QZFC00000000	CP032161	RAHP00000000

Notes.
[a]Plasmids: The presence of plasmid in genome is indicated by plasmid name, while absent is represented by a minus (-) sign.

the genome of SA H32 harbor only one intact prophage (phiH32.1). The *lukF-PV* and *lukM* genes (Bicomponent leukotoxins), and *plc* gene (Phospholipase C) were identified in the prophages of phiG6.4, phiG8.4, phiH27.2, and phiH32.1. The prophages of phiG6.3, phiG8.4, and phiH27.2 carried *sak* gene (staphylokinase) and *scn* gene (staphylococcal complement inhibitor). Chemotaxis inhibitory protein encoded by *chp* gene was identified in phiG8.4 and phiH27.2 prophages. Enterotoxin A encoded by *sea* gene was harbored by the prophages of phiG6.3 and phiG8.4. Hemolysin genes such as *hlb* (β-hemolysin), and *hlgB* (-hemolysin B) were found in the prophages of phiH27.2, and phiH32.1. In addition to virulence factors, phiG6.4 prophage carried ARGs genes that conferred resistance to beta-lactamase (*blaZ*), aminoglycoside (*ant(6)-Ia* and *aph(3′)-III*) and nucleoside (*sat-4*) antibiotics. The comparative analysis of VFGs associated with putative prophages was summarized in Table S2.

In-silico analysis of antimicrobial resistance and associated genes in the genomes

Four study genomes of *S. aureus* isolates shared 63.3% (19/30) of antibiotic resistance and associated genes (Fig. 2). The shared genes comprise methicillin-resistant PBP2a (*mecA* and *mecR1*); multidrug resistance efflux (*ygaD*); fluoroquinolone (*norA* and *gyrA*); fluoroquinolone and acridine dye (*arlS* and *arlR*); glycylcycline (*mepA*); tetracycline (*tet-38*); tetracycline, penam, cephalosporin, glycylcycline, rifamycin, phenicol, triclosan, fluoroquinolone (*mgrA* and *marR*); lipopeptide (*pgsA*, *clsA* and *rpoC*); rifampicin (*rpoB*); aminocoumarin (*gyrB* and *parE*); dihydrofolate reductase (*dfrA/folA*) and defensin (*mprF/fmtC*, multiple peptide resistance factor) that play roles in resistance mechanism including antibiotic efflux, antibiotic target alteration, and antibiotic target replacement.
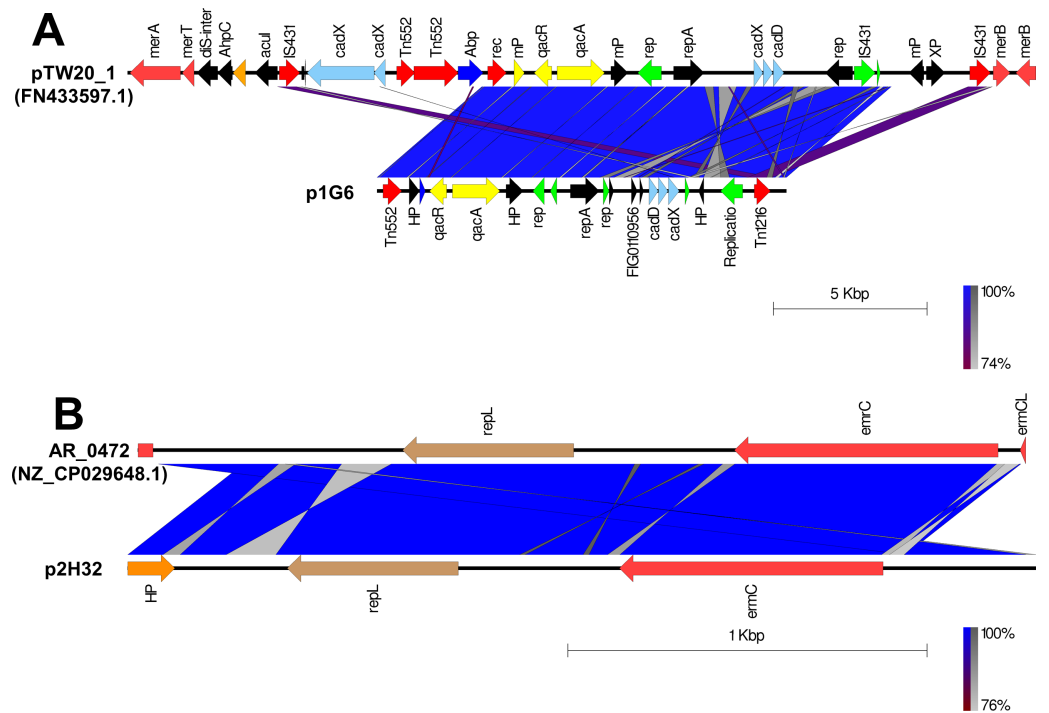
Figure 1 Comparison of linear plasmid maps by Easyfig alignment. Coding Sequences are represented by colored arrows. Blue lines between the plasmids indicate the shared similarity regions according to BLASTn identity. CDS are characterized by functions as follows: antiseptic resistance genes (yellow), erythromycin resistance gene (light red), DNA replication (green), transposons/integrases (red), replication A gene/ hypothetical proteins/others (black), replication L gene (brown), and cadmium resistance gene (cyan). The outer scale is marked in kilobases. (A) Sequence alignment of p1G6 plasmid of SA G6 isolates with the reference pTW20_1 (FN433597.1) plasmid. (B) Sequence alignment of p2H32 plasmid of SA H32 with the reference AR_0472 (NZ_CP029648.1) plasmid.

Full-size 🖾 DOI: 10.7717/peerj.10185/fig-1

The comparative analysis of ARGs revealed that the genome of SA G6 isolate acquired additional ARGs responsible for the resistance of aminoglycoside (*aph (3′)-IIIa, ant (6′)-I* and *aac (6′)-II*), nucleoside (*sat*), fluoroquinolone (*qacA*). The macrolide, lincosamide, streptogramin (MLS) erythromycin antibiotic resistance genes (*emrA*) were detected in the genomes of SA G6 and SA G8 isolates while the genome of SA H32 isolate present *emrC* gene. The penicillin resistance gene (*blaZ*) was found absent in SA G8 isolate. This in-silico identification and our previous antibiotic susceptibility test results were correlated with beta-lactam, erythromycin (MLS), and vancomycin antibiotic resistance analysis (*Naorem et al., 2020*). The secondary metabolite biosynthetic gene clusters identified among the genomes were staphylobactin, aureusimine, bacteriocin, and staphyloferrin A. The auto-inducing peptide (AIP)-II gene was identified in SA G6 and SA G8 genomes while AIP-I gene was identified in SA H27 and SA H32 genomes.

In-silico analysis of virulence-factors encoding genes in the genomes
The VFGs predicated against the VFDB revealed 59 VFGs were shared in all strains that are responsible for adherence, toxin, anti-phagocytosis immune evasion, secretion system,
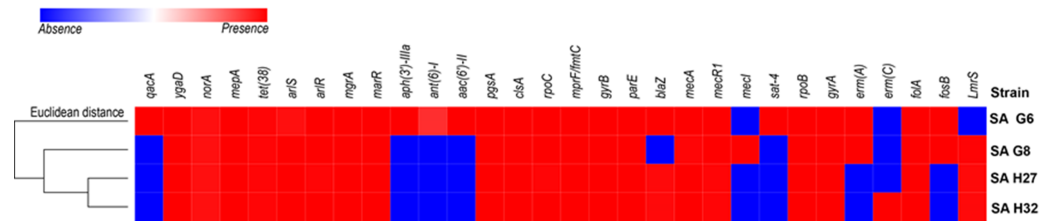
Figure 2 **Heat map showing the presence (red color) and absence (blue color) of antibiotic resistance genes.** The labels on top indicate the gene names and the label on the left indicates the strains.
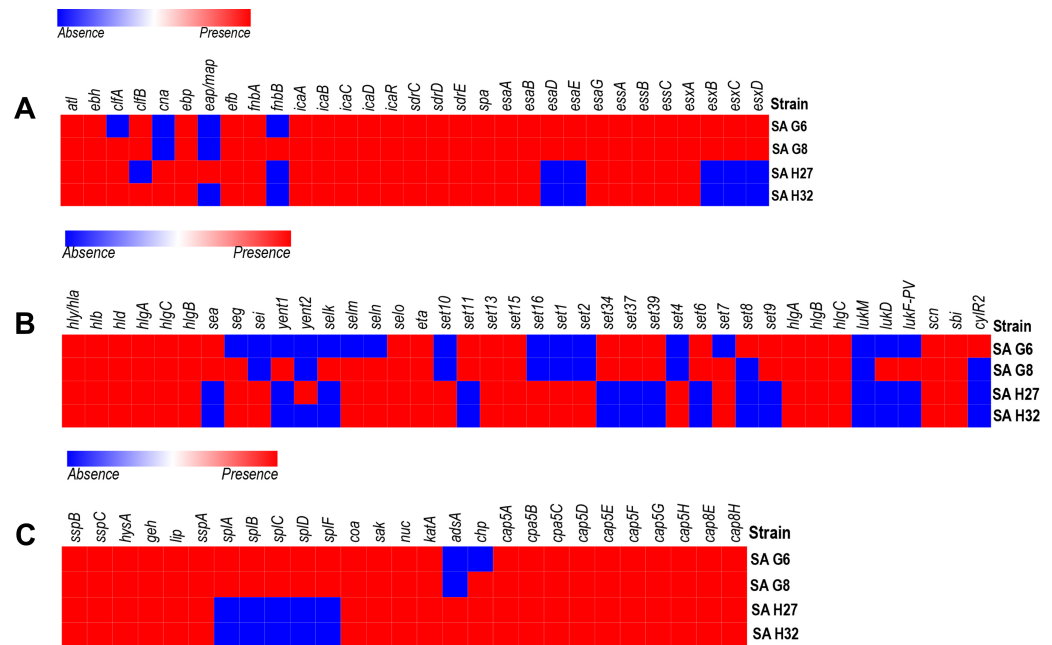
Full-size 🖼 DOI: 10.7717/peerj.10185/fig-2

Figure 3 **Heat map showing the presence (red color) and absence (blue color) of virulence factors encoding genes.** (A) Adherence and secretary factors. (B) Toxins. (C) Enzymes and anti-phagocytosis (capsules) factors. The labels on top indicate the gene names and the label on the left indicates the strains.

Full-size 🖼 DOI: 10.7717/peerj.10185/fig-3

exoenzyme activity, and iron uptake (Fig. 3). The genome of SA G8 isolate has occupied 3.40% of VFGs against its CDS, whereas the genome of SA H32 isolate has 2.97% of VFGs against its CDS.

Adherence associated genes shared in all genomes of *S. aureus* isolates were 63.63% (14/22) such as autolysin (*atl*), cell wall-associated fibronectin-binding protein (*ebh*), elastin binding protein (*ebp*), fibrinogen binding protein (*efb*), fibronectin-binding proteins A (*fnbA*), intercellular adhesin (*icaA*, *icaB*, *icaC*, *icaD*, *icaR*), ser-Asp rich fibrinogen-binding proteins (*sdrC*, *sdrD*, sdrE), staphylococcal protein A (*spa*) (Fig. 3A). The genome of SA G8 isolates present 77.27% (17/22) of adherence associated genes with additional genes of clumping factor A (*clfA*), clumping factor B (*clfB*), and fibronectin-binding proteins (*fnbB*).

Type VII secretion system involves in membrane-associated proteins (*esaA, essA, essB, and essC*), soluble cytosolic (*esaB, esaE, esaG*), and secreted virulence factors (*esxA, esxB, esxC, esxD,* and *esaD*) were identified in the genomes of SAG6 and SA G8 isolates while the SAH27 and SAH32 isolates absence *esaD, esaE, esxB, esxC,* and *esxD* genes (Fig. 3A).

The genomes of all isolates shared 29.41% (10/34) of toxin genes such as alpha-hemolysin gene (*hla*), beta-hemolysin gene (*hlb*), delta hemolysin gene (*hld*), gamma hemolysin A (*hlgA*), gamma hemolysin B (*hlgB*), gamma hemolysin C (*hlgC*), enterotoxin-like O (*selo*), exfoliative toxin type A (*eta*), and exotoxin (*set13, set15*) (Fig. 3B). The highest number of toxin genes were identified in the genome of SA G8 *i.e.,* 73.52% (25/34), and in addition to shared genes, the extra genes were enterotoxin A (*sea*), enterotoxin B (*seg*), enterotoxin Yent1 (*yent1*), enterotoxin-like K (*selk*), enterotoxin-like M (*selm*), enterotoxin-like N (*seln*), exotoxin (*set6, set7, set9, set11, set34, set37, set39*), leukotoxin D (*lukD*), Panton-Valentine leukocidin (*lukF-PV*).

The genes involve in anti-phagocytosis namely capsular polysaccharide synthesis genes belong to stereotype 5 and 8 predominantly present in all the genomes of isolates were capsular polysaccharide synthesis enzyme Cap5A (*cap8A*), capsular polysaccharide synthesis enzyme Cap5B (*cap8B*), capsular polysaccharide synthesis enzyme Cap5C (cap8C), probable polysaccharide biosynthesis protein EpsC (*cap8D*), capsular polysaccharide synthesis enzyme Cap8E (*cap8E*), capsular polysaccharide synthesis enzyme Cap5F (*cap8F*), UDP-N-acetyl-L-fucosamine synthase (*cap8G*), capsular polysaccharide synthesis enzyme Cap5L (*cap8L*), capsular polysaccharide synthesis enzyme Cap8M (*cap8M*), capsular polysaccharide synthesis enzyme Cap8N (*cap8N*), UDP-N-acetyl-D-mannosamine dehydrogenase (*cap8O*) and UDP-N-acetylglucosamine 2-epimerase (*cap8P*) (Fig. 3C). Other genes responsible for the host immune evasion such as IgG-binding protein (*sbi*), staphylococcal complement inhibitor (*scn*), and chemotaxis inhibiting protein (*chp*) were identified in all isolates.

Several exoenzymes encoding genes namely cysteine protease/ staphopain (*sspB, sspC*), hyaluronate lyase (*hysA*), lipase (*geh, lip*) serine V8 protease (*sspa*), staphylocoagulase (*coa*), staphylokinase (*sak*), and thermonuclease (*nuc*) were present in the genomes of all isolates. However, five genes cluster for serine protease (*splA, splB, splC, splD, splF*) were absent in the genomes of SA H27 and SA H32 isolates (Fig. 3C).

Eight genes involved in iron uptake mechanism including cell surface protein (*isdA*), cell surface receptor (*isdB*) and cell wall anchor proteins (*isdC*), heme transporter component (*isdD*), high-affinity heme uptake system protein (*isdE*), heme-iron transport system permease protein (*isdF*), sortase B (*srtB*), heme-degrading monooxygenase; staphylobilin-producing (*isdG*) were identified in all the genomes of isolates.

Comparative genome analysis

The genome comparative analysis based on ANIb matrices results indicated that the genome of SA G8 isolate exhibits the nearest identities to all genomes. SA G8 genome showed ∼99.5% identities to SA G6 genome. The genomes of SA G6 and SA G8 exhibited ∼99.7% identities to *S. aureus subsp. aureus* ST228 (HE579071.1). Also, the genomes of SA H27 and SA H32 showed 99.9% identities to each other and these two genomes displayed

the highest identities (99.9%) to *S. aureus subsp. aureus* HO 5096 0412 (HE681097.1) (Fig. S1). The digital DDH values between the genome of *S. aureus* isolates and the closest relative genomes were 90.7–100% (using GBDP distance formula d0), 77.1–97.0% (using GBDP distance formula d4), and 91.1–99.9% (using GBDP distance formula d6). SA H27 and SA H32 genomes exhibit the nearest identities of 99.9%, 99.8%, and 100% using the formula d0, d4, and d6 respectively and displayed G+C difference of 0.01%. However, the SA G6 genome showed less identity to all the comparative genomes based on dDDH. The high G+C constituent difference (0.1%) was observed in the case of isolated strains of SA G8 and SA H32 genomes.

A whole-genome circular comparative map of four *S. aureus* genomes and their close reference genomes was generated against *S. aureus subsp. aureus* HO 5096 0412 (HE681097.1) genome using CGView server based on BLAST sequence similarities. Each genome was indicated by a different color, and the darker areas in the circular genome showed a 100% sequence similarity with the reference genome, while the lighter (gray) areas showed a 70% sequence similarity (Fig. 4). The map revealed less gap between the SA H27 (CP032161) and SA H32 (RAHP00000000) genomes showing high proximity between them when compared to other genomes. SA G6 (RAHA00000000) genome has many gaps with white color than the other genomes showing a distant relationship.

The SEED subsystem categories identified by RAST revealed that the genomes of all the isolates possessed "amino acids and derivatives" was the largest subsystem, followed by "carbohydrates", "protein metabolism", and "cofactors, vitamins, prosthetic groups, pigments" (Fig. 5A). The "carbohydrate", and "protein metabolism" subsystems were found largest in SA H27 (12.76%) and SA H32 (10.47%) genomes, respectively. The subsystem belongs to "phages, prophages, pathogenicity island" (2.5%) was identified as highest in the SA G8 genome. Amongst the genomes, the SA G6 genome has the largest subsystem of "amino acids and derivatives" (15.9%), and "virulence, disease, and defense" (4.43%). In the "virulence, disease, and defense" subsystem of SA G6, 93 genes associated with adhesion, bacitracin stress response, colicin v, and bacteriocin production cluster, copper homeostasis, bile hydrolysis, cobalt-zinc-cadmium resistances, multidrug resistances, 2-protein, mercuric reductase, mercury resistance operon, streptothricin resistance, teicoplanin-resistances, aminoglycoside adenylyltransferases, fluoroquinolone resistances, arsenic resistance, fosfomycin resistance, beta-lactamase, cadmium resistance, multidrug resistance efflux pumps, and invasion and intracellular resistances. In the comparative eggNOG function study of *S. aureus* genomes, "amino acid transport and metabolism" was observed as for the majority of COGs, followed by those COGs related to "translation, ribosomal structure, and biogenesis", "transcription", and "cell wall/membrane/envelope". The eggNOG analysis results revealed that the SA G6 genome has the highest number of COGs associated with defense mechanisms (Fig. 5B). In the core genomes, 9.46%, 7.21%, and 6.9% of COGs had functions associated to "amino acid transport and metabolism (E)", "translation, ribosomal structure, and biogenesis (J)", and "transcription (K)", respectively. Amongst the functional prediction of genomes, most COGs were associated with "function unknown (S)".
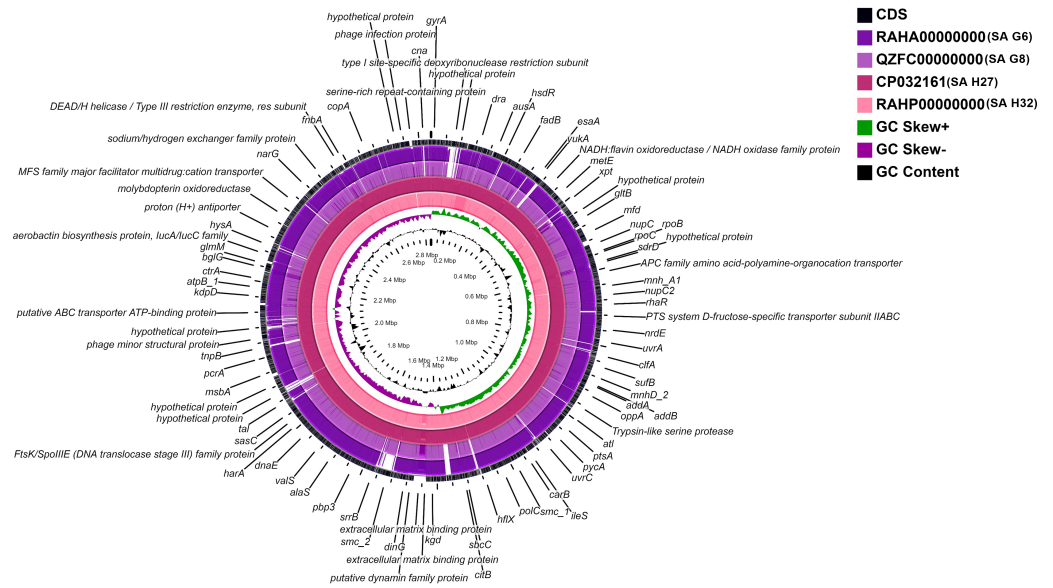
Figure 4 Circular genome comparison map showing homologous chromosome segment of four *S. aureus* genomes with the reference genome of *S. aureus subsp. aureus HO 50960412* (HE681097.1) strain using CGviewer. The inner scales designate the coordinates in kilobase pairs (kbp). White spaces indicate regions with no identity to the reference genome.
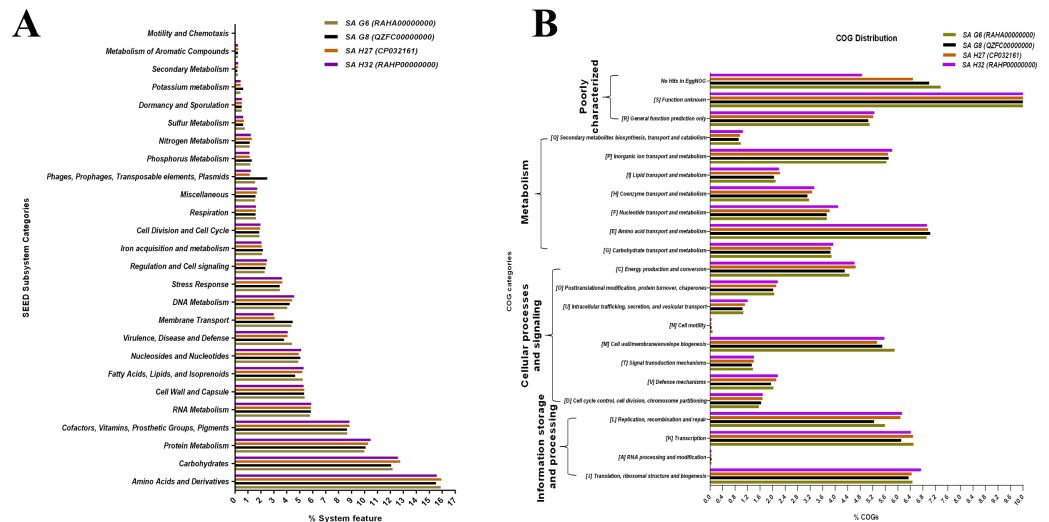
Figure 5 Comparative functional categorization of all predicted ORFs in the genomes of the *S. aureus* isolates. (A) Percentage distribution of subsystem categories based on the SEED database. (B) Percentage distribution of COGs based on EggNOG.

Pan-genome, core-genome, and singletons analysis

The orthologous groups are categories into three groups based on the pan-genome distribution such as core (present in all genomes of *S. aureus* strains), dispensable (present in at least two strains, but not all), and singleton genes (present no orthologs

in any other genomes). The comparison of four study *S. aureus* genomes generated a pan-genome size of 3265 genes, of which 2304 (70.6%) genes were core genome, 462 (14.2%) genes were dispensable, and 499 (15.3%) genes were singletons. The Rcp value for the genomes of *S. aureus* isolates was calculated and the ratio Rcp was 0.70 and it is indicated that the genomes of *S. aureus* isolates were high inter-species diversity. A total of 499 singleton genes were calculated across the genomes of four *S. aureus* isolates, of which SA G6 genome acquired the highest number of singleton genes (220) that constitute the genes encode for proteins *viz.* aminoglycoside 3-phosphotransferase, aminoglycoside 6-phosphotransferase, aminoglycoside N(6)-acetyltransferase, streptothricin acetyltransferase, antiseptic resistance protein, cadmium resistance proteins, cadmium efflux system accessory protein, cadmium-transporting ATPase, ferric siderophore transport system, mercuric ion reductase, anti-adhesin, Tn552 transposase, pathogenicity islands (SaPI and SaPIn2), prophage-like elements, mobile elements, phage associated hypothetical proteins, hypothetical proteins, *etc.* The identified singleton genes of SA G6 genome were present within the genomic island (GI). This GI region is located between 2804353–2873411 base pair sequence region of the genomic sequence. While the SA H27 genome has the least singleton genes (6) constituting the genes encode for hypothetical proteins and phage proteins. The difference in the genomic constituents between the genome of SA H27 and SA H32 isolates revealed that SA H32 acquired the genes encoding for 23S rRNA (adenine(2058)-N(6))-dimethyltransferase, replication and maintenance protein, hypothetical proteins, phage-like elements, and mobile element protein. The genes shared by four study genomes and their respective singletons is represented in Fig. 6A. When the three reference *S. aureus* genomes were included in the pan-genome analysis, the core/pan-genome ratio drop down by 18.97% with inflation of pan-genome to 3415 genes and deflation of core-genome to 1762 genes. The core-genome and singleton genes formed by seven genomes of *S. aureus* strains is represented in flower-plot (Fig. 6B). When the three reference genomes of *S. aureus* strains were included in the pan-genome analysis, SA G6 isolate occupied the highest number of singleton genes (104) while SA H27 isolate has the lowest singleton gene (2) (Fig. 6B). In the pan-genome development analysis of seven *S. aureus* strains, α value (the power-law co-efficient) was estimated as 0.141 which corresponds to the growing and open pan-genome model (Fig. S2). The pan *vs.* core development plot appeared the progression of the pan and core-genomes as additional genomes are added for analysis, and showing that the sharp decline of the core-genome size with the introduction of *S. aureus subsp. aureus* ST228 (HE579073) (Fig. 6C). In the plot of core-genome development, the core-genome size approach (Ω) value revealed that the core-genome size of seven *S. aureus* would be declined to 1404.8 (Fig. S3). The singleton development analysis suggested that the pan-genome size will continue to expand at the rate of 35.9 genes per novel, representative genome (Fig. S4). The shape of the pan-genome *vs.* core-genome curve showed fluctuation in their gene numbers when different order of the genomes was set, even so, pan-genome and core genome developmental plots result remained unaffected by the genomes order.
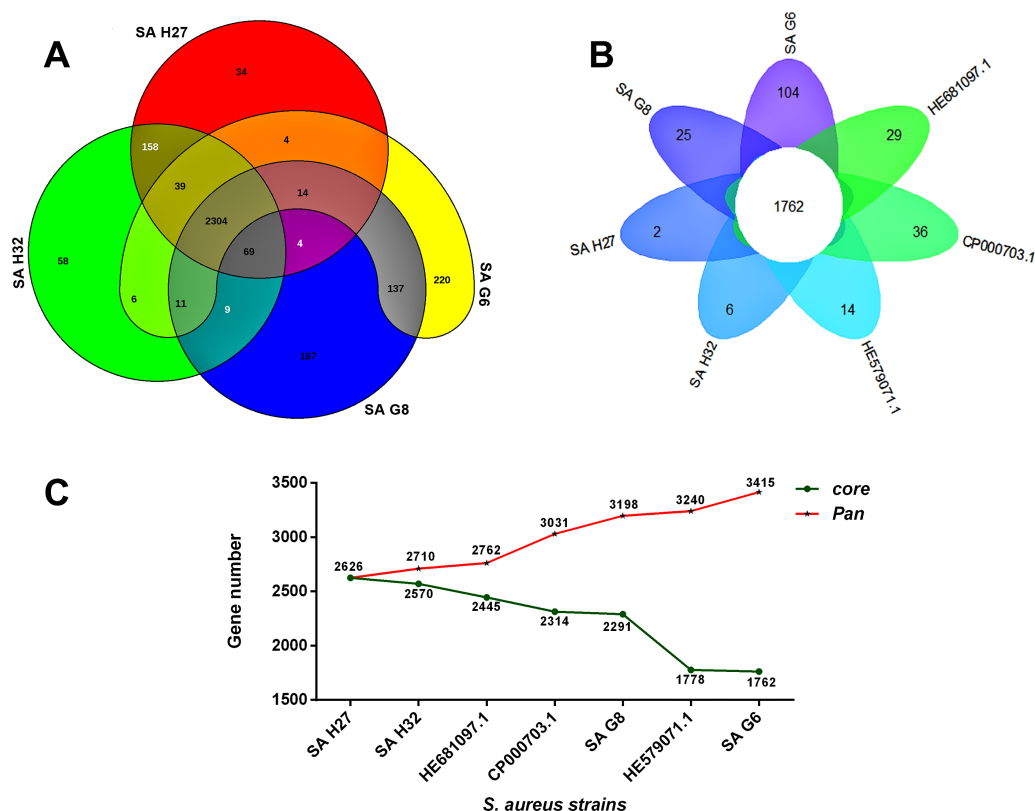
Naorem et al. (2021), *PeerJ*, DOI 10.7717/peerj.10185

14/33

Figure 6 **Pan-genome analysis of _S. aureus_ strains.** (A) The Venn diagram represents the pan and core-genomes of four study genomes (SA G6, SA G8, and SA H27 and SA H32) based on orthology analysis. Overlapping regions represent common CDSs shared between the genomes. The numbers outside the overlapping regions signify the singletons of each genome. (B) Flower plot diagram representing the four study isolates and three reference strains. The core-genes of 1762 was represented in the center of the flower and the petals represent the singletons of concern genomes. (C) Core vs. pan-genome plot of the seven genomes.

Full-size ⬛ DOI: 10.7717/peerj.10185/fig-6

Comparative phylogenetic tree analysis

The phylogenomic analysis of _S. aureus_ isolates provide the tree into three major clades (Fig. 7A). The clade A consists of 5 strains that belonged to CC22 and showing that Hungarian isolates, SA H27 and SA H32 have the highest proximity. Germany isolates, SA G6 and SA G8 isolate and other strains belonged to CC5 were clustered in the clade C, showing that SA G6 isolate has closely relatedness to _S. aureus subsp. aureus_ ST228 (HE579071.1) and the SA G8 isolate has a higher relatedness to _S. aureus subsp. aureus_ JH9 (CP000703.1) than SA G6 isolate.

The phylogenetic relationship inferred from core-genome SNPs holds a similar agreement with the whole genome-based phylogenetic analysis, and these methods could be useful in distinguishing the genomes even in the strain level and phylogenetic trees are illustrated in Fig. 7B.
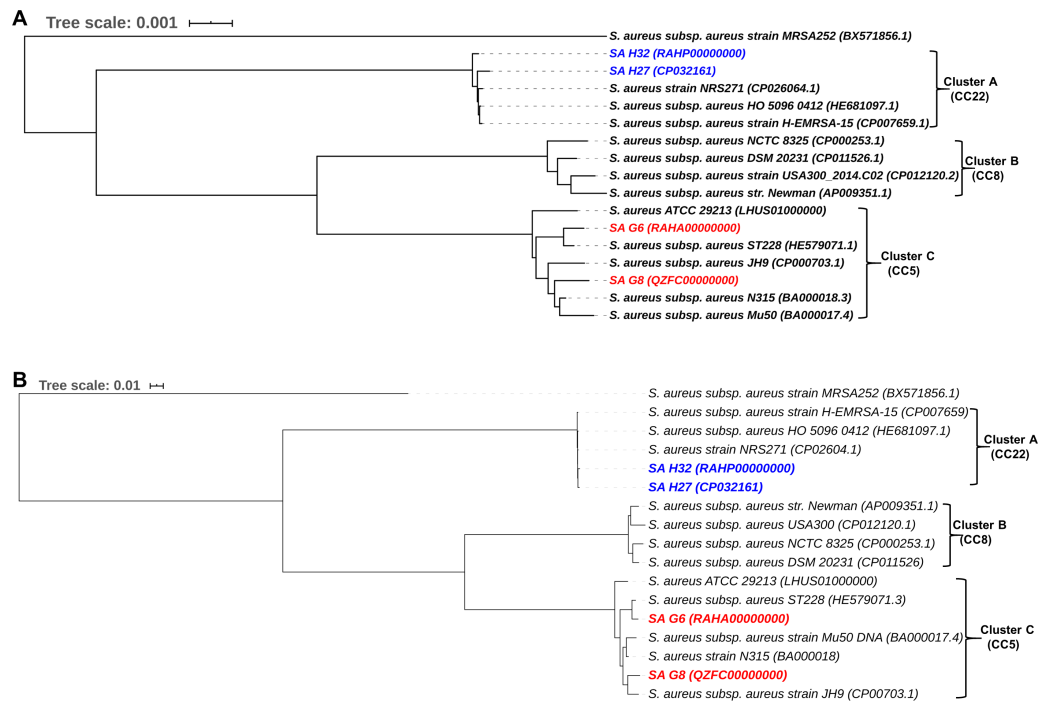
A Tree scale: 0.001

- *S. aureus subsp. aureus strain MRSA252 (BX571856.1)*
- *SA H32 (RAHP00000000)*
- *SA H27 (CP032161)*
- *S. aureus strain NRS271 (CP026064.1)*
- *S. aureus subsp. aureus HO 5096 0412 (HE681097.1)*
- *S. aureus subsp. aureus strain H-EMRSA-15 (CP007659.1)*

Cluster A (CC22)

- *S. aureus subsp. aureus NCTC 8325 (CP000253.1)*
- *S. aureus subsp. aureus DSM 20231 (CP011526.1)*
- *S. aureus subsp. aureus strain USA300_2014.C02 (CP012120.2)*
- *S. aureus subsp. aureus str. Newman (AP009351.1)*

Cluster B (CC8)

- *S. aureus ATCC 29213 (LHUS01000000)*
- *SA G6 (RAHA00000000)*
- *S. aureus subsp. aureus ST228 (HE579071.1)*
- *S. aureus subsp. aureus JH9 (CP000703.1)*
- *SA G8 (QZFC00000000)*
- *S. aureus subsp. aureus N315 (BA000018.3)*
- *S. aureus subsp. aureus Mu50 (BA000017.4)*

Cluster C (CC5)

B Tree scale: 0.01

- *S. aureus subsp. aureus strain MRSA252 (BX571856.1)*
- *S. aureus subsp. aureus strain H-EMRSA-15 (CP007659)*
- *S. aureus subsp. aureus HO 5096 0412 (HE681097.1)*
- *S. aureus strain NRS271 (CP02604.1)*
- *SA H32 (RAHP00000000)*
- *SA H27 (CP032161)*

Cluster A (CC22)

- *S. aureus subsp. aureus str. Newman (AP009351.1)*
- *S. aureus subsp. aureus USA300 (CP012120.1)*
- *S. aureus subsp. aureus NCTC 8325 (CP000253.1)*
- *S. aureus subsp. aureus DSM 20231 (CP011526)*

Cluster B (CC8)

- *S. aureus ATCC 29213 (LHUS01000000)*
- *S. aureus subsp. aureus ST228 (HE579071.3)*
- *SA G6 (RAHA00000000)*
- *S. aureus subsp. aureus strain Mu50 DNA (BA000017.4)*
- *S. aureus strain N315 (BA000018)*
- *SA G8 (QZFC00000000)*
- *S. aureus subsp. aureus strain JH9 (CP00703.1)*

Cluster C (CC5)

Figure 7 **Comparative phylogenetic analysis of *S. aureus* isolates strains with their closely related *S. aureus* strains.** Phylogenomic tree generated using closely related genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5 (A), and core-genome SNP tree generated using the alignment of the high-quality SNPs and PhyML+SMS module was applied (B).

Full-size 🖾 DOI: 10.7717/peerj.10185/fig-7

DISCUSSION

S. aureus is a significant causative agent of both hospital and community-associated infections (*Chambers & Deleo, 2009*). The study of such pathogen at a molecular level through genome comparative analysis improve the ideas of pathogenesis and evolution. Further, such a study provides advantages in diagnosis, treatments, and infection controls (*Kwong et al., 2015*). In the present study, we used whole-genome sequencing (WGS) and in-silico analysis to determine the comparative ARGs, VRGs, pH tolerance associated genes, and evolutionary relationship of four *S. aureus* isolated from different sites of human infection such as skin, nostril, trachea, and others.

The molecular epidemiology study of MRSA helps to find the risk factors associated with MRSA infections and able to differentiate the several MRSA strains (*Mistry et al., 2016*). The genome-based molecular epidemiology studies found that Germany isolates exhibit *SCCmec* type I with ST228, and *SCCmec* type II with ST225 while Hungarian isolates hold *SCCmec* type IVa with ST22. Also, *agr* type II and I were owned by Germany and Hungarian isolates, respectively (Table 1). According to previous studies suggested that MRSA strains with *SCCmec* types I or II or III are dominant among the HA-MRSA, while *SCCmec* types IV or V are the characteristic of CA-MRSA (*Monecke et al., 2011*; *Chua et al., 2014*). The STs of Germany isolates belonged to CC5 which is typical of HA-MRSA, while Hungarian isolates suggest its relationship CA-MRSA. In hospitals, the multidrug-resistance *SCCmec*

type III was replaced by the multidrug-susceptible *SCCmec* type IV (ST22) strains slowly (*D'Souza et al., 2010*). The Hungarian isolates were found positive to Panton-Valentine Leukocidin (PVL) toxin, which is commonly used as a marker of CA-MRSA (*Shukla et al., 2012*; *Bhutia et al., 2015*) besides this toxin has shown to play a role necrosis, accelerating apoptosis and polynuclear—and mononuclear cells lysis, thereby contributing morbidity and mortality (*Barrera-Rivas et al., 2017*; *Lina et al., 1999*).

Staphylococcal β-lactamase encoded by *blaZ* gene is carried by the transposon Tn*552* or Tn*552*-like elements located on a large plasmid and can be non-inducible or inducible with antibiotics (*Maddux, 1991*). It was noticed that *blaZ* gene was absent in SA G8 isolate, probably due to the curing of *blaZ* positive plasmid (*Pugazhendhi et al., 2020*). Erythromycin resistance gene (*ermA*) was detected in the chromosome of SA G6, and SA G8 isolates, however, *emrC* gene was found in the plasmid of SA H32 (Fig. 1B). It was suggested that these genes may not be involved in the loss of specific ARGs for environmental adaptation, but it is expected to be essential for these isolates (*Lim et al., 2015*).

MRSA is responsible for causing biofilm infections that are more difficult to treat and need more intensive care as compared to *Staphylococcus epidermidis* biofilm (*Reffuveille et al., 2017*). The principal component of biofilm formation is PIA which consists of different intracellular adhesion (*ica*) genes (*Cramton et al., 1999*) and play a crucial role in the initial stage of bacterial cell adherence to surfaces and intercellular adhesion for the cells to aggregate (*Farran et al., 2013*). These genes were detected in all isolates however, the biofilm production ability varies from weak to strong were observed in our previous study (*Naorem et al., 2020*). Our previous study identified that SA G6 isolate obtained from skin infection showed a weak biofilm-forming ability (*Naorem et al., 2020*). The low biofilm formation in SA G6 might be degraded the biofilm by DNase enzyme found in skin cells (*Eckhart et al., 2007*). The previous study revealed that the presence of the *ica* genes did not always correlate with biofilm (*Møretrøet al., 2003*; *Nasr, Abushady & Hussein, 2012*). Some authors reported that despite the presence of *ica* operon, some staphylococcal isolates produce weak biofilm production due to the inactivation of *icaA* by insertion of IS256 (*Cho et al., 2002*; *Kiem et al., 2004*). Further reported that the insertion of IS256 inactivates *mutS* and contributes to vancomycin resistance development in vancomycin-intermediate *S. aureus* strains (*Kleinert et al., 2016*). Also, the upregulation of *icaA* and *icaD* genes during acidic stress promotes biofilm formation which in-turn plays a role to resist it from acidic and alkaline environments and establishes the niche adaptation in Staphylococcus strains (*Lindsay et al., 2006*). In addition to *ica* locus, the presence of *clfA*, *clfB*, and *epbs* genes initiates the biofilm formation (*Ghasemian et al., 2015*), however in the present study, the SA H27 isolate carried *clfA*, and *epbs* genes and showed strong biofilm formation in our previous study (*Naorem et al., 2020*) compared to other isolates while SA G8 and SA H32 isolates carried *clfA*, *clfB*, and *epbs* genes though their biofilm formation was relatively low, suggesting that presence or absence of such genes have no significant in biofilm formation. A recent study reported that *sdrC* mutant exhibited significantly inhibited biofilm formation (*Chen et al., 2019*) and the expression of the *ica* operon and *sdrC* are highly responsive to biofilm formation (*Shin et al., 2013*). Our study revealed

the sequence variation in *sdrC* in Hungarian isolates, this variation might influence the biofilm formation. The global regulatory gene, *agr* repression has been associated with biofilm formation and its induction through AIP results in seeding dispersal in mature biofilm (*Boles & Horswill, 2008*). CA-MRSA strains showed higher activity of *agr*, which controls and enhance the virulence (*Aires-De-Sousa, 2017*). It was reported that *S. aureus* strains belonged to *agr* I group exhibited a strong biofilm-forming ability than the strains belonged to *agr* IV group (*Zhang et al., 2018*) and a similar result was observed in one of our isolate SA H27. In addition to this extracellular adherence protein (encoded by *eap* gene), and beta toxin (encoded by *hlb* gene) play a role in biofilm maturation (*Huseby et al., 2010*; *Sugimoto et al., 2013*). In our finding showed that *eap* gene was present only in the SA H27 isolate and this gene might be attributed to high biofilm formation. Since biofilm formation involves many factors/ genes that take part in PIA dependent or independent biofilm, biofilm formation by regulator genes and eDNA (*Archer et al., 2011*). Also, the presence of such genes in *S. aureus* may not provide much impact on biofilm formation profiling. There was a difference in the prevalence of biofilm-associated genes between the isolated strains and suggests that the presence of genes encoding biofilm formation is not an absolute determinant of biofilm formation ability observed in our previous study (*Naorem et al., 2020*). Thus, our future studies will focus on the expression profiling of such relevant genes which may be necessary to determine the key genes involved in biofilm formation.

The high survival rates were observed in both acidic and alkaline pH conditions in all isolates was evaluated by the genomic study, it is elucidated that all the isolates possessed the arginine deiminase and urease operon that aids in the generation of ammonia due to the hydrolysis of L-arginine and urea by arginine deiminase and urease. The released ammonia and urea counteract the acidic environment (*Cotter & Hill, 2003*; *Valenzuela, Cerda & Toledo, 2003*). Further, the proton efflux pump (F_0F_1 ATPase) plays a role to extrude H^+ out of the cells and maintains the pH homeostasis (*Foster, 2004*; *Maurer et al., 2005*). However, in the case of alkaline tolerance, it was reported that the *S. aureus* genome encodes a unique Ktr-like system where the cytoplasmic gating protein KtrC regulates the uptake of K^+ that is essential for maintaining cytoplasmic pH and supporting H^+ efflux under alkaline conditions (*Gries et al., 2016*).

The ability of *S. aureus* as a pathogen can be accredited to its arsenal of virulence factors among which secreted pore-forming toxins (PFTs), exfoliative toxins (ETs), ESAT-6-like proteins, exoenzymes, and superantigens (SAgs) play a significant role in the pathogenesis of invading infections in healthy individuals (*Otto, 2014*; *Bartlett & Hulten, 2010*). The presence of *hlb* gene in the isolates contributes to the phagosomal escape of *S. aureus* and influences biofilm development (*Huseby et al., 2010*; *Periasamy et al., 2012*). The PVL toxin was identified in the prophages of Hungarian isolates and expressing Sa2 integrase. These isolates have cytolytic activity against blood cells and leukocytes, contributing to the *S. aureus* pathogenicity (*Vandenesch et al., 2003*). Staphylococcal enterotoxins (SEs) or staphylococcal superantigens proteins (SAgs) are well-known for causing food poisoning, localized epidermal infections (bullous impetigo), and generalized diseases (Staphylococcal scalded skin syndrome) (*Grumann, Nübel & Bröker, 2014*; *Argudín, Mendoza & Rodicio,*

Naorem et al. (2021), *PeerJ*, DOI 10.7717/peerj.10185

18/33

2010). SEs encoding genes are located on mobile elements including bacteriophages, pathogenicity islands (SaPI), or plasmids. In this study, SEs encoding genes such as *sea, seg, sei, yent1, yent2, selk, selm, seln, and selo* were identified. Hungarian isolates, SA H27, and SA H32 acquired *seg* and *sei* genes, however *sei* gene was absent in Germany isolates, SA G6, and SA G8. These *seg* and *sei* genes belong to *egc* (enterotoxin gene cluster), involve in staphylococcal food poisoning TSS, and SSF (*Jarraud et al., 2001*; *Chen, Chiou & Tsen, 2004*) and *egc* was distributed widely in clinical isolates and playing a role in pathogenesis (*Jarraud et al., 2001*). Exfoliative toxins (ETs) are known as epidermolytic toxins that induce skin shedding and blister formation (*Melish & Glasgow, 1970*). In this study, *eta* gene encoded for ETA toxin was found in all the isolates and responsible for causing human skin damage, and most prevalent in Europe (*Ladhani, 2001*). Capsular polysaccharide synthesis genes are almost all detected in clinical isolates *S. aureus* showing significant virulence by targeting the antibodies that protect against Staphylococcal infections Suetal1997. Type VII secretion system (T7SS) was present in Germany isolates (Fig. 3A) and promoting them to persist in their hosts (*Tchoupa et al., 2019*). The *esxA* and *esxB* gene show a significant role in the distribution and colonization of *S. aureus*, and activation of the cell-mediated immune responses, boost the pathogenesis (*Burts et al., 2005*). Also, *esaD* gene found only in Germany isolates suggesting that this gene can inhibit the growth of other closely related *S. aureus* strains and playing a role in an intra-species competition (*Cao et al., 2016*). The family of beta-hemolysin converting phage encodes proteins such as SCIN (staphylococcal complement inhibitor) and CHIPS (chemotaxis inhibiting protein of staphylococcus) involved in host-pathogen interaction and contribute to evading human innate immune response (*Wamel et al., 2006*), these proteins were identified in intact prophages of SA G8 and SA G27 genomes but CHIPs was absent in the prophages of SA H32 genome. Therefore, prophages were the reservoir of virulence and resistance factors that play a role in the evolution of virulence strains and causing a major threat to human and animal health (*Barrera-Rivas et al., 2017*). The presence of ARGs and VFGs in the prophage regions of SA G6 genome differentiates it from the other *S. aureus* isolates and may determine its greater pathogenic potential by modifying its antigenicity (*Barrera-Rivas et al., 2017*). Also, plasmid p1G6 carried *qacA* gene, which is known to decrease chlorhexidine (antiseptic) susceptibility and giving an event of MGEs transfer evidence of *qacA* across the *S. aureus* strains (*Labreck et al., 2018*). The harbor of MGEs (mosaic features of prophages and plasmids) contributes to the tremendous distribution of ARGs and VFGs among the *S. aureus* isolates (*McCarthy & Lindsay, 2012*; *McCarthy et al., 2014*). This MGEs transfer event could be useful for the survival of *S. aureus* in different ecological niches (*Lindsay, 2010*).

The pangenome described here is composed of 3415 genes, of these, 1762 genes are shared among *S. aureus* isolates (Fig. 6B). Functional annotation of the core-genome revelaled that they are mostly associated transcription and translation, and different metabolism categories, such similar result was reported previously (*Bosi et al., 2016*; *Sharma et al., 2018*). The core-genome and accessory genome functional characterizations revealed that *S. aureus* isolates required amino acids than carbohydrates as the energy source and suggests that these isolates adapted to grow in a protein-rich medium than carbohydrates

(Figs. 5A and 5B). It was suggested that the survival of *S. aureus* can be maintained by the catabolism of amino acids (*Halsey et al., 2017*). The core-genome has 51.6% of genes and validated that *S. aureus* is a clonal species (*Feil et al., 2003*; *Bosi et al., 2016*). The mutation event that occurred in the core-genome of closely related *S. aureus* provides important roles in virulence and persistence of *S. aureus* strains (*Kennedy et al., 2008*). So, an in-depth analysis of strain-specific genetic variation is required for further understanding of the pathogenicity. The inflation of pan-genome and deflation of core-genome was seen after the introduction of reference genomes and its regression analysis revealed that the pan-genome is open, suggesting that the gene repertoire of this species is theoretically limitless. A similar finding was observed in the DNA microarray experiment of thirty-six *S. aureus* isolates (*Fitzgerald et al., 2001*). The drastic decline of the core/pan- genome ratio after the introduction HE579071.1 (*S. aureus subsp. aureus* ST228) and SA G6 suggested that these two strains have distinct genomic contents (Fig. 6C). The genomic content variation between the genomes is due to the acquisition of certain genes that encode for virulence and resistance factors, pathogenicity islands, prophage-like elements, plasmids, mobile element proteins, and hypothetical proteins in the GIs. These GIs are mobilized across organisms via HGT events (*Schmidt & Hensel, 2004*). This finding was supported by gaps that appeared in the genome ring of SA G6 genome and suggesting that this isolate showed a distant relationship to others (Fig. 4). The gaps that appeared in the map are due to the GC% content difference in the comparative genomes, and it results from the event of MGEs transfer via HGT and the GC skewed regions indicated the regions where HGT occurred (*Hayek, 2013*).

We specifically analyzed the presence of ARGs and VFGs in the core genomes and pangenomes. Some genes involved in multidrug resistance or drug efflux such *ygaD*, *arlR*, *arlS*, and *mepA* are components of the core-genome (Fig. 2). The large repertoire of genes (29%) in the accessory genome gives advantages in adaptation and that can contribute to pathogenicity or niche specificity of strains (*Medini et al., 2005*). The analysis of pangenome is essential to understand the event of MGEs transfer and *S. aureus* evolution (*Ozer, 2018*). The interpretation from the dispersible and singleton genes content analysis of *S. aureus* genomes allows us to understand the genetic variation among the CC5 and CC22. Juhas et al. reported that most dispensable and singleton genes were acquired through HGT and operate an important role in drug resistance or virulence (*Juhas, Eberl & Church, 2012*). A high portion of unique genes or singletons in *S. aureus* genomes were related to MGEs, which could drive the gaining of novel functional elements especially drug resistance and virulence. These singletons are the main drivers of the phenotypic variation within *S. aureus* strains and the evolution of *S. aureus* (*Carvalho et al., 2019*).

The phylogenetic trees based on whole-genome and core-genome SNP methods support each other and revealed that these methods were able to distinguish between strains at a higher resolution in terms of the geographic origin of strains and phylogenetic trees are illustrated in Fig. 7. The phylogenomic analysis revealed that the strains with ST225 (Germany), ST228 (Germany, Switzerland), ST105 (USA), and ST5 (Japan) were clustered in the same CC5 clade (Cluster C), and a different clade (Cluster A) was noticed among the UK origin ST22 (CC22) and diverged from Germany origin strains (Fig. 7A), this

finding was in good agreement with the previously published article (*Aanensen et al., 2016*). The CC5 (ST225) and CC22 (ST22) were found to be the most dominant clones circulating in Europe (*Grundmann et al., 2014*). The comparative genome analysis revealed that Germany and Hungarian isolates are genetically diverse and showing variation among them due to the gain or loss of MGEs such as *SCCmec*, plasmid, phage elements, or the insertion of transposase. The event of MGEs transfer was observed in ST5, ST225, and ST228 (Fig. 7A), and similar results were also reported previously (*Nubel et al., 2008*; *Nübel et al., 2010*; *Vogel et al., 2012*). The SNPs located in the core-genome define as the element present in *S. aureus* strains, these SNPs based phylogenetic tree was constructed to avoid the HGT of MGEs misuse phylogenetic interpretation, as well as this tree, resolved the subdivision within cluster C of Fig. 7A indicating that SA G6 isolates and *S. aureus subsp. aureus* ST228 exhibits the closest strains (Fig. 7B). These strains shared the genetic background (ST228/*SCCmec*-I) and revealing 99.8% OrthoANIu similarity value in their genomes, likewise, Hungarian isolates (SA H27 and SA H32) in clade A (Fig. 7B) shared molecular epidemiological background in terms of *SCCmec*-IVa, and ST-22 and showing 99.8% OrthoANIu value. However, SA G8 isolate and *S. aureus subsp. aureus* ST228 belongs to ST225 and ST105, respectively were clustered together (Fig. 7B). The strains with the same genetic background were clustered together in both phylogenetic trees which suggest that these strains are highly alike, however comparative genome analysis exposed that the acquisition of phage elements and plasmids through the events of MGEs transfer contribute to differences in their phenotypic characters. Such events provide an impact on the fitness or pathogenicity or epidemicity of the strains.

CONCLUSIONS

Using WGS, we characterized the four clinical MRSA isolates that infect the skin, nostrils, trachea, and other sites. The data generated from the WGS confirmed the diversity of MRSA among the same CC5 and CC22. It is clearly stated that the biofilm-forming ability of MRSA was not correlated with the presence of biofilm-forming encoding genes, also the genetic constituents have no information regarding the infection sites. So, expression profiling of biofilm-related genes is required to define the key genes involved in biofilm formation. The comparative genome study allowed the segregation of isolates of geographical origin, and differentiation of clinical isolates from the commensal isolates. An interesting finding is the addition of SA G6 genome responsible for open pan-genome and diversity among genomes. The openness of pan-genomes of *S. aureus* isolates relies on the acquisition of MGEs. The evidence of MGEs transfer event especially in SA G6 is supported by the drastic drop of the core/pan-genome ratio curve, and gaps and GC skewed regions in comparative genome map. The presence of *ant(6)-Ia, aph(3′)-III)* and *sat-4* in the GI region of SA G6 are likely acquired and these genes may provide fitness and a selective advantage during host-adaptation and colonization. Phylogenetic analysis suggests that SA G6 and *S. aureus subsp. aureus* ST228 strains are distinct from its group. The acquisition of plasmid and prophage functional modules such as ARGs and VFGs in *S. aureus* isolates contributes a major role in the rapid evolution of pathogenic *S. aureus* lineages and that confer specific

advantages in a defined host under environmental conditions. Through this comparative genome analysis would improve the knowledge about the pathogenic *S. aureus* strain's characterization, adaptation, and dynamic evolutionary process in the transmission of infections globally.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Grant Disclosures

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Romen Singh Naorem conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Jochen Blom analyzed the data, prepared figures and/or tables, and approved the final draft.
- Csaba Fekete conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

Data are available at the NCBI database under the accession numbers: RAHA00000000, QZFC00000000, CP032161, RAHP00000000.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.10185#supplemental-information.

REFERENCES

Aanensen DM, Feil EJ, Holden MTG, Dordel J, Yeats CA, Fedosejev A, Goater R, Castillo-Ramírez S, Corander J, Colijn C, Chlebowicz MA, Schouls L, Heck M, Pluister G, Ruimy R, Kahlmeter G, Åhman J, Matuschek E, Friedrich AW, Parkhill J, Bentley SD, Spratt BG, Grundmann H. 2016. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive *staphylococcus aureus* in Europe. *mBio* **7**(**3**):e00444-16 DOI 10.1128/mbio.00444-16.

Aires-De-Sousa M. 2017. Methicillin-resistant *staphylococcus aureus* among animals: current overview. *Clinical Microbiology and Infection* **23**:373–380 DOI 10.1016/j.cmi.2016.11.002.

Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen A-LV, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran H-K, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar GV, Mcarthur AG. 2019. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research* **48**:D517–D525 DOI 10.1093/nar/gkz935.

Archer NK, Mazaitis MJ, Costerton JW, Leid JG, Powers ME, Shirtliff ME. 2011. *staphylococcus aureus* biofilms. *Virulence* **2**:445–459 DOI 10.4161/viru.2.5.17724.

Argudín MÁ, Mendoza MC, Rodicio MR. 2010. Food poisoning and *Staphylococcus aureus* enterotoxins. *Toxins* **2**:1751–1773 DOI 10.3390/toxins2071751.

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research* **44**:W16–W21 DOI 10.1093/nar/gkw387.

Aziz RK, Bartels D, Best AA, Dejongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, Mcneil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* **9**:75 DOI 10.1186/1471-2164-9-75.

Barrera-Rivas CI, Valle-Hurtado NA, González-Lugo GM, Baizabal-Aguirre VM, Bravo-Patiño A, Cajero-Juárez M, Valdez-Alarcón JJ. 2017. Bacteriophage therapy:

an alternative for the treatment of *staphylococcus aureus* infections in animals and animal models. In: Enany S, Crotty Alexander LE, eds. *Frontiers in Staphylococcus aureus*. London: IntechOpen, 179–201 DOI 10.5772/65761.

Bartels MD, Petersen A, Worning P, Nielsen JB, Larner-Svensson H, Johansen HK, Andersen LP, Jarlov JO, Boye K, Larsen AR, Westh H. 2014. Comparing whole-genome sequencing with sanger sequencing for spa typing of methicillin-resistant *staphylococcus aureus*. *Journal of Clinical Microbiology* **52**:4305–4308 DOI 10.1128/jcm.01979-14.

Bartlett AH, Hulten KG. 2010. *Staphylococcus aureus* pathogenesis. *The Pediatric Infectious Disease Journal* **29**:860–861 DOI 10.1097/inf.0b013e3181ef2477.

Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing Group, Lau BY, Hoad G, Winsor GL, Brinkman FS. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research* **45**:W30–W35 DOI 10.1093/nar/gkx343.

Bikandi J, Millan RS, Rementeria A, Garaizar J. 2004. In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction. *Bioinformatics* **20**:798–799 DOI 10.1093/bioinformatics/btg491.

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research* **47**:W81–W87 DOI 10.1093/nar/gkz310.

Blom J, Kreis J, Spänig S, Juhre T, Bertelli C, Ernst C, Goesmann A. 2016. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Research* **44** DOI 10.1093/nar/gkw255.

Boles BR, Horswill AR. 2008. agr-mediated dispersal of *staphylococcus aureus* biofilms. *PLOS Pathogens* **4**:e1000052 DOI 10.1371/journal.ppat.1000052.

Bosi E, Donati B, Galardini M, Brunetti S, Sagot M-F, Lió P, Crescenzi P, Fani R, Fondi M. 2015. MeDuSa: a multi-draft based scaffolder. *Bioinformatics* **31**:2443–2451 DOI 10.1093/bioinformatics/btv171.

Bosi E, Monk JM, Aziz RK, Fondi M, Nizet V, Palsson BØ. 2016. Comparative genome-scale modelling of *staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity. *Proceedings of the National Academy of Sciences of the United States of America* **113**:E3801–E3809 DOI 10.1073/pnas.1523199113.

Burts ML, Williams WA, Debord K, Missiakas DM. 2005. EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *staphylococcus aureus* infections. *Proceedings of the National Academy of Sciences of the United States of America* **102**:1169–1174 DOI 10.1073/pnas.0405620102.

Cao Z, Casabona MG, Kneuper H, Chalmers JD, Palmer T. 2016. The type VII secretion system of *Staphylococcus aureus* secretes a nuclease toxin that targets competitor bacteria. *Nature Microbiology* **2**:16183 DOI 10.1038/nmicrobiol.2016.183.

Carattoli A, Zankari E, García-Fernández A, Larsen MV, Lund O, Villa L, Aarestrup FM, Hasman H. 2014. In silicodetection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy* **58**:3895–3903 DOI 10.1128/aac.02412-14.

Carvalho SPD, Almeida JBD, Freitas LMD, Guimarães AMS, Nascimento NCD, Santos APD, Campos GB, Messick JB, Timenetsky J, Marques LM. 2019. Genomic profile of Brazilian methicillin-resistant *staphylococcus aureus* resembles clones dispersed worldwide. *Journal of Medical Microbiology* **68**:693–702 DOI 10.1099/jmm.0.000956.

Chaieb K, Mahdouani K, Bakhrouf A. 2005. Detection of icaA and icaD loci by polymerase chain reaction and biofilm formation by Staphylococcus epidermidis isolated from dialysate and needles in a dialysis unit. *Journal of Hospital Infection* **61**:225–230 DOI 10.1016/j.jhin.2005.05.014.

Chambers HF, Deleo FR. 2009. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nature Reviews Microbiology* **7**:629–641 DOI 10.1038/nrmicro2200.

Chan CX, Beiko RG, Ragan MA. 2011. Lateral transfer of genes and gene fragments in staphylococcus extends beyond mobile elements. *Journal of Bacteriology* **193**:3964–3977 DOI 10.1128/jb.01524-10.

Chen Q, Xie S, Lou X, Cheng S, Liu X, Zheng W, Zheng Z, Wang H. 2019. Biofilm formation and prevalence of adhesion genes among *Staphylococcus aureus* isolates from different food sources. *Microbiology Open* **9** DOI 10.1002/mbo3.946.

Chen T-R, Chiou C-S, Tsen H-Y. 2004. Use of novel PCR primers specific to the genes of staphylococcal enterotoxin G, H, I for the survey of *Staphylococcus aureus* strains isolated from food-poisoning cases and food samples in Taiwan. *International Journal of Food Microbiology* **92**:189–197 DOI 10.1016/j.ijfoodmicro.2003.10.002.

Cho S-H, Naber K, Hacker J, Ziebuhr W. 2002. Detection of the icaADBC gene cluster and biofilm formation in Staphylococcus epidermidis isolates from catheter-related urinary tract infections. *International Journal of Antimicrobial Agents* **19**:570–575 DOI 10.1016/s0924-8579(02)00101-2.

Chua KY, Howden BP, Jiang J-H, Stinear T, Peleg AY. 2014. Population genetics and the evolution of virulence in *staphylococcus aureus*. *Infection, Genetics and Evolution* **21**:554–562 DOI 10.1016/j.meegid.2013.04.026.

Cotter PD, Hill C. 2003. Surviving the acid test: responses of gram-positive bacteria to low pH. *Microbiology and Molecular Biology Reviews* **67**:429–453 DOI 10.1128/mmbr.67.3.429-453.2003.

Cramton SE, Gerke C, Schnell NF, Nichols WW, Friedrich G. 1999. The Intercellular Adhesion (ica) locus is present in *staphylococcus aureus* and is required for biofilm formation. *Infection and Immunity* **67**:5427–5433 DOI 10.1128/iai.67.10.5427-5433.1999.

Darling AE, Mau B, Perna NT. 2010. progressivemauve: multiple genome alignment with gene gain, loss and rearrangement. *PLOS ONE* **5**:e11147 DOI 10.1371/journal.pone.0011147.

Deleo FR, Diep BA, Otto M. 2009. Host defense and pathogenesis in *staphylococcus aureus* infections. *Infectious Disease Clinics of North America* **23**:17–34 DOI 10.1016/j.idc.2008.10.003.

Du J, Chen C, Ding B, Tu J, Qin Z, Parsons C, Salgado C, Cai Q, Song Y, Bao Q, Zhang L, Pan J, Wang L, Yu F. 2011. Molecular characterization and antimicrobial

susceptibility of nasal *staphylococcus aureus* isolates from a chinese medical college campus. *PLOS ONE* **6**:e27328 DOI 10.1371/journal.pone.0027328.

D'souza N, Rodrigues C, Mehta A. 2010. Molecular Characterization of Methicillin-Resistant Staphylococcus aureus with Emergence of Epidemic Clones of Sequence Type (ST) 22 and ST 772 in Mumbai, India. *Journal of Clinical Microbiology* **48**:1806–1811 DOI 10.1128/jcm.01867-09.

Eckhart L, Fischer H, Barken K, Tolker-Nielsen T, Tschachler E. 2007. DNase1L2 suppresses biofilm formation by Pseudomonas aeruginosa and *staphylococcus aureus*. *British Journal of Dermatology* **156**:1342–1345 DOI 10.1111/j.1365-2133.2007.07886.x.

Farran CE, Sekar A, Balakrishnan A, Shanmugam S, Arumugam P, Gopalswamy J. 2013. Prevalence of biofilm-producing Staphylococcus epidermidis in the healthy skin of individuals in Tamil Nadu, India. *Indian Journal of Medical Microbiology* **31(1)**:19–23 DOI 10.4103/0255-0857.108712.

Farris JS. 1972. Estimating phylogenetic trees from distance matrices. *The American Naturalist* **106**:645–668 DOI 10.1086/282802.

Feil EJ, Cooper JE, Grundmann H, Robinson DA, Enright MC, Berendt T, Peacock SJ, Smith JM, Murphy M, Spratt BG, Moore CE, Day NPJ. 2003. How clonal is *staphylococcus aureus*? *Journal of Bacteriology* **185**:3307–3316 DOI 10.1128/jb.185.11.3307-3316.2003.

Fitzgerald JR, Sturdevant DE, Mackie SM, Gill SR, Musser JM. 2001. Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. *Proceedings of the National Academy of Sciences of the United States of America* **98**:8821–8826 DOI 10.1073/pnas.161098098.

Foster JW. 2004. Escherichia coli acid resistance: tales of an amateur acidophile. *Nature Reviews Microbiology* **2**:898–907 DOI 10.1038/nrmicro1021.

Galata V, Fehlmann T, Backes C, Keller A. 2018. PLSDB: a resource of complete bacterial plasmids. *Nucleic Acids Research* **47**:D195–D202 DOI 10.1093/nar/gky1050.

Ghasemian A, Najar Peerayeh S, Bakhshi B, Mirzaee M. 2015. The microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) Genes among clinical isolates of *staphylococcus aureus* from hospitalized children. *Iranian Journal of Pathology* **10(4)**:258–264.

Ghatak S, Blom J, Das S, Sanjukta R, Puro K, Mawlong M, Shakuntala I, Sen A, Goesmann A, Kumar A, Ngachan SV. 2016. Pan-genome analysis of Aeromonas hydrophila, Aeromonas veronii and Aeromonas caviae indicates phylogenomic diversity and greater pathogenic potential for Aeromonas hydrophila. *Antonie van Leeuwenhoek* **109**:945–956 DOI 10.1007/s10482-016-0693-6.

Götz F, Bannerman T, Schleifer K-H. 2006. The genera staphylococcus and macrococcus. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, eds. *The prokaryotes*. New York: Springer, 5–75 DOI 10.1007/0-387-30744-3_1.

Goudarzi M, Mobarez AM, Najar-Peerayeh S, Mirzaee M. 2018. Prevalence of biofilm formation and vancomycin-resistant genes among *Enterococcus faecium* isolated

from clinical and environmental specimens in Lorestan hospitals. *Iranian Journal of Microbiology* **10(2)**:74–81.

Gould I. 2005. The clinical significance of methicillin-resistant *Staphylococcus aureus*. *Journal of Hospital Infection* **61**:277–282 DOI 10.1016/j.jhin.2005.06.014.

Grant JR, Stothard P. 2008. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Research* **36**:W181–W184 DOI 10.1093/nar/gkn179.

Gries CM, Sadykov MR, Bulock LL, Chaudhari SS, Thomas VC, Bose JL, Bayles KW. 2016. Potassium uptake modulates *staphylococcus aureus* metabolism. *mSphere* **1(3)**:e00125-16 DOI 10.1128/msphere.00125-16.

Grumann D, Nübel U, Bröker BM. 2014. *staphylococcus aureus* toxins –Their functions and genetics. *Infection, Genetics and Evolution* **21**:583–592 DOI 10.1016/j.meegid.2013.03.013.

Grundmann H, Schouls LM, Aanensen DM, Pluister GN, Tami A, Chlebowicz M, Glasner C, Sabat AJ, Weist K, Heuer O, Friedrich AW. 2014. The dynamic changes of dominant clones of *Staphylococcus aureus* causing bloodstream infections in the European region: Results of a second structured survey. *Eurosurveillance* **19**:es2014.19.49.20987 DOI 10.2807/1560-7917.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**:1072–1075 DOI 10.1093/bioinformatics/btt086.

Halsey CR, Lei S, Wax JK, Lehman MK, Nuxoll AS, Steinke L, Sadykov M, Powers R, Fey PD. 2017. amino acid catabolism in *Staphylococcus aureus* and the function of carbon catabolite repression. *mBio* **8(1)**:e01434–16 DOI 10.1128/mBio.01434-16.

Hayek N. 2013. Lateral transfer and GC content of bacterial resistant genes. *Frontiers in Microbiology* **4**:41 DOI 10.3389/fmicb.2013.00041.

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, Von Mering C, Bork P. 2018. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research* **47** DOI 10.1093/nar/gky1085.

Hughes AL, Friedman R. 2005. Nucleotide substitution and recombination at orthologous loci in *staphylococcus aureus*. *Journal of Bacteriology* **187**:2698–2704 DOI 10.1128/jb.187.8.2698-2704.2005.

Huseby MJ, Kruse AC, Digre J, Kohler PL, Vocke JA, Mann EE, Bayles KW, Bohach GA, Schlievert PM, Ohlendorf DH, Earhart CA. 2010. Beta toxin catalyzes formation of nucleoprotein matrix in staphylococcal biofilms. *Proceedings of the National Academy of Sciences of the United States of America* **107**:14407–14412 DOI 10.1073/pnas.0911032107.

Jansen W, Vanderbruggen J, Verhoef J, Fluit A. 2006. Bacterial resistance: a sensitive issue Complexity of the challenge and containment strategy in Europe. *Drug Resistance Updates* **9**:123–133 DOI 10.1016/j.drup.2006.06.002.

Jarraud S, Peyrat MA, Lim A, Tristan A, Bes M, Mougel C, Etienne J, Vandenesch F, Bonneville M, Lina G. 2001. egc, a highly prevalent operon of enterotoxin gene,

forms a putative nursery of superantigens in *staphylococcus aureus*. *The Journal of Immunology* **166**:669–677 DOI 10.4049/jimmunol.166.1.669.

Juhas M, Eberl L, Church GM. 2012. Essential genes as antimicrobial targets and cornerstones of synthetic biology. *Trends in Biotechnology* **30**:601–607 DOI 10.1016/j.tibtech.2012.08.002.

Kaya H, Hasman H, Larsen J, Stegger M, Johannesen TB, Allesøe RL, Lemvigh CK, Aarestrup FM, Lund O, Larsen AR. 2018. SCCmecFinder, a web-based tool for typing of staphylococcal cassette chromosome mec in *staphylococcus aureus* using whole-genome sequence data. *mSphere* **3(1)**:e00612-17 DOI 10.1128/msphere.00612-17.

Kennedy AD, Otto M, Braughton KR, Whitney AR, Chen L, Mathema B, Mediavilla JR, Byrne KA, Parkins LD, Tenover FC, Kreiswirth BN, Musser JM, Deleo FR. 2008. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: recent clonal expansion and diversification. *Proceedings of the National Academy of Sciences of the United States of America* **105**:1327–1332 DOI 10.1073/pnas.0710217105.

Kiem S, Oh WS, Peck KR, Lee NY, Lee J-Y, Song J-H, Hwang ES, Kim E-C, Cha CY, Choe K-W. 2004. Phase variation of biofilm formation in *staphylococcus aureus* by IS256Insertion and its impact on the capacity adhering to polyurethane surface. *Journal of Korean Medical Science* **19(6)**:779–782 DOI 10.3346/jkms.2004.19.6.779.

Kleinert F, Kallies R, Zweynert A, Bierbaum G. 2016. Draft genome sequences of three northern german epidemic *staphylococcus aureus* (ST247) strains containing multiple copies of IS 256: TABLE 1. *Genome Announcements* **4(5)**:e00936-16 DOI 10.1128/genomea.00936-16.

Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, Holden MTG, Dougan G, Bentley SD, Parkhill J, Peacock SJ. 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLOS Pathogens* **8**:e1002824 DOI 10.1371/journal.ppat.1002824.

Kwong J, Mccallum N, Sintchenko V, Howden B. 2015. Whole genome sequencing in clinical and public health microbiology. *Pathology* **47**:199–210 DOI 10.1097/pat.0000000000000235.

Labreck PT, Rice GK, Paskey AC, Elassal EM, Cer RZ, Law NN, Schlett CD, Bennett JW, Millar EV, Ellis MW, Hamilton T, Bishop-Lilly KA, Merrell DS. 2018. Conjugative transfer of a novel staphylococcal plasmid encoding the biocide resistance gene, qacA. *Frontiers in Microbiology* **9**:2664 DOI 10.3389/fmicb.2018.02664.

Ladhani S. 2001. Recent developments in staphylococcal scalded skin syndrome. *Clinical Microbiology and Infection* **7**:301–307 DOI 10.1046/j.1198-743x.2001.00258.x.

Laing C, Buchanan C, Taboada EN, Zhang Y, Kropinski A, Villegas A, Thomas JE, Gannon VP. 2010. Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions. *BMC Bioinformatics* **11**:461 DOI 10.1186/1471-2105-11-461.

Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Ponten T, Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus

sequence typing of total-genome-sequenced bacteria. *Journal of Clinical Microbiology* **50**:1355–1361 DOI 10.1128/jcm.06094-11.

Larsen MV, Cosentino S, Lukjancenko O, Saputra D, Rasmussen S, Hasman H, Sicheritz-Ponten T, Aarestrup FM, Ussery DW, Lund O. 2014. Benchmarking of methods for genomic taxonomy. *Journal of Clinical Microbiology* **52**:1529–1539 DOI 10.1128/jcm.02981-13.

Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program: table 1. *Molecular Biology and Evolution* **32**:2798–2800 DOI 10.1093/molbev/msv150.

Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, Gascuel O. 2019. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Research* **47**:W256–W259 DOI 10.1093/nar/gkz303.

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* **47** DOI 10.1093/nar/gkz239.

Lim S, Lee D-H, Kwak W, Shin H, Ku H-J, Lee J-E, Lee GE, Kim H, Choi S-H, Ryu S, Lee J-H. 2015. Comparative genomic analysis of *Staphylococcus aureus* FORC_001 and *S. aureus* MRSA252 reveals the characteristics of antibiotic resistance and virulence factors for human infection. *Journal of Microbiology and Biotechnology* **25**:98–108 DOI 10.4014/jmb.1410.10005.

Lina G, Piemont Y, Godail-Gamot F, Bes M, Peter M-O, Gauduchon V, Vandenesch F, Etienne J. 1999. Involvement of panton-valentine leukocidin–producing *staphylococcus aureus* in primary skin infections and pneumonia. *Clinical Infectious Diseases* **29**:1128–1132 DOI 10.1086/313461.

Lindsay JA. 2010. Genomic variation and evolution of *Staphylococcus aureus*. *International Journal of Medical Microbiology* **300**:98–103 DOI 10.1016/j.ijmm.2009.08.013.

Lindsay JA, Moore CE, Day NP, Peacock SJ, Witney AA, Stabler RA, Husain SE, Butcher PD, Hinds J. 2006. Microarrays reveal that each of the ten dominant lineages of *staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. *Journal of Bacteriology* **188**:669–676 DOI 10.1128/jb.188.2.669-676.2006.

Lister JL, Horswill AR. 2014. *staphylococcus aureus* biofilms: recent developments in biofilm dispersal. *Frontiers in Cellular and Infection Microbiology* **4**:178 DOI 10.3389/fcimb.2014.00178.

Liu B, Zheng D, Jin Q, Chen L, Yang J. 2018. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Research* **47**:D687–D692 DOI 10.1093/nar/gky1080.

Maddux MS. 1991. Effects of ß-lactamase-mediated antimicrobial resistance: the role of ß-lactamase inhibitors. *Journal of Human Pharmacology and Drug Therapy* **11**:40S–50S.

Maurer LM, Yohannes E, Bondurant SS, Radmacher M, Slonczewski JL. 2005. pH regulates genes for flagellar motility, catabolism, and oxidative stress in escherichia coli K-12. *Journal of Bacteriology* **187**:304–319 DOI 10.1128/jb.187.1.304-319.2005.

Mazmanian SK. 1999. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**:760–763 DOI 10.1126/science.285.5428.760.

McCarthy AJ, Lindsay JA. 2012. The distribution of plasmids that carry virulence and resistance genes in *staphylococcus aureus* is lineage associated. *BMC Microbiology* **12**:104 DOI 10.1186/1471-2180-12-104.

McCarthy AJ, Loeffler A, Witney AA, Gould KA, Lloyd DH, Lindsay JA. 2014. Extensive horizontal gene transfer during *staphylococcus aureus* co-colonization in vivo. *Genome Biology and Evolution* **6**:2697–2708 DOI 10.1093/gbe/evu214.

McNair K, Bailey BA, Edwards RA. 2012. PHACTS, a computational approach to classifying the lifestyle of phages. *Bioinformatics* **28**:614–618 DOI 10.1093/bioinformatics/bts014.

Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R. 2005. The microbial pan-genome. *Current Opinion in Genetic Development* **15**:589–594 DOI 10.1016/j.gde.2005.09.006.

Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**:60 DOI 10.1186/1471-2105-14-60.

Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications* **10**:2182 DOI 10.1038/s41467-019-10210-3.

Melish ME, Glasgow LA. 1970. The staphylococcal scalded-skin syndrome. *New England Journal of Medicine* **282**:1114–1119 DOI 10.1056/nejm197005142822002.

Mistry H, Sharma P, Mahato S, Saravanan R, Kumar PA, Bhandari V. 2016. Prevalence and characterization of oxacillin susceptible mecA-positive clinical isolates of *Staphylococcus aureus* causing bovine mastitis in India. *PLOS ONE* **11**:e0162256 DOI 10.1371/journal.pone.0162256.

Monecke S, Coombs G, Shore AC, Coleman DC, Akpaka P, Borg M, Chow H, Ip M, Jatzwauk L, Jonas D, Kadlec K, Kearns A, Laurent F, O'brien FG, Pearson J, Ruppelt A, Schwarz S, Scicluna E, Slickers P, Tan H-L, Weber S, Ehricht R. 2011. A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *staphylococcus aureus*. *PLOS ONE* **6**:e17936 DOI 10.1371/journal.pone.0017936.

Morens DM, Fauci AS. 2013. Emerging infectious diseases: threats to human health and global stability. *PLOS Pathogens* **9**:e1003467 DOI 10.1371/journal.ppat.1003467.

Møretrø T, Hermansen L, Holck AL, Sidhu MS, Rudi K, Langsrud S. 2003. Biofilm formation and the presence of the intercellular adhesion locus ica among staphylococci from food and food processing environments. *Applied and Environmental Microbiology* **69**:5648–5655 DOI 10.1128/aem.69.9.5648-5655.2003.

Mottola C, Matias CS, Mendes JJ, Melo-Cristino J, Tavares L, Cavaco-Silva P, Oliveira M. 2016. Susceptibility patterns of *staphylococcus aureus* biofilms in diabetic foot infections. *BMC Microbiology* **16** DOI 10.1186/s12866-016-0737-0.

Naorem RS, Urban P, Goswami G, Fekete C. 2020. Characterization of methicillin-resistant *Staphylococcus aureus* through genomics approach. *3Biotech* **10**:401 DOI 10.1007/s13205-020-02387-y.

Nasr RA, Abushady HM, Hussein HS. 2012. Biofilm formation and presence of icaAD gene in clinical isolates of staphylococci. *Egyptian Journal of Medical Human Genetics* **13**:269–274 DOI 10.1016/j.ejmhg.2012.04.007.

Nübel U, Dordel J, Kurt K, Strommenger B, Westh H, Shukla SK, Žemličková H, Leblois R, Wirth T, Jombart T, Balloux F, Witte W. 2010. A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *staphylococcus aureus*. *PLOS Pathogens* **6**:119 DOI 10.1371/journal.ppat.1000855.

Nubel U, Roumagnac P, Feldkamp M, Song J-H, Ko KS, Huang Y-C, Coombs G, Ip M, Westh H, Skov R, Struelens MJ, Goering RV, Strommenger B, Weller A, Witte W, Achtman M. 2008. Frequent emergence and limited geographic dispersal of methicillin-resistant *staphylococcus aureus*. *Proceedings of the National Academy of Sciences of the United States of America* **105**:14130–14135 DOI 10.1073/pnas.0804178105.

Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, Mclean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *Journal of Computational Biology* **20**:714–737 DOI 10.1089/cmb.2013.0084.

O'Gara JP. 2007. ica and beyond: biofilm mechanisms and regulation in Staphylococcus epidermidis and *staphylococcus aureus*. *FEMS Microbiology Letters* **270**:179–188 DOI 10.1111/j.1574-6968.2007.00688.x.

Otto M. 2014. *Staphylococcus aureus* toxins. *Current Opinion in Microbiology* **17**:32–37 DOI 10.1016/j.mib.2013.11.004.

Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2013. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Research* **42**:D206–D214 DOI 10.1093/nar/gkt1226.

Ozer EA. 2018. ClustAGE: a tool for clustering and distribution analysis of bacterial accessory genomic elements. *BMC Bioinformatics* **19**:150 DOI 10.1186/s12859-018-2154-x.

Payne DE, Boles BR. 2015. Emerging interactions between matrix components during biofilm development. *Current Genetics* **62**:137–141 DOI 10.1007/s00294-015-0527-5.

Periasamy S, Joo H-S, Duong AC, Bach T-HL, Tan VY, Chatterjee SS, Cheung GYC, Otto M. 2012. How *staphylococcus aureus* biofilms develop their characteristic structure. *Proceedings of the National Academy of Sciences of the United States of America* **109**:1281–1286 DOI 10.1073/pnas.1115006109.

Pugazhendhi A, Michael D, Prakash D, Krishnamurthy PP, Shanmuganathan R, Al-Dhabi NA, Duraipandiyan V, Arasu MV, Kaliannan T. 2020. Antibiogram and plasmid profiling of beta-lactamase producing multi drug resistant *staphylococcus aureus* isolated from poultry litter. *Journal of King Saud University - Science* 2723–2727 DOI 10.1016/j.jksus.2020.06.007.

Reffuveille F, Josse J, Vallé Q, Mongaret C, Gangloff SC. 2017. *staphylococcus aureus* biofilms and their impact on the medical field. In: Enany S, Crotty Alexander LE, eds. *The rise of virulence and antibiotic resistance in staphylococcus aureus.* London: IntechOpen, 187–214 DOI 10.5772/66380.

Richter M, Rosselló-Móra R, Glöckner FO, Peplies J. 2015. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **32**:929–931 DOI 10.1093/bioinformatics/btv681.

Schmidt H, Hensel M. 2004. Pathogenicity islands in bacterial pathogenesis. *Clinical Microbiology Reviews* **17(1)**:14–56 DOI 10.1128/cmr.17.1.14-56.2004.

Sharma S, Chaudhry V, Kumar S, Patil PB. 2018. Phylogenomic based comparative studies on indian and american commensal staphylococcus epidermidis isolates. *Frontiers in Microbiology* **9**:333 DOI 10.3389/fmicb.2018.00333.

Shin K, Yun Y, Yi S, Lee HG, Cho J-C, Suh K-D, Lee J, Park J. 2013. Biofilm-forming ability of *Staphylococcus aureus* strains isolated from human skin. *Journal of Dermatological Science* **71**:130–137 DOI 10.1016/j.jdermsci.2013.04.004.

Shopsin B, Mathema B, Alcabes P, Said-Salim B, Lina G, Matsuka A, Martinez J, Kreiswirth BN. 2003. Prevalence of agr specificity groups among *staphylococcus aureus* strains colonizing children and their guardians. *Journal of Clinical Microbiology* **41**:456–459 DOI 10.1128/jcm.41.1.456-459.2003.

Shukla SK, Pantrang M, Stahl B, Briska AM, Stemper ME, Wagner TK, Zentz EB, Callister SM, Lovrich SD, Henkhaus JK, Dykes CW. 2012. Comparative whole-genome mapping to determine *staphylococcus aureus* genome size, virulence motifs, and clonality. *Journal of Clinical Microbiology* **50**:3526–3533 DOI 10.1128/jcm.01168-12.

Bhutia KO, Singh T, Adhikari L, Biswas S. 2015. Molecular characterization of community- & hospital-acquired methicillin-resistant & methicillin-sensitive *staphylococcus aureus* isolates in Sikkim. *Indian Journal of Medical Research* **142**:330–335 DOI 10.4103/0971-5916.166600.

Sugimoto S, Iwamoto T, Takada K, Okuda K-I, Tajima A, Iwase T, Mizunoe Y. 2013. Staphylococcus epidermidis Esp degrades specific proteins associated with *staphylococcus aureus* biofilm formation and host-pathogen interaction. *Journal of Bacteriology* **195**:1645–1655 DOI 10.1128/jb.01672-12.

Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* **27(7)**:1009–1010 DOI 10.1093/bioinformatics/btr039.

Tchoupa AK, Watkins KE, Jones RA, Kuroki A, Alam MT, Perrier S, Chen Y, Unnikrishnan M. 2019. The type VII secretion system protects *Staphylococcus aureus* against antimicrobial host fatty acids. *Scientific Reports* **10**:14838 DOI 10.1038/s41598-020-71653-z.

Tettelin H, Riley D, Cattuto C, Medini D. 2008. Comparative genomics: the bacterial pan-genome. *Current Opinion in Microbiology* **11**:472–477 DOI 10.1016/j.mib.2008.09.006.

Valenzuela M, Cerda O, Toledo H. 2003. Overview on chemotaxis and acid resistance in Helicobacter pylori. *Biological Research* **36(3–4)**:1399–1404 DOI 10.4067/s0716-97602003000300014.

Van Heel AJ, De Jong A, Song C, Viel JH, Kok J, Kuipers OP. 2018. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Research* **46** DOI 10.1093/nar/gky383.

Vandenesch F, Naimi T, Enright MC, Lina G, Nimmo GR, Heffernan H, Liassine N, Bes M, Greenland T, Reverdy M-E, Etienne J. 2003. Community-acquired methicillin-resistant *staphylococcus aureus* carrying panton-valentine leuco-cidin genes: worldwide emergence. *Emerging Infectious Diseases* **9**:978–984 DOI 10.3201/eid0908.030089.

Vogel V, Falquet L, Calderon-Copete SP, Basset P, Blanc DS. 2012. Short term evolution of a highly transmissible methicillin-resistant *staphylococcus aureus* clone (ST228) in a tertiary care hospital. *PLOS ONE* **7**(**6**):e38969 DOI 10.1371/journal.pone.0038969.

Wamel WJBV, Rooijakkers SHM, Ruyken M, Kessel KPMV, Strijp JAGV. 2006. The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of *staphylococcus aureus* are located on β-hemolysin-converting bacteriophages. *Journal of Bacteriology* **188**:1310–1315 DOI 10.1128/jb.188.4.1310-1315.2006.

Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJ, Yoo HS, Zhang C, Zhang Y, Sobral BW. 2013. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research* **42**:D581–D591 DOI 10.1093/nar/gkt1099.

Zhang K, Mcclure J-A, Conly JM. 2012. Enhanced multiplex PCR assay for typing of staphylococcal cassette chromosome mec types I to V in methicillin-resistant *staphylococcus aureus. Molecular and Cellular Probes* **26**:218–221 DOI 10.1016/j.mcp.2012.04.002.

Zhang Y, Xu D, Shi L, Cai R, Li C, Yan H. 2018. Association between agr type, virulence factors, biofilm formation and antibiotic resistance of *staphylococcus aureus* isolates from pork production. *Frontiers in Microbiology* **9**:1876 DOI 10.3389/fmicb.2018.01876.